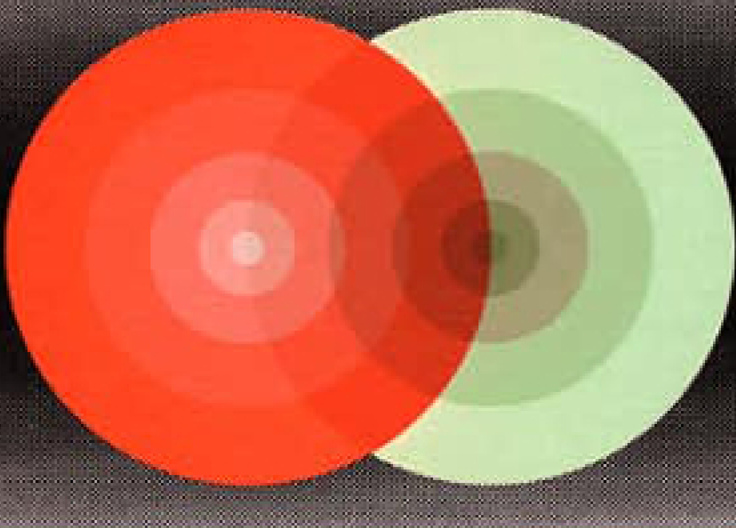
Black Hole Physics

Basic Concepts and New Developments

by

Valeri P. Frolov and Igor D. Novikov

Kluwer Academic Publishers



Fundamental Theories of Physics

Black Hole Physics:

Basic Concepts and New Developments

Valeri P. Frolov   Igor D. Novikov

November 25, 1997

To our parents.

Authors

Contents

Preface

It is not an exaggeration to say that one of the most exciting predictions of Einstein's theory of gravitation is that there may exist "black holes": putative objects whose gravitational fields are so strong that no physical bodies or signals can break free of their pull and escape. The proof that black holes do exist, and an analysis of their properties, would have a significance going far beyond astrophysics. Indeed, what is involved is not just the discovery of yet another even if extremely remarkable, astrophysical object, but a test of the correctness of our understanding of the properties of space and time in extremely strong gravitational fields.

Theoretical research into the properties of black holes, and into the possible corollaries of the hypothesis that they exist, has been carried out with special vigor since the beginning of the 1970's. In addition to those specific features of black holes that are important for the interpretation of their possible astrophysical manifestations, the theory has revealed a number of unexpected characteristics of physical interactions involving black holes. By the middle of the 1980's a fairly detailed understanding had been achieved of the properties of the black holes, their possible astrophysical manifestations, and the specifics of the various physical processes involved. Even though a completely reliable detection of a black hole had not yet been made at that time, several objects among those scrutinized by astrophysicists were considered as strong candidates to be confirmed as being black holes. Furthermore, profound links were found between black hole theory and such seemingly very distant fields as thermodynamics, information theory, and quantum theory. The branch of physics that is now referred to as black hole physics was born and actually took shape as a full-blooded scientific discipline during the past two decades at the junction of the theory of gravitation, astrophysics, and classical and quantum field theories.

In 1986 we published in Russian a book "Physics of Black Holes" devoted to this relatively young and rapidly developing branch of physics. In 1989 the book was translated and published in English. During the years that have passed since then, the physics of black holes enjoyed a period of rapid growth. First of all, great progress was achieved in the observational astrophysics of black holes. Ten years ago there were only a few black hole candidates in binary systems in which the companion was a normal star. In spite of enormous efforts, no unequivocal signature of black holes had been found in these systems. Today after great improvements in the quantity and quality of the observational material, our confidence that we observe the manifestation of black holes in at least a few binaries is almost a hundred per cent. Moreover, impressive progress in optical, radio, and X-ray astronomy greatly bolstered the evidence for supermassive black holes (up to several billion solar mass) in the centers of galaxies. During the same period of time, great progress was made also in the development of the theoretical and mathematical aspects of black hole physics. Black hole collisions might be one of the most powerful sources of gravitational radiation in the Universe. The gravitational wave observations that might become possible after construction of the LIGO and other gravitational antennas of the new generation required developments in numerical relativity and analytical methods for the analysis of the gravitational radiation generated by black holes.

More than five years ago we were requested to prepare a second edition of the book to reflect the enormous progress in our knowledge concerning black holes. We started work on this project, and quite soon we realized that it is virtually impossible to keep the book unchanged and restrict ourselves to tiny "cosmetic" improvements. As a result, in order to update the book, we practically rewrote it though we preserved its original structure and certainly used much of the old material.

This monograph is the result of our attempts to understand the modern status of black hole physics. This new book is twice longer than the old one; it contains four new chapters. We also added a lot of new material in the form of appendices which cover more mathematical subject.

This book is written to introduce the reader to the physics of black holes and the methods employed in it, and to review the main results of this branch of physics. But attention is focused primarily on questions that were answered relatively recently, and thus could not be adequately reflected in earlier textbooks and reviews. Those aspects that are relatively familiar are presented concisely (but as clearly as possible).

We have tried to make the representation lucid not only for a specialist, but also for a broad spectrum of physicists and astrophysicists who do not have a special knowledge of black hole physics. An attempt is made to explain, first and foremost, the physical essentials of the phenomena, and only after this do we pass on to the mathematical means of describing them. These objectives determined both the spectrum of selected topics and the style of presentation. A conscious attempt has been made to avoid excessive rigor in the formulations and proofs of theorems on black holes. Quite often, only the principal idea is given instead of a complete proof; then the successive stages of the proof are outlined, and references to the original papers are supplied where the reader will find the required details. This approach was chosen not only because the excellent monographs of Penrose (1968), Hawking and Ellis (1973), Chandrasekhar (1983), and Thorne, Price and Macdonald (1986) cover most of the material omitted in this book, but also because we are of the opinion that excessive rigor stands in the way of understanding the physical ideas that are at the foundation of the specific properties of black holes.

The material in the book is partly based on the courses of lectures which the authors have presented during a number of years at the University of Copenhagen

and University of Alberta. One of the authors (V.F.) is grateful to the Killam trust for its financial support during the work on this book. The authors also would like to thank the Danish National Research Foundation for the support of this project through its establishment of the Theoretical Astrophysics Center, and the Natural Sciences and Engineering Research Council of Canada for its financial support.

While working on this monograph, we enjoyed the unfailing support of our colleagues from the P.N. Lebedev Physical Institute (Moscow), University Observatory, NORDITA, and Theoretical Astrophysics Center (Copenhagen), and the University of Alberta (Edmonton). We wish to express our sincere thanks to Roger Blanford, Bernard Carr, Vitaly Ginzburg, David Kirzhnitz, Vladimir Lipunov, Draza Markovic, Michael Nowak, Don Page, Christopher Pethick, Martin Rees, Kip Thorne, and John Wheeler for helpful discussions of the problems of black holes. We are much indebted to Werner Israel for numerous stimulating discussions and for his enormous work in reading the draft of the book. His advice concerning the contents and style were very important for us in our work. We want also to mention particularly Nils Andersson, who is a co-author of Chapter 4 and Section 9.9. We also appreciate the help of Vasily Beskin in preparing the new version of Chapter 8. Their help was invaluable. Additional thanks go to Patrick Sutton for his help with proofreading.

This book would never have appeared if we had not received support and encouragements from our friends and colleagues. We owe the deepest debt of gratitude to our wives Raya Frolova and Eleonora Kotok. They were not only patient, but contributed countless hours helping us in the preparation of the electronic version of the book. We would also like to thank Andrei Frolov. His skills as the computer guru helped us to overcome numerous problems which are inevitable in any project of such a scale.

Copenhagen and Edmonton

October 1997

Valeri Frolov and Igor Novikov

Part I

Basic Concepts

Chapter 1

Introduction: Brief History of Black Hole Physics

A *black hole* is, by definition, a region in spacetime in which the gravitational field is so strong that it precludes even light from escaping to infinity.

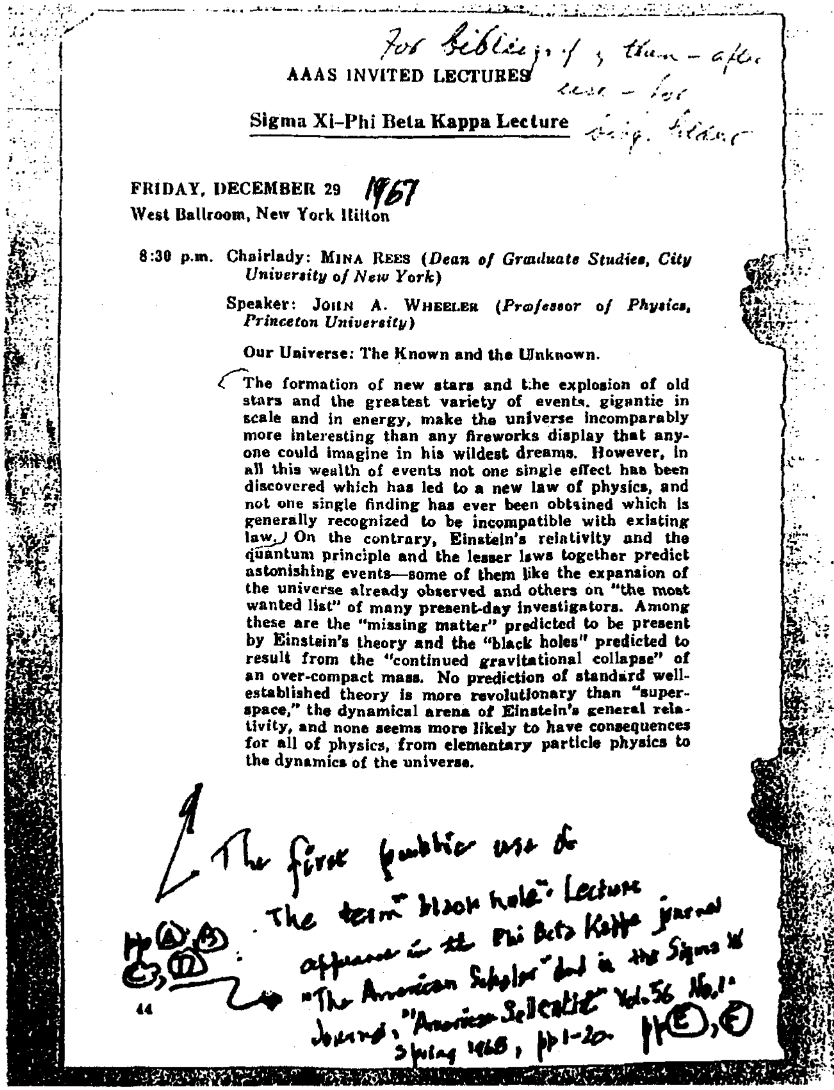
A black hole is formed when a body of mass M contracts to a size less than the so-called *gravitational radius* $r_g = 2GM/c^2$ (G is Newton's gravitational constant, and c is the speed of light). The velocity required to leave the boundary of the black hole and move away to infinity (the escape velocity) equals the speed of light. One easily concludes then that neither signals nor particles can escape from the region inside the black hole since the speed of light is the limiting propagation velocity for physical signals. This conclusion is of absolute nature in Einstein's theory of gravitation because the gravitational interaction is universal. The role of gravitational charge is played by mass whose value is proportional to the total energy of the system. Hence, all objects with nonzero energy participate in the gravitational interaction.

Einstein's theory of gravitation, alias *general relativity*, is employed to the full in the description of black holes. It may appear at first glance that one cannot hope to obtain an acceptably complete description of black holes, owing to the complexity of the equations involved and, among other factors, their essential nonlinearity. Fortunately, it was found that shortly after its formation, any black hole becomes stationary, and its field is determined in a unique manner by a small number of parameters; namely, its mass and angular momentum, and its electric charge (if it is charged). The physical reason for this striking property of black holes is the fact that in the extremely strong field of a black hole in empty space, only very special types of configuration of physical fields (including the gravitational field) can be stationary.

Since signals cannot escape from a black hole, while physical objects and radiation can fall into it, the surface bounding the black hole in spacetime (called the *event horizon*) is a lightlike surface. The birth of a black hole signifies the formation of a non-trivial causal structure in spacetime. As a result of these specific features, new methods had to be developed to analyze the interaction of black holes with physical fields and matter, and with other black holes.

The term "black hole" was introduced by Wheeler in 1967 although the theoretical study of these object has quite a long history. The very possibility of the existence of such objects was first discussed by Michell and Laplace within the framework of the Newtonian theory at the end of the 18th century [see Barrow and Silk (1983), Israel (1987), Novikov (1990)]. In fact, in general relativity, the problem arose within a year after the theory had been developed, i.e., after Schwarzschild (1916) obtained the first exact (spherically symmetric) solution of Einstein's equations in vacuum. In addition to a singularity at the center of symmetry (at $r = 0$), this solution had an additional singularity on the gravitational-radius surface (at $r = r_g$). More than a third of a century elapsed before a profound understanding of the structure of spacetime in strong gravitational fields was achieved as a result of analyses of the "unexpected" features of the Schwarzschild solution by Flamm (1916), Weyl (1917), Eddington (1924), Lemaître (1933), Einstein and Rosen (1935), and the complete solution of the formulated problem was obtained [Synge (1950), Finkelstein (1958), Fronsdal (1959), Kruskal (1960), Szekeres (1960), Novikov (1963, 1964a)]. The length of this interval may have been influenced by the general belief that nature could not admit a body whose size would be comparable to its gravitational radius; this viewpoint was shared by the creator of general relativity himself [see e.g., Israel (1987) and references therein]. Some interest in the properties of very compact gravitational systems was stimulated in the thirties after Chandrasekhar's (1931) work on white dwarfs and the works of Landau (1932), Baade and Zwicky (1934), and Oppenheimer and Volkoff (1939) who showed that neutron stars are possible, with a radius only a few times that of the gravitational radius. The gravitational collapse of a massive star which produces a black hole was first described by Oppenheimer and Snyder (1939).

The next period began in the middle sixties when intensive theoretical studies were initiated on the general properties of black holes and their classical interactions, after the work of Synge (1950), Kruskal (1960) and others who obtained the complete solution for the Schwarzschild problem, and of Kerr (1963) who discovered a solution describing the gravitational field of a rotating black hole. Before this period specialists had considered black holes as dead objects, ultimate final stage of the evolution of massive stars, and probably of more massive objects. The names "frozen" or "collapsed" stars used by specialists until the end of sixties for the description of these objects reflected such an attitude. This point of view proved to be rather restrictive when processes in the close vicinity of these objects became the focus of interest. Moreover, it prevents from the outset even formulating any ideas concerning the physical processes inside them. In the middle of the sixties a new approach (paradigm) gained ascendancy in the community of theorists working in general relativity (but not among astronomers who mainly were very far from the problems of black holes, and even discussions of these problems were not welcomed in the "decent society"!). This new point of view was a historical development of the



AAAS INVITED LECTURES

Sigma Xi–Phi Beta Kappa Lecture

FRIDAY, DECEMBER 29 1967
West Ballroom, New York Hilton

8:30 p.m. Chairlady: MINA REES (*Dean of Graduate Studies, City University of New York*)

Speaker: JOHN A. WHEELER (*Professor of Physics, Princeton University*)

Our Universe: The Known and the Unknown.

The formation of new stars and the explosion of old stars and the greatest variety of events, gigantic in scale and in energy, make the universe incomparably more interesting than any fireworks display that anyone could imagine in his wildest dreams. However, in all this wealth of events not one single effect has been discovered which has led to a new law of physics, and not one single finding has ever been obtained which is generally recognized to be incompatible with existing law. On the contrary, Einstein's relativity and the quantum principle and the lesser laws together predict astonishing events—some of them like the expansion of the universe already observed and others on "the most wanted list" of many present-day investigators. Among these are the "missing matter" predicted to be present by Einstein's theory and the "black holes" predicted to result from the "continued gravitational collapse" of an over-compact mass. No prediction of standard well-established theory is more revolutionary than "superspace," the dynamical arena of Einstein's general relativity, and none seems more likely to have consequences for all of physics, from elementary particle physics to the dynamics of the universe.

The first public use of the term "black hole". Lecture appeared in the Phi Beta Kappa journal "The American Scholar" and in the Sigma Xi journal, "American Scientist" Vol.56 No.1 Spring 1968, pp1-20.

44

Figure 1.1: The first public use of the term "black hole". Lecture (given by J.A.Wheeler) appeared in the Phi Beta Kapper journal "The American Scholar" (Vol.37, No.2, Spring 1968, pp.248) and in the Sigma Xi journal, "American Scientist" Vol.56 No.1 Spring 1968, pp. 1-20. This page was sent to the authors by John Wheeler with his handwritten notes.

route initiated by the work of Oppenheimer and Snyder (1939), where it was shown that an observer on the surface of a collapsing star sees no "freezing" at all, but can register events both outside and inside the gravitational radius. This point of view implied that an object formed after the gravitational collapse could be considered in some sense as a "hole" in spacetime.

Just at this time Wheeler (1968) coined the name "black hole". When in 1992 we started working on this book in Copenhagen, John Wheeler visited us and in very lively fashion recalled his lecture in late 1967 when he used this name for the first time (see Figure 1.1). Soon after that this name was adopted enthusiastically by everybody. We believe that readers would agree that this graphic expression reflects very picturesquely the remarkable properties of the object.

The now-classic theorems stating that "black holes have no hair" (that is, no external individual attributes except mass, angular momentum, and charge), that a black hole contains a singularity inside it, and that the black hole area cannot decrease were proved during this period. These and other results made it possible to construct a qualitative picture of the formation of a black hole, its possible further evolution, and its interaction with matter and classical physical fields. Many of these results were summarized in the well-known monographs of Zel'dovich and Novikov (1967b, 1971a,b), Misner, Thorne, and Wheeler (1973), Hawking and Ellis (1973), Thorne, Price, and Macdonald (1986), and Novikov and Frolov (1989).

After pulsars (neutron stars) were discovered at the end of the sixties, astrophysicists had to examine the prospects for the observational detection of black holes. The analysis of the accretion of matter onto isolated black holes and onto black holes in binary systems predicted that accreting black holes may constitute powerful sources of X-rays [Novikov and Zel'dovich (1966), Shklovsky (1967b), Burbidge (1972)]. The progress of X-ray astronomy and the studies using X-ray satellites that began in the 1970's led to the discovery of a number of X-ray sources. The hypothesis was proposed that some of them are black holes in binary stellar systems.

More than 25 years of constant study of these objects provided confirmation of the initial hypothesis. Now at the end of nineties we are sure that black holes with stellar masses do exist in a number of binaries in our Galaxy [Thorne (1994b)]. There is also good reason to believe that the nuclei of active galaxies (and possibly of any galaxy) and quasars contain massive or supermassive black holes [see Blandford and Thorne (1979), Rees (1982), Begelman and Rees (1996)]. Two recent discoveries – one by astronomers using the Hubble Telescope [Ford *et al.* (1994), Harms *et al.* (1994)], the other by radioastronomers [Miyoshi *et al.* (1995)] – gave clear evidence for huge black holes in the centers of galaxies. In both cases observations revealed disks of gas orbiting the central objects, and it is possible to give robust arguments that these objects can be nothing but supermassive black holes.

The discussion of the possible observational aspects of black hole study drew considerable attention to the problem of the motion of particles and propagation of physical fields in the spacetime of stationary black holes. This problem, which is predominantly mathematical and involves the integration of the equations of geodesics and the solution (by expansion in eigenfunctions) of the wave equations in the Kerr metric, has now been completely solved. Numerous relevant results are summarized in the monograph by Chandrasekhar (1983).

The sensational "news" of the possible discovery of a black hole in an X-ray binary (*Cygnus X-1*) had scarcely died down when a new unexpected result obtained by Hawking (1974, 1975) again focussed physicists' attention on black holes. It was found that as a result of the instability of the vacuum in the strong gravitational field of a black hole, these objects are sources of quantum radiation. The most intriguing property of this radiation is that it has a thermal spectrum. In other words, if one neglects the scattering of the radiation by the external gravitational field, a black hole radiates like a heated black body. If the black hole mass is small, it decays over a time shorter than the age of the Universe. Such small black holes, now called primordial black holes, may have been formed only at a very early stage of the Universe's evolution [Zel'dovich and Novikov (1966, 1967b), Hawking (1971)]. In principle, the discovery of primordial black holes or of their decay products would supply valuable information on the physical processes occurring in the Universe at that period.

Hawking's discovery stimulated a large number of papers which analyzed specific features of quantum effects in black holes. In addition to a detailed description of the effects due to the creation of real particles escaping to infinity, substantial progress has been achieved in the understanding of the effect of vacuum polarization in the vicinity of a black hole. This effect is important for the construction of a complete quantum description of an "evaporating" black hole.[1]

It is quite interesting that not so many (say 15–20) years ago, black holes were considered as highly exotic objects, and the general attitude in the wider physical and astrophysical community (i.e., among the scientists who were not working on this subject) to these objects was quite cautious. Now the situation has changed drastically. It happened both because of new astrophysical data and because of the development of the theory.

In binary systems and in galactic centers, accretion of gas onto a black hole generates radiation of light or X-rays. The efficiency of this process is so high that the accretion of matter onto a black hole is one of the most powerful energy sources in the Universe. That is why black holes recently became the favored hypothesis for trying to explain processes with huge energy release from compact regions of space. This is what gives black holes their current importance in astrophysics.

[1]There were a number of review articles written in the 1970's and early 1980's which had summarized main results obtained during this "heroic" period of black hole physics. These are references to some of them: Penrose (1972), Carter (1973a, 1976), Sexl (1975), Israel (1983), Markov (1970, 1973), De Witt (1975), Sciama (1976), Dymnikova (1986), Bekenstein (1980), Ruffini (1979), Frolov (1976b, 1978b, 1983b). See also a remarkable review article by Israel (1987) which describes the history of evolution of the black hole idea.

More recently another aspect of black hole physics became very important for astrophysical applications. The collision of a black hole with a neutron star or coalescence of a pair of black holes in binary systems is a powerful source of gravitational radiation which might be strong enough to reach the Earth and be observed in a new generation of gravitational wave experiments (LIGO, LISA, and others). The detection of gravitational waves from these sources requires a detailed description of the gravitational field of a black hole during the collision. In principle, gravitational astronomy opens remarkable opportunities to test gravitational field theory in the limit of very strong gravitational fields. In order to be able to do this, besides the construction of the gravitational antennas, it is also necessary to obtain the solution of the gravitational equations describing this type of situation. Until now there exist no analytical tools which allow this to be done. Under these conditions one of the important problems is the numerical study of colliding black holes.

Besides its direct astrophysical application, the physics of black holes has a more general importance. The existence of black holes introduces into physics a new concept which can be called the concept of *invisibility*. In the presence of a black hole, physically important classes of observers, at rest or moving in the black hole exterior (external observers), would agree that there exists a spacetime region which is in principle unobservable in their reference frame. Since exactly these frames are used in astrophysics, we have a situation which never occurred before. Is there any sense in discussing what happens inside a black hole if there is no way to compare our predictions with observations outside a black hole or to transfer our knowledge to the external observer if we decide to dive into a black hole? Perhaps the general answer to this almost philosophical question lies somewhere outside physics itself.[2]

The very existence of "invisible" regions ("holes") in spacetime has a number of important physical consequences. One of them is the thermal nature of the Hawking radiation. In the process of black hole formation the information concerning the state of quantum fields inside the horizon is lost for an external observer. Even if at the beginning (before the collapse) a system was in a pure quantum state, after the black hole formation the state outside the horizon is mixed, and it is described by a density matrix. According to Hawking, this density matrix is thermal. In other words, the combination of gravity and quantum mechanics in the presence of black holes requires for its consistency the introduction of thermodynamical methods. The statistical mechanical foundations of the thermodynamics of black holes still remain one of the most intriguing problems in theoretical physics.

An important new development of black hole physics is connected with the attempts to construct a unified theory of all interactions. Unification of gravity with other gauge theories made it interesting to study black-hole-like solutions, describing a black hole with "colored" and "quantum hair". Modern superstring theory,

[2] We should add here that the standard assertion about the impossibility of seeing what happens inside a black hole is to be taken with care. A gedanken experiment has been proposed [Frolov and Novikov (1993a)] where one can probe a black hole interior.

which explains gravity as some collective state of fundamental string excitations, reproduces general relativity in the low energy limit. This theory requires additional fundamental fields (e.g., the dilaton field), which inevitably violate the equivalence principle. The study of black-hole-like solutions in string-generated gravity attracted the attention of theorists, many of whom have been working in high energy physics and superstring theory.

Besides the main solved and unsolved problems in black hole physics listed above, there are a lot of other questions connected with black hole physics and its applications. It is now virtually impossible to write a book where all these problems and questions are discussed in detail. Every month new issues of Physical Review D, Astrophysical Journal, and other physical and astrophysical journals add scores of new publications on the subject of black holes. In writing a book on this subject we were restricted by space and time. We tried to include material that is connected to the basic concepts of black hole physics and their recent development. Both of the authors have been working in the area of black hole physics for more than 30 years. Certainly we have our favorite topics. The reader should excuse the authors if some subjects are treated in more detail than other that are more important from his or her point of view.

This book presents a systematic exposition of black hole physics. The first part of the book (Chapters 2 to 9) contains what might be called the "classical physics of black holes" and its applications. In order to introduce certain important notions descriptively and to place emphasis on the fundamental problems, the authors have attempted to make the presentation in the opening chapters particularly simple and clear. This is especially true of Chapter 2 which describes the properties of the simplest spherical black hole. This chapter also presents the properties of the spacetime within a black hole.

Chapter 3 introduces the most important properties of rotating black holes and of those having a nonzero electric charge. Special attention in Chapters 2 and 3 is paid to the special reference frames used for the study of properties of black holes and to other tools which allow one to visualize the properties of these unusual objects.

Chapter 4 (written jointly with N. Andersson) treats the propagation of weak physical fields near black holes. The main focus is on the evolution of weak gravitational fields. This is especially important in the problem of the stability of a black hole under external perturbations, and in the problem of the emission of gravitational waves by bodies (and fields) moving in the neighborhood of a black hole. Here, we also treat the formation of a black hole in the collapse of a slightly non-spherical body.

Chapter 5 presents the general theory of non-stationary black holes and results on the existence of singularities inside black holes. We also discuss the cosmic censorship conjecture and critical behavior in black hole formation.

Chapter 6 is devoted to properties of stationary black holes. We give a formulation and proof of the uniqueness theorems for stationary black holes in the Einstein-

Maxwell theory.

Chapter 7 describes various physical effects in the gravitational field of a black hole: superradiance, shift in the self-energy of a charged particle, the transformation of electromagnetic waves into gravitational waves, and the inverse process. The motion and deformation of black holes in an external field, interaction between black holes, and interaction of cosmic strings with a black hole are also discussed in this chapter. We conclude the chapter by presenting the recent numerical results on the head-on collision of two black holes.

Chapter 8 treats black hole electrodynamics. The membrane approach is used to reformulate the electrodynamics in (3+1)-form. As the most important application of these formulation we discuss the structure of black hole magnetospheres.

Chapter 9 contains a review of black hole astrophysics. Primary attention is paid to observational evidence for the existence of black holes.

The second part of the book (beginning with Chapter 10) concerns more advanced subjects of black hole physics and its quantum generalization.

The quantum physics of black holes is the subject of Chapters 10 and 11. Chapter 10 gives the general solution to the problem of the creation of particles in the field of a stationary black hole. Chapter 11 summarizes the results on vacuum polarization in the neighborhood of a black hole and discusses properties of black holes as quantum objects.

The thermodynamics of black holes and the problem of its statistical-mechanical foundations is discussed in Chapter 12.

Black holes in unified theories and in theories of gravity in spacetimes with dimensions less or higher that four are considered in Chapter 13.

Chapter 14 treats various aspects of the structure of spacetime inside black holes, including the instability of Cauchy horizons and so-called mass inflation.

The problem of the final state of an evaporating black hole is the subject of Chapter 15. We discuss different possible outcomes of this process and the information loss puzzle.

Chapter 16 analyzes some general aspects of the relation between the topology of spacetime and causality. It contains a brief introduction into wormholes and time machines.

The book has a number of appendices which collect results from Riemannian geometry and general relativity. We included in the appendices the derivation of some of the results used in the main text. They also contain useful formulas which might be needed to verify some of the relations in the main part of the book that are given without derivation. In particular, they set out some mathematical relations for the Kerr-Newman geometry. We also included in the appendices a brief summary of quantum field theory in curved spacetime and its application to quantum theory in the Kerr metric.

In Chapter 2 all formulas are written with dimensional physical constants c and G. Beginning with Chapter 3, where more complicated material is treated and the use of dimensional constants would yield unwieldy expressions, we employ (except in the final formulas or where specified otherwise) the system of units $c = G = \hbar = k = 1$. The signs in the definitions of metric, curvature tensor, and Ricci tensor are chosen as in the monograph of Misner, Thorne, and Wheeler (1973).

Chapter 2

Spherically Symmetric Black Holes

2.1 Spherically Symmetric Gravitational Field

We begin the analysis of the physical properties of black holes with the simplest case in which both the black hole and its gravitational field are spherically symmetric. The *spherically symmetric gravitational field* (spacetime with spherical three-dimensional space) is described in every textbook on general relativity [see, e.g., Landau and Lifshitz (1975), Misner, Thorne, and Wheeler (1973)]. Therefore, here we will only reproduce the necessary results. Mathematical details connected with definition and properties of a spherically symmetric gravitational field can be found in Appendix B.

Let us write the expression for the metric[1] in a region far from strong gravitational fields (i.e., where special relativity is valid), using the spherical spatial coordinate system (r, θ, ϕ):

$$ds^2 = -c^2\,dt^2 + dl^2 = -c^2\,dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)\,, \tag{2.1.1}$$

where c is the speed of light, and dl is the distance in three-dimensional space.

Consider now a curved spacetime but preserve the condition of spatial spherical symmetry. Spacetime is not necessarily empty, it may contain matter and physical fields (which are, of course, also spherically symmetric if their gravitation is considered). It can be shown (see, e.g., Misner, Thorne, and Wheeler (1973), and Appendix B) that there exist coordinates (x^0, x^1, θ, ϕ) in a spherically symmetric spacetime such that its metric is of the form

$$ds^2 = g_{00}(x^0, x^1)\,dx^{0^2} + 2g_{01}(x^0, x^1)\,dx^0\,dx^1 + g_{11}(x^0, x^1)\,dx^{1^2}$$

$$+ g_{22}(x^0, x^1)(d\theta^2 + \sin^2\theta\,d\phi^2)\,. \tag{2.1.2}$$

[1] We use $(-,+,+,+)$ signature convention for the spacetime metric $ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu$. Greek indices μ, ν take values 0,1,2,3, while small Latin indices $i, j = 1, 2, 3$ enumerate "spatial" coordinates. For definition and properties of other geometrical quantities used in the book see Appendix A.

It is evident that there is a freedom in a choice of coordinates (x^0, x^1, θ, ϕ). The form of the metric (2.1.2) remains unchanged under the coordinate transformations

$$\tilde{x}^0 = \tilde{x}^0(x^0, x^1), \quad \tilde{x}^1 = \tilde{x}^1(x^0, x^1), \quad \tilde{\theta} = \theta, \quad \tilde{\phi} = \phi. \tag{2.1.3}$$

Quite often it is convenient to use coordinate transformation (2.1.3) to banish g_{01} coefficient of the metric (2.1.2) so that the metric takes the form

$$ds^2 = g_{00}(x^0, x^1)\, dx^{0^2} + g_{11}(x^0, x^1)\, dx^{1^2} + g_{22}(x^0, x^1)(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{2.1.4}$$

The remaining coordinate freedom after this is one function of two variables. This freedom can be used for further simplification of the metric. For example, if the gradient of the function $g_{22}(x^0, x^1)$ does not vanish in some region, this function can be chosen there as a new coordinate (say x^1). In these coordinates the spherically symmetric metric takes the form

$$ds^2 = g_{00}(x^0, x^1)\, dx^{0^2} + g_{11}(x^0, x^1)\, dx^{1^2} + (x^1)^2(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{2.1.5}$$

For fixed values of x^0 and x^1 the metric (2.1.5) is the metric on a two-dimensional sphere. Its surface area is $4\pi(x^1)^2$. For this reason, the coordinate x^1 is invariantly defined and has a well-defined meaning. The coordinates in which the expression for g_{22} is written in the form $(x^1)^2$ are called *curvature coordinates.* Usually by analogy to (2.1.1) the x^1 coordinate is denoted by r ("radius") and $x^0/c \equiv t$ ("time"). We will see that this choice of symbols is not always logically justified inside a black hole (see Section 2.4).

If a spherical gravitational field is considered not in vacuum, then in the general case matter moves radially in the three-dimensional coordinate system defined by the coordinates x^1, θ, ϕ; that is, energy flows exist. Sometimes it is more convenient to choose a different frame of reference; for example, a one comoving to the matter, but which is also spherically symmetric. All such reference frames where the metric is of the form (2.1.4) possess the following property. The points of which one such reference frame is composed move radially with respect to some other frame. A point $x^1 =$ const which is at rest in the older reference frame moves with respect to a new reference frame. If the coordinates in the new reference frame are marked with a tilde, then this motion is described by the equation $\tilde{x}^1 = \tilde{x}^1(x^0, x^1)$. Once the function $\tilde{x}^1(x^0, x^1)$ has been chosen, thus defining the new frame of reference, it is always possible to choose $\tilde{x}^0 = \tilde{x}^0(x^0, x^1)$, which defines the time coordinate $\tilde{x}^0$ in the new system in such a way that the component $\tilde{g}_{01}$ would not arise, and the general expression for ds^2 would have the same form as (2.1.4)

$$ds^2 = \tilde{g}_{00}(\tilde{x}^0, \tilde{x}^1)\, d\tilde{x}^{0^2} + \tilde{g}_{11}(\tilde{x}^0, \tilde{x}^1)\, d\tilde{x}^{1^2} + \tilde{g}_{22}(\tilde{x}^0, \tilde{x}^1)(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{2.1.6}$$

If we start with metric (2.1.5) in the curvature coordinates, then the expression for $\tilde{g}_{22}$ can be written in the form

$$\sqrt{\tilde{g}_{22}} = x^1(\tilde{x}^0, \tilde{x}^1), \tag{2.1.7}$$

where $x^1 = x^1(\tilde{x}^0, \tilde{x}^1)$ is the solution of (2.1.3) for x^1. It describes the radial motion of the points of the new reference frame (with the coordinate $\tilde{x}^1 = \text{const}$) with respect to the older one.

2.2 Spherically Symmetric Gravitational Field in Vacuum

2.2.1 Schwarzschild metric

Consider a spherical gravitational field in vacuum. The solution to Einstein's equations for this case was found by Schwarzschild (1916). It has the following form (see Appendix B.4.1):

$$ds^2 = -\left(1 - \frac{2GM}{c^2 r}\right) c^2\, dt^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2\theta\, d\phi^2). \quad (2.2.1)$$

Here G is Newton's gravitational constant, and M is the mass of the field source.

An important property of this solution is that it is independent of the temporal coordinate t. The solution is determined by a single parameter M; that is, the total mass of the gravitational source which produces the field. This interpretation of the parameter M immediately follows from the asymptotic form of the metric. Far from the center of gravity (as $r \to \infty$), spacetime approaches the flat Minkowski spacetime with metric (2.1.1), and the gravitational field can be described by using the *weak field approximation*. In this approximation $-g_{tt} = 1 + 2\varphi/c^2$, where $\varphi = -GM/r$ is the *Newtonian gravitational potential*. By comparing this result of the weak field approximation with the metric (2.2.1), one concludes that M is the mass of the gravitating system.

Even if the field source involves radial motions (which preserve spherical symmetry), the field beyond the region occupied by matter remains constant, and it is always described by metric (2.2.1). This statement is known as *Birkhoff's theorem* (Birkhoff (1923); see also Appendix B.3).

2.2.2 Schwarzschild reference frame

The coordinates (t, r, θ, ϕ) in which (2.2.1) is written are called *Schwarzschild coordinates*, and the frame of reference which they form is called the *Schwarzschild reference frame*.

For ordinary measurement of length in a small neighborhood of each spatial point, we can use a local Cartesian coordinate system (x, y, z). Let $(\mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be three spatial orthonormal vectors at a chosen point p. Then the displacement vector $\delta\mathbf{r}$ for the position of a point in a small neighborhood of p can be characterized by three numbers $(\delta x, \delta y, \delta z)$: $\delta \mathbf{r} = \delta x \, \mathbf{e}_{\hat{1}} + \delta y \, \mathbf{e}_{\hat{2}} + \delta z \, \mathbf{e}_{\hat{3}}$. If $\mathbf{e}_{\hat{1}}$, $\mathbf{e}_{\hat{2}}$, and $\mathbf{e}_{\hat{3}}$ are directed along r, θ, and ϕ, respectively, then

$$\delta x = \sqrt{g_{11}} \, dr = \left(1 - \frac{2GM}{c^2 r}\right)^{-1/2} dr \,, \tag{2.2.2}$$

$$\delta y = \sqrt{g_{22}} \, d\theta = r \, d\theta \,, \tag{2.2.3}$$

$$\delta z = \sqrt{g_{33}} \, d\phi = r \, \sin\theta \, d\phi \,. \tag{2.2.4}$$

The factor $(1 - 2GM/c^2 r)^{-1/2}$ in (2.2.2) reflects the curvature of the three-dimensional space.

The physical time τ at a point r is given by the expression

$$d\tau = c^{-1} \sqrt{-g_{00}} \, dx^0 = \sqrt{-g_{00}} \, dt = \left(1 - \frac{2GM}{c^2 r}\right)^{1/2} dt \,, \tag{2.2.5}$$

where $x^0 = ct$. If $\mathbf{e}_{\hat{0}}$ is a unit time-directed vector orthogonal to $\mathbf{e}_{\hat{m}}$ $(m = 1, 2, 3)$, then the position of any event with respect to the local frame can be presented as $c \, \delta\tau \, \mathbf{e}_{\hat{0}} + \delta \mathbf{r}$.

Far from the gravitational center (as $r \to \infty$), we have $d\tau = dt$; that is, t is the physical time of the observer located at infinity. At smaller r, the time τ runs progressively slower in comparison with the time t at infinity. As $r \to 2GM/c^2$, we find $d\tau \to 0$.

Let us now calculate the acceleration of free fall of a body which is initially at rest (or moves at a low velocity $v \ll c$) in the Schwarzschild reference frame. Using formulae (A.42) and (A.43) of the Appendix A, we find

$$a = \sqrt{a_i \, a^i} = \frac{GM}{r^2 (1 - 2GM/c^2 r)^{1/2}} \,. \tag{2.2.6}$$

The acceleration points along the radius and is directed toward the center. As $r \to 2GM/c^2$, the acceleration tends to infinity. The singularity in the flow of time arising as $r \to 2GM/c^2$ [see (2.2.5)] and in the expression for acceleration a [see (2.2.6)] demonstrates that at this value of r, the Schwarzschild reference frame has a physical singularity.[2] The quantity $r = r_g = 2GM/c^2$ is called the *Schwarzschild radius*

[2] If a test body is at rest with respect to the Schwarzschild reference frame, its motion is not geodesic. Expression (2.2.6) gives the acceleration, and $\mathcal{F}_0 = m \, a$ gives the force acting on a body of mass m, measured by an observer located near this body at a point r_0. If the body is suspended by a weightless absolutely rigid string, then the force applied to the free end of the string at the point r_1 is

$$\mathcal{F}_1 = \mathcal{F}_0 \sqrt{\frac{g_{00}(r_0)}{g_{00}(r_1)}} \,.$$

As r_0 tends to $2GM/c^2$, $\mathcal{F}_0 \to \infty$, while $\mathcal{F}_1$ remains finite.

(or *gravitational radius*), and the sphere of radius r_g is said to be the *Schwarzschild sphere*. We will later give a detailed analysis of the physical meaning of the singularity at $r = r_g$.

The Schwarzschild reference frame is static, and hence rigid [$D_{ik} = 0$; see (A.44)]. The three-dimensional geometry of the space g_{ij} and the gravitational potential g_{00} in it do not depend on time t. The generator of this time-symmetry transformation is known as the *Killing vector* $\boldsymbol{\xi}_{(t)}$: $\xi^{\mu}_{(t)} = \delta^{\mu}_{0}$. The Schwarzschild reference frame can be thought of as a coordinate lattice "welded" out of weightless rigid rods which fill the space around the black hole. We can study the motion of particles relative to this lattice, the evolution of physical fields at different points of this lattice, and so on. The Schwarzschild lattice thus, to some extent, resembles the lattice of fixed coordinates in the absolute Newtonian space of non-relativistic physics. When we speak of the motion of particles or the evolution of a field in the Schwarzschild spacetime, we always mean the motion and evolution of the field in this analogue of absolute Newtonian space.[3] This similarity of the Schwarzschild and Newtonian frames is a great help to our intuition. Of course, the geometry of the three-dimensional Schwarzschild space around a gravitational center is non-Euclidean, in contrast to the Euclidean Newtonian space of non-relativistic physics. Indeed, the 3-geometry of a space section $t = \text{const}$ is

$$dl^2 = \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,. \tag{2.2.7}$$

To "visualize" the properties of this three-dimensional curved space, one may consider its various two-dimensional sections. For example, Figure 2.1 shows the embedding diagram representing the geometry of the equatorial plane $\theta = \pi/2$.

As a result of the presence of the critical radius, $r_g = 2GM/c^2$, in the spherical field in vacuum, where the free-fall acceleration a with respect to the Schwarzschild frame becomes infinite, such a rigid non-deformable lattice cannot be extended to $r \le r_g$, since this region contains no non-deformable space (no analogue of the Newtonian space). The fact that the free fall acceleration a tends to infinity at r_g is an indication that for $r \le r_g$ all systems must be nonrigid in the sense that $g_{\alpha\beta}$ must be a function of time and all systems must be deformed (all bodies must fall to the center). We will see that this is indeed the case.

Note that these peculiarities at $r = r_g$ do not indicate that there is a singularity of infinite curvature, or something similar, in the geometry of the four-dimensional spacetime. We shall see later that the spacetime near the gravitational radius is quite regular, and the singularity at r_g points to a physical singularity only in the Schwarzschild reference frame; that is, it signifies the impossibility of extending this reference frame as a rigid and non-deformable one (not falling on the center) to $r \le r_g$.

[3] Recall that, in the general case of nonspherical time-dependent gravitational fields, it is not possible to introduce a rigid three-dimensional space; this fact quite often stands in the way of clear intuitive concepts and inhibits calculations.

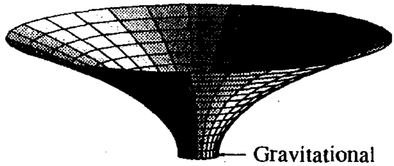


Figure 2.1: Embedding diagram for the geometry on the equatorial plane in the Schwarzschild spacetime. The two-geometry $dl^2 = \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2\, d\phi^2$ (obtained from the Schwarzschild metric (2.2.1) by putting there $dt = 0$ and $\theta = \pi/2$) is isometric to the internal geometry on the two-dimensional surface of rotation $z = 2\sqrt{2M(r - 2m)}$ in a flat three-dimensional space $ds^2 = dz^2 + dr^2 + r^2\, d\phi^2$ (see Appendix B.5.4). A coordinate lattice "welded" from rigid rods which fill the space around the black hole is schematically shown.

Note in conclusion that r_g is extremely small even for celestial bodies. Thus, $r_g \approx 0.9\,\text{cm}$ for the Earth's mass and $r_g \approx 3\,\text{km}$ for the solar mass. If $r \gg r_g$, the gravitational field is weak, and it can be described by the Newtonian approximation with a very high accuracy. The free fall acceleration $a = GM/r^2$, and the curvature of the three-dimensional space is negligible. Outside typical celestial bodies (and all ordinary bodies as well), the gravitational field is the Newtonian field because their sizes are typically much greater than r_g. (The only known exceptions are neutron stars and black holes.) The Schwarzschild solution is not valid within these bodies , but, obviously, the gravitational field is still Newtonian with enormously high accuracy.

We will see later that a spherical black hole is formed when a non-rotating spherical body collapses to a size below its gravitational radius. But before we discuss this process of the birth of a black hole, we need to look at the laws of radial motion of test particles in the Schwarzschild field.

2.3 Radial Motion of Test Particles in the Schwarzschild Field

2.3.1 Radial motion of light

We begin with a radial motion of a photon, which always propagates at the fundamental velocity c. This is also true for any other ultra-relativistic particle. For this particle, $ds = 0$. For radial motion, $d\theta = d\phi = 0$. Substituting $ds = d\theta = d\phi = 0$ into (2.2.1), we find the equation of motion

$$\frac{dr}{dt} = \pm c\left(1 - \frac{r_g}{r}\right). \tag{2.3.1}$$

Recall that dr/dt is the rate at which the coordinate r changes with the time t of a distant observer; that is, this is the coordinate (not the physical) velocity. The physical velocity is the rate of change of physical distance, dx [see (2.2.2)], in the physical time τ [see (2.2.5)]

$$\frac{dx}{d\tau} = \pm \frac{\sqrt{g_{11}}\, dr}{\sqrt{-g_{00}}\, dt} = \pm c\,. \tag{2.3.2}$$

Of course, the physical velocity of the photon (in any reference frame) is always equal to c.

From the standpoint of a distant observer (and according to his clock), the change dx in the physical radial distance with t is

$$\frac{dx}{dt} = \pm c \left(1 - \frac{r_g}{r}\right)^{1/2}. \tag{2.3.3}$$

Therefore, a distant observer finds that a light ray close to r_g moves slower, and as $r \to r_g$, $dx/dt \to 0$. Obviously, this behavior reflects the slowing down of time close to r_g [see (2.2.5)].

How long does a photon take, by the clock of a distant observer, to reach the point r_g if the motion starts radially from $r = r_1$? We integrate equation (2.3.1) and obtain

$$t = \frac{r_1 - r}{c} + \frac{r_g}{c} \ln\left(\frac{r_1 - r_g}{r - r_g}\right) + t_1\,, \tag{2.3.4}$$

where r_1 is the position occupied by the photon at the moment t_1. Expression (2.3.4) shows that $t \to \infty$ as $r \to r_g$. Whatever the coordinate r_1 from which the photon starts its fall, by the clock of the distant observer, the time t taken by the photon to reach r_g is infinite.

How does the photon energy change in the course of the radial motion? Energy is proportional to frequency. Let us look at the evolution of frequency. Suppose that light flashes are emitted from a point $r = r_1$ at an interval Δt. The field being static, the flashes will reach the observer at $r = r_2$ after the same interval Δt. The ratio of the proper time intervals at these two points is

$$\frac{\Delta\tau_1}{\Delta\tau_2} = \frac{\sqrt{-g_{00}(r_1)}\,\Delta t}{\sqrt{-g_{00}(r_2)}\,\Delta t}; \tag{2.3.5}$$

hence, the ratio of frequencies is

$$\frac{\omega_1}{\omega_2} = \frac{\Delta\tau_2}{\Delta\tau_1} = \sqrt{\frac{g_{00}(r_2)}{g_{00}(r_1)}} = \sqrt{\frac{1 - r_g/r_2}{1 - r_g/r_1}}\,. \tag{2.3.6}$$

The frequency of a quantum decreases as it leaves the gravitational field and increases as it moves to smaller values of r. These effects are called the gravitational *redshift* and *blueshift*, respectively.

2.4 Spacetime Within the Schwarzschild Sphere

2.4.1 Lemaître reference frame

The fact that the proper time of fall to the Schwarzschild sphere is finite, suggests a method of constructing a reference frame which can be extended to $r < r_g$.[5] The reference frame must be fixed to the freely falling particles. No infinite accelerations **a** and no corresponding infinite forces will arise at the gravitational radius of this system because the particles of the system fall freely, and **a** is identically zero everywhere. The simplest such frame of reference, so-called *Lemaître reference frame*, consists of freely falling particles that have zero velocity at spatial infinity [Lemaître (1933), see also Rylov (1961)]. The motion of these particles is described by equation (2.3.12).

In order to introduce this reference frame, we choose for the time coordinate the time T measured by a clock fixed to the falling particles. At a certain instant of time T (say at $T = 0$), the different freely falling particles of the ensemble are located at different r_1. We can choose these values of r_1, which mark the particles and remain unchanged for each of them, as the new radial coordinate in this reference frame.

The metric in the frame of freely falling particles is written in the form

$$ds^2 = -c^2\,dT^2 + \frac{r_1\,dr_1^2}{r_g\,B} + B^2\,r_g^2\,(d\theta^2 + \sin^2\theta\,d\phi^2)\,, \tag{2.4.1}$$

where $B = \left[(r_1/r_g)^{3/2} - (3cT)/(2r_g)\right]^{2/3}$. It is more convenient to use instead of r_1 a new radial coordinate R

$$R = \frac{2}{3}\,r_g\left(\frac{r_1}{r_g}\right)^{3/2}. \tag{2.4.2}$$

The metric (2.4.1) is now transformed to the form

$$ds^2 = -c^2\,dT^2 + \frac{dR^2}{B} + B^2\,r_g^2\,(d\theta^2 + \sin^2\theta\,d\phi^2)\,, \tag{2.4.3}$$

[5] Historically the first system of coordinates without singularities on the gravitational radius r_g was constructed by Painlevé (1921) and Gullstrand (1922). It is obtained by transformation

$$cT = ct + f(r)\,, \qquad \frac{df}{dr} = \pm\frac{1}{1 - r_g}\sqrt{\frac{r_g}{r}}\,.$$

The Schwarzschild metric in coordinates (T, r, θ, ϕ) is

$$ds^2 = -\left(1 - \frac{r_g}{r}\right)c^2\,dT^2 \pm 2\sqrt{\frac{r_g}{r}}\,dr\,c\,dT + dr^2 + r^2\,d\omega^2\,.$$

This metric is evidently regular at the gravitational radius. [For more details see historical articles by Eisenstaedt (1982) and Israel (1987).]

$$B = \left[\frac{3(R-cT)}{2r_g}\right]^{2/3}.$$

The reference frame with interval (2.4.3) (the *Lemaître reference frame*) indeed has no singularity on the Schwarzschild sphere. In order to show this, we write the explicit relation between the Schwarzschild and Lemaître coordinates

$$r = B\,r_g = {r_g}^{1/3}\left[\frac{3}{2}(R-cT)\right]^{2/3}, \tag{2.4.4}$$

$$t = \frac{r_g}{c}\left\{-\frac{2}{3}\left(\frac{r}{r_g}\right)^{3/2} - 2\left(\frac{r}{r_g}\right)^{1/2} + \ln\left|\frac{(r/r_g)^{1/2}+1}{(r/r_g)^{1/2}-1}\right| + \frac{R}{r_g}\right\}. \tag{2.4.5}$$

Setting $r = r_g$ in (2.4.4), we obtain an equation for the position of the Schwarzschild sphere in the Lemaître reference frame:

$$B = 1 \quad \text{and} \quad r_g = \frac{3}{2}(R-cT). \tag{2.4.6}$$

Components of the metric $g_{\alpha\beta}$ in (2.4.3) on the Schwarzschild sphere are regular, without any singularity. The calculation of invariants of the curvature of four-dimensional spacetime also reveals no singularities on the Schwarzschild sphere. The Lemaître reference frame extends to $r < r_g$. The spacetime in the Lemaître (T, R)-coordinates is shown in Figure 2.2 (by virtue of symmetry, the angular coordinates θ and ϕ are irrelevant).

This reference frame can be extended up to $r = 0$, or up to $R = cT$ in the Lemaître coordinates [see (2.4.4)]. Here we find the true singularity of spacetime; namely, infinite curvature. For example, the curvature invariant $R_{\alpha\beta\gamma\delta}\,R^{\alpha\beta\gamma\delta}$, which for the Schwarzschild metric is $12r_g^2/r^6$, is finite at the gravitational radius, but grows infinitely as $(R-cT) \to 0$. The infinity in this invariant signifies an infinity of gravitational tidal forces.

As shown in Figure 2.2, each freely falling particle with $R = \text{const}$ in the Lemaître reference frame moves with time T to smaller r. The particle reaches r_g at a time $T = c^{-1}(R - \frac{2}{3}r_g)$, keeps falling, and reaches the true singularity $r = 0$ at $T = c^{-1}R$.[6] Spacetime cannot be extended beyond the singularity at which gravitational tidal forces grow infinitely and particles would be destroyed. In the neighborhood of $r = 0$, quantum effects of the gravitational field become important; this aspect will be discussed in Chapter 14.

[6]The fact that the lines $r = \text{const}$ are plotted in (R, T)-coordinates by straight lines is a corollary of the special choice of R [see (2.4.2)]. This is one of the reasons why we choose the coordinate R instead of r_1.

2.3.2 Radial motion of particles

Let us now look at the radial motion of non-relativistic particles in vacuum. We begin with free motion in which no non-gravitational forces act on a particle (free fall, motion along a geodesic). The integration of the equation for a geodesic in the case $d\theta = d\phi = 0$ [see Bogorodsky (1962)] yields the expression

$$\frac{dr}{dt} = \pm \frac{(1 - r_g/r)[(E/mc^2)^2 - 1 + r_g/r]^{1/2}}{E/mc^2} c, \tag{2.3.7}$$

where E is the constant of motion[4] describing the total energy of a particle, including its mass m. If the particle is at rest at infinity where the gravitational field vanishes, then $E = mc^2$. In the general case, the value E/mc^2 may be greater or smaller than unity, but E is invariably positive for a particle moving outside the sphere of radius r_g.

At a large distance $r \gg r_g$, we find that for non-relativistic particles $|1 - E/mc^2| \ll 1$, and expression (2.3.7) takes the form

$$\frac{1}{2} m \left(\frac{dr}{dt}\right)^2 = (E - mc^2) + \frac{GmM}{r}. \tag{2.3.8}$$

The quantity $\mathcal{E} = E - mc^2$ is the energy of a particle in Newtonian theory (where the rest, or proper, mass is not included in the energy), and thus expression (2.3.8) reduces to the energy conservation law in Newtonian theory.

Recall again that dr/dt in (2.3.7) is the coordinate (not the physical) velocity. The physical velocity v measured by an observer who is at rest in the Schwarzschild reference frame situated in the neighborhood of the freely moving body is

$$v = \frac{dx}{d\tau} = \sqrt{\frac{g_{11}}{|g_{00}|}} \frac{dr}{dt} = \pm \frac{[(E/mc^2)^2 - 1 + r_g/r]^{1/2}}{E/mc^2} c. \tag{2.3.9}$$

---

[4] The energy of a particle or a field propagating in a stationary (i.e., time independent) gravitational field is conserved. This conservation law is a consequence of Noether's theorem. If u^μ is the four-velocity of a particle of mass m freely moving in a stationary gravitational field ($\partial g_{\mu\nu}/\partial x^0 = 0$), then the conserved energy is $E = -mc^2 u_\mu \xi^\mu_{(t)} = -mc^2 u_0$. Here $\xi^\mu_{(t)}$ is the Killing vector generating time-symmetry transformations. In a static gravitational field (i.e., when $g_{0i} = 0$), the energy $E = mc^2(-g_{00})u^0$ is always positive in the region where $g_{00} < 0$. The energy in a static spacetime can be rewritten in the form

$$E = mc^2 \sqrt{-g_{00}} / \sqrt{1 - v^2/c^2},$$

where v is the physical three-velocity, defined as the rate of change of the physical distance in the physical time: $v^2 = g_{ij}\, dx^i\, dx^j/(-g_{00})\, dt^2$. For radial motion in the Schwarzschild metric

$$E = mc^2 \sqrt{A} / \sqrt{(1 - \dot{r}^2/A^2)},$$

where $A = 1 - r_g/r$, and $\dot{r} = dr/(c\, dt)$. This relation implies (2.3.7)

If the falling body approaches r_g, the physical velocity $v = dx/d\tau$ constantly increases: $v \to c$ as $r \to r_g$. By the clock of the distant observer, the velocity dx/dt tends to zero as $r \to r_g$, as in the case of the photon. This fact reflects the slowing down of time as $r \to r_g$.

What is the time required for a body falling from a point $r = r_1$ to reach the gravitational radius r_g (by the clock of the distant observer)? The time of motion from r_1 to r_g is given by the integral of (2.3.7). This integral diverges as $r \to r_g$. This result is not surprising because $\Delta t \to \infty$ as $r \to r_g$ even for light, and nothing is allowed to move faster than light. Furthermore, the divergence of Δt for a falling body is of the same type as for light because the physical velocity of the body, v, always tends to c as $r \to r_g$. Obviously, whatever the force acting on a particle, the time Δt to reach r_g is always infinite because in this case again $v < c$. We conclude that both free fall and motion towards r_g with any acceleration always take an infinite time measured by the clock of the distant observer.

Let us return to a freely moving particle. What is the time ΔT to reach r_g measured by the clock of the falling particle itself? It is found from the formula

$$\Delta T = \frac{1}{c} \int_{r_g}^{r_1} |\, ds \,| , \tag{2.3.10}$$

where ds is taken along the world line of the particle. Here, using the expression for ds from (2.2.1), for $d\theta = d\phi = 0$, we find

$$\Delta T = \frac{1}{c} \int_{r_g}^{r_1} \sqrt{\left| \frac{g_{00}}{(dr/c\,dt)^2} + g_{11} \right|}\, dr \, . \tag{2.3.11}$$

In order to calculate ΔT, we substitute into (2.3.11) the expression for dr/dt from (2.3.7). It is easy to show that the integral converges and the interval ΔT is finite. In the particular case $E = mc^2$, when the particle falls at the parabolic (escape) velocity (i.e., $dr/dt = 0$ at $r \to \infty$), we find for the time of fall from r_1 to r

$$\Delta T = \frac{2}{3} \frac{r_g}{c} \left[\left(\frac{r_1}{r_g} \right)^{3/2} - \left(\frac{r}{r_g} \right)^{3/2} \right] . \tag{2.3.12}$$

We thus conclude that while the duration Δt of falling is infinite for the distant observer, the time ΔT measured by the clock of the particle itself is finite. This result, which at first glance seems quite unexpected, can be given the following physical interpretation. The clock on the particle falling toward r_g is slowed down relative to the clock at infinity, first, because time is slowed down in the gravitational field [see (2.2.5)], and second,because of the Lorentz contraction of time when the velocity of the clock $v \to c$ as $r \to r_g$. As a result, the interval in t is infinite, but it becomes finite when clocked in T.

forward physical interpretation of this system if $r < r_g$? Indeed, there is [Novikov (1961)]. As was demonstrated above, the r coordinate cannot be the radial spatial coordinate in the region $r < r_g$. However, it can play the role of the temporal coordinate, as follows from (2.2.1) where the coefficient with dr^2 reverses its sign on crossing the Schwarzschild sphere and is negative where $r < r_g$. On the other hand, now the coordinate t can be used as the spatial coordinate, the coefficient of dt^2 being positive for $r < r_g$. The coordinates r and t thus change their roles when r becomes less than r_g. We change the variables, $r = -c\widetilde{T}$, $t = \widetilde{R}/c$, and rewrite (2.2.1) in the form

$$ds^2 = -\left[\frac{r_g}{(-c\widetilde{T})} - 1\right]^{-1} c^2\, d\widetilde{T}^2 + \left[\frac{r_g}{(-c\widetilde{T})} - 1\right] d\widetilde{R}^2 + \\ + c^2\, \widetilde{T}^2 (d\theta^2 + \sin^2\theta\, d\phi^2)\,, \tag{2.4.8}$$

$$0 < -c\widetilde{T} < r_g\,, \qquad -\infty < \widetilde{R} < \infty\,. \tag{2.4.9}$$

The frame of reference (2.4.8)–(2.4.9) can be realized by free test particles moving along geodesics inside the sphere $r = r_g$. A three-dimensional section $\widetilde{T} = \text{const}$ has an infinite spatial extension along the coordinate $\widetilde{R}$, while along the coordinates θ and ϕ it is closed, constituting on the whole a topological product of the sphere S^2 by a straight line R^1. The three-dimensional volume of this section is infinite. The system is nonstationary; it contracts along θ and ϕ (the radius of the sphere decreases from r_g to 0) and expands along $\widetilde{R}$. Its proper lifetime is finite:

$$\tau = \int_{-r_g/c}^{0} \sqrt{-g_{00}}\, d\widetilde{T} = \frac{\pi}{2}\, r_g\,. \tag{2.4.10}$$

The world lines of the particles with $\widetilde{R} = \text{const}$ that form the frame are plotted in Lemaître coordinates (2.4.3) in Figure 2.3. This figure shows that the particles move within the Schwarzschild sphere, and the system is by no means an extension of the Schwarzschild system to $r < r_g$ (its world lines $r = \text{const}$ are shown in the same figure). Time and the spatial radial direction undergo a peculiar change of roles in these systems.

2.4.3 Eddington-Finkelstein coordinates

Now we describe another reference frame without singularities on r_g constructed by Eddington (1924) and Finkelstein (1958). This reference frame is fixed to radially moving photons. The equation of motion of photons is given by (2.3.4). In the region $r > r_g$ the photons moving towards the center are characterized by r decreasing with t. Expression (2.3.4) for such photons can be rewritten in the form

$$ct = v - r_*\,, \qquad r_* = r + r_g \ln\left|\frac{r}{r_g} - 1\right| + \frac{v}{c}\,. \tag{2.4.11}$$

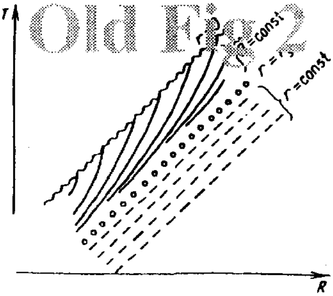


Figure 2.3: $\widetilde{R}$ = const world lines of particles, forming the frame of reference (2.4.9) in Lemaître coordinates.

Here r_* is a so-called *tortoise coordinate*, and v is a constant characterizing the radial coordinate of the photon at a fixed instant t.

The logarithm in r_* involves the modulus of the difference $r/r_g - 1$ so that r_* is defined both for $r > r_g$ and $r < r_g$. If we take a set of photons at a fixed t and assign to each photon a number v which remains unchanged during the motion of the photon, this v can be chosen as another new coordinate. This is similar to our choice of r_1 for a new radial coordinate in the case of a non-relativistic particle [see (2.3.12)]. However, an essential difference must not be overlooked. Namely, no observer can move together with a photon, so that in this sense the new frame does not fall, strictly speaking, under the definition of a reference frame. Nevertheless, such a "system" of test photons proves to be convenient. One needs to remember, though, that v (usually called an *advanced time*) is a lightlike coordinate (neither spatial nor temporal). For a second coordinate, we can choose the familiar coordinate r. Differentiating (2.4.11) and substituting the obtained expression for dt into (2.2.1), we find

$$ds^2 = -\left(1 - \frac{r_g}{r}\right) dv^2 + 2\,dv\,dr + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,. \tag{2.4.12}$$

Expression (2.4.12) is regular on $r = r_g$. Indeed, the coefficient with dv^2 vanishes on r_g, but the presence of the term $2\,dv\,dr$ ensures that the metric (and hence the coordinate system) remains non-degenerate. The spacetime in the coordinates (v, r) is shown in Figure 2.4. Coordinate lines of constant v, representing ingoing radial null rays, are plotted on a 45-degree slant, just as they would be in flat spacetime.

Later we shall use another type of the Eddington-Finkelstein coordinates (u, r, θ, ϕ), where $u = c\,t - r_*$ is the *retarded time*. In this coordinates the Schwarzschild metric takes the form

$$ds^2 = -\left(1 - \frac{r_g}{r}\right) du^2 - 2\,du\,dr + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,. \tag{2.4.13}$$

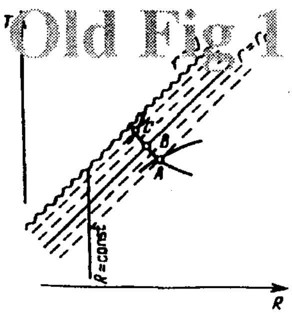


Figure 2.2: Schwarzschild spacetime in Lemaître coordinates. Dashed lines are lines of $r = \text{const}$; ABC is the world line of a photon falling toward a black hole. Segments of the world lines of photons moving in the opposite direction are shown at points A, B, C.

Figure 2.2 also shows the world lines of radial light rays. They are found from (2.4.3) by imposing the conditions $ds = 0$, $d\theta = d\phi = 0$:

$$c\frac{dT}{dR} = \pm\left[\frac{r_g}{\frac{3}{2}(R - cT)}\right]^{1/3}. \tag{2.4.7}$$

The position of light cones in Figure 2.2 immediately demonstrates why the Schwarzsch sphere plays a special role in the spherical vacuum gravitational field, in general, and in the Schwarzschild reference frame, in particular. Indeed, consider radial lines $\theta = \text{const}$, $\phi = \text{const}$. If $r > r_g$, then the $r = \text{const}$ world lines lie within the light cone, and they are timelike. The $r = r_g$ line coincides with the outgoing photon's world line, and it is lightlike. Finally, if $r < r_g$, the $r = \text{const}$ world lines are spacelike. This is why the Schwarzschild reference frame formed by particles with $r = \text{const}$ cannot be extended to $r < r_g$.

This situation is found to be typical for general relativity and constitutes the difference between it and ordinary field theory in flat space. Special coordinates must be chosen for solving Einstein's equations. Correspondingly, additional conditions are to be introduced in order to fix the form of metrics. In general, it is impossible to guarantee that the chosen coordinates cover the entire spacetime because the spacetime of general relativity may have a non-trivial global (topological and causal) structure. That was the situation encountered above in the attempts to describe the entire spherically symmetric spacetime in curvature coordinates (2.1.5). The general method of establishing whether the obtained solution indeed describes the entire spacetime or only part of it is to analyze the motion of test particles and light rays. If some of the particles reach the "boundary" of the chosen coordinate system in a finite proper time (or for a finite value of the affine parameter for photons), and there are no physical singularities at the "final" points of particle trajectories, then this coordinate system is incomplete. By changing the coordinates and switching to metric (2.4.1), we were able to cover a greater part of the spacetime and, among other things, describe attainable events below the gravitational radius. A discussion of whether the Lemaître coordinate system is indeed complete and whether metric (2.4.1) describes the entire spacetime will be deferred to Section 2.7.

2.4.2 *R*- and *T*-regions

After these general remarks we return to consider the properties of the Schwarzschild sphere and the region of spacetime within it. The most striking feature of the Schwarzschild sphere is the following. An outgoing light ray (a ray moving to the right in Figure 2.2) from a point with $r > r_g$ travels to greater r and escapes to spatial infinity. For points with $r < r_g$, both rays (moving to the right and to the left in Figure 2.2), travel toward smaller r; they do not escape to spatial infinity but are "stopped" at the singularity $r = 0$. The world line of any particle necessarily lies within the light cone. For this reason, if $r < r_g$, all particles have to move toward $r = 0$: this is the direction into the future. Motion toward greater r is impossible in the region $r < r_g$ [see Finkelstein (1958)]. It should be emphasized that this is true not only for freely falling particles (i.e., particles moving along geodesics) but also for particles moving with arbitrary accelerations. Neither radiation nor particles can escape from within the Schwarzschild sphere to the distant observer.

In (2.2.1), we defined r as the radial coordinate in the curvature coordinate system so that $g_{22} = r^2$. Formally, r within the Schwarzschild sphere is defined in the same manner, and $4\pi r^2$ still is the area of two-dimensional sphere $(r, t) =$ const. But world lines of "observers" at constant r are no longer timelike. If $r < r_g$, the quantity g_{22} is always a function of time, and a monotone function in any reference frame determined by relation (2.1.4). At $r < r_g$ all reference frames are nonstatic, and both radial rays travel only to smaller r (and hence to smaller g_{22}). Spacetime regions possessing this property are referred to as *T-regions* [Novikov (1962a,b, 1964a)]. The spacetime region outside the Schwarzschild sphere is said to form the *R-region*.

R- and T-regions can be defined in an arbitrary (not necessarily empty) spherically symmetric spacetime. By the definition of a spherically symmetric gravitational field, its metric can locally be written in the form (2.1.5). If the $x^1 =$ const, $\theta =$ const, $\phi =$ const world line in the neighborhood of a given point is timelike, this point belongs to the R-region. If this line is spacelike, the point in question belongs to the T-region.

Let us return to the case of a spherically symmetric gravitational field in vacuum. Apart from the already described Lemaître reference frame, other reference frames are employed for analyzing regions both inside and outside the Schwarzschild sphere. Here and in the sections that follow, we describe some of these frames.

First of all, we again turn to coordinate system (2.2.1). As we have shown in Section 2.2, this system has a singularity on the Schwarzschild sphere. But if r is strictly less than r_g, the metric coefficients are again regular. Is there a straight-

$$\dot{r}^2 = f(R) + \frac{F(R)}{r}, \tag{2.6.2}$$

$$g_{11}(T, R) = \frac{(r')^2}{1 + f(R)}, \tag{2.6.3}$$

$$\frac{8\pi G\rho}{c^2} = \frac{F'(R)}{r' r^2}. \tag{2.6.4}$$

Here a dot denotes differentiation with respect to cT, and a prime, differentiation with respect to R; $f(R)$ and $F(R)$ are two arbitrary functions of R (subject to the condition $1 + f(R) > 0$). These functions are specified by fixing initial conditions. Three functions $r(T_0, R)$, $\dot{r}(T_0, R)$, and $\rho(T_0, R)$ must be given at an initial moment of time T_0. Only two of these three functions are really important because the choice of the coordinate R at the moment T_0 is arbitrary and the coordinate transformation $R \to \tilde{R} = \tilde{R}(R)$ does not change physical observables. These two functions describe the initial distribution and radial velocity of the matter.

It is convenient to choose the coordinate R so that $R = 0$ at the center of the cloud, and $R = R_1$ at its boundary. Equations (2.6.2) and (2.6.4) being considered at $T = T_0$ determine functions $f(R)$ and $F(R)$. In particular, one has

$$F(R) = \frac{8\pi G}{3c^2} \int_0^R dR'\, \rho(T_0, R') \frac{d}{dR'} \left[r^3(T_0, R')\right].$$

The boundary condition $F(0) = 0$ follows from the condition $\dot{r}(0, T) = 0$ which must be satisfied at the center $r = 0$ of the cloud. For our problem $r(R)$ is a monotonic positive function. That is why $F(R) \geq 0$.

By solving equation (2.6.2), one obtains $r(T, R)$. Equations (2.6.3) and (2.6.4) determine $g_{11}(T, R)$ and $\rho(T, R)$, respectively. Thus, one gets a complete solution of the problem.

Differentiating (2.6.2) with respect to the time T, one has

$$\ddot{r} = -\frac{F}{2r^2},$$

and since $F \geq 0$, $\ddot{r}$ is non-positive. For this reason, each particle with fixed R and $\dot{r} < 0$ reaches the point $r = 0$, where the spacetime has a true singularity, in a finite time T.

If the coordinate R_1 determines the boundary of the sphere, then beyond the sphere (at $R > R_1$) we find $\rho = 0$ and $F = \text{const}$. Outside the matter, the metric of spacetime is determined in a unique manner by the value of F at the boundary R_1. In vacuum this metric is the Schwarzschild metric (see Section 2.2). The other function $f(R)$ which enters the Tolman solution can be chosen arbitrarily. In the absence of matter the change of $f(R)$ simply corresponds to the change of the reference frame without changing the gravitational field.

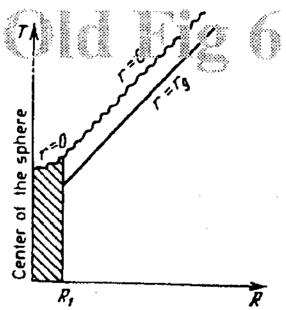


Figure 2.7: Spacetime of a contracting spherical cloud creating a black hole: Lemaître coordinates. The region inside the sphere is hatched.

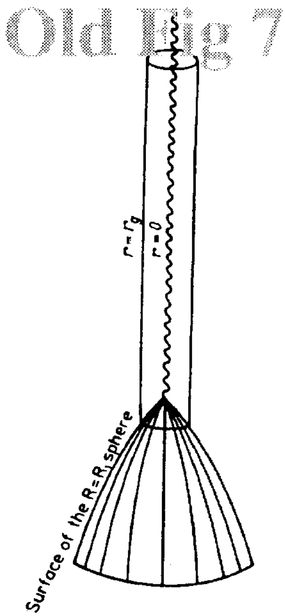


Figure 2.8: Spacetime of a contracting spherical cloud creating a black hole: Eddington-Finkelstein coordinates.

The particles on the boundary surface $R = R_1$ are freely falling in the outer metric so that their motion can also be described as motion along radial geodesics in the Schwarzschild metric [see (2.3.7)]. As a special case we can consider the contraction of a spherical cloud on whose boundary surface the particles fall at the parabolic (escape) velocity. The motion of such particles is described by especially simple formulas [see (2.3.12)]. In the Lemaître reference frame the equation of such a boundary is $R = R_1$. In the Eddington-Finkelstein reference frame the equation of the same boundary is given parametrically by expressions (2.4.11), (2.4.4), (2.4.5) provided we set $R = R_1$ in the last two formulas.

Figures 2.7 and 2.8 represent the spacetime in the case of the contracting spherical dust cloud in Lemaître and Eddington-Finkelstein coordinates, respectively. Figure 2.8, which also shows one of the angular coordinates, is especially illustrative. The surface of the contracting spherical cloud reaches the Schwarzschild sphere $r = r_g$ in a finite proper time and then contracts to a point at $r = 0$. This process is known as relativistic *gravitational collapse.* As a result of the collapse, a spacetime region is formed within the Schwarzschild sphere from which no signals can escape to spatial infinity. This region is defined as a *black hole.* The relativistic gravitational collapse

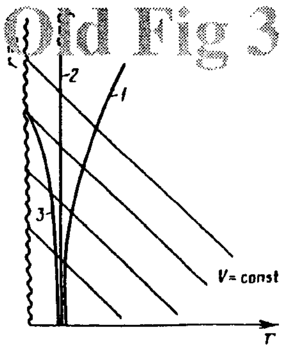


Figure 2.4: Schwarzschild spacetime in Eddington-Finkelstein coordinates (2.4.12) . The equation $v = \text{const}$ describes the world lines of photons falling toward $r = 0$. Lines 1, 2, 3 are the world lines of outgoing photons moving in the direction opposite to that of $v = \text{const}$.

2.5 Contracting and Expanding T-Regions

The properties, discussed above, of reference frames within the Schwarzschild sphere in the T-region are quite peculiar. Indeed, we notice that all these coordinate systems must contract along the θ and ϕ directions, and the coefficient g_{22} must decrease in time (this is equivalent to r decreasing with time). This fact can be rephrased as the inevitable motion of all light rays and all particles in the T-region toward the singularity. We know that Einstein's equations are invariant under time reversal. All the formulas given above remain a solution of Einstein's equations if the following change of variables is made: $t \to -t$, $T \to -T$, $\widetilde{T} \to -\widetilde{T}$, $v \to -u$, where u enumerates the outgoing rays ($u = 2t - v$). However, this change is equivalent to time reversal. Hence, reference frames are possible which are similar to the Lemaître and Eddington reference frames but which expand from below the Schwarzschild sphere and are formed by particles dashing out the singularity in the T-region, later intersecting the Schwarzschild sphere and escaping to infinity (Figures 2.5 and 2.6).

Is this conclusion of the escape of particles from below the Schwarzschild sphere compatible with the statement, emphasized on several occasions above, that no particle can escape from this sphere? The situation is as follows. No particle can escape from a T-region (or from within the Schwarzschild sphere) if this (or some other) particle *entered* this region from the external space ($r > r_g$) before. In other words, if the Schwarzschild sphere *can be entered*, it *cannot be escaped*. The T-region which appears in the solution with reversed time is a *quite different* T-region, with very different properties. While only contraction was possible in the former T-region, only expansion is possible in the latter one so that nothing can fall into it (this is very clear in Figures 2.5 and 2.6).

Note that the external space (beyond the $r = r_g$ sphere) is essentially the same

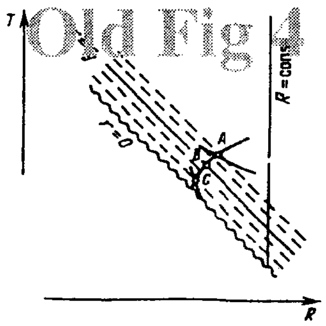


Figure 2.5: Schwarzschild spacetime in expanding Lemaître coordinates. The time arrow is reversed in comparison with Figure 2.3.

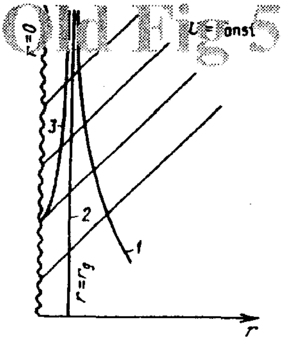


Figure 2.6: Schwarzschild spacetime in expanding Eddington-Finkelstein coordinates. The time arrow is reversed in comparison with Figure 2.4.

in the two cases. Its metric is reduced to (2.2.1) by a transformation of coordinates, but it can be extended within the Schwarzschild sphere *in two ways*: either as a contracting T-region, or as an expanding T-region (but these modes are incompatible !). It depends on boundary or initial conditions which type of T-region is realized in a specific situation. This aspect is treated in detail in the next section. The contracting T-region is usually denoted by T_-, and the expanding one by T_+.

2.6 Formation of a Black Hole in a Gravitational Collapse

2.6.1 Gravitational collapse

In this section we analyze the process of formation of a black hole as a result of the contraction of a spherical mass to a size less than r_g. In order to eliminate effects which are not directly relevant to the formation of a black hole and can only make the solution more difficult to obtain, we consider the contraction of a spherical cloud of matter at zero pressure $p = 0$ (dust cloud). There is no need then to include in the analysis the hydrodynamic phenomena due to the pressure gradient. All dust particles move along geodesics, being subjected only to the gravitational field. The solution of Einstein's equations for this case was obtained by Tolman (1934) (see Appendix B.6). In the solution below, the reference frame is comoving with the matter; that is, dust particles have constant R, θ, ϕ:

$$ds^2 = -c^2\, dT^2 + g_{11}(T,R)\, dR^2 + r^2(T,R)(d\theta^2 + \sin^2\theta\, d\phi^2)\,, \tag{2.6.1}$$

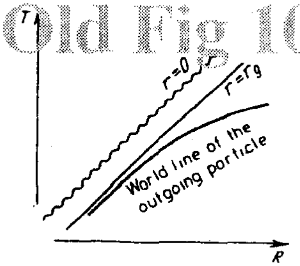


Figure 2.11: World line of a particle escaping from the Schwarzschild sphere, in contracting Lemaître coordinates.

sphere. Its world line in the contracting Lemaître coordinate system is given by the expression

$$\frac{2cT - R}{r_g} - 4\left[\frac{3(R - cT)}{2r_g}\right]^{1/3} + 2\ln\left\{\frac{\left[\frac{3}{2}(R - cT)\right]^{1/3} + r_g^{1/3}}{\left[\frac{3}{2}(R - cT)\right]^{1/3} - r_g^{1/3}}\right\} = \text{const} \qquad (2.7.1)$$

and is plotted in Figure 2.11. Continued into the past, this line asymptotically approaches the line $r = r_g$, without intersecting it. In the time T of the Lemaître frame the particle exists beginning from $T = -\infty$. But we know that the path from r_g to any finite r takes a finite interval of proper time. Therefore, Figure 2.11 does not cover the entire past of the particle in question from $\tau = -\infty$ in its proper clock. The history of a free particle does not terminate on the Schwarzschild sphere. The world line of such a particle must either continue indefinitely in its proper time or must terminate at the true singularity of spacetime, where new physical laws take over. Consequently, the map is incomplete and does not cover the entire space.

2.7.2 Complete empty spherically symmetric spacetime

Is it possible to construct an everywhere empty spacetime with an eternal black hole which is complete in the sense that it covers the histories of all particles moving in this space? The answer was found to be affirmative although it includes not only an eternal black hole but also an eternal white hole. In order to approach this construction in a natural way, consider a white hole with an expanding dust ball. Assume that the energy of motion of particles in the dust ball is such that the surface of the cloud does not escape to infinity but reaches a maximal radius and then again contracts to the size r_g and subsequently collapses to $r = 0$.

According to formula (2.3.7), the specific energy E/mc^2 of a particle on the surface of the dust cloud must be less than unity to ensure that $dr/dt = 0$ for a certain r. In Tolman's solution, (2.6.1)–(2.6.4), this expansion to a finite radius corresponds to the choice $f(R) < 0$. A qualitative representation of the spacetime with an expanding

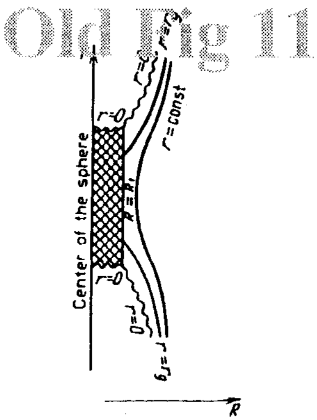


Figure 2.12: Expansion of a spherical cloud from within the Schwarzschild sphere, followed by contraction back into the sphere. The region inside the spherical cloud is hatched.

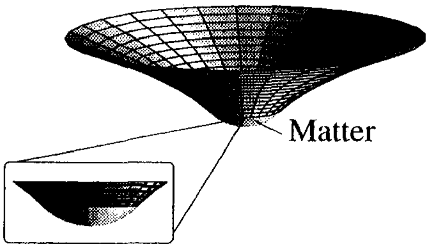


Figure 2.13: The embedding diagram for the geometry on the equatorial plane of the slice taken at the moment $\eta = \pi$ of maximal expansion of the dust ball in the spacetime shown in Figure 2.12.

and then contracting dust ball is shown in Figure 2.12. This spacetime first contains a white hole and then develops a black hole. Note that the lines $r = 0$ and $r = r_g$ are represented in this solution on this figure by lines which are not straight as we had in the case of motion at a parabolic velocity (see Figures 2.7 and 2.9).

Let us begin to reduce the specific energy E/mc^2 of the particles on the surface of the dust ball, assuming the total mass M of the ball, and hence the value of r_g to be fixed. In other words, by reducing E/mc^2, we reduce the share of the kinetic energy of outwards motion in the total energy Mc^2 of the dust ball. As a result, the cloud will expand to gradually smaller radii. Finally, the cloud expands to $r = r_g$ when $E/mc^2 = 0$. A qualitative map of the spacetime is then seen in Figure 2.14.

What if the constant E is reduced further and made negative? At first glance, this is physically meaningless; formally, it leads to enhanced maximum expansion radius r which is found by equating dr/dt in (2.3.7) to zero. Actually, there is nothing meaningless in this operation. In order to clarify the situation, let us look again at formulas (2.6.1)–(2.6.4) (see also Appendix B.6). Assume that the dust cloud whose evolution we analyze is homogeneous. Then the spacetime metric within the cloud corresponds to the metric of a homogeneous isotropic Universe. In the solution (2.6.1)–(2.6.4), this metric corresponds to choosing the functions

$$f(R) = -\sin^2 R, \tag{2.7.2}$$

$$F(R) = a\sin^3 R, \tag{2.7.3}$$

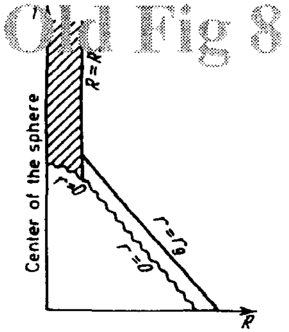


Figure 2.9: Expansion of a spherical cloud from within the Schwarzschild sphere in expanding Lemaître coordinates.

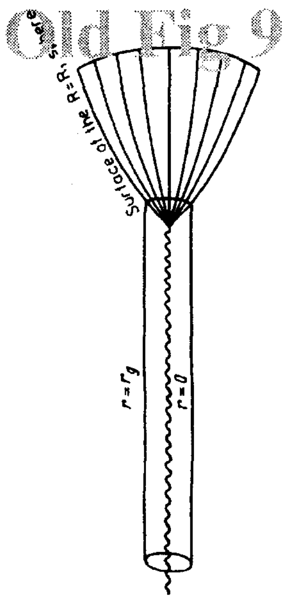


Figure 2.10: Expansion of a spherical cloud from within the Schwarzschild sphere in Eddington-Finkelstein coordinates.

of a spherical non-rotating body thus generates a spherical black hole.

Note now that the assumption made above on the absence of pressure produces no qualitative changes in the picture of the birth of a spherical black hole. We find the same behavior in the general case of the contraction of a sphere at a nonzero pressure ($p \neq 0$). When the surface of the contracting spherical cloud approaches the Schwarzschild sphere, no pressure can prevent the formation of the black hole [for details, see Zel'dovich and Novikov (1971b)]. These aspects connected with the role of pressure do not directly concern us here so the specifics are omitted.

The gravitational collapse produces a contracting T_{-}-region within the Schwarzschild sphere. We know that r on the boundary surface (or the coefficient g_{22} of the angular spatial term of the metric) decreases with time. But the regularity of the spacetime requires that r is a continuous function in the vicinity of the boundary. Therefore, r will decrease, owing to the continuity, also outside the spherical cloud (for $r < r_g$). The region within the $r = r_g$ sphere is the contracting T_{-}-region.

2.6.2 White holes

What are the conditions necessary for the formation of an expanding T_+-region? Reversing the time arrow in Figures 2.7 and 2.8, we obtain Figures 2.9 and 2.10. They represent the expansion of a spherical cloud from within the Schwarzschild sphere. Now the continuity of g_{22} at the boundary of the sphere implies that the vacuum spacetime beyond the spherical cloud but within the Schwarzschild sphere $r = r_g$ contains the expanding T_+-region. Figure 2.9 clearly illustrates the general situation. Recall that the $r = 0$ line is spacelike so that a reference frame exists in which all events on this line are simultaneous. Therefore, one cannot say (as one would be tempted to conclude at first glance) (see Figures 2.9 and 2.10) that *first* the singularity $r = 0$ existed in vacuum, and *then* the matter of the spherical cloud began to expand from the singularity. These events are not connected by a timelike interval. It is more correct to say that the nature of the spacelike singularity at $r = 0$ is such that it produces the expansion in vacuum (expanding T_+-region) to the right of R_1 and the expanding matter of spherical cloud to the left of R_1 (see Figure 2.9). Note that no particle coming from the spatial infinity (or from any region of $r > r_g$) can penetrate the expanding T_+-region. Such regions of spacetime are called *white holes* [Novikov (1964b), Ne'eman (1965)]. These objects cannot appear in the Universe as a result of collapse of some body, but could be formed, in principle, in the expanding Universe at the moment the expansion set in. This range of problem is discussed in detail in Section 15.2.

To conclude the section, we again emphasize that it is mathematically impossible to extend the solution beyond the true spacetime singularity at $r = 0$. Therefore, general relativity cannot answer the question what will happen after the contraction to $r = 0$ in a T_--region, or what was there before the start of the expansion from $r = 0$ in a T_+-region (or even say whether these questions are correctly formulated). It is physically clear that in the neighborhood of $r = 0$ quantum processes become essential for the spacetime itself (this effect is not described by general relativity); we return to this phenomenon in Chapter 14.

2.7 Eternal Black and White Holes

2.7.1 Incompleteness of the Lemaître frame

At first glance, it seems that an *eternal black hole* might exist perpetually in empty space in the form shown in Figures 2.2 and 2.4. Such a spacetime would contain a Schwarzschild sphere but would have no contracting material cloud. Surprisingly, the existence of such a "pure" eternal black hole is forbidden in principle. The reason is as follows. The picture (or rather, map) of the spacetime shown in Figure 2.2 (or Figure 2.4) does not cover the entire spacetime. In order to demonstrate this, consider a particle which moves freely along the radius away from the Schwarzschild

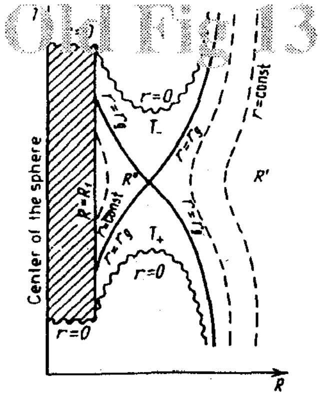


Figure 2.16: Expansion and contraction of a semi-closed world.

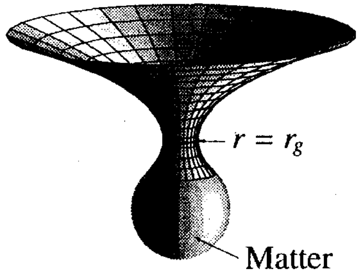


Figure 2.17: The corresponding embedding diagram for the moment of maximal expansion of the dust ball.

As $R_1 \to \pi$, the boundary R_1 gradually shifts to the left in Figure 2.16, leaving a progressively greater fraction of R'' free. The ratio $r_{\max}/r_g$ tends to infinity, and the ratio M/M_*, to zero. In the limit $R_1 = \pi$, the region occupied by matter vanishes, leaving the entire spacetime empty (Figure 2.18). It contains a white hole T_+, a black hole T_-, and two identical outer spaces R' and R'' which are asymptotically flat at their spatial infinities. This spacetime is complete in the sense that any geodesic now either continues infinitely or terminates at the true singularity.

The reference frame covering the entire spacetime in Figure 2.18 is described by a solution of type (2.7.6), where now it is convenient to place the origin of R at the minimum of the function $f(R)$. Then

$$f(R) = -\frac{1}{R^2+1}, \qquad F = r_g, \qquad -\infty < R < \infty. \tag{2.7.12}$$

The complete, everywhere empty spacetime shown in Figure 2.18 was first constructed by Synge (1950) and then by Fronsdal (1959), Kruskal (1960), and Szekeres (1960). The physical arguments given above and the solution (2.7.12) were obtained by Novikov (1962b, 1963).

We have thus obtained an everywhere empty space with a white and a black holes (essentially together). These holes can be described as "eternal" because for distant observers which are at rest in R' and R'' these holes are eternal. The physical meaning of the second "outer space" R'' has become clear above, where we described the evolution of a spherical cloud with progressively lower specific energy M/M_*. In Section 15.2 we discuss whether eternal black and white holes (similar to those

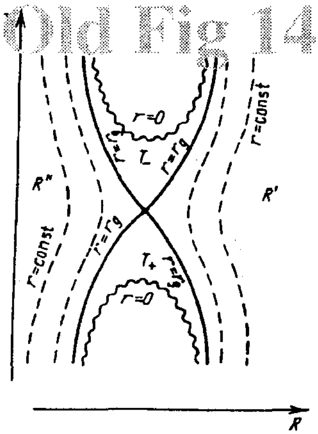


Figure 2.18: Everywhere empty spacetime with a white hole and a black hole.

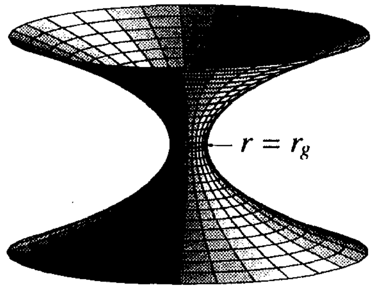


Figure 2.19: The corresponding embedding diagram for the moment of maximal expansion of the reference frame. This geometry is known as the *Einstein-Rosen bridge.*

in Figure 2.18) completely devoid of matter can really exist; this is related to the problem of stability of white holes.

2.7.3 Kruskal coordinates

To conclude this section, we give the coordinate system suggested by Kruskal (1960) and Szekeres (1960) (see Appendix B.5). Like the system (2.7.12), this one covers the entire spacetime of eternal white and black holes. In these coordinates, the metric is written in the form

$$ds^2 = \frac{4r_{\mathrm{g}}^3}{r}\, e^{-(r/r_{\mathrm{g}}-1)}\,(-d\tilde{T}^2 + d\tilde{R}^2) + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,, \tag{2.7.13}$$

where r is a function of $\tilde{T}$ and $\tilde{R}$:

$$\left(\frac{r}{r_{\mathrm{g}}} - 1\right) e^{(r/r_{\mathrm{g}}-1)} = \tilde{R}^2 - \tilde{T}^2\,. \tag{2.7.14}$$

In these coordinates the R' (R'') region is defined by the condition $\tilde{R} > |\,\tilde{T}\,|$ ($-\tilde{R} > |\,\tilde{T}\,|$), while the T_- (T_+)-region is defined by the condition $\tilde{T} > |\,\tilde{R}\,|$ ($-\tilde{T} > |\,\tilde{R}\,|$). The curvature singularity $r = 0$ is located at $\tilde{T}^2 - \tilde{R}^2 = e^{-1}$.

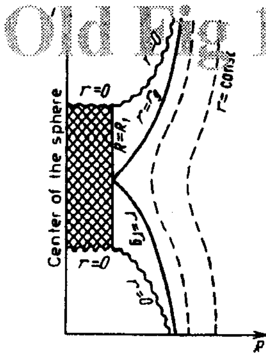


Figure 2.14: The boundary of the spherical cloud expands only up to the Schwarzschild sphere and then contracts.

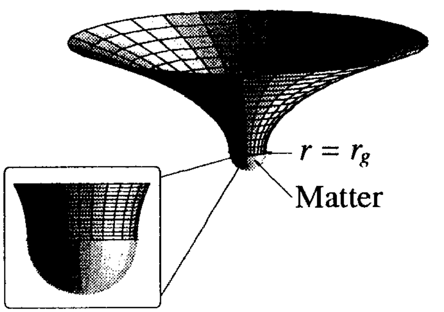


Figure 2.15: The corresponding embedding diagram for the moment of maximal expansion of the dust ball.

where a is the scale factor determined by the density ρ_0 within the spherical cloud at the moment of its maximum expansion:

$$a = \sqrt{\frac{3c^2}{8\pi G\rho_0}}. \qquad (2.7.4)$$

The matter of the cloud extends from $R = 0$ up to the boundary value of the coordinate $R = R_1$. The value R_1 may be in the range $0 < R_1 \leq \pi$.

The time evolution of the cloud can be presented in the following parametric form

$$r = \frac{a}{2}\sin R\,(1 - \cos\eta)\,, \qquad \tau = \frac{a}{2}(\eta - \sin\eta)\,. \qquad (2.7.5)$$

At the moment $\tau = 0$ of proper time τ in the comoving frame the cloud begins its expansion from $r = 0$. The cloud reaches its maximum size at $\tau = a\pi/2$, and at this moment $r = a\sin R$. After this the cloud collapses and all its particles simultaneously reach $r = 0$ at $\tau = a\pi$.

In the vacuum outside the cloud (at $R > R_1$), the particles that constitute the reference frame move freely along radial geodesics. The metric is determined by the following functions [Novikov (1963, 1964a)] (see also Appendix B.6):

$$f(R) = -\frac{1}{(R + \cot R_1 - R_1)^2 + 1}\,, \qquad F(R) = r_g\,. \qquad (2.7.6)$$

In this situation,

$$r_g \equiv \frac{2GM}{c^2} = a\sin^3 R_1\,. \qquad (2.7.7)$$

The quantity M (the gravitational mass) characterizes the total energy of the particles in the cloud, the gravitational energy included. It can be measured at far distances from a gravitating system by registering the acceleration of test particles.

In the general case the gravitational mass M differs from the proper mass M_* defined as the sum of the masses of the particles making up the cloud. The latter equals the product of density by the ball volume:

$$M_* = \frac{3}{4}\frac{ac^2}{G}\left(R_1 - \frac{1}{2}\sin 2R_1\right). \tag{2.7.8}$$

If the boundary coordinate R_1 lies in the range $\pi/2 < R_1 < \pi$, the inner region of the sphere is the so-called *semi-closed world* [Klein (1961), see also Novikov (1962a, 1963, 1964a), Zel'dovich (1962c), Zel'dovich and Novikov (1975, 1983)]. In these conditions an increase in R_1 (addition of new layers of matter) increases M_* but diminishes M (because of a large gravitational defect of mass).

Our objective is to analyze the sequence of spacetimes when we supply progressively smaller and smaller specific energy to the cloud particles. This means that we will take progressively smaller ratios M/M_*. In order to find the result of this change, we can take different values of the ratio M/M_*, while fixing either M or M_*. This choice is of methodological significance only. When we are interested in the metric outside the dust cloud, we fix M which determines the outer metric.

The ratio M/M_* is determined by (2.7.7) and (2.7.8):

$$\frac{M}{M_*} = \frac{2}{3}\sin^3 R_1 \left(R_1 - \frac{1}{2}\sin 2R_1\right)^{-1}. \tag{2.7.9}$$

The ratio of the maximum radius of expansion of the cloud boundary , $r_{\max}$, to the gravitational radius r_g is:

$$\frac{r_{\max}}{r_g} = (\sin R_1)^{-2}. \tag{2.7.10}$$

When $R_1 \ll \pi/2$, the ratio M/M_* is only slightly less than unity; a qualitative picture of the evolution is shown in Figure 2.12 ($r_{\max}/r_g \gg 1$). If $R_1 = \pi/2$, then

$$\frac{M}{M_*} = \frac{4}{3\pi} \quad \text{and} \quad r_{\max}/r_g = 1; \tag{2.7.11}$$

the corresponding situation is shown in Figure 2.14. If $R_1 > \pi/2$, the dust ball is a semi-closed world, and the ratio M/M_* decreases as R_1 approaches π. Now the metric looks as shown in Figure 2.16. In the latter case a qualitatively new feature has emerged. The ratio $r_{\max}/r_g$ is again greater than unity. But now the boundary of the cloud does not emerge from under the sphere of radius r_g into the space of the distant observer R'. A new region R'' has appeared outside the spherical cloud which is perfectly identical to R' .

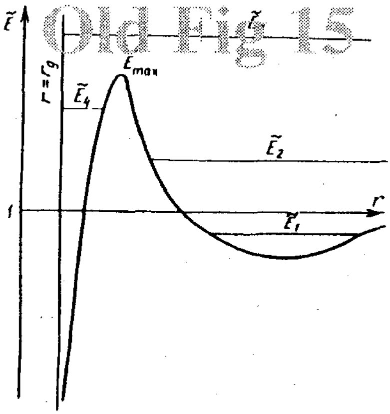


Figure 2.21: Effective black hole potential.

The qualitative features of the motion are revealed in the following way. Setting dr/dt equal to zero, we find the turning points; that is, the points of maximum approach of the particle to the black hole and the maximum distance from it. The right-hand side (2.8.1) vanishes when the equality

$$\widetilde{E}^2 = V^2(r)\,, \qquad V^2(r) = (1 - r_g/r)\left(1 + \tilde{L}^2 r_g^2/r^2\right) \tag{2.8.3}$$

is satisfied. The function $V(r)$ is called the *effective potential.* A typical form of the effective potential $V(r)$ for a fixed $\tilde{L}$ is plotted in Figure 2.21.

The specific energy of the moving particle remains constant; in Figure 2.21 this motion is shown by a horizontal line. Since the numerator of (2.8.1) must be positive, the horizontal segment representing the motion of the particle lies above the curve representing the effective potential. The intersection of the horizontal line with the effective potential determines the turning points. Figure 2.21 shows horizontal lines for typical motions. The horizontal line $\widetilde{E}_1 < 1$ corresponds to the motion in a bounded region in space between r_1 and r_2; this is an analogue of elliptic motion in Newtonian theory (an example of such a trajectory is shown in Figure 2.22a). It should be stressed that the corresponding trajectory is not a conic section, and, in general, is not closed. If the orbit as a whole lies far from the black hole, it is an ellipse which slowly rotates in the plane of motion.

The segment $\widetilde{E}_2 > 1$ corresponds to a particle arriving from infinity and again moving away to infinity (an analogue of the hyperbolic motion). An example of this trajectory is plotted in Figure 2.22b. Finally, the segment $\widetilde{E}_3$ does not intersect the

that $v^2/c^2 = A^{-2}\dot{r}^2 + A^{-1}r^2\dot{\phi}^2$, where $A = 1 - r_g/r$, and $\dot{r} = dr/c\,dt$. Note also that the physical velocity v of the particle, measured by a local observer by his clock, τ, is directly related to the specific energy $\widetilde{E}$: $\widetilde{E}^2 = (1 - r_g/r)(1 - v^2/c^2)^{-1}$.

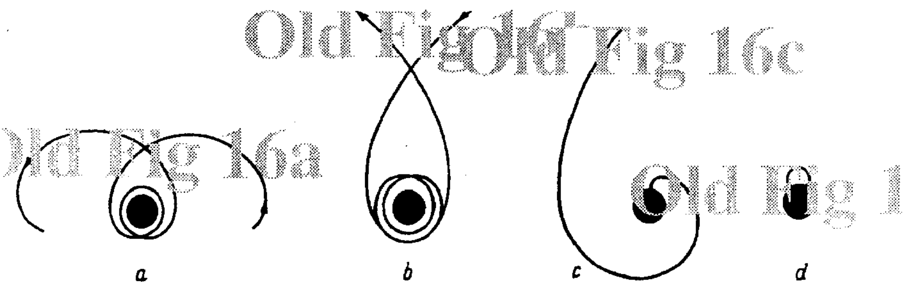


Figure 2.22: Trajectories of particles with energies (a) $\widetilde{E}_1$, (b) $\widetilde{E}_2$, (c) $\widetilde{E}_3$, and (d) $\widetilde{E}_4$.

potential curve but passes above its maximum $\widetilde{E}_{\max}$. It corresponds to a particle falling into the black hole (*gravitational capture*). This type of motion is impossible in Newtonian theory and is characteristic for the black hole. The trajectory of this motion is shown in Figure 2.22c. Gravitational capture becomes possible because the effective potential has a maximum. No such maximum appears in the effective potential of Newtonian theory.

In addition, another type of motion is possible in the neighborhood of a black hole; namely, a motion corresponding to the horizontal segment $\widetilde{E}_4$ in Figure 2.21. This line may lie below or above unity (in the latter case, for $\widetilde{E}_{\max} > 1$), stretching from r_g to the intersection with the curve $V(r)$. This segment represents the motion of a particle which first recedes from the black hole and reaches $r_{\max}$ [at the point of intersection of $\widetilde{E}_4$ and $V(r)$], and then again falls toward the black hole and is absorbed by it (Figure 2.22d).

A body can escape to infinity if its specific energy $\widetilde{E} \geq 1$. From the equation $\widetilde{E}^2 = (1 - r_g/r)(1 - v^2/c^2)^{-1} = 1$ we find that the escape velocity $v_{\rm esc}$ is

$$v_{\rm esc} = c\sqrt{r_g/r} = \sqrt{2GM/r}\,, \tag{2.8.4}$$

which coincides with the expression given by Newtonian theory.

Note that in Newtonian theory, in the gravitational field of a point-like mass, the escape velocity guarantees escape to infinity regardless of the direction of motion. The case of the black hole is different. Even if a particle has the escape velocity, it can be trapped by the black hole (trajectories of type $\widetilde{E}_4$ or $\widetilde{E}_3$ in Figure 2.21, the latter occurring if the particle moves towards the black hole). We have already mentioned this effect called *gravitational capture*.

2.8.2 Circular motion

Circular motion around a black hole is an important particular case of motion of a particle, in which $dr/dt = 0$. This motion is represented in Figure 2.21 by a point at

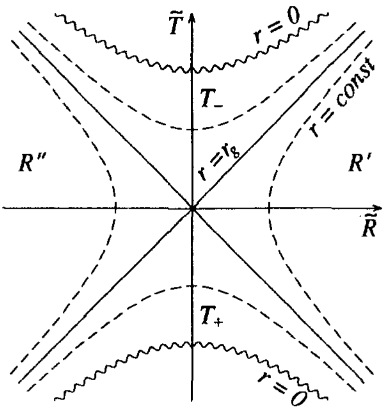


Figure 2.20: Spherically symmetric vacuum spacetime in Kruskal coordinates.

The following formulas give the relation between the $(\tilde{T}, \tilde{R})$-coordinates and (r, t)-coordinates in the regions R' and T_-:

$$\text{In } R' \quad (\text{for} \quad r > r_g): \quad \begin{cases} \tilde{R} = (r/r_g - 1)^{1/2} \, e^{(r-r_g)/2r_g} \cosh(ct/2r_g), \\ \tilde{T} = (r/r_g - 1)^{1/2} \, e^{(r-r_g)/2r_g} \sinh(ct/2r_g), \end{cases} \tag{2.7.15}$$

$$\text{In } T_- \quad (\text{for} \quad r < r_g): \quad \begin{cases} \tilde{R} = (1 - r/r_g)^{1/2} \, e^{(r-r_g)/2r_g} \sinh(ct/2r_g), \\ \tilde{T} = (1 - r/r_g)^{1/2} \, e^{(r-r_g)/2r_g} \cosh(ct/2r_g). \end{cases} \tag{2.7.16}$$

Similar relations in the regions R'' and T_+ are obtained by the change of variables $\tilde{R} \to -\tilde{R}$, $\tilde{T} \to -\tilde{T}$. The convinience of the Kruskal coordinate system lies in that the radial null geodesics are always represented by straight lines inclined at an angle of 45^0 to coordinate axes (see Figure 2.20).

Quite often null coordinates $U = \tilde{T} - \tilde{R}$ and $V = \tilde{T} + \tilde{R}$ are used instead of $(\tilde{R}, \tilde{T})$. In these coordinates the Kruskal metric takes the form

$$ds^2 = -\frac{4r_g^3}{r} \, e^{-(r/r_g - 1)} \, dV \, dU + r^2(d\theta^2 + \sin^2\theta \, d\phi^2), \tag{2.7.17}$$

$$-UV = \left(\frac{r}{r_g} - 1\right) \exp\left(\frac{r}{r_g} - 1\right). \tag{2.7.18}$$

Usually the Kruskal metric is obtained as a result of the analytical continuation of the Schwarzschild metric (see, e.g., Misner, Thorne, and Wheeler (1973) and Appendix B.5.3). The direct proof that the Kruskal metric (2.7.17)–(2.7.18) is a spherically symmetric vacuum solution of the Einstein equations can be found in the Appendix B.5.1.

It is easy to see that the relation $r = r_g$ in Kruskal coordinates implies that either $U = 0$ or $V = 0$. The null surface $U = 0$ which separates the exterior region R' and the black hole region T_- is called the *event horizon*. The null surface $V = 0$ which separates the exterior region and white hole region is called the *past horizon*.

2.8 Celestial Mechanics in the Gravitational Field of the Black Hole

2.8.1 Equations of motion of a free test particle

Let us return to the discussion of processes in the black hole exterior, in the space outside the Schwarzschild sphere. In this section, we consider the motion of test particles along geodesics in the gravitational field of a black hole. These phenomena were analyzed in detail a long time ago and included in text books and monographs [see e.g., Zel'dovich and Novikov (1971b), Misner, Thorne, and Wheeler (1973)]. Here we will briefly discuss those features of the motion which are specific for black holes, not just for strong gravitational fields (say, the field around a neutron star).

As before, we describe the motion of particles with respect to the Schwarzschild reference frame, clocked by an observer at infinity (Section 2.2). The gravitational field being spherically symmetric, the trajectory of a particle is planar; we can assume it to lie in the plane $\theta = \pi/2$. The equations of motion have the form

$$\left(\frac{dr}{c\,dt}\right)^2 = \frac{(1-r_g/r)^2[\widetilde{E}^2 - (1-r_g/r)(1+\tilde{L}^2 r_g^2/r^2)]}{\widetilde{E}^2}\,, \tag{2.8.1}$$

$$\frac{d\phi}{c\,dt} = \frac{(1-r_g/r)\tilde{L}\,r_g}{\widetilde{E}\,r^2}\,. \tag{2.8.2}$$

Here $\widetilde{E}$ is the specific energy of a particle ($\widetilde{E} = E/mc^2$, E being the energy, and m being the mass of the particle). $\tilde{L}$ is the specific angular momentum ($\tilde{L} = L/mcr_g$, L being the angular momentum). These two quantities are conserved in the course of motion.[7]

[7] As we discussed earlier, the energy E is conserved because the gravitational field is static. The conserved energy can be rewritten as $E = mc^2\sqrt{-g_{00}}\,/\sqrt{1-v^2/c^2}$, where v is the physical three-velocity, defined as the rate of change of the physical distance in the physical time: $v^2 = g_{ij}\,dx^i\,dx^j/(-g_{00})dt^2$. The spherically symmetric metric does not depend on the angular coordinate ϕ ($\partial g_{\mu\nu}/\partial\phi = 0$). The conservation of the angular momentum L is the consequence of this symmetry generated by the Killing vector $\xi_{(\phi)}$: $\xi^\mu_{(\phi)} = \delta^\mu_\phi$. The conserved angular momentum is $L = mu_\mu\,\xi^\mu_{(\phi)} = mcu_\phi = mc^2r^2\dot{\phi}\,(-g_{00})^{-1/2}\,(1-v^2)^{-1/2}$, where $\dot{\phi} = d\phi/c\,dt$. Equation (2.8.2) directly follows from this relation. Equation (2.8.1) can be easily obtained from the expression for the energy if we note

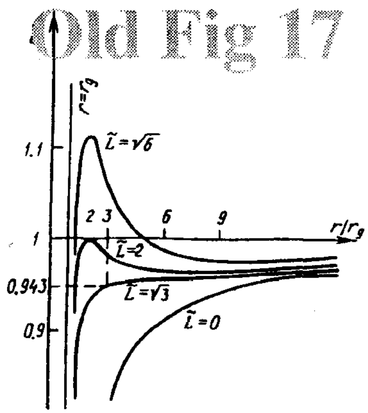


Figure 2.23: Effective potentials for different values of $\tilde{L}$.

the extremum of the effective potential curve. A point at the minimum corresponds to a stable motion; and that at the maximum, to an unstable one. The latter motion has no analogue in Newtonian theory. It is specific to black holes. Generally speaking, motion along an unstable trajectory is impossible. However, if the motion of a particle is represented by a horizontal line $\widetilde{E}$ = const very close to $\widetilde{E}_{\rm max}$, then the particle makes many turns around the black hole at radii close to r corresponding to $\widetilde{E}_{\rm max}$ before the orbit moves far away from this value of r. An example of this motion is shown by the orbit of Figure 2.22b. The shape and position of the potential $V(r)$ are different for different $\tilde{L}$: the corresponding curves for some values of $\tilde{L}$ are shown in Figure 2.23.

The maximum and minimum appear on the $V(r)$ curves when $\tilde{L} > \sqrt{3}$. If $\tilde{L} < \sqrt{3}$, the $V(r)$ curve is monotone. Hence, motion on circular orbits is possible only if $\tilde{L} > \sqrt{3}$. The minima of the curves then lie at $r > 3r_{\rm g}$. Stable circular orbits thus exist only for $r > 3r_{\rm g}$ [Hagihara (1931)]. At smaller distance, there are only unstable circular orbits corresponding to the maxima of the $\widetilde{E}_{\rm max}$ curves. If $\tilde{L} \to \infty$, the coordinates of the maxima on the $\widetilde{E}_{\rm max}$ curves decrease to $r = 1.5\,r_{\rm g}$. Even unstable inertial circular motion becomes impossible for r less than $1.5\,r_{\rm g}$.[8]

The critical circular orbit that separates stable motions from unstable ones corresponds to $r = 3r_{\rm g}$. Particles move along it at a velocity $v = c/2$, the energy of a particle being $\widetilde{E} = \sqrt{8/9} \approx 0.943$. This is the motion with the maximum possible binding energy $E \approx 0.057\,mc^2$. The velocity of motion on (unstable) orbits with $r < 3r_{\rm g}$ increases as r decreases, from $c/2$ to c on the last circular orbit with $r = 1.5\,r_{\rm g}$. When $r = 2r_{\rm g}$, the particle's energy is $\widetilde{E} = 1$; that is, the circular velocity

[8] If there is an external non-gravitational force acting on a particle, then it could move along a circular orbit inside $r = 1.5\,r_{\rm g}$. This type of motion will be considered in Section 2.8.4.

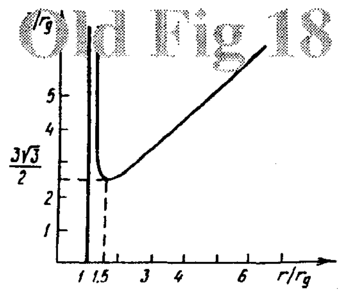


Figure 2.24: The position of extrema in r on the trajectory of an ultra-relativistic particle as a function of impact parameter b.

is equal to the escape velocity. If r is still smaller, the escape velocity is smaller than the circular velocity. There is no paradox in this since the circular motion here is unstable and even the tiniest perturbation (supplying momentum away from the black hole) transfers the particle to an orbit removing it to infinity; that is, an orbit corresponding to hyperbolic motion.

2.8.3 Motion of an ultra-relativistic particle

Let us consider the motion of an ultra-relativistic particle. In (2.8.1), (2.8.2) it corresponds to the limit $v \to c$, so that $\tilde{E} \to \infty$ and $\tilde{L} \to \infty$. One must keep in mind that in this limit the ratio $\tilde{L}/\tilde{E}$ is equal to $b/r_{\rm g}$, where b is the impact parameter of particle at infinity. In view of this remark, we obtain, instead of (2.8.1), (2.8.2),

$$\left(\frac{dr}{c\,dt}\right)^2 = \left(1 - \frac{r_{\rm g}}{r}\right)^2 \left[1 - \frac{b^2}{r^2}\left(1 - \frac{r_{\rm g}}{r}\right)\right], \tag{2.8.5}$$

$$\frac{d\phi}{c\,dt} = \left(1 - \frac{r_{\rm g}}{r}\right)\frac{b}{r^2}. \tag{2.8.6}$$

Formulas (2.8.5), (2.8.6) describe the bending of the trajectories of an ultra-relativistic particle and a light beam moving close to the black hole. Setting the expression in the square brackets in (2.8.5) equal to zero, we find the relation between the position of the radial turning point on the trajectory and the impact parameter b. The corresponding $b(r)$ curve is shown in Figure 2.24. The sign of b depends on the sense of motion; we assume that b is positive. In this figure the motion of an ultra-relativistic particle with a given b is represented by a horizontal line $b = \text{const}$. A particle approaches the black hole, passes by it at the minimal distance corresponding to the point of intersection of $b = \text{const}$ with the right-hand branch of the $b(r)$ curve, and again recedes to infinity. If the intersection occurs close to the minimum $b_{\min} = 3\sqrt{3} \times r_{\rm g}/2$, the particle may go through a large number of turns before it flies away to infinity. The exact minimum of the curve $b(r)$ corresponds to (unstable) motion on a circle of radius $r = 1.5\,r_g$ at the velocity $v = c$. Note that the left-hand branch of $b(r)$ in Figure 2.24 corresponds to the maximum distance between the ultra-relativistic particle and the black hole; the particle first recedes to $r < 1.5\,r_g$ but then again falls toward the black hole. Obviously, for such a motion the parameter b does not have the literal meaning of an impact parameter at infinity since the particle never recedes to infinity. For a given coordinate r, this parameter can be found as a function of the angle ψ between the trajectory of the particle and the direction to the center of the black hole:

$$b = \frac{r|\tan\psi|}{\sqrt{(1 - r_g/r)(1 + \tan^2\psi)}}. \tag{2.8.7}$$

If an ultra-relativistic particle approaches the black hole on the way from infinity and the parameter b is less than the critical value $b_{\rm min} = 3\sqrt{3}\,r_g/2$, this particle falls into the black hole.

2.8.4 Non-inertial circular motion

We consider now another special type of motion; namely, a circular motion of a test particle around a black hole with a constant angular velocity $\Omega \equiv u^\phi/u^t$, where u^ϕ and u^t are non-vanishing constant components of the four-velocity of the particle. This angular velocity can have an arbitrary value and may differ from the Keplerian value Ω_K calculated for the chosen radius of the orbit. In the general case such a motion is not geodesic (non-inertial), and in order to support the particle in the orbit, there must be a non-gravitational force acting on the particle. Such a motion could probably occur in some astrophysical phenomena (for example, in disk accretion onto a relativistic black hole) as well as in gedanken experiments with astronauts moving near a black hole.

The four-acceleration $w^\mu = u^\nu u^\mu{}_{;\nu}$ for a circular motion can be easily calculated. The radial component of the acceleration w_r is

$$w_r = \frac{GM/r^2 - \Omega^2 r}{1 - 2GM/c^2 r - \Omega^2 r^2/c^2}. \tag{2.8.8}$$

Note that the physical value of the circular velocity is

$$v_{\rm circ} \equiv \frac{r\,d\phi}{\sqrt{-g_{00}}\,dt} = \frac{\Omega r}{(1 - 2GM/c^2 r)^{1/2}}. \tag{2.8.9}$$

This velocity satisfies the relation $v_{\rm circ} \le c$. Thus, the denominator in (2.8.8) is always positive. It tends to zero when $v_{\rm circ} \to c$. The circular motion along the geodesic line corresponds to the case $w_r = 0$ (free motion). This gives the expression for the Keplerian angular velocity Ω_K

$$\Omega_K \equiv \left(\frac{GM}{r^3}\right)^{1/2}, \tag{2.8.10}$$

and the physical value of the Keplerian velocity is

$$v_K = \left(\frac{GM}{r}\right)^{1/2}\left(1 - \frac{2GM}{c^2 r}\right)^{-1/2} . \tag{2.8.11}$$

From this equation and the inequality $v_K \leq c$ one can conclude that motion with the Keplerian circular velocity is possible only in the region $r \geq 1.5r_g$. We already mentioned this fact in Section 2.8.2. If $\Omega < \Omega_K$, then an outward force (for example, rocket thrust) is needed to hold the particle on a circular orbit. For orbits located at $r > 1.5r_g$ an increase of Ω leads to decrease of w_r. This is in accordance with our intuition: The increase of Ω results in increase of the centrifugal force, and hence in decrease of the required rocket thrust.

Abramowicz and collaborators [Abramowicz and Lasota (1986), Abramowicz and Prasanna (1990), Abramowicz (1990, 1992)] drew attention to the fact that in the region $r < 1.5r_g$ the situation is different: An increase of Ω requires *increase* of the needed rocket thrust. Indeed, differentiation of (2.8.8) with respect to Ω gives

$$\frac{\partial w_r}{\partial \Omega} = -\frac{2\Omega r(1 - 3GM/c^2 r)}{(1 - 2GM/c^2 r - \Omega^2 r^2/c^2)^2} . \tag{2.8.12}$$

Thus, for $r > 1.5r_g$ the signs of variations of w_r and Ω are opposite, while for $r < 1.5r_g$ they are the same. Abramowicz and co-authors interpreted this result in terms of reversal of the action of the centrifugal force at $r = 1.5r_g$.

Other authors [for example, see de Felice (1991, 1994), Page (1993a), Semerák (1994), and Barrabès, Boisseau, and Israel (1995)] gave another interpretation. They emphasize that the effect depends crucially on the dependence of the denominator in the expression (2.8.8) on Ω. This dependence describes the special-relativistic effect of increase of mass with velocity. This increase, literally speaking, outweighs the effect of the centrifugal force in the region $r < 1.5r_g$. Of course, the effect itself does not depend on the particular interpretation, but still we believe that the latter interpretation is more physical. The following arguments are in favor of this point of view.

Centrifugal force in a flat spacetime is absent for a body moving along a straight line. It is reasonable to define centrifugal force in a curved space in such a way that it also vanishes when a body is moving with a constant velocity along a three-dimensional geodesic, so that

$$\frac{d^2x^i}{dl^2} + \Gamma^i_{jk}\frac{dx^j}{dl}\frac{dx^k}{dl} = 0 , \tag{2.8.13}$$

where l is a proper length. For such a motion in a static spacetime with the line element

$$-c^2\, d\tau^2 = -Fc^2\, dt^2 + dl^2 \tag{2.8.14}$$

the spatial components of the four-acceleration w^μ can be written as

$$w^i = \Gamma^i_{00}\left(\frac{dt}{d\tau}\right)^2 + \frac{dx^i}{cd\tau}\frac{d^2l}{d\tau^2}\left(\frac{dl}{d\tau}\right)^{-1} . \tag{2.8.15}$$

Define a three-dimensional (coordinate) velocity $dl/dt \equiv v$, then using the relations

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{F-(v/c)^2}}, \qquad \frac{dl}{d\tau} = \frac{v}{\sqrt{F-(v/c)^2}}, \tag{2.8.16}$$

we get

$$w^i = \frac{1}{2}\frac{\nabla^i F}{F-(v/c)^2} + \frac{v^i}{v}\frac{d}{d\tau}\frac{v}{\sqrt{F-(v/c)^2}} . \tag{2.8.17}$$

Here $v^i = dx^i/dt$. Now consider a motion along a three-dimensional geodesic with a constant proper velocity $dl/d\tau$. Then $d^2l/d\tau^2 = 0$, and the second terms on the right-hand sides of (2.8.15) and (2.8.17) vanish so that w^i takes the form

$$w^i = \frac{1}{2}\frac{\nabla^i F}{F-(v/c)^2} . \tag{2.8.18}$$

It is reasonable to interpret this acceleration as the "gravitational acceleration" in a static spacetime because there is neither a centrifugal acceleration (for motion along the three-dimensional geodesics) nor an inertial acceleration (for motion with constant proper velocity). By comparing the expression (2.8.18) with the relation (2.8.8), we can identify the term

$$w_{\text{centr},r} = -\frac{\Omega^2 r}{1-2GM/c^2r-\Omega^2r^2/c^2} \tag{2.8.19}$$

with the centrifugal acceleration. For this definition of the centrifugal acceleration we have

$$\frac{\partial w_{\text{centr},r}}{\partial\Omega} = -2\Omega r\,\frac{1-2GM/c^2r}{(1-2GM/c^2r-\Omega^2r^2/c^2)^2} \tag{2.8.20}$$

so that everywhere outside the gravitational radius the sign of $\partial w_{\text{centr},r}/\partial\Omega$ remains the same as in the usual non-relativistic theory.

2.9 Gravitational Capture

In this section we deal with a motion of a test particle such that its trajectory terminates in the black hole. Two types of such motion are possible. First, the trajectory of the particle starts at infinity and ends in the black hole; second, the trajectory starts and ends in the black hole. Of course, a particle cannot be ejected from the black hole. Hence, the motion on the second-type trajectory becomes possible either if the particle was placed on this trajectory via a non-geodesic curve or the particle was created close to the black hole.[9]

The gravitational capture of a particle coming from infinity is of special interest. Let us have a closer look at this case. It is clear from the analysis of motion given in the preceding section that a particle coming from infinity can be captured if its specific energy $\widetilde{E}$ is greater, for a given $\tilde{L}$, than the maximum $\widetilde{E}_{\max}$ on the curve $V(r)$. Let us consider the gravitational capture in two limiting cases, one for a particle whose velocity at infinity is much lower than the speed of light ($v_\infty/c \ll 1$) and the other for a particle which is ultra-relativistic at infinity.

In the former case, $\widetilde{E} \approx 1$. The curve $V(r)$, which has $\widetilde{E}_{\max} = 1$, corresponds to $\tilde{L}_{\rm cr} = 2$ (see Figure 2.23). The maximum of this curve lies at $r = 2r_g$. Hence, this radius is minimal for the periastra of the orbits of particles with $v_\infty = 0$ which approach the black hole and again recede to infinity. If $\tilde{L} \leq 2$, gravitational capture takes place. The angular momentum of a particle moving with the velocity v_∞ at infinity is $L = mv_\infty b$, where b is an impact parameter. The condition $\tilde{L} \equiv L/mcr_g = 2$ defines the critical value $b_{\rm cr,nonrel} = 2r_g(c/v_\infty)$ of the impact parameter for which the capture takes place. The capture cross-section for a non-relativistic particle is

$$\sigma_{\rm nonrel} = \pi b_{\rm cr}^2 = 4\pi (c/v_\infty)^2 r_g^2 . \tag{2.9.1}$$

For an ultra-relativistic particle, $b_{\rm cr} = 3\sqrt{3}\, r_g/2$, and the capture cross-section is

$$\sigma_{\rm rel} = \frac{27}{4}\,\pi\, r_g^2 . \tag{2.9.2}$$

Owing to a possible gravitational capture, not every particle whose velocity exceeds the escape limit flies away to infinity. In addition, it is necessary that the angle ψ between the direction to the black hole center and the trajectory be larger than a certain critical value $\psi_{\rm cr}$. For the velocity equal to the escape threshold this critical angle is given by the expression

$$\tan\psi_{\rm cr,esc} = \pm\,\frac{2\sqrt{(1-r_g/r)r_g/r}}{\sqrt{1-4r_g/r(1-r_g/r)}} . \tag{2.9.3}$$

The plus sign is chosen for $r > 2r_g$ ($\psi_{\rm cr} < 90^0$), and the minus sign is chosen for $r < 2r_g$ ($\psi_{\rm cr} > 90^0$).

For an ultra-relativistic particle, the critical angle is given by the formula

$$\tan\psi_{\rm cr,rel} = \pm\,\frac{\sqrt{1-r_g/r}}{\sqrt{r_g/r - 1 + \frac{4}{27}\,(r/r_g)^2}} . \tag{2.9.4}$$

The plus sign is taken for $r > 1.5\, r_g$, and the minus, for $r < 1.5\, r_g$.

[9] Of course, a particle may escape from a white hole and fall into a black one; examples of this kind were discussed in Section 2.7.

2.10 The Motion of Particles Corrected for Gravitational Radiation

In relativistic theory, celestial mechanics differs from Newtonian theory in an additional factor not yet discussed above: emission of gravitational waves by accelerated bodies. As a result, the energy $\widetilde{E}$ and angular momentum $\tilde{L}$ are not strictly integrals of the motion. The emission of gravitational waves decelerates the moving particle (it loses energy and angular momentum). The force of deceleration is related to the interaction of the test particle of mass m with its own gravitational field and is proportional to m^2, while the interaction with the external field is proportional to mM. Therefore, if m/M is small, the force of "radiative friction" is a small correction to the main force, and the motion of the test particle is almost indistinguishable from motion along a geodesic. Nevertheless, these small corrections may accumulate over long periods of time, and thus cause appreciable deviations of the motion from the initial trajectory.

Let us evaluate the change in the cross-section of the capture of a test particle approaching the black hole from infinity, taking into account the emission of gravitational waves and the process of gradual capture of the body circling the center [Zel'dovich and Novikov (1964a, 1971b)]. The gravitational radiation can be calculated by analyzing the small perturbations of the Schwarzschild metric. We shall discuss this type of calculations in detail in Chapter 4. Here we just mention that the analysis shows that an evaluation of the changes in motion within the framework of the weak field theory and for non-relativistic velocities gives a good approximation in all interesting cases.

Let us consider a particle approaching a black hole from infinity where its velocity was small. As a result of the loss of energy due to the gravitational radiation, which mostly occurs at the periastron of the orbit, the orbit changes. If the loss of energy is big enough, the particle cannot escape to infinity after it passes by the black hole (as would happen without radiation). It switches to a bound elongated orbit, which brings the particle back to the black hole. Emission will occur again at the periastron, etc., until the particle falls into the hole. After the first passage past the black hole at a distance r_1 at periastron, the particle recedes to a maximum distance $r_{\max}$ (apoastron) given by the approximate formula

$$r_{\max} \approx r_g/2 \left[\frac{m}{M} \left(\frac{r_g}{r_1} \right)^{3/5} - \frac{v_\infty^2}{2c^2} \right]^{-1} . \tag{2.10.1}$$

If r_1 is small, $r_{\max}$ rapidly decreases after subsequent passages until the particle falls into the black hole.

Taking this into consideration, one arrives at the following approximate formula for the capture cross-section of a particle of mass m and velocity v_∞ at infinity:

$$\sigma_{\text{grav,rad}} \approx \pi \left(\frac{c}{v_\infty}\right)^2 (2x)^{2/7} r_g^2 . \tag{2.10.2}$$

This formula is valid for $x \equiv (c^2/v_\infty^2)m/M \gg 2^6$. If $x \leq 2^6$, the cross-section practically coincides with (2.9.1) for non-relativistic particles.

Let us look now at the effect of gravitational radiation on the circular motion of particles (for more detailed discussion, see Section 4.10.1). If a particle moves at $r \gg r_g$, the gradual decrease of the orbital radius obeys the following law [Landau and Lifshitz (1975)]:

$$\frac{dr}{dt} = \frac{8}{5} c \left(\frac{m}{M}\right) \left(\frac{r_g}{r}\right)^3 . \tag{2.10.3}$$

This process lasts until the limiting stable circular orbit at $r = 3r_g$ is reached. At this orbit the binding energy $e \approx 0.057mc^2$ (see page 43). This energy is emitted during the entire preceding motion. The energy emitted by the particle during one revolution on the critical circle $r = 3r_g$ is $\delta e \approx 0.1\, mc^2(m/M)$. Then the particle slips into a spiraling infall into the black hole; this takes about $(M/m)^{1/3}$ additional revolutions. The amount of energy radiated away at this stage is much less than that lost before $r = 3r_g$ was reached.

2.11 Tidal Interaction of Extended Bodies with a Black Hole

2.11.1 Equations of motion

In the above consideration of the motion of test bodies in the gravitational field of a black hole up to now we neglected their size, and described such a body as a test point-like particle. In other words, we neglected all effects connected with the finite size of a body. In this section we consider the tidal interaction of an extended body (a star) with a black hole. We again assume that the mass of the body is small compared with the mass of the black hole so that its gravitational effects can be neglected. We suppose also that the backreaction of the tidal forces on the motion of the body (*tidal friction*) is small, and the trajectory of the center of mass of the extended body is practically a geodesic. We analyze properties of the relativistic tidal forces and compare them with the non-relativistic ones. The tidal interaction can be important for close encounters of stars with supermassive black holes in the centers of galaxies (see Section 9.6). Bearing in mind this application, we focus our attention on the case where the velocity of a test body at infinity is much smaller then it is during the close encounter. Under this condition, with high accuracy the motion of the body near the black hole can be described as "parabolic", for which the specific energy $\tilde{E} \equiv E/mc^2 = 1$. In the subsequent discussion we follow the work of Frolov *et al.* (1994).

To describe a motion of an extended test body moving in the Schwarzschild spacetime, we consider first a trajectory representing the motion of its center of mass. We choose coordinates in such a way that this trajectory lies in the equatorial plane $\theta = \pi/2$. Its equations are given by (2.8.1) and (2.8.2). Besides these equations, we shall use also the equation

$$\frac{dt}{d\tau} = \frac{\tilde{E}}{1 - r_g/r} \tag{2.11.1}$$

relating t and the proper time τ. As before, $\tilde{E}$ and $\tilde{L}$ are specific energy and specific angular momentum. For parabolic motion $\tilde{E} = 1$, and trajectories are characterized by one dimensionless constant $\tilde{L}$. It is convenient to introduce the following notations:

$$R_p = \tilde{L}^2 r_g\,, \qquad t_p = \tilde{L}^3 r_g/c\,, \qquad \mu = \tilde{L}^{-1} = (r_g/R_p)^{1/2}\,, \tag{2.11.2}$$

$$\tau^* = \tau/t_p\,, \qquad x(\tau^*) = r(\tau)/R_p\,, \qquad t^* = t/t_p\,. \tag{2.11.3}$$

In these notations the equations of motion take the form

$$\left(\frac{dx}{d\tau^*}\right)^2 = \frac{1}{x} - \frac{1}{x^2} + \mu^2 \frac{1}{x^3}\,, \tag{2.11.4}$$

$$\frac{d\phi}{d\tau^*} = \frac{1}{x^2}\,, \qquad \frac{dt^*}{d\tau^*} = \frac{1}{1 - \mu^2/x}\,. \tag{2.11.5}$$

We choose $\tau^* = 0$ at the point of periastron $r = r_p = R_p x_p$, where $x_p = \frac{1}{2}(1 + \sqrt{1 - 4\mu^2}\,)$. At this point $dx/d\tau^*$ changes sign from negative (for $\tau^* < 0$) to positive (for $\tau^* > 0$). We also choose $\phi(\tau^* = 0) = 0$ so that ϕ is negative (positive) for $\tau^* < 0$ ($\tau^* > 0$).

The parameter $\mu = (r_g/R_p)^{1/2}$ characterizes relativistic corrections. If it vanishes, $\mu = 0$, the equations of motion take their non-relativistic form, and R_p coincides with the radius of periastron in the non-relativistic problem.

The function $x(\tau^*)$ is even, while $\phi(\tau^*)$ is an odd function of τ^*. That is why it is sufficient to find their values for $\tau^* \geq 0$. The equations of motion can be integrated in terms of elliptic integrals. In particular, we have (for $\tau^* \geq 0$)

$$\tau^* = \frac{2}{3}\,\frac{\sqrt{x}\,\sqrt{x^2 - x + \mu^2}\,(x + 2 - \mu)}{x - \mu} + \frac{2}{3\sqrt{1 + 2\mu}}\,\left[(1 - \mu^2)\,F(\varphi\backslash\alpha) - (1 + 2\mu)\,E(\varphi\backslash\alpha)\right]\,, \tag{2.11.6}$$

$$\phi = \frac{2}{\sqrt{1 + 2\mu}}\,F(\varphi\backslash\alpha)\,, \tag{2.11.7}$$

where F and E are elliptic integrals of the first and second kind, respectively, and

$$\sin\varphi = \frac{\sqrt{x^2 - x + \mu^2}}{x - \mu}, \qquad \sin\alpha = \sqrt{\frac{4\mu}{1 + 2\mu}}. \tag{2.11.8}$$

Equation (2.11.7) determines the trajectory, while equation (2.11.6) determines the position of the body on the trajectory as the function of proper time.

2.11.2 Tidal forces

To determine tidal forces acting on an extended body moving in the gravitational field, we shall use a local comoving non-rotating reference frame (see Appendix A.11). This reference frame is a natural generalization of the usual inertial reference frames of a flat spacetime to the case when the spacetime is curved. For a given geodesic line representing the motion of the body a local reference frame is defined by an orthonormal tetrad $\mathbf{e}_{\hat m}$ $(m = 0, \dots, 3)$ along the world line. The first of the vectors $(\mathbf{e}_{\hat 0})$ is chosen to coincide with the four-velocity of the body, and the tetrad $\mathbf{e}_{\hat m}$ is chosen to be parallelly propagated along the trajectory. Under these conditions it is uniquely defined if it is fixed at some initial point. The equation of parallel transport in a vacuum stationary black hole metric can be solved analytically [Marck (1983)] and for the Schwarzschild black hole one has [Luminet and Marck (1985)]

$$e^{\mu}_{\hat 0} = \frac{1}{\sqrt{1 - \mu^2/x}} \left(\frac{1}{\sqrt{1 - \mu^2/x}}\, \delta^{\mu}_{t} + \frac{\mu}{x}\frac{1}{r}\, \delta^{\mu}_{\phi} \right), \qquad e^{\mu}_{\hat 2} = \frac{1}{r}\, \delta^{\mu}_{\theta},$$

$$e^{\mu}_{\hat 1} = \cos\psi\, \lambda^{\mu}_{\hat 1} - \sin\psi\, \lambda^{\mu}_{\hat 3}, \qquad e^{\mu}_{\hat 3} = \sin\psi\, \lambda^{\mu}_{\hat 1} + \cos\psi\, \lambda^{\mu}_{\hat 3}, \tag{2.11.9}$$

where

$$\lambda^{\mu}_{\hat 1} = \frac{1}{\sqrt{1 - \dot r^2}} \left[\frac{\dot r}{\sqrt{1 - \mu^2/x}}\, \delta^{\mu}_{t} + \sqrt{1 - \mu^2/x}\, \delta^{\mu}_{r} \right], \tag{2.11.10}$$

$$\lambda^{\mu}_{\hat 3} = \frac{1}{\sqrt{1 - \dot r^2}} \frac{\mu}{x} \left[\frac{1}{\sqrt{1 - \mu^2/x}}\, \delta^{\mu}_{t} + \dot r \sqrt{1 - \mu^2/x}\, \delta^{\mu}_{r} \right] + \sqrt{1 + \mu^2/x^2}\, \frac{1}{r}\, \delta^{\mu}_{\phi},$$

and

$$\dot r = \pm \mu \sqrt{\frac{1}{x} - \frac{1}{x^2} + \frac{\mu^2}{x^3}}. \tag{2.11.11}$$

The function $\psi = \psi(\tau^*)$ satisfies the equation

$$\frac{d\psi}{d\tau^*} = \frac{1}{x^2 + \mu^2} \tag{2.11.12}$$

and the initial condition $\psi(\tau^* = 0) = 0$. The solution of this equation can be written in terms of elliptic integrals:

$$\psi = \arctan\sqrt{x - 1 + \mu^2/x} + \frac{1}{\sqrt{1+2\mu}} F(\varphi\backslash\alpha) . \qquad (2.11.13)$$

The tetrad is chosen in such a way that the vector $\mathbf{e}_{\hat{1}}$ at the point of periastron $\tau = 0$ is directed toward the black hole center.

The geodesic line γ representing the motion of a test particle together with the tetrad $\mathbf{e}_{\hat{m}}$ form a comoving local reference frame. The position of a point of a star with respect to the chosen reference frame can be described by a three-vector

$$\mathbf{X} = \sum_{i=1}^{3} X^i \, \mathbf{e}_i . \qquad (2.11.14)$$

Consider first a test body which is formed from pressure-free matter (dust), and let $X^i = X^i(\tau)$ be the coordinates of a chosen volume element of matter in the comoving reference frame. The geodesic deviation equation (see Appendix A.9) can be used to show that

$$\frac{d^2 X^i}{d\tau^2} = -\frac{\partial U}{\partial X^i} , \qquad (2.11.15)$$

where

$$U = \frac{1}{2} C_{ij} \, X^i X^j , \qquad C_{ij} \equiv R_{\alpha\beta\gamma\delta} \, e^{\alpha}_{\hat{0}} \, e^{\beta}_{\hat{i}} \, e^{\gamma}_{\hat{0}} \, e^{\delta}_{\hat{j}} , \qquad (2.11.16)$$

and $R_{\alpha\beta\gamma\delta}$ are the components of the Riemann curvature tensor calculated on the world line representing the moving body (i.e., at the beginning of the comoving reference frame). The non-vanishing components of the tensor C_{ij} in the comoving reference frame are [Marck (1983), Luminet and Marck (1985)]

$$C_{ij} = \frac{GM}{R_p^3} \frac{1}{x^3} \sigma_{ij} , \qquad (2.11.17)$$

$$\sigma_{11} = 1 - 3\,\frac{x^2+\mu^2}{x^2} \cos^2\psi , \qquad \sigma_{22} = 1 + 3\,\frac{\mu^2}{x^2} ,$$

$$\sigma_{33} = 1 - 3\,\frac{x^2+\mu^2}{x^2} \sin^2\psi , \qquad \sigma_{13} = \sigma_{31} = -3\,\frac{x^2+\mu^2}{x^2} \sin\psi\cos\psi . \qquad (2.11.18)$$

The quantities $x = x(\tau^*)$ and $\psi = \psi(\tau^*)$ are to be calculated on the chosen trajectory.

The tidal gravitational potential U takes quite a simple form if one uses spherical coordinates ρ, Θ, and Φ instead of the Cartesian comoving coordinates X^i :

$$X^1 = \rho \sin\Theta \cos\Phi , \qquad X^2 = -\rho\cos\Theta , \qquad X^3 = \rho\sin\Theta\sin\Phi . \qquad (2.11.19)$$

In these coordinates

$$U = \frac{1}{2}\frac{GM}{R_p^3}\frac{\rho^2}{x^3}\left(\sigma_1 + \frac{3\mu^2}{x^2}\sigma_2\right), \tag{2.11.20}$$

where

$$\sigma_1 = 1 - 3\sin^2\Theta\cos^2(\Phi - \psi), \quad \sigma_2 = \cos^2\Theta - \sin^2\Theta\cos^2(\Phi - \psi). \tag{2.11.21}$$

In the presence of pressure, equation (2.11.15) is to be modified by the pressure forces:

$$\frac{d^2X^i}{d\tau^2} = -\frac{\partial U}{\partial X^i} + \nabla^i P. \tag{2.11.22}$$

To compare the motion of a relativistic star with a non-relativistic one, we fix in both problems the same conserved quantity – the angular momentum of the body, which for a parabolic motion unambiguously fixes the trajectory. The following relativistic effects are important:

1. *Relativistic shift of the periastron radius.* The difference Δx_p between x_p and its non-relativistic limit $x_{p,\text{nonrel}} = 1$ is $\Delta x_p = -1/2(1 - \sqrt{1-4\mu^2})$. For small μ this difference is $\Delta x_p \simeq -\mu^2$. In other words, the relativistic periastron lies at the radius which is closer to the gravitating center by the value $\Delta r_p \simeq r_g$.
2. *Relativistic time delay.* It takes a longer time (as measured by the clock of a distant observer) for the relativistic body to pass near a black hole and to return to the same initial radius than for the non-relativistic motion with the same angular momentum. The time delay is $\Delta T = 3\pi/4\left(r_g^3/R_p\right)^{1/2}$.
3. *Relativistic precession.* Let a non-rotating test body have three orthogonal axes rigidly attached to it. The orientation of these axes in space, after a test body moving in the gravitational field passes near a black hole and then recedes, will slightly differ from their initial orientation. This effect remains even if the body is absolutely rigid. In the case under consideration this effect results in a difference of the angle ψ which enters equation (2.11.9) from the angle ϕ. The total precession angle is

$$\Delta\phi_{\text{prec}} = 2\lim_{x\to\infty}(\phi - \psi) = \frac{2}{\sqrt{1+2\mu}}K - \pi, \tag{2.11.23}$$

where $K = K(4\mu/(1+2\mu))$ is the complete elliptic integral of the first kind. For small μ one has

$$\Delta\phi_{\text{prec}} \simeq \frac{3\pi}{4}\frac{r_g}{R_p}. \tag{2.11.24}$$

4. The ***tensorial structure of tidal forces*** is different in the relativistic and non-relativistic cases. In both (relativistic and non-relativistic) cases the tidal force tensor written in a parallelly propagated frame is non-diagonal. This tensor can be easily diagonalized by using a frame $(\boldsymbol{\lambda}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \boldsymbol{\lambda}_{\hat{3}})$. The diagonal elements of the tensor σ in this new frame are

$$\sigma_{\hat{1}\hat{1}} = -2 - 3\frac{\mu^2}{x^2}, \qquad \sigma_{\hat{2}\hat{2}} = 1 + 3\frac{\mu^2}{x^2}, \qquad \sigma_{\hat{3}\hat{3}} = 1. \tag{2.11.25}$$

The extra term $3\mu^2/x^2$ which enters these relations describes the relativistic effects. These effects lead to an additional increase of the tidal acceleration in the orbital plane, without changing the acceleration in the perpendicular direction. The effect is maximized for the limiting value $\mu = 1/2$, when at periastron the star is stretched in the radial direction with a strength 5/2 times larger, and compressed in the orthogonal direction with a strength 4 times larger than in the non-relativistic case.

The formulas obtained in this section can be directly applied to study the distortion and disruption of stars passing near a black hole. Results of the numerical computations and astrophysical consequences of the process are discussed in the paper by Frolov *et al.* (1994). Tidal interaction of a star with a massive rotating black hole was considered by Diener *et al.* (1997).

Chapter 3

Rotating Black Holes

3.1 Formation of a Rotating Black Hole

In the preceding chapter we have demonstrated that the gravitational collapse of a spherical non-rotating mass produces a spherically symmetric black hole when the radius of the collapsing body becomes less than the gravitational radius. Suppose the collapsing body deviates considerably from spherical symmetry and/or its angular momentum and electromagnetic field are large. Will a black hole form? If it does, what will its properties be?

In order to answer these questions, we are to be able to generalize the definition of a black hole so that it can be applied to non-stationary spacetimes without symmetry. This general theory will be developed in Chapter 5. It will be argued that if contraction to a sufficiently small size of an arbitrary mass possessing rotation and an electromagnetic field produces a black hole, then such a black hole rapidly becomes stationary and after this all its properties and its external gravitational field are completely determined by three parameters: mass M, angular momentum J, and electric charge Q.[1] Other properties of the collapsing body, such as its composition, asymmetry in the distribution of mass and electric charge, the magnetic field and its characteristics and so forth, do not influence the properties of the resulting stationary black hole.[2]

This conclusion follows qualitatively from an analysis of the behavior of small perturbations in the course of the formation of the spherical black hole (Section 4.2). In the collapse of a body slightly deviating from spherical symmetry, all deviations from

[1] It will be clear hereafter that the black hole forms only if the inequality $M^2 \geq J^2/M^2 + Q^2$ is satisfied. This inequality is written in the system of units where $G = c = 1$.

[2] Strictly speaking, this conclusion is valid only in the framework of the Einstein-Maxwell theory. We shall see in Chapter 13 that in more general unified theories that include Einstein gravity and Maxwell field as its part stationary black holes might have some additional characteristics. The Einstein-Maxwell theory certainly is valid with high accuracy in astrophysical applications. Bearing this in mind, we postpone till Chapter 13 the discussion of modified theories that might be important for microscopically small black holes.

spherical symmetry rapidly disappear, except those due to small angular momentum J. Gravitational radiation carries off a part of the energy and angular momentum of the collapsing mass in the course of the gravitational collapse when the deviations from symmetry are large. As a result, M and J for the black hole become slightly smaller than those the body had before the collapse (this will be discussed later). This reduction could not be found in the analysis of small perturbations because the backreaction of the perturbations on the metric was assumed to be negligible.

If the collapsing body has an electric charge, then after the formation of a black hole radiative multipoles of the electromagnetic field also rapidly disappear, and only a non-radiative mode determined by the electric charge Q survives. Macroscopic physics knows no other "classical" physical fields besides gravitational and electromagnetic ones with other (non-radiative) modes. That is why a stationary metric describing a "final" state of an isolated black hole is determined by the three parameters M, J, Q. In astrophysics, the total electric charge of a body can typically be treated as small and accordingly neglected. Therefore, we first consider the case of $Q = 0$. The case of nonzero charge is discussed in Section 3.6.

3.2 Gravitational Field of a Rotating Black Hole

3.2.1 Kerr metric

What is the gravitational field of a stationary black hole with nonzero angular momentum J? It will be shown in Section 6.4 that the most general solution describing such a black hole is axisymmetric, and it coincides with the vacuum stationary axisymmetric solution of Einstein's equations which was found by Kerr (1963). We consider now the Kerr metric and describe the physical properties of the external space of a rotating black hole.

In the case of a rotating black hole (Kerr metric), the spacetime outside it is stationary, and one can choose a time-independent reference frame which asymptotically tends to the Lorentz frame at infinity. The coordinates proposed by Boyer and Lindquist (1967) represent such a frame of reference. The Kerr metric in these coordinates is:

$$ds^2 = -\left(1 - \frac{2Mr}{\Sigma}\right) dt^2 - \frac{4Mra\sin^2\theta}{\Sigma}\, dt\, d\phi + \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \frac{A\sin^2\theta}{\Sigma}\, d\phi^2, \quad (3.2.1)$$

where

$$\Sigma \equiv r^2 + a^2\cos^2\theta\,, \qquad \Delta \equiv r^2 - 2Mr + a^2\,,$$

$$A = \left(r^2 + a^2\right)^2 - a^2\Delta\sin^2\theta\,. \quad (3.2.2)$$

Mathematical properties of the Kerr metric and its generalization with electric charge included (the Kerr-Newman metric) are discussed in the Appendix D. For our purpose, now it is sufficient to mention only its simplest properties.

The Kerr metric depends on two parameters M and a. The physical meaning of them can be easily obtained by analyzing the asymptotic behavior of the metric. At large distances ($r \to \infty$) the metric has the following asymptotic form

$$ds^2 \approx -\left(1 - \frac{2M}{r}\right) dt^2 - \frac{4Ma\sin^2\theta}{r} dt\, d\phi$$
$$+ \left[1 + O(r^{-1})\right] \left[dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right]. \tag{3.2.3}$$

Comparing this with the results of calculations in the weak field approximation, one may conclude that a is the specific angular momentum ($a = J/M$, J is the angular momentum) and M is the black hole mass.[3] Presumably, the physically meaningful solutions are those with $M^2 \geq J$ (see note 1 to page 56).

In the absence of rotation, $a = 0$, the Kerr metric obviously reduces to the Schwarzschild metric. On the other hand, if we take the limit $J = 0$ and $M = 0$ but keep the parameter a constant, the metric (3.2.1) takes the form

$$ds^2 = -dt^2 + (r^2 + a^2\cos^2\theta)\left(\frac{dr^2}{r^2 + a^2} + d\theta^2\right) + (r^2 + a^2)\sin^2\theta\, d\phi^2. \tag{3.2.4}$$

It can be shown that the curvature of this metric vanishes identically, and hence this is the metric of a flat spacetime

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2. \tag{3.2.5}$$

The coordinates (r, θ, ϕ) are connected with Cartesian coordinates (x, y, z) by

$$x = \sqrt{r^2 + a^2}\sin\theta\cos\phi, \quad y = \sqrt{r^2 + a^2}\sin\theta\sin\phi, \quad z = r\cos\theta. \tag{3.2.6}$$

The surfaces $r =$ const are oblate ellipsoids of rotation

$$\frac{x^2 + y^2}{r^2 + a^2} + \frac{z^2}{r^2} = 1. \tag{3.2.7}$$

The Kerr metric is stationary (time-independent) and axisymmetric. We denote by $\boldsymbol{\xi}_{(t)}$ and $\boldsymbol{\xi}_{(\phi)}$ the Killing vectors which are the generators of the corresponding symmetry transformations. We choose the Killing vector field $\boldsymbol{\xi}_{(t)}$ in such a way that far from the black hole (at infinity) the Killing vector is directed along the time lines of the Lorentz coordinate system defined by the asymptotic form of the metric (3.2.1). This clarification is necessary because a linear combination of two Killing vectors is again a Killing vector (i.e., in our case we can combine a shift in

[3] We recall the reader that beginning from this chapter we shall use the system of units in which $c = G = 1$. This is standard practice because keeping c, G, and other dimensional constants makes complicated formulas practically unreadable. It is not difficult to restore the correct dimensions in the final formulas. From time to time we shall write some of the final important relations in their complete form with all the dimensional constants restored.

time with a rotation around a spatial axis). We single out a Killing vector $\boldsymbol{\xi}_{(t)}$ that corresponds to the absence of any rotation of the space grid around the symmetry axis at large distances from the black hole ($r \to \infty$). In a similar manner the Killing vector $\boldsymbol{\xi}_{(\phi)}$ is unambiguously singled out by the condition that its integral lines are closed. For this choice $\xi^{\mu}_{(t)} = \delta^{\mu}_{t}$ and $\xi^{\mu}_{(\phi)} = \delta^{\mu}_{\phi}$. The coefficients g_{tt}, $g_{t\phi}$, and $g_{\phi\phi}$ of the Kerr metric coincide with the scalar products $\boldsymbol{\xi}_{(t)} \cdot \boldsymbol{\xi}_{(t)}$, $\boldsymbol{\xi}_{(t)} \cdot \boldsymbol{\xi}_{(\phi)}$, and $\boldsymbol{\xi}_{(\phi)} \cdot \boldsymbol{\xi}_{(\phi)}$, respectively. Since the metric coefficient $g_{t\phi}$ does not vanish, the Kerr metric (3.2.1) is not invariant under reflection of time $t \to -t$. In other words, the Kerr metric is stationary but not static. The symmetry is restored if the inversion of time $t \to -t$ (changing the sign of the angular momentum) is accompanied by the transformation $\phi \to -\phi$ (changing the direction of rotation).

3.2.2 The (3 + 1)-split of the spacetime outside the black hole

To discuss the physical properties of the gravitational field of a rotating black hole and of matter propagating in its background, it is important to choose a convenient reference frame. In the absence of rotation such a reference frame was discussed in the previous chapter. The Schwarzschild reference frame used there is static, independent of time, and uniquely defined for each black hole. Recall the main properties of this frame. At large distances from the black hole, this reference frame turns into the Lorentz frame in which the black hole is at rest. It can be thought of as a lattice "welded" out of weightless rigid rods. The motion of particles was defined with respect to this lattice. For the time variable, we used the time t of an observer placed at infinity. True, the rate of the flow of the physical (proper) time τ at each point of our lattice did not coincide with that of t (time is slowed down in the neighborhood of a black hole), but in this "parameterization" t = const meant simultaneity in our entire frame of reference. The Schwarzschild reference frame in a certain sense resembles the absolute Newtonian space in which objects move, and t resembles the absolute Newtonian time of the equations of motion. Of course, important differences exist. Our "absolute" space is curved (curved very strongly close to the black hole), and the "time" t is not the physical time.

This reference frame is used not only to facilitate mathematical manipulations in solving, say, the equations of motion, but also to increase pictorial clarity. We make use of the familiar concepts of the Newtonian physics (the "absolute" rigid space as the scene on which events take place, and absolute time) and thereby help our intuition. Although the Schwarzschild reference frame has a singularity at r_g, we choose this reference frame for the spacetime outside the black hole and not, say, the Lemaître frame which has no singularity at r_g but is everywhere deformable. Obviously, a rigid reference frame can be chosen only because the metric outside the black hole is time independent. In the general case of a variable gravitational field, this choice is impossible since the spatial grid would be deformed with time.

The natural generalization of the Schwarzschild reference frame to the case of a rotating black hole is easily obtained if we use the fact that the Kerr metric is stationary. The properties of the three-dimensional space t = const in the metric (3.2.1), which is external with respect to the black hole, do not change with time. Shifting the spatial section along the Killing vector field $\boldsymbol{\xi}_{(t)}$ (a generator of time symmetry transformations), we pass from one section to another identical to it. We can thus "trace" in the space a grid which remains invariant in the transition from one section to another along the Killing vector field $\boldsymbol{\xi}_{(t)}$. The variable t, the time of distant observers, can serve as a universal "time" labeling the spatial sections, as was the case for the Schwarzschild spacetime.

Important differences must be mentioned, of course:

1. In the case of the Schwarzschild field, the transition from one three-dimensional section to another, preserving the coordinate grid, is carried out by shifting along the time lines perpendicular to the spatial section. The situation in the Kerr field is different. Since the component $g_{t\phi}$ of the metric does not vanish, the Killing vector field $\boldsymbol{\xi}_{(t)}$ is tilted with respect to the section t = const. The tilting angle is different for different r and θ.

2. The Killing vector $\boldsymbol{\xi}_{(t)}$ that generates the transition from one section to another becomes spacelike at points close to the boundary of the black hole. This happens in the region where $\xi_{(t)\mu}\xi^{\mu}_{(t)} \equiv g_{tt} > 0$. The part of this region lying outside the black hole is called *ergosphere.* The external boundary of the ergosphere is the surface $r = M + \sqrt{M^2 - a^2\cos^2\theta}$. This means that a three-dimensional rigid grid extended from infinity and constructed from material bodies ("welded" out of rods) *cannot be continued inside the ergosphere* . Inside the ergosphere, this grid would move at superluminal velocity with respect to any observer (on a timelike world line).

Despite these specifics, we can still operate with our space sections t = const as with "absolute" rigid space (resembling the Newtonian case) and with t, as with a "time" which is universal in the entire "space" (of course, subject to all the qualifications given above).

We conclude this section with the following general remark. In general relativity, the splitting of spacetime in an arbitrary gravitational field into a family of three-dimensional spatial sections (in general, their geometries vary from section to section) and a universal "time" that labels these sections is referred to as a (3+1)-split of spacetime,[4] or the *kinemetric method* [Vladimirov (1982)]. This method is

[4]Another approach to the (3+1)-split is possible, where one chooses first not three-dimensional sections but congruences of timelike lines. This approach is known as the *chronometric method* [Zel'manov (1956)]. In a stationary spacetime it is convenient to use a congruence of Killing trajectories. These trajectories can be considered as points of a new three-dimensional space so that four-dimensional tensor fields and field equations can be reduced to three-dimensional ones. For more details concerning this approach, see e.g., [Geroch (1971), Perjés (1993), Boersma and Dray (1995a-c)] and Appendices A.9 and A.10.

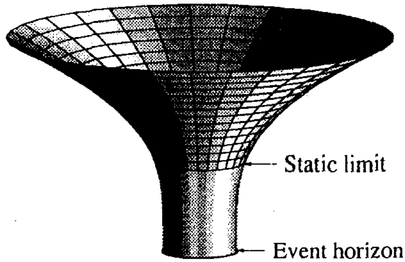


Figure 3.1: Embedding diagram for the geometry on the equatorial plane in the Kerr spacetime with $a = 0.999M$. A coordinate lattice "welded" of rigid rods, which fill the space around the black hole, is schematically shown. The rigid coordinate frame constructed from these "rods" cannot be continued beyond the static limit.

especially useful when all spatial sections are identical and the motion of particles, electromagnetic processes, etc., that unfold on this invariant "scene" can be described in terms of a universal "time" t. We have already mentioned that in this case our intuition is supported by our "visual" images of space and time supplied by everyday experience. Studying the processes in the vicinity of stationary black holes, we employ the kinemetric method. As spatial sections, we choose the $t =$ const sections in metric (3.2.1); t is the time coordinate.

3.3 Reference Frames. Event Horizon

3.3.1 Chronometric reference frame

First we consider the geometric properties of our "absolute" space. They are described by a three-dimensional metric obtained from (3.2.1) by setting $dt = 0$. The 3-geometry of a space section $t =$ const is

$$dl^2 = g_{ij}\, dx^i\, dx^j = \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \frac{A\sin^2\theta}{\Sigma}\, d\phi^2. \tag{3.3.1}$$

To "visualize" the properties of this three-dimensional curved space, one may consider its different two-dimensional sections. For example, Figure 3.1 shows the embedding diagram representing the geometry on the equatorial plane $\theta = \pi/2$. It should be stressed that other sections have a geometry different from that shown in Figure 3.1.

Consider now the reference frame of observers who are at rest in the "absolute" space $t =$ const; that is, observers who "sit still" on our rigid non-deformable lattice. This frame of reference is called the chronometric [Vladimirov (1982)], Lagrange [Thorne and Macdonald (1982), Macdonald and Thorne (1982), Thorne, Price, and Macdonald (1986)] or Killing reference frame. The latter name is connected with the fact that the world lines of the observers in this reference frame are tangent to the Killing vector field $\boldsymbol{\xi}_{(t)}$.

Let us look at the forces acting in this frame due to the presence of a rotating black hole. The four-velocity of the observers is $u^\mu = \xi^\mu_{(t)} / |\xi^2_{(t)}|^{1/2}$, and their four-acceleration $w^\mu \equiv u^\nu u^\mu{}_{;\nu}$ is $w_\mu = (1/2)[\ln(-\xi^2_{(t)})]_{;\mu}$ (see Appendix A.10.1). A freely falling body which is moving along a geodesic will have an acceleration $a^\mu = -w^\mu$ with respect to the frame. The three-dimensional components of the free fall acceleration vector $\tilde{a}_i$ in the coordinates (r, θ, ϕ) can be easily calculated (see Vladimirov (1982) and Appendix D.5.1). The result is

$$\tilde{a}_r = \frac{M(\Sigma - 2r^2)}{\Sigma(\Sigma - 2Mr)}, \qquad \tilde{a}_\theta = \frac{Mra^2 \sin 2\theta}{\Sigma(\Sigma - 2Mr)}, \qquad \tilde{a}_\phi = 0. \tag{3.3.2}$$

We mark all quantities in this chronometric reference frame by a tilde to avoid confusing them with the quantities used hereafter.

The physical components of acceleration are[5]

$$\tilde{a}_{\hat{r}} = \frac{M(\Sigma - 2r^2)\sqrt{\Delta}}{\Sigma^{3/2}(\Sigma - 2Mr)}, \quad \tilde{a}_{\hat{\theta}} = \frac{Mra^2 \sin 2\theta}{\Sigma^{3/2}(\Sigma - 2Mr)}, \quad \tilde{a}_{\hat{\phi}} = 0. \tag{3.3.3}$$

The reference frame is rigid so that the deformation rate tensor vanishes [see (A.53)]:

$$\widetilde{D}_{ik} = 0. \tag{3.3.4}$$

The vorticity tensor (A.60) calculated for the Kerr metric is (see Appendix D.5.1):

$$\tilde{\omega}_{r\phi} = -\frac{Ma(\Sigma - 2r^2)\sin^2\theta}{\Sigma^{1/2}(\Sigma - 2Mr)^{3/2}}, \quad \tilde{\omega}_{\theta\phi} = -\frac{Mra \sin 2\theta \Delta}{\Sigma^{1/2}(\Sigma - 2Mr)^{3/2}}, \quad \tilde{\omega}_{r\theta} = 0. \tag{3.3.5}$$

The non-vanishing tensor $\tilde{\omega}_{ik}$ signifies that gyroscopes which are at rest in the reference frame are precessing with respect to it, and hence with respect to distant objects because at a large distance our rigid reference frame becomes Lorentzian. The tensor $\tilde{\omega}_{ik}$ is proportional to the specific angular momentum of the black hole and reflects the presence of a "vortex" gravitational field due to its rotation.

The following important difference between the external fields of a rotating and a non-rotating black holes must be emphasized. If a black hole is non-rotating, the condition $t =$ const signifies physical simultaneity in the entire external space for observers that are at rest in it (with respect to a rigid reference frame). In the case of a rotating black hole, a non-vanishing component $g_{t\phi}$ in the rigid reference frame forbids [see Landau and Lifshitz (1975)] the introduction of the concept of simultaneity. Usually, the events with equal t are said to be simultaneous in the time of a distant observer. But this does not at all mean the physical simultaneity of these events for the chosen set of observers which is determined by the synchronization of clocks by exchange of light signals between these observers.

---

[5] $\tilde{a}_{\hat{m}}$ are the components of the acceleration vector measured directly by an observer at rest in the given reference frame. In formula (3.3.3), they are given in a local Cartesian coordinate system $\mathbf{e}_{\hat{m}}$ whose axes are aligned along the directions of r, θ, and ϕ: $\mathbf{a} = \tilde{a}^{\hat{m}}\mathbf{e}_{\hat{m}}$.

3.3.2 Ergosphere. Event horizon

We already noted that the Killing vector $\xi^{\mu}_{(t)}$ becomes null at the boundary of the ergosphere. This happens at a so-called *static limit surface* $S_{\rm st}$ where the component g_{tt} in (3.2.1) (which determines the rate of flow of time) vanishes. The equation of this surface is

$$\Sigma - 2Mr = r^2 + a^2\cos^2\theta - 2Mr = 0\,. \tag{3.3.6}$$

It can be also rewritten in the explicit form $r = r_{\rm st}$, where $r_{\rm st}$ is given by the relation

$$r_{\rm st} = M + \sqrt{M^2 - a^2\cos^2\theta}\,. \tag{3.3.7}$$

Note that the components $\tilde{a}_r$, $\tilde{a}_\theta$, the components $\tilde{\omega}_{ik}$, and the angular velocity of precession $\widetilde{\Omega}_{\rm pr}$ of a gyroscope, calculated using these components [see (A.46)], tend to infinity at $S_{\rm st}$. These properties signify that a physical singularity exists at this surface in the reference frame, and this frame cannot be extended closer to the black hole; that is, observers cannot be at rest relative to our grid.[6] Formally, the reason for this is the same as in the Schwarzschild field at $r = r_{\rm g}$. Namely, the world line of observer, $r = \text{const}$, $\theta = \text{const}$, $\phi = \text{const}$, ceases to be timelike, as indicated by the reversal of sign of g_{tt} at $r < r_{\rm st}$. However, an essential difference in comparison with the Schwarzschild field must be emphasized.

In order to obtain a world line inside the light cone in a non-rotating black hole for $r < r_{\rm g}$, it was sufficient to perform the transformation

$$r = r(\tilde{r}, t)\,, \qquad \frac{\partial r}{\partial t} = v \neq 0\,. \tag{3.3.8}$$

With a suitable choice of $v = v(r) < 0$, the coordinate line $(\tilde{r}, \phi, \theta = \text{const})$ becomes timelike. This means that at $r < r_{\rm g}$, a body necessarily moves to the center *along the radius*, and that $r_{\rm g}$ is the boundary of an isolated black hole.

In the case of a rotating black hole [we assume $\Delta > 0$; see (3.2.1)] a transformation of the type of (3.3.8) cannot generate a timelike world line in the region $r > r_{\rm st}$. But a transformation of the type

$$\phi = \phi(\tilde{\phi}, r, \theta, t)\,, \qquad \frac{\partial \phi}{\partial t} = \omega \neq 0 \tag{3.3.9}$$

makes this possible (ω is a function of r and θ). To illustrate this, consider a special type of motion; namely, circular motion with $r = \text{const}$ and $\theta = \text{const}$. For this motion $\boldsymbol{u} \sim \boldsymbol{\xi}_{(t)} + \omega\boldsymbol{\xi}_{(\phi)}$, and $\omega = d\phi/dt = u^\phi/u^t$ is the angular velocity of rotation. The condition that the vector $\boldsymbol{u}$ is timelike requires $\omega_- < \omega < \omega_+$, where

$$\omega_\pm = \frac{-g_{t\phi} \pm \sqrt{(g_{t\phi})^2 - g_{tt}\,g_{\phi\phi}}}{g_{\phi\phi}}\,. \tag{3.3.10}$$

[6] Of course, the coordinate grid can be extended closer to the black hole, but it cannot be constructed of material bodies.

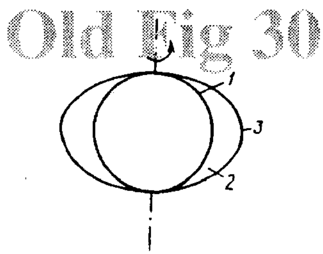


Figure 3.2: A rotating black hole: 1—horizon, 2—ergosphere, 3—static limit.

The limiting values $\omega = \omega_{\pm}$ correspond to motion with the velocity of light. Outside the static limit surface (where $g_{tt} < 0$), ω_+ is positive, while ω_- is negative. It means that in this region both types of motion (with increase and with decrease of the angular coordinate ϕ) are possible. In the ergosphere only motion with $\omega > 0$ is allowed. For circular motion r remains constant, but by a small deformation we can get a motion directed outward (toward increasing r). This fact signifies that if $r < r_{\rm st}$ and $\Delta > 0$, all bodies necessarily participate in the rotation around the black hole (the direction of rotation is determined by the sign of a; see below) with respect to a rigid coordinate grid that stretches to infinity. As for motion along the radius r, bodies can move in the range $r < r_{\rm st}$, $\Delta > 0$ both increasing and decreasing the value of r. Therefore, the *static limit* $r_{\rm st}$ has a quite different nature for a rotating black hole than in the Schwarzschild field. Inside it, bodies are unavoidably dragged into rotation although $r_{\rm st}$ is not the event horizon because particles or light propagating inside the ergosphere can escape from this region and reach infinity (see Section 3.4).

For $(g_{t\phi})^2 - g_{tt}g_{\phi\phi} = 0$, $\omega_+ = \omega_-$. Using equation (D.9), one may conclude that this occurs when $\Delta = 0$. This equation defines the *event horizon* which lies at $r = r_+$, where

$$r_+ = M + \sqrt{M^2 - a^2}\,. \tag{3.3.11}$$

For $r < r_+$, where $\Delta < 0$, it is impossible to have motion with $\boldsymbol{u}^2 \leq 0$ without falling to the center (i.e., without decrease of r). In the absence of rotation ($a = 0$) r_+ coincides with the gravitational radius $r_{r_g} = 2M$. The name "event horizon" is used for the boundary of any black hole. The exact definition of the event horizon in the general case will be given in Chapter 5.

A rigid, static frame of reference which is at rest relative to a distant observer and which is made of material bodies does not extend to r_+ (see Figure 3.1). The static limit lies beyond the horizon and coincides with it at the poles (Figure 3.2). An important feature of a static reference frame is the precession of gyroscopes in it, as we have mentioned above. Our reference frame rotates at each of its points with respect to the local Lorentz frame.[7] This is, of course, a reflection of the fact that the rotation of the black hole changes the state of motion of local Lorentz frames, dragging them into rotation around the black hole. This effect has been known qualitatively for quite a long time in the case of the weak gravitational field of a rotating body [Thirring and Lense (1918)].

3.3.3 Reference frame of locally non-rotating observers

Now we introduce into the external space of a rotating black hole another reference frame which does not rotate, in the sense given above. Obviously, such a frame cannot be rigid. To introduce it, we trace a congruence of world lines which are everywhere orthogonal to the spatial sections $t = \text{const}$ that we choose. A future-directed unit vector $\boldsymbol{u}$ tangent to these world lines is (see Appendix D.5.2)

$$u_\mu = -\alpha\,\delta^0_\mu\,, \quad \alpha = \sqrt{\frac{\Sigma\Delta}{A}}\,. \tag{3.3.12}$$

By definition, these timelike lines are not twisted and form the sought reference frame. Observers that are at rest in this frame are said to be *locally non-rotating observers* [sometimes this reference frame is referred to as Eulerian; see Thorne and Macdonald (1982)]. Since $u_\mu\,\xi^\mu_{(\phi)} = 0$, the angular momentum for particles moving with the four-velocity $\boldsymbol{u}$ vanishes. For this reason, the reference frame is also known as the reference frame of *zero angular momentum observers* (ZAMO) [see Thorne, Price, and Macdonald (1986)]. (The discussion of more general reference frames based on stationary congruences can be found in [Page (1997)].)

These observers move with respect to the Boyer-Lindquist coordinate system; that is, they move in the "absolute" space.[8] This motion takes place at constant r and θ,

[7]This point requires some explanation. If a motion of an origin of a local reference frame is fixed, there still exists ambigutity in the directions of its three mutually orthogonal basic vectors connected with the spatial orientation of the reference frame. The following gedanken experiment can be used to single out frames where the orientation is not changed with time so that the frames are non-rotating. Consider light emitted at the origin of the frame, and place at some small distance a mirror orthogonal to the light ray. It is easy to verify that for this experiment performed in an inertial frame in a flat spacetime the reflected ray returns to the origin if and only if the reference frame in which the mirror is at rest is non-rotating. The same experiment can be used to define a non-rotating frame in a curved spacetime for an arbitrary (not necessary inertial) motion. It can be shown that this definition is equivalent to the requirement that the tetrad vectors forming the frame are Fermi-transported along the trajectory (for more discussions, see Synge (1960) and Appendices A.7 and A.11). We call such a non-rotating inertial local frame a local Lorentz frame.

[8]Let us make several remarks on terminology. In fact, it is not uniform among different authors and even in different papers of one author. Thus, Macdonald and Thorne (1982) refer to the space that we describe in Section 3.2 as "absolutely rigid" and specify that locally non-rotating observers move in this space [see Section 2 of their paper prior to formula (2.6)]. In another paper, Thorne (1985) refers to the space comoving with locally non-rotating observers as absolute, and says that the observers are at rest in the "absolute" space [see Thorne (1985), p. 11; see also Thorne, Price, and

at a constant (in time) angular velocity in ϕ. If the angular velocity ω is determined with respect to the universal time t (time of the distant observer), then

$$\omega \equiv \frac{d\phi}{dt} = -\frac{g_{\phi t}}{g_{\phi\phi}} = \frac{2Mar}{(r^2+a^2)^2 - \Delta a^2 \sin^2\theta}, \tag{3.3.13}$$

where $g_{\phi t}$ and $g_{\phi\phi}$ are taken from (3.2.1). Comparing this expression with equation (3.3.10), we can also write ω as $\omega = (\omega_+ + \omega_-)/2$.

If angular velocity is measured by the clock of a locally non-rotating observer, then

$$\Omega_\tau = \frac{\omega}{\sqrt{-g_{tt} - 2\omega\, g_{t\phi} - \omega^2\, g_{\phi\phi}}}. \tag{3.3.14}$$

The physical linear velocity of locally non-rotating observers with respect to a rigid reference frame is

$$v_\phi = \frac{2Mra\sin\theta}{\Sigma\sqrt{\Delta}}. \tag{3.3.15}$$

At the static limit surface $\Sigma = 2Mr$ and $\Delta = a^2\sin^2\theta$. Hence, as would be expected, this velocity becomes equal to the speed of light at the static limit $r = r_{\rm st}$ and exceeds it in the ergosphere.

We again emphasize that the proper time τ of locally non-rotating observers is not equal to the universal "time" t. Their ratio is equal to the "lapse" function α:

$$\left(\frac{d\tau}{dt}\right)_{\rm ZAMO} \equiv \alpha = \left(\frac{\Sigma\,\Delta}{A}\right)^{1/2}. \tag{3.3.16}$$

The components of the vector $\mathbf{a}$ of acceleration of free fall in the reference frame of locally non-rotating observers are (see Appendix D.5.2):

$$a_r = \frac{M}{\Sigma\,\Delta\,A}\left[(r^2+a^2)^2(a^2\cos^2\theta - r^2) + 4Mr^3a^2\sin^2\theta\right],$$

$$a_\theta = a^2\sin 2\theta\,\frac{Mr(r^2+a^2)}{\Sigma\,\Delta}, \qquad a_\phi = 0. \tag{3.3.17}$$

This vector is related to α as follows:

$$\mathbf{a} = -\boldsymbol{\nabla}\ln\alpha. \tag{3.3.18}$$

The tensor of deformation rates of the reference frame can be written in the form (see Appendix D.5.2)

$$D_{rr} = D_{r\theta} = D_{\theta\theta} = D_{\phi\phi} = 0,$$

Macdonald (1986)]. We invariably hold to the former viewpoint. We have mentioned that several terms are used to indicate the chronometric reference frame (see page 61). Partly this "discord" has historical roots, but it continues to the present.

$$D_{r\phi} = -Ma\left[2r^2(r^2+a^2) + \Sigma(r^2-a^2)\right]\sin^2\theta\,(\Sigma^3\,\Delta\,A\,)^{-1/2}, \tag{3.3.19}$$

$$D_{\theta\phi} = 2Mr\,a^3\sin^3\theta\cos\theta\,\Delta^{1/2}\,(\Sigma^3\,A\,)^{-1/2},$$

and the vorticity tensor $\omega_{ik} = 0$.

The introduced frame of reference has no singularities at the static limit surface and extends into the ergosphere up to the boundary of the black hole, $r = r_+$. At $r \le r_+$, descent along r necessarily occurs in addition to rotation around the black hole. At $r = r_+$, the reference frame of locally non-rotating observers has a physical singularity: $a_r \to \infty$ as $r \to r_+$ [see formula (3.3.17)].

As we approach the event horizon, the angular velocity ω of locally non-rotating observers tends to the limit

$$\omega_+ = \frac{a}{2M\,r_+} = \frac{a}{r_+^2 + a^2}\,. \tag{3.3.20}$$

This limit is constant at the horizon, being independent of θ. It is called the angular velocity of rotation of the black hole (or horizon), Ω^H. Restoring the dimensional constants c and G, we have

$$\Omega^H = \frac{c^3\,a}{2GM\,r_+} = \frac{ac}{r_+^2 + a^2}\,, \tag{3.3.21}$$

$$r_+ = \frac{GM}{c^2} + \sqrt{\left(\frac{GM}{c^2}\right)^2 - a^2}\,, \qquad a = \frac{J}{Mc}\,. \tag{3.3.22}$$

At spatial infinity, the reference frame of locally non-rotating observers transforms into the same Lorentz frame as the Boyer-Lindquist coordinate system (the chronometric reference frame) does.

To conclude this section, consider the precession of gyroscopes in the reference frame of locally non-rotating observers. Because of the absence of vorticity ($\omega_{ik} = 0$), this set of observers has the following property. Single out an observer γ, and consider a local non-rotating frame K. The mutually orthogonal unit vectors forming the frame K are Fermi-transported along the world line of the chosen observer (see, e.g., Synge (1960) and Appendix A.11). Consider now another observer γ' from the set. Then the direction to this observer γ' determined in the local frame K always remains the same. Certainly, the distance between γ and γ' may depend on time. It remains unchanged only if the system of observers is rigid, $D_{ij} = 0$. Evidently, there is no precession of gyroscopes in the reference frame K.

On the other hand, it is said in the monograph of Misner, Thorne, and Wheeler (1973), that gyroscopes precess with respect to locally non-rotating observers at an angular velocity

$$\Omega_{\rm pr} = \frac{1}{2}\sqrt{\frac{g_{\phi\phi}}{g_{tt} - \omega^2\,g_{\phi\phi}}}\left[\frac{\omega_{,\theta}}{\rho}\,\mathbf{e}_{\hat r} - \frac{\Delta^{1/2}\,\omega_{,r}}{\rho}\,\mathbf{e}_{\hat\theta}\right], \tag{3.3.23}$$

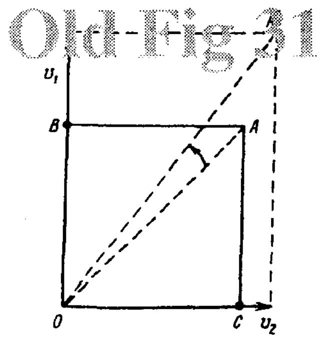


Figure 3.3: Tilting of the diagonal OA owing to anisotropic deformation of a volume element along the directions OB and OC.

where $\mathbf{e}_{\hat{r}}$ and $\mathbf{e}_{\hat{\theta}}$ are unit vectors along r and θ, respectively, and the quantities $g_{\alpha\beta}$ are taken from (3.2.1). Is it possible to make these statements compatible?

The paradox is solved in the following way. Recall that the motion of a small element of an arbitrary frame of reference with respect to a chosen frame K consists of a rotation around the instantaneous rotation axis and of a deformation along the principal axes of the deformation rate tensor. In the case of no rotation ($\omega_{ik} = 0$), we have only deformation. A gyroscope whose center of mass is at rest in the reference frame K does not precess with respect to the principal axes of the deformation rate tensor. If lines comoving with the frame of reference (lines "glued" to it) are traced along these directions, a gyroscope cannot change its orientation with respect to these lines. But this does not mean that the gyroscope does not change its orientation with respect to any line traced in a given element of volume in the comoving reference frame. Indeed, Figure 3.3 shows that anisotropic deformation tilts the lines traced, say, at an angle of 45^o to the principal axes of the deformation tensor, so that they turn closer to the direction of greatest extension. The gyroscope precesses with respect to these lines even though $\omega_{ik} = 0$. It is this situation that we find in the case of locally non-rotating observers in the Kerr metric.

Consider locally non-rotating observers in the equatorial plane. Everywhere $\omega_{ik} = 0$, and formulas (3.3.19) imply that only the component $D_{r\phi}$ is nonzero. This means that the instantaneous orientations of the principal axes of the deformation tensor lie at an angle of 45^o to the vectors $\mathbf{e}_{\hat{r}}$ and $\mathbf{e}_{\hat{\phi}}$. Note that the coordinate lines are "glued" to the reference frame. A gyroscope does not rotate with respect to the principal axes but, in view of the remark made above, does rotate relative to the ϕ coordinate line, and hence relative to $\mathbf{e}_{\hat{\phi}}$ (and therefore relative to the vector $\mathbf{e}_{\hat{r}}$, perpendicular to $\mathbf{e}_{\hat{\phi}}$, which is not "glued" to the reference frame; see below).

If a locally non-rotating observer always orients his frame vectors along the directions $\mathbf{e}_{\hat{r}}$, $\mathbf{e}_{\hat{\phi}}$, and $\mathbf{e}_{\hat{\theta}}$, the gyroscope thus precesses with respect to this frame as given by formula (3.3.23) although in the observers frame of reference we have $\omega_{ik} = 0$. The $\mathbf{e}_{\hat{r}}$, $\mathbf{e}_{\hat{\phi}}$, $\mathbf{e}_{\hat{\theta}}$ frame is a natural one; the precession of a gyroscope must be determined with respect to this frame because it is dictated by the symmetry of the space around the observer. Nevertheless, we could introduce a different frame, a frame which is also fixed to locally non-rotating observers but does not rotate with respect to the instantaneously comoving Lorentz frame. Obviously, gyroscopes do not precess in such a frame.

Finally, we notice that if we choose, at some instant of time, one system of coordinate lines "glued" to locally non-rotating observers and oriented strictly along r, and another system oriented along ϕ, the coordinate lines directed along ϕ slide with time in the "absolute" space along themselves, while the lines perpendicular to them "wind" on the black hole and become helical because they are dragged by a faster motion of the observers located closer to the black hole; hence, these lines rotate with respect to the ϕ lines.

3.4 Celestial Mechanics Near a Rotating Black Hole

3.4.1 Equations of motion. First integrals

Consider the motion of test particles along geodesics in the gravitational field of a rotating black hole. In the general case, the trajectories are fairly complicated because the field has no spherical symmetry. For detailed analysis, see Bardeen *et al.* (1972), Stewart and Walker (1973), Ruffini and Wheeler (1971b), Misner, Thorne, and Wheeler (1973), Sharp (1979), Chandrasekhar (1983), Shapiro and Teukolsky (1983), Dymnikova (1986), and Bičák *et al.* (1993). Important aspects of the gravitational capture of particles by a rotating black hole were treated by Dymnikova (1982) and Bičák and Stuchlík (1976). The references given above cite numerous original publications.

We consider the motion of test particles with respect to the "absolute" space introduced in Section 3.2, that is, with respect to the rigid lattice of the chronometric reference frame described by the Boyer-Lindquist coordinates (see Section 3.3). Because of the symmetries of the Kerr spacetime, there exist four integrals of motion, and equations of motion can be written in the form (see Appendix D.4)

$$\Sigma \frac{dr}{d\lambda} = \pm \mathcal{R}^{1/2}, \tag{3.4.1}$$

$$\Sigma \frac{d\theta}{d\lambda} = \pm \Theta^{1/2}, \tag{3.4.2}$$

$$\Sigma \frac{d\phi}{d\lambda} = \frac{L_z}{\sin^2\theta} - aE + \frac{a}{\Delta}\left[E(r^2+a^2) - L_z\, a\right], \tag{3.4.3}$$

$$\Sigma \frac{dt}{d\lambda} = a(L_z - aE\sin^2\theta) + \frac{r^2 + a^2}{\Delta}\left[E(r^2 + a^2) - L_z\, a\right], \tag{3.4.4}$$

where

$$\mathcal{R} = \left[E(r^2 + a^2) - L_z\, a\right]^2 - \Delta\left[m^2 r^2 + (L_z - aE)^2 + \mathcal{Q}\right], \tag{3.4.5}$$

$$\Theta = \mathcal{Q} - \cos^2\theta\left[a^2(m^2 - E^2) + \frac{L_z^2}{\sin^2\theta}\right]. \tag{3.4.6}$$

Here, m is the test particle mass; $\lambda = \tau/m$; τ is the proper time of the particle; E is the constant energy of the test particle; L_z is the constant projection of the angular momentum of a particle on the rotation axis of the black hole. The quantity $\mathcal{Q}$ is the integral of the motion found by Carter (1968a):[9]

$$\mathcal{Q} = p_\theta^2 + \cos^2\theta\left[a^2(m^2 - E^2) + \sin^{-2}\theta\, L_z^2\right], \tag{3.4.7}$$

where p_θ is the θ component of the four-momentum of the test particle. The motion of an ultra-relativistic particle corresponds to the limit as $m \to 0$. It should be noted that the signs $\pm$ which enter (3.4.1) and (3.4.2) are independent from one another.

The physical meaning of the Carter integral of motion $\mathcal{Q}$ can be obtained by considering motion of a particle at a large distance from the black hole, where the metric takes the asymptotic form (3.2.3). The total angular momentum L of such a particle is

$$L^2 = m^2 r^4\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right]. \tag{3.4.8}$$

In the asymptotic region equations (3.4.2) and (3.4.3) take the form

$$mr^2\sin\theta_\infty \frac{d\phi}{d\tau} = \frac{L_z}{\sin\theta_\infty}, \tag{3.4.9}$$

$$mr^2\frac{d\theta}{d\tau} = \pm\left[\mathcal{Q} + a^2(E^2 - m^2)\cos^2\theta_\infty - L_z^2\cot^2\theta_\infty\right]^{1/2}. \tag{3.4.10}$$

Using these relations, we get

$$\mathcal{Q} = L^2 - L_z^2 - a^2(E^2 - m^2)\cos^2\theta_\infty. \tag{3.4.11}$$

In the Schwarzschild spacetime $\mathcal{Q} + L_z^2$ is the square of the conserved total angular momentum of the particle.

[9] E and L_z are constant since the spacetime is stationary and axisymmetric. Integral of motion $\mathcal{Q}$ is implied by the existence of a Killing tensor field $\xi_{\mu\nu}$ in the Kerr metric. For the particle with the four-velocity u^μ the conserved quantity is $\mathcal{K} = m^2\xi_{\mu\nu}\, u^\mu u^\nu$. This quantity is connected with the integral of motion $\mathcal{Q}$ as $\mathcal{Q} = \mathcal{K} - (Ea - L_z)^2$. (See Carter (1968a, 1973a, 1977), Walker and Penrose (1970) and Appendix D.4).

3.4.2 General properties

The geodesic world line of a particle in the Kerr metric is completely determined by the first integrals of motion E, L_z, and $\mathcal{Q}$. Consider $\mathcal{R}$ given by (3.4.5) as the function of r for fixed values of other parameters

$$\mathcal{R} = (E^2 - m^2)r^4 + 2Mm^2r^3 + \left[(E^2 - m^2)a^2 - L_z^2 - \mathcal{Q}\right] r^2$$
$$+ 2M \left[\mathcal{Q} + (Ea - L_z)^2\right] r - a^2\mathcal{Q}. \quad (3.4.12)$$

At large distances the leading term ($\sim r^4$) on the right-hand side is positive if $E^2 > m^2$. Only in this case can the orbit extend to infinity. It is possible to show [Wilkins (1972)] that, in fact, all orbits with $E/m > 1$ are unbound except for a "measure-zero" set of unstable orbits. For $E^2 < m^2$ the orbit is always bounded, i.e., the particle cannot reach infinity.

To study qualitative characteristics of the motion of test particles in the Kerr metric, it is convenient to use the *effective potential*. Let us rewrite $\mathcal{R}$ as

$$\mathcal{R} = \alpha E^2 - 2\beta E + \gamma\,, \quad (3.4.13)$$

where

$$\alpha = \left[r^4 + a^2(r^2 + 2Mr)\right]\,, \quad \beta = 2aML_z r\,,$$
$$\gamma = L_z^2 a^2 - (m^2r^2 + L_z^2 + \mathcal{Q})\Delta\,. \quad (3.4.14)$$

The radial turning points $\mathcal{R} = 0$ are defined by the condition $E = V(r)$, where

$$V_\pm = \frac{\beta \pm \sqrt{\beta^2 - \alpha\gamma}}{\alpha}\,. \quad (3.4.15)$$

The quantities $V_\pm$ are known as the effective potentials. They are functions of r, and of integrals of motion L_z and $\mathcal{Q}$ as well as the parameters M and a characterizing the black hole. A general discussion of the topology of sets of the parameters for which the motion is possible can be found in [Zakharov (1986, 1989)].

The limiting values of the effective potentials $V_\pm$ at the infinity and at the horizon are

$$V_\pm(r = \infty) = \pm\, m \qquad V_\pm(r_+) = aL_z/2Mr_+ = \omega_+ L_z\,, \quad (3.4.16)$$

where ω_+ is the angular velocity (3.3.20) of the black hole. The motion of a particle with energy E is possible only in the regions where either $E \geq V_+$ or $E \leq V_-$. Expression (3.4.12) for $\mathcal{R}$ remains invariant under transformations $E \to -E$, $L_z \to -L_z$ relating these regions. In the Schwarzschild geometry the second region $E \leq V_-$ is excluded because $E \geq 0$ in the exterior of the black hole, while $V_- < 0$.

Simple analysis shows that for the motion in the Kerr geometry with $E^2 > m^2$ there can exist at most two turning points, while for $E^2 < m^2$ there cannot be more than three turning points [Dymnikova (1986)].

Consider now the equation (3.4.2). Since $\Theta \geq 0$, bounded motion with $E^2 < m^2$ is possible only if $\mathcal{Q} \geq 0$. The bounded orbit is characterized by the value $\mathcal{Q} = 0$ if and only if it is restricted to the equatorial plane. Non-equatorial *bounded* orbits with $\theta =$ const do not exist in the Kerr metric. For $\mathcal{Q} \geq 0$, there exist both bounded and unbounded trajectories. They intersect the equatorial plane or (for $\mathcal{Q} = 0$) are entirely situated in it. The particles with $\mathcal{Q} < 0$ never cross the equatorial plane and move between two cones $\theta = \theta_+$ and $\theta = \theta_-$, where $\theta_\pm$ are solutions of the equation $\Theta = 0$.

3.4.3 Motion in the equatorial plane

First, we consider the characteristic features of the motion of particles in the equatorial plane of a rotating black hole. In this case, the expressions for $dr/d\lambda$ and $d\phi/d\lambda$ can be written in the form

$$r^3\left(\frac{dr}{d\lambda}\right)^2 = E^2\left(r^3 + a^2 r + 2Ma^2\right) - 4aMEL_z - (r-2M)L_z^2 - m^2 r\Delta\,, \tag{3.4.17}$$

$$\frac{d\phi}{d\lambda} = \frac{(r-2M)L_z + 2aME}{r\Delta}\,. \tag{3.4.18}$$

These expressions are analogues of equations (2.8.1)–(2.8.2) for a Schwarzschild black hole. An analysis of the characteristics of the motion is performed in the same way as in Section 2.8 by using the effective potential (3.4.15).

The most important class of orbits are circular orbits. For given E and L_z the radius r_0 of a circular orbit can be found by solving simultaneously the equations

$$\mathcal{R}(r_0) = 0\,, \qquad \left.\frac{d\mathcal{R}}{dr}\right|_{r_0} = 0\,. \tag{3.4.19}$$

One can also use these equations to obtain the expressions for the energy E_{circ} and angular momentum L_{circ} as functions of the radius r of the circular motion [Bardeen, Press, and Teukolsky (1972)]

$$E_{\text{circ}}/m = \frac{r^2 - 2Mr \pm a\sqrt{Mr}}{r(r^2 - 3Mr \pm 2a\sqrt{Mr}\,)^{1/2}}\,, \tag{3.4.20}$$

$$L_{\text{circ}}/m = \pm\frac{\sqrt{Mr}\,(r^2 \mp 2a\sqrt{Mr} + a^2)}{r(r^2 - 3Mr \pm 2a\sqrt{Mr}\,)^{1/2}}\,. \tag{3.4.21}$$

The upper signs in these and subsequent formulas correspond to direct orbits (i.e., corotating with $L_z > 0$), and the lower signs correspond to retrograde orbits (counter-rotating with $L_z < 0$). We always assume that $a \geq 0$. The coordinate angular velocity of a particle at the circular orbit is

$$\omega_{\text{circ}} = \frac{d\phi}{dt} = \frac{\pm\sqrt{Mr}}{r^2 \pm a\sqrt{Mr}}. \tag{3.4.22}$$

Circular orbits can exist only for those values of r for which the denominator in (3.4.20) and (3.4.21) is real so that

$$r^2 - 3Mr \pm 2a\sqrt{Mr} \geq 0. \tag{3.4.23}$$

The radius of the circular orbit closest to the black hole (the motion along it being at the speed of light) is

$$r_{\text{photon}} = 2M\left\{1 + \cos\left[\frac{2}{3}\arccos\left(\mp\frac{a}{M}\right)\right]\right\}. \tag{3.4.24}$$

This orbit is unstable. For $a = 0$, $r_{\text{photon}} = 3M$, while for $a = M$, $r_{\text{photon}} = M$ (direct motion) or $r_{\text{photon}} = 4M$ (retrograde motion).

The circular orbits with $r > r_{\text{photon}}$ and $E/m \geq 1$ are unstable. A small perturbation directed outward forces this particle to leave its orbit and escape to infinity on an asymptotically hyperbolic trajectory.

The unstable circular orbit on which $E_{\text{circ}} = m$ is given by the expression

$$r_{\text{bind}} = 2M \mp a + 2M^{1/2}(M \mp a)^{1/2}. \tag{3.4.25}$$

These values of the radius are the minima of periastra of all parabolic orbits. If a particle moving in the equatorial plane comes in from infinity (where its velocity $v_\infty \ll c$) and passes within a radius r_{bind}, it will be captured.

Finally, the radius of the boundary circle separating stable circular orbits from unstable ones is given by the expression

$$r_{\text{bound}} = M\{3 + Z_2 \mp [(3 - Z_1)(3 + Z_1 + 2Z_2)]^{1/2}\}, \tag{3.4.26}$$

where

$$Z_1 = 1 + (1 - a^2/M^2)^{1/3}\left[(1 + a/M)^{1/3} + (1 - a/M)^{1/3}\right],$$

$$Z_2 = (3a^2/M^2 + Z_1^2)^{1/2}. \tag{3.4.27}$$

The quantities r_{photon}, r_{bind}, and r_{bound} as the functions of the rotation parameter a/M are shown in Figure 3.4.

Table 3.4.3 lists r_{photon}, r_{bind}, and r_{bound} (in units of $r_g = 2GM/c^2$) for a black hole rotating at the limiting angular velocity, $a = M$, and gives a comparison with

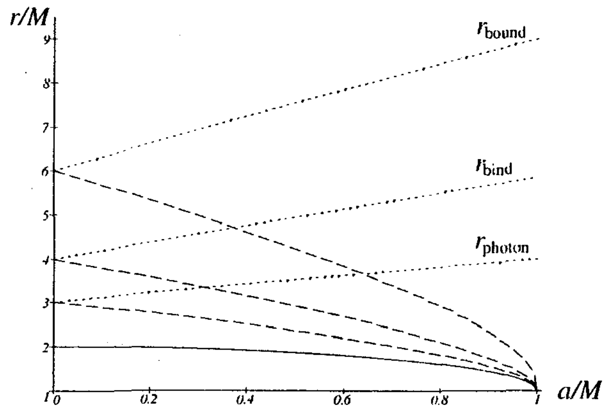


Figure 3.4: r_{photon}, r_{bind}, and r_{bound} as the functions of the rotation parameter a/M. The quantities corresponding to the direct and retrograde motions are shown by dashed and dotted lines, respectively.

the case of $a = 0$. Note that as $a \to M$, the invariant distance from a point r to the horizon r_+, equal to

$$\int_{r_+}^{r} \frac{r'\,dr'}{\Delta^{1/2}(r')},$$

diverges. As a result, it is not true that all three orbits coincide in this limit and lie at the horizon even though for $L > 0$ the radii r of all three orbits tend to the same limit r_+ [see Bardeen *et al.* (1972)].

Finally, we will give the values of specific energy E/m, specific binding energy $(m - E)/m$, and specific angular momentum $|L|/mM$ of a test particle at the last

Orbit	$a = 0$	$a = M$	
		$L > 0$	$L < 0$
r_{photon}	1.5	0.5	2.0
r_{bind}	2.0	0.5	2.92
r_{bound}	3.0	0.5	4.5

Table 3.1: r_{photon}, r_{bind}, and r_{bound} (in units of $r_g = 2GM/c^2$) for a non-rotating ($a = 0$) and extremely rotating ($a = M$) black holes.

stable circular orbit, r_{bound} (see Table 3.4.3).

	$a=0$	$a=M$	
		$L>0$	$L<0$
E/m	$\sqrt{8/9}$	$\sqrt{1/3}$	$\sqrt{25/27}$
$(m-E)/m$	0.0572	0.4236	0.0377
$\mid L\mid/mM$	$2\sqrt{3}$	$2/\sqrt{3}$	$22/3\sqrt{3}$

Table 3.2: Specific energy E/m, specific binding energy $(m-E)/m$, and specific angular momentum $|L|/mM$ of a test particle at the last stable circular orbit.

Equation (3.4.17) shows that particle motion with negative E is possible in the neighborhood of a rotating black hole. Let us solve this equation for E:

$$E = \frac{2aML + [L^2r^2\Delta + (r^3 + a^2r + 2Ma^2)(m^2r\Delta + r^3(dr/d\lambda)^2)]^{1/2}}{r^3 + a^2r + 2Ma^2}. \quad (3.4.28)$$

The positive sign of the radical was chosen in order to have the four-momentum of the particle pointing into future light cone [Misner, Thorne, and Wheeler (1973)]. In fact, for motion in the equatorial plane equation (3.4.4) implies that $dt/d\lambda > 0$ only if $E > 2aML_z/(r^3 + a^2r + 2Ma^2)$, so that the negative sign of the radical in (3.4.28) is excluded. The numerator in (3.4.28) is negative if $L < 0$ and the first term is of greater magnitude than the square root of the expression in brackets. The terms in the last parentheses can be made arbitrarily small ($m \to 0$ corresponds to the transition to an ultra-relativistic particle, and $dr/d\lambda \to 0$ is the transition to the motion in the azimuthal direction). Then E may become negative if we choose points inside the ergosphere, $r < r_{\text{st}}$. Additional constraints appear if $m \neq 0$ and $dr/d\lambda \neq 0$.

The expression (3.4.28) holds only for $\theta = \pi/2$. It is easy to show that orbits with negative E are possible within the ergosphere for any $\theta \neq 0$ and $\theta \neq \pi$. This follows from the fact that the Killing vector $\boldsymbol{\xi}_{(t)}$ is spacelike inside the ergosphere. The energy E is defined as $E = -p_\mu \xi^\mu_{(t)}$. Local analysis shows that for a fixed spacelike vector $\boldsymbol{\xi}_{(t)}$ it is always possible to find a timelike or null vector p^μ representing the momentum of a particle or a photon so that E is negative. Orbits with $E < 0$ make it possible to devise processes that extract the "rotational energy" of the black hole. Such processes were discovered by Penrose (1969). This phenomenon and its physical implications are discussed in detail in Section 7.1.

3.4.4 Motion off the equatorial plane

Now we consider some forms of motion of test particles off the equatorial plane. The simplest case is when particles are moving quasi-radially along trajectories on which the value of the polar angle θ remains constant, $\theta = \theta_0$. The relation between the integrals of motions corresponding to this type of motion can be found by solving simultaneously the equations

$$\Theta(\theta_0) = 0\,, \quad \left.\frac{d\Theta}{d\theta}\right|_{\theta_0} = 0\,. \tag{3.4.29}$$

There always exist trivial solutions: $\theta_0 = 0$, $\theta_0 = \pi$, and $\theta_0 = \pi/2$. The first two solutions correspond to motion along the axis of symmetry, passing either through the "north" or "south" pole. The last solution corresponds to motion in the equatorial plane, which we already discussed. If we exclude these cases, the relations between the integrals of motion can be written in the form

$$L_z^2 = a^2(E^2 - m^2)\sin^4\theta_0\,, \tag{3.4.30}$$

$$Q = -a^2(E^2 - m^2)\cos^4\theta_0\,. \tag{3.4.31}$$

Equation (3.4.30) shows that motion with constant $\theta = \theta_0$ is possible only when $E/m > 1$ (unbounded motion). For this motion $Q < 0$. We mentioned this fact already earlier.

Non-relativistic particles moving at parabolic velocity ($v_\infty = 0$) and zero angular momentum ($L_z = 0$) form a special limiting case. Such particles fall at constant θ and are dragged into the rotation around the black hole in the latitudinal direction at angular velocity (3.3.14). Therefore, these particles fall radially at each point in the reference frame of locally "non-rotating observers".

Another important limiting case is the infall of ultra-relativistic particles (photons) which move at infinity with $d\theta/d\lambda = 0$ and $L_z = aE\sin^2\theta$. Equations (3.4.1)–(3.4.4) reduce for such particles to

$$\frac{dr}{d\lambda} = -E\,, \quad \frac{d\theta}{d\lambda} = 0\,, \quad \frac{d\phi}{d\lambda} = \frac{aE}{\Delta}\,, \quad \frac{dt}{d\lambda} = \frac{(r^2 + a^2)E}{\Delta}\,. \tag{3.4.32}$$

The world lines of these photons will be used in constructing the Kerr coordinate system (Section 3.5).

3.4.5 Gravitational capture

By analogy to Section 2.9, consider the gravitational capture of particles by a rotating black hole [this topic is reviewed in Dymnikova (1986)].

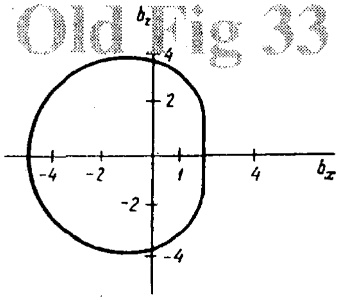


Figure 3.5: The capture impact parameter for particles moving with $v_\infty \to 0$ at right angles to the rotation axis of a black hole with $a = M$. Coordinate axes are marked off in units of M/v_∞. The x-direction is in the equatorial plane and z-direction is perpendicular to it.

The impact parameter $b_\perp$ for capture of a non-relativistic particle moving in the equatorial plane is given by the expression

$$b_\perp = \pm\, 2M \frac{1}{v_\infty}\left(1 + \sqrt{1 \mp \frac{a}{M}}\right). \tag{3.4.33}$$

The capture impact parameter for non-relativistic particles falling perpendicularly to the rotation axis of the black hole with $a = M$ plotted in Figure 3.5 [Young (1976)]. The corresponding capture cross-section is

$$\sigma_\perp \approx 14.2\, \pi (1/v_\infty)^2 M^2. \tag{3.4.34}$$

The impact parameter of particles falling parallel to the rotation axis, $b_\parallel$, can be found in the following manner. Let us define $\tilde{b}_\parallel = b_\parallel/M$, $\tilde{a} = a/M$. Then $\tilde{b}_\parallel$ is found as the solution of the equation

$$(1 - \tilde{a}^2)q_0^4 + 4(5\tilde{a}^2 - 4)q_0^3 - 8\tilde{a}^2(6 + \tilde{a}^2)q_0^2 - 48\tilde{a}^4 q_0 - 16\tilde{a}^6 = 0\,, \tag{3.4.35}$$

where $q_0 = v_\infty^2(\tilde{b}_\parallel^2 - \tilde{a}^2)$. If $\tilde{a} = 1$, then

$$\tilde{b}_\parallel \approx 3.85\left(\frac{1}{v_\infty}\right)M\,, \quad \sigma_\parallel \approx 14.8\, \pi \left(\frac{1}{v_\infty}\right)^2 M^2. \tag{3.4.36}$$

Consider now ultra-relativistic particles. The impact parameter of capture, $b_\perp$, for motion in the equatorial plane is

$$\frac{b_\perp}{M} = \begin{cases} 8\cos^3\left[\frac{1}{3}(\pi - \arccos\tilde{a})\right] + \tilde{a}\,, & L_z > 0, \\ -\,8\cos^3\left(\frac{1}{3}\arccos|\,\tilde{a}\,|\right) + \tilde{a}\,, & L_z < 0\,. \end{cases} \tag{3.4.37}$$

The capture cross-section for photons falling perpendicularly to the rotation axis of the black hole with $a = M$ is

$$\sigma_\perp \approx 24.3\,\pi M^2. \tag{3.4.38}$$

For photons propagating parallel to the rotation axis of the black hole with $a = M$, we have

$$\frac{b_\parallel}{M} = 2(1+\sqrt{2}\,), \qquad \sigma_\parallel \approx 23.3\,\pi M^2. \tag{3.4.39}$$

A comparison of cross-sections given in this section with those of Section 2.9 for $a = 0$ demonstrates that a rotating black hole captures incident particles with lower efficiency that a non-rotating black hole of the same mass.

3.5 Spacetime of a Rotating Black Hole

Let us consider some general properties of the spacetime of a rotating black hole described by the solution of (3.2.1). At first, we introduce a coordinate system which does not have coordinate singularities at the event horizon r_+ in the same manner as was done in the Schwarzschild spacetime.[10] In that case we could use the world lines of photons moving the center along the radii as coordinate lines [see (2.4.11)]. The world lines of photons moving toward a rotating black hole can also be chosen, but now the trajectories of the photons wind around the black hole in its neighborhood because they are dragged into rotation by the "vortex" gravitational field. Therefore, if the black hole rotates, we have to supplement the substitution of new coordinates [like (2.4.11)] with a "twist" in the coordinate ϕ.

The simplest expression for the metric is obtained if we use the world lines of photons that move at infinity at constant θ and whose projection of the angular momentum on the rotation axis of the black hole is $L_z = aE\sin^2\theta$ [see equation (3.4.30)], where E is the photon energy at infinity. It can be shown that a transition to such a reference frame of "freely falling" photons is achieved by a change of coordinates:

$$dv = dt + (r^2+a^2)\,\frac{dr}{\Delta}, \qquad d\tilde{\phi} = d\phi + a\,\frac{dr}{\Delta}. \tag{3.5.1}$$

The resulting coordinates are known as the *Kerr ingoing coordinates* [Kerr (1963)]:

$$\begin{aligned} ds^2 = &-\left[1-\Sigma^{-1}(2Mr)\right]\,dv^2 + 2\,dr\,dv + \Sigma\,d\theta^2 \\ &+\Sigma^{-1}\left[(r^2+a^2)^2 - \Delta\,a^2\sin^2\theta\right]\sin^2\theta\,d\tilde{\phi}^2 \\ &-2a\sin^2\theta\,d\tilde{\phi}\,dr - 4a\,\Sigma^{-1}Mr\sin^2\theta\,d\tilde{\phi}\,dv\,. \end{aligned} \tag{3.5.2}$$

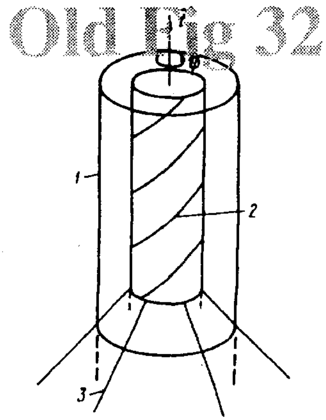


Figure 3.6: Spacetime of a rotating black hole: 1—null world line along the static limit, 2—"outgoing" photons forming the horizon, 3—photons falling into the black hole.

These coordinates are regular on the future event horizon.

The general properties of the geometry of a rotating black hole are best seen on a spacetime diagram in Kerr ingoing coordinates shown in Figure 3.6. Here the time coordinate $\tilde{t} = v - r$ is substituted for the coordinates v. We have already employed such diagrams in the Eddington coordinates in Chapter 2. The case we are considering now is essentially different in that the Kerr metric has axial but not spherical spatial symmetry.

Since one of the rotational degrees of freedom (the rotation translating a point along the "meridians" of θ) is not shown in these diagrams, they display information only on one chosen section (e.g., the equatorial plane $\theta = \pi/2$, as we see on Figure 3.6). This figure plots several world lines of photons that are important for describing the properties of the Kerr geometry. The first thing to remember is that the closer the coordinates are to the horizon, the more they are twisted around the black hole. In these coordinates the world lines of photons falling into the black hole are mapped by straight lines. In Boyer-Lindquist coordinates (a rigid grid; see Sections 3.2 and 3.3), they would appear twisted. Here the coordinate lines are twisted just as the photon trajectories are, so that these trajectories appear as straight with respect to the coordinate lines (in fact, we choose the coordinate lines precisely to have them coincide with the trajectories of the falling photons). At the static limit r_{st} [see (3.3.7)], the r, θ, $\tilde{\phi}$ = const world line is a null line tangent to the light cone. At $r < r_{st}$, all photons and particles necessarily participate in the rotational motion around the black hole, moving at $d\tilde{\phi}/d\tilde{t} > 0$. But they can escape from below the

[10] Of course, we ignore the trivial coordinate singularity at the pole of the spherical coordinate system: everyone is used to it and its meaning is obvious.

static limit to $r > r_{\rm st}$.

At the horizon, all timelike and null world lines point into the black hole, except a single null line, unique for each point of the horizon, of an "escaping" photon; this null line is tangent to the horizon. This family of world lines "winds up" on the horizon (see Figure 3.6), always staying on it. In Kerr coordinates, the equation of these null geodesics is

$$r = r_+ , \qquad \theta = \text{const} , \qquad \tilde{\phi} = \frac{av}{r_+^2 + a^2} . \tag{3.5.3}$$

All other photons and particles have to continue falling into the black hole after they reach the horizon.

If we put $dv = dr = 0$ and $r = r_+$ into the metric (3.5.2), we get the following two-dimensional metric describing the geometry of a section $v =$ const of the event horizon

$$dl^2 = (r_+^2 + a^2 \cos^2 \theta)\, d\theta^2 + \frac{(r_+^2 + a^2)^2}{r_+^2 + a^2 \cos^2 \theta} \sin^2 \theta \, d\tilde{\phi}^2. \tag{3.5.4}$$

The Gaussian curvature K for this metric is [Smarr (1973b)]

$$K = (r_+^2 + a^2) \frac{r_+^2 - 3a^2 \cos^2 \theta}{(r_+^2 + a^2 \cos^2 \theta)^3} . \tag{3.5.5}$$

For $a < \sqrt{3}\, M/2$ the Gaussian curvature is positive. The embedding diagram for this geometry is shown in Figure 3.7 . For $a > \sqrt{3}\, M/2$, regions near the poles $\theta = 0$ and $\theta = \pi$ have negative Gaussian curvature, and it is impossible to realize this geometry on a surface of rotation embedded in a three-dimensional flat space. The *surface area of the black hole* is

$$\mathcal{A} = \int \sqrt{g_{\theta\theta}\, g_{\phi\phi}}\, d\theta \, d\phi = 4\pi (r_+^2 + a^2) . \tag{3.5.6}$$

The Kerr metric is invariant under the transformation $t \to -t$, $\phi \to -\phi$ which transforms incoming light rays into outgoing ones; hence, this transformation can be performed in (3.5.1). The *Kerr outgoing coordinates* $(u, r, \theta, \bar{\phi})$, defined as

$$du = dt - (r^2 + a^2) \frac{dr}{\Delta} , \qquad d\bar{\phi} = d\phi - a \frac{dr}{\Delta} , \tag{3.5.7}$$

are regular on the past horizon. The equations $u =$ const, $\bar{\phi} =$ const describe the family of outgoing light rays, and the coordinate u at infinity coincides with the ordinary coordinate of retarded time. The Kerr metric is obtained in these coordinates from (3.5.2) by the transformation $v = -u$, $\tilde{\phi} = -\bar{\phi}$.

In contrast to the Schwarzschild metric, here we do not consider the continuation of the Kerr metric into the region within the horizon.[11] The reason is as follows. In

[11] The structure of the maximal analytic continuation of the Kerr-Newman metric is discussed in Section 6.6 .

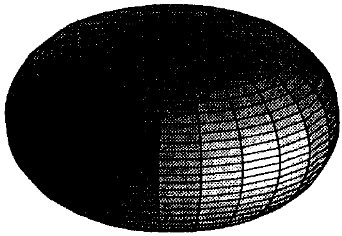


Figure 3.7: The embedding diagram for a two-dimensional section of the event horizon (v = const, $r = r_+$). The diagram is constructed for the critical value $a/M = \sqrt{3}/2$ of the rotation parameter so that the Gaussian curvature vanishes at the poles $K(\theta = 0) = K(\theta = \pi) = 0$.

the collapse of a spherical body (generating a Schwarzschild black hole), the spacetime metric beyond the collapsing body is exactly Schwarzschild one both inside and outside the black hole. In the collapse of a non-rotating slightly non-spherical body, the metric outside the black hole rapidly tends to the Schwarzschild metric as $t \to \infty$. We will see in Chapter 14 that the same property holds inside the Schwarzschild black hole. The inner region of the Schwarzschild metric thus describes the real "interior" of a non-rotating black hole.

These arguments do not hold for the Kerr metric. First, when an arbitrary rotating body contracts and turns into a black hole, the metric outside the body cannot become stationary immediately (and hence cannot be a Kerr metric) because gravitational waves are emitted in the course of the collapse. This statement holds both for the region outside the horizon and for that inside it. Outside the horizon all derivations from the Kerr metric are radiated away via gravitational waves (see Chapter 4), and the limiting metric at $t \to \infty$ is the Kerr solution (Chapter 6). In the external spacetime, therefore, the Kerr metric describes the real rotating black hole.

In the region $r \sim r_- = M - \sqrt{M^2 - a^2}$ inside the horizon, however, the metric does not tend to the Kerr solution either immediately after the collapse or at later stages. For this reason, this solution does not describe (inside the horizon) the inner structure of real rotating black hole (the detailed structure of black hole region inside the horizon is treated in Chapter 14). Note that all the above-discussed properties of the black hole spacetime are valid only if $M \geq |a|$. Otherwise, the horizon vanishes from the solution, and it ceases to describe the black hole. Pathological features appear [Hawking and Ellis (1973)] so that this solution may hardly relate to reality. From a physical standpoint, the formation of an object with $M < |a|$ requires the compression of a rotating body with such a high angular momentum that at $r \approx M$ the linear velocity of rotation inevitably exceeds the speed of light. Hereafter we always assume (for uncharged black holes) that $M \geq |a|$.

3.6 Charged Rotating Black Holes

3.6.1 Kerr-Newman geometry

In any realistic situation, the electric charge of a black hole is negligible. As a rule, the ratio of the charge Q to mass M of a black hole cannot exceed 10^{-18} [Wald (1984)]. For example, the charge-to-mass ratio of the electron and the proton are $(q/m)_e = 10^{21}$ and $(q/m)_p = 10^{18}$, respectively. The ratio of the gravitational force to the electrostatic one in the interaction of these particles with the black hole of charge Q and mass M are, in order of magnitude, qQ/mM. The ratio Q/M cannot be greater than $(q/m)^{-1}$; otherwise, charges of like sign would be repelled from the black hole, while charges of the opposite sign would fall into it and neutralize the electric charge of the black hole.

However, on theoretical grounds, it would be interesting to discuss, at least briefly, the general case of a rotating charged black hole.

The *Kerr-Newman metric* describing the geometry of a rotating charged black hole written in *Boyer-Lindquist coordinates* $x^\mu = (t, r, \theta, \phi)$ (similar to (3.2.1) is

$$ds^2 = -\left(1 - \frac{2Mr - Q^2}{\Sigma}\right) dt^2 - \frac{(2Mr - Q^2)2a\sin^2\theta}{\Sigma}\, dt\, d\phi$$

$$+ \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \left(r^2 + a^2 + \frac{(2Mr - Q^2)a^2\sin^2\theta}{\Sigma}\right)\sin^2\theta\, d\phi^2. \qquad (3.6.1)$$

Here we use the standard notations

$$\Delta = r^2 - 2Mr + a^2 + Q^2, \qquad \Sigma = r^2 + a^2\cos^2\theta\,. \qquad (3.6.2)$$

In addition to the gravitational field, the black hole is now surrounded with a stationary electromagnetic field which is completely determined by the charge Q and rotation parameter a. The vector potential of this field in the Boyer-Lindquist coordinates is written in the form

$$A_\alpha\, dx^\alpha = -\frac{Qr}{\Sigma}\,(dt - a\sin^2\theta\, d\phi)\,. \qquad (3.6.3)$$

If $a = 0$, the black hole does not rotate and the metric represents a spherically symmetric charged black hole with a spherically symmetric electric field (the Reissner (1916)-Nordström (1918) solution).

If the black hole rotates ($a \neq 0$), the electric field is supplemented by a magnetic field due to the dragging of the inertial reference frames into rotational motion around the black hole. At large distances from a black hole in the "rigid" reference frame (chronometric frame; see Section 3.3), which transforms at infinity into the Lorentz frame, the largest components of the electromagnetic field correspond to a monopole electric field with charge Q and a dipole magnetic field with magnetic moment $\mu^* =$

Qa. Other moments of the field are also expressed in terms of Q and a in a unique manner [for details, see Cohen and Wald (1971), Hanni and Ruffini (1973)]. For non-vanishing electric charge, the Kerr-Newman solution has a horizon only if $M^2 \geq Q^2 + a^2$. We consider only such solutions (cf. the discussion in Section 3.4).

The event horizon ($\Delta = 0$) is defined by the relation $r = r_+$, where

$$r_+ = M + \sqrt{M^2 - Q^2 - a^2}\,. \tag{3.6.4}$$

The surface area of a charged rotating black hole is

$$\mathcal{A} = 4\pi(r_+^2 + a^2)\,. \tag{3.6.5}$$

A static limit surface ($g_{tt} = 0$) is defined as $r = r_{\rm st}$, with

$$r_{\rm st} = M + \sqrt{M^2 - Q^2 - a^2\cos^2\theta}\,. \tag{3.6.6}$$

As for the uncharged black hole, the static limit surface is located outside the event horizon and crosses it at two polar points $\theta = 0$ and $\theta = \pi$. As before, the region $r_+ < r < r_{\rm st}$ between the static limit surface and the horizon is called the *ergosphere*.

3.6.2 Motion of test particles

The motion of a test particle in the Kerr-Newman metric can be written in a form similar to (3.4.1)–(3.4.4). We denote by E the conserved energy of the particle with charge q and mass m, and by L_z, the conserved projection of angular momentum on the black hole axis. Then we have

$$E = -(p_t + q\,A_t)\,, \qquad L_z = p_\phi + q\,A_\phi\,, \tag{3.6.7}$$

where p_α is the four-momentum of the particle. The equations of motion are written in the form

$$\Sigma\,\frac{dr}{d\lambda} = \pm\left\{[E(r^2+a^2) - L_z\,a - qQr]^2 - \Delta\,[m^2r^2 + (L_z - a\,E)^2 + \mathcal{Q}]\right\}^{1/2}, \tag{3.6.8}$$

$$\Sigma\,\frac{d\theta}{d\lambda} = \pm\left\{\mathcal{Q} - \cos^2\theta\left[a^2(m^2 - E^2) + \frac{L_z^2}{\sin^2\theta}\right]\right\}^{1/2}, \tag{3.6.9}$$

$$\Sigma\,\frac{d\phi}{d\lambda} = -\left(aE - \frac{L_z}{\sin^2\theta}\right) + \frac{a}{\Delta}\left[E(r^2+a^2) - L_z\,a - qQr\right], \tag{3.6.10}$$

$$\Sigma\,\frac{dt}{d\lambda} = -a(aE\sin^2\theta - L_z) + (r^2+a^2)\,\Delta^{-1}\left[E(r^2+a^2) - L_z\,a - qQr\right] \tag{3.6.11}$$

[the expression for $\mathcal{Q}$ is given in (3.4.7)]. The signs $\pm$ which enter (3.6.8) and (3.6.9) are independent of each another.

It must be emphasized that in this general form, the equations describe not only phenomena specific to black holes (these were mostly discussed in the preceding sections) but also their combination with ordinary effects caused by the motion of a test particle in an electromagnetic field. Special types of motion of charged test particles in the Kerr-Newman geometry were analyzed by Johnston and Ruffini (1974), Young (1976), Bičák *et al.* (1989), and Balek *et al.* (1989).

3.7 Problem of "visualization" of Black Holes and the Membrane Paradigm

In this as well as in the previous chapter we used different reference frames and various diagrams for visualizing numerous properties of black holes. Each diagram helps our intuition to catch one or another feature of these unusual objects. But probably neither of them is well suited for describing all of the details. This happens because the spacetime of a black hole is four-dimensional and highly curved, which makes the problem of its "visualization" very complicated. Fortunately, there are different tricks which can help. For example, to study spherically symmetric collapse it is possible to omit angular coordinates and to reduce the problem to a two-dimensional one. For different black hole problems different approaches could be specially useful. To study the interior of a black hole, it is useful to "look" at the processes through a "freely-falling observer's eyes" (Section 2.4). On the other hand, for many aspects of physics in the black hole exterior the point of view of a distant observer can be convenient. In the vicinity of the horizon the point of view of the "Rindler observer" is appropriate (see Appendix C). Penrose-Carter conformal diagrams, which will be described in Chapter 5 , allow one easily to get information about the global causal structure of spacetime in the presence of a black hole.

For many types of gedanken experiments with black holes and especially in the numerous astrophysical applications it is very convenient and fruitful to treat a black hole as some "object" in space and time, which has physical properties more or less similar to ordinary astrophysical objects (planets, stars, etc.), and hence which are familiar to us. The natural idea is to surround a black hole by some surface which is considered as the "boundary" of a black hole. The interaction of the black hole with the outside world can then be described in terms of special boundary conditions at this "boundary" surface. One of the possibilities is to identify the event horizon with such a surface. The corresponding formalism was developed by Znajek (1978) and Damour (1978). This formalism is very useful, but there is one property by which this "boundary" differs from the boundaries of ordinary bodies: The event horizon is null, while the boundaries of ordinary bodies are timelike. Thorne and collaborators [Thorne *et al.* (1986)] proposed to choose the "boundary" surface representing a black hole to be located slightly outside the event horizon. This surface is known as a *stretched horizon*. It is timelike and possesses a number of useful properties which allowed Thorne and collaborators to develop a powerful approach to black hole physics which they called the *membrane paradigm*. In the framework of this paradigm black holes are described as bodies which not only have mass, charge, and angular momentum, but have definite entropy, surface temperature, surface densities of charge and electric current, surface viscosity, and so on.

The membrane approach uses a (3+1)-split of spacetime outside the black hole, discussed in Section 3.2. The most striking property of a black hole as an isolated object is that from the point of view of an external observer all physical processes are infinitely slowed down near the event horizon. This is connected with the delay of signal propagation and the "freezing" of time in the regions close to the horizon in a system of an external observer (lapse function $\alpha \to 0$; see (3.3.16)). For a rotating black hole we identify the reference frame of an external observer with the ZAMO frame (see Section 3.3). In the non-rotating limit it reduces to the standard Schwarzschild frame. In both cases in the vicinity of the horizon these frames coincide with very high accuracy with the Rindler reference frame (see Appendix C).

A freely falling particle in the spacetime of a rotating black hole crosses the horizon in a finite proper time. From the point of view of a distant observer it takes an infinite time. Therefore, its velocity is slowed down, while its angular velocity measured by a distant observer tends to a universal limit (angular velocity of the black hole): $\Omega^H = c^3 a/(2GMr_+)$ (see (3.3.21)). This is the same angular velocity as for ZAMO observers near the horizon. In this sense freely falling particles become "frozen" at the horizon. The same happens with physical fields (see Section 8.3). This means that from the point of view of an external observer the structure of the field in the very thin layer near the horizon is very complicated. In fact, all the previous history of the field evolution and particles which have fallen down is imprinted in this layer. The data in this layer practically do not influence the events in the exterior region. If we are not interested in all these details in the immediate vicinity of the horizon, we can restrict our reference frame by a stretched horizon. The position of the stretched horizon is chosen so that it coincides with an equipotential surface of the lapse function α =const. The exact value of α is not specified, but it is chosen to be sufficiently small. As a result of the "freezing" effect, basic physical characteristics near the stretched horizon depend mainly on α. It is convenient to "normalize" them by multiplying by a suitable power of α. The boundary conditions for these "normalized" quantities at the stretched horizon to leading order are independent of the position of the stretched horizon. For small values of α the stretched horizon is practically a membrane, which can be penetrated by fields and particles from the exterior, but practically nothing can go through it from inside.

The properties of this membrane are quite attractive. The stretched horizon is a timelike surface as it is for an ordinary body. For concrete physical problems one can simply exclude the black hole interior from our considerations, by imposing special boundary conditions at the stretched horizon. In the process of stretching the horizon the entire distant past history of the fields at t = const close to the event horizon is cut off and ignored (see Figure 3.8). The concept of the stretched horizon works perfectly not only for stationary black holes but also in fully dynamical situations; for example, for the collision of a body of finite mass with a black hole. The membrane paradigm is discussed in detail in the book by Thorne *et al.* (1986), where the relevant bibliography can be found. In our book we shall discuss some aspects of the membrane approach in Chapter 8 in connection with astrophysical applications of black hole physics. A special attraction of the membrane approach for the consideration of astrophysical problems is connected with the possibility of

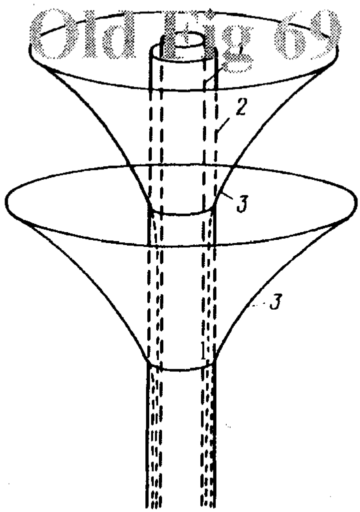


Figure 3.8: (3+1)-split of spacetime close to the black hole horizon: 1 – black hole event horizon, 2 – stretched horizon, 3 – t =const sections.

transferring directly to black hole problems methods of electrodynamics and plasma physics developed earlier for pulsar physics.

In concluding this section, we would like to emphasize that the membrane approach is a useful technique, but certainly there are no physical membranes outside a black hole. A freely falling observer will see no "special membrane". This surface in spacetime as well as the event horizon have no special local properties which allow one to distinguish their points from other points of the spacetime. A free falling observer discovers that the membrane with its "familiar properties" is fictitious, while an observer who remains outside the stretched horizon might insist that it (as well as surface charges, currents, and so on) is "perfectly real". Let us stress once again that the membrane approach is useful only for the description of physics in the exterior region. When we are interested in physics in the black hole interior, we should use other methods.

Chapter 4

Black hole Perturbations[1]

4.1 Introduction

This chapter is devoted to problems involving perturbed black holes. We describe the evolution of physical fields in the exterior of Schwarzschild and Kerr black holes. As we shall see, an understanding of wave propagation in black hole geometries opens the door to studies of:

- The radiation from a slightly non-spherical gravitational collapse
- The radiation generated by bodies falling into black holes
- The late-time (after black hole formation) behavior of the gravitational field
- Scattering and absorption of waves by black holes
- The stability of the Schwarzschild and the Kerr solutions
- The generation of gravitational waves by coalescing binary systems

and several other relevant problems.

The physical fields considered here are always assumed to be weak in the sense that the effect of their energy-momentum on the background metric of a black hole can be neglected. In essence, we consider only linear perturbations of Schwarzschild and Kerr black holes. One may object that such studies can have little relevance since general relativity is a manifestly nonlinear theory. However, it turns out that the linear approximation is surprisingly robust. It provides useful approximations that serve as benchmark tests for fully nonlinear studies.

Theoretical studies of perturbed black holes were pioneered by Regge and Wheeler in the late 1950's. Their aim was to find whether a small perturbation of a black hole would become unbounded if evolved according to the linearized version of Einstein's equations. If that were the case, black holes could clearly not be considered as

[1] Written jointly with N. Andersson

astrophysically relevant. In their original work, Regge and Wheeler (1957) studied perturbations of the metric directly. In principle, this is done by introducing

$$g_{\mu\nu} = g_{\mu\nu}^{\text{background}} + h_{\mu\nu} , \tag{4.1.1}$$

where $|h_{\mu\nu}|$ is considered (in some sense) small. Then only terms linear in $h_{\mu\nu}$ are retained in all calculations. This way it can be shown that a wave equation with an effective potential governs linear perturbations of a Schwarzschild black hole. This equation is (for obvious reasons) now known as the *Regge-Wheeler equation.*

An infinitesimal coordinate transformation $x^\mu \to x^\mu + \xi^\mu(x)$ in the decomposition (4.1.1) implies that the metric perturbations will transform as $h_{\mu\nu} \to h_{\mu\nu} + 2\,\xi_{(\mu;\nu)}$. One would expect that the quantity determined by the Regge-Wheeler equation depends on the chosen gauge. However, although the original derivation of Regge and Wheeler was carried out in a specific gauge, one can show that the quantity governed by the final wave equation is gauge invariant [Moncrief (1974a)]. Thus, all the information contained in the Regge-Wheeler equation is physically relevant.

One of the first applications of the theory of small gravitational perturbations was an analysis of a collapse of slightly non-spherical rotating bodies [Doroshkevich, Zel'dovich, and Novikov (1965) and Novikov (1969)]. This analysis shows that only rotational perturbations, are not "dying", while all other deflections from sphericity disappear.

The approach of Regge and Wheeler was later extended to the case of electrically charged (Reissner-Nordström) black holes [Zerilli (1974), Moncrief (1974bc, 1975), Bičák (1972, 1980b, 1982)], but the problem was found to be far more complicated in the case of rotating (Kerr) black holes. That it is possible to reduce also the Kerr problem to a single wave equation was shown by Teukolsky (1972, 1973) who used the Newman-Penrose formalism (see Appendix E). This leads to an elegant – gauge and tetrad invariant – description of perturbed black holes.

This chapter will deal only briefly with the derivation of the perturbation equations for black holes. We will focus more on their application in various physical scenarios. Nevertheless, it is important to understand that the formulation of these problems required an enormous effort by many physicists. An interesting account that puts some of this work into its proper historical context has recently been given by Thorne (1994b). Here the emphasis will be on the formulation of various physically relevant problems. We will outline the method of solution and highlight the most important conclusions. We refer the interested reader to the original texts for further details. A more complete mathematical treatment can be found in Chandrasekhar's book (1983). A physically clear representation of the main aspects can be found in the monograph of Misner, Thorne, and Wheeler (1973). Perturbation theory is reviewed and used in the context of gravitational synchrotron radiation in Breuer's book (1975). The scattering theory in the context of black holes is discussed in great detail in the book by Futterman, Handler, and Matzner (1988). Review articles that contain relevant material cover gravitational collapse and power-law tails [Thorne

(1972)], the astrophysics of black holes [Eardley and Press (1975)], gravitational waves [Thorne (1978)], quasinormal modes [Detweiler (1979)], the mathematical description of the perturbation problem [Chandrasekhar (1979b)] and test bodies falling into black holes [Nakamura, Oohara, and Kojima (1987)]. These texts provide useful complements and alternatives to the present material [see also Andersson (1997)].

4.2 Weak Fields in the Schwarzschild Metric

4.2.1 Scalar field near a spherically symmetric black hole

In order to demonstrate essentials of the perturbation problem, we begin by considering a test massless scalar field in the Schwarzschild geometry. Even though this case is somewhat unphysical, we will use it to illustrate the basic concepts in black-hole perturbation theory throughout this chapter. We do so because the analytical complexity of physically more relevant fields (spin 1/2, 1, 2) often obscures the simple underlying ideas. This is especially true for problems involving Kerr black holes.

A massless scalar field Φ evolves according to the Klein-Gordon equation

$$\Box\Phi \equiv (-g)^{-1/2}\,\partial_\mu\left[(-g)^{1/2}\,g^{\mu\nu}\,\partial_\nu\Phi\right] = 0\,, \tag{4.2.1}$$

where g is the determinant of the metric tensor $g_{\mu\nu}$. The components of $g_{\mu\nu}$ follow immediately from the Schwarzschild line element

$$ds^2 = -\left(1-\frac{2M}{r}\right)dt^2 + \left(1-\frac{2M}{r}\right)^{-1}dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\,. \tag{4.2.2}$$

Because the metric is spherically symmetric, it is natural to introduce the mode decomposition

$$\Phi_{\ell m} = \frac{u_\ell(r,t)}{r}Y_{\ell m}(\theta,\phi)\,, \tag{4.2.3}$$

where $Y_{\ell m}$ are the standard *spherical harmonics*. The function $u_\ell(r,t)$ then solves the wave equation

$$\left[\frac{\partial^2}{\partial r_*^2} - \frac{\partial^2}{\partial t^2} - V_\ell(r)\right]u_\ell(r,t) = 0\,. \tag{4.2.4}$$

In the above equation we have used the so-called *tortoise coordinate* r_* that was first introduced by Wheeler (1955) and that is related to the standard Schwarzschild radial coordinate r by

$$\frac{d}{dr_*} = \left(1-\frac{2M}{r}\right)\frac{d}{dr}\,, \tag{4.2.5}$$

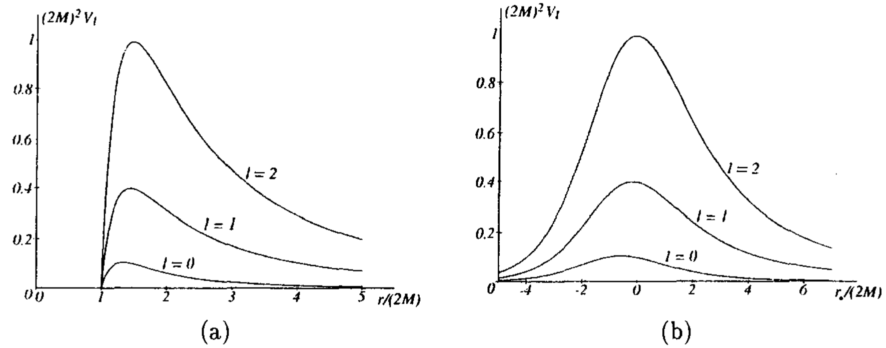


Figure 4.1: The effective potential V_ℓ for $\ell = 0, 1, 2$ as the function of r (a) and r_* (b). The constant in the definition (4.2.7) of r_* is fixed so that $r_* = 0$ for $r = 3M$.

or

$$r_* = r + 2M \log \left(\frac{r}{2M} - 1 \right) + \text{constant} \,. \tag{4.2.6}$$

We already met this coordinate in Section 2.4.3. It is the natural variable to use in the perturbation problem because $dr_* = \pm\, dt$ are null radial lines in the background geometry. The choice of integration constant in (4.2.6) is irrelevant in most situations, but a specific choice can prove useful in considerations of wave scattering by black holes [Futterman, Handler, and Matzner (1988)].

If we further assume a harmonic time dependence, $u_\ell(r,t) = \hat{u}_\ell(r,\omega)\, e^{-i\omega t}$, we get the ordinary differential equation

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - V_\ell(r) \right] \hat{u}_\ell(r,\omega) = 0 \,. \tag{4.2.7}$$

As we shall see in Section 4.2.4, this equation has the same form as that derived by Regge and Wheeler for gravitational perturbations in 1957. Hence, we will from now on refer to (4.2.7) as the *Regge-Wheeler equation*.

The effective potential in (4.2.7) can be written

$$V_\ell(r) = \left(1 - \frac{2M}{r} \right) \left[\frac{\ell(\ell+1)}{r^2} + \frac{2M}{r^3} \right] , \tag{4.2.8}$$

where M is the mass of the black hole. Thus, the scalar field problem is, in many ways, remarkably simple. The effective potential $V_\ell(r)$ corresponds to a single potential barrier, the maximum of which is at roughly the location of the unstable circular photon orbit ($r = 3M$) (see Figure 4.1). Hence, most problems concerning perturbed black holes involve elements familiar from potential scattering in quantum mechanics. For example, one would expect waves of short wavelength $\lambda \ll 2M$ to be easily transmitted through the barrier. Waves with $\lambda \sim 2M$ will be partly transmitted and partly reflected, and finally waves with $\lambda \gg 2M$ should be completely reflected by the black hole barrier.

Since the potential vanishes at both spatial infinity and the event horizon of the black hole, it follows immediately that two linearly independent solutions to (4.2.7) behave asymptotically as

$$\hat{u}_\ell(r,\omega) \sim e^{\pm i\omega r_*} \tag{4.2.9}$$

both as $r \to +\infty$ and $r \to 2M$. The tortoise coordinate differs from what might be an ordinary radial variable by a logarithmic term. This has the effect that $r_* \to +\infty$ as $r \to +\infty$ (at spatial infinity), but $r_* \to -\infty$ as $r \to 2M$ (at the event horizon). Thus, the semi-infinite interval $(2M, +\infty)$, representing the spacetime external to the black hole, is mapped into the infinite interval $(-\infty, +\infty)$. The event horizon has been "pushed all the way to $-\infty$". For large r the logarithmic behavior in (4.2.6) acts in the same way as the Coulomb logarithmic phase in quantum mechanical problems. It introduces a phase-shift that arises even at great distances from the black hole. This effect of the long-range nature of Newtonian gravity (or the Coulomb interaction) is evident in (4.2.9).

4.2.2 A useful set of basic solutions

In virtually, all applications of the theory of black holes (and certainly in the astrophysical ones) it is implicitly assumed that a black hole is formed as a result of the gravitational collapse of matter. The Schwarzschild metric is, of course, valid only in a region outside the collapsing matter. Similarly, the Kerr metric describes spacetime accurately only some time after the collapse of a rotating non-spherical body. When working in the late-time region, it is often convenient to "forget" about the details of the collapse. For example, to describe the wave scattering by black holes or the radiation from a moving test particle, it is sufficient to work in the corresponding Schwarzschild or Kerr metric. But one should keep in mind that these are identical to the "real" geometry only in the late-time region.

If we want to construct a simple basis in the space of solutions to wave equations in the black hole spacetime, the following trick proves to be very useful. Consider an eternal black hole[2] with the same parameters as the metric in the late-time region. We call this the *eternal version of the black hole*. To specify useful basic solutions, one can then impose boundary conditions at past infinity and the past horizon in the spacetime of the eternal version of the black hole. When propagated to the future

[2]Following adopted tradition, we use the notion "eternal black hole" as the synonym for the complete empty Schwarzschild or Kerr spacetime. Strictly speaking, this terminology is not very accurate since, as we have seen earlier (in Section 2.7), such a spacetime always contains both black and white holes.

according to the field equations, these solutions evidently form a basis in the late-time region.

Given the asymptotic behavior (4.2.9), it is easy to define solutions to (4.2.7) that satisfy simple boundary conditions. One such solution satisfies the causal condition that no waves should emerge from the black hole. In the notation of Chrzanowski and Misner (1974) this is the *IN*-mode, which is defined by

$$\hat{u}_\ell^{\rm in}(r_*,\omega) \sim \begin{cases} e^{-i\omega r_*} , & r_* \to -\infty , \\ A_{\rm out}(\omega)\, e^{i\omega r_*} + A_{\rm in}(\omega)\, e^{-i\omega r_*} , & r_* \to +\infty , \end{cases} \tag{4.2.10}$$

where ω is assumed to be real and positive. Note that this leaves the actual normalization of the solution unspecified. In the problems that we discuss in the present chapter we never need to specify this normalization. The situation is different when one considers quantum processes. We will discuss this in more detail in Chapter 10.

It is well known (and easy to demonstrate) that the Wronskian of any two linearly independent solutions to (4.2.7) must be a constant. Hence, using $\hat{u}_\ell^{\rm in}$ and its complex conjugate (the *OUT*-mode) and evaluating the Wronskian at $r_* = \pm\infty$, one can show that

$$1 + |A_{\rm out}|^2 = |A_{\rm in}|^2 . \tag{4.2.11}$$

Introducing the transmission (T) and reflection (R) amplitudes

$$T = \frac{1}{A_{\rm in}} , \qquad R = \frac{A_{\rm out}}{A_{\rm in}} , \tag{4.2.12}$$

this can be recognized as the standard scattering relation

$$|T|^2 + |R|^2 = 1 . \tag{4.2.13}$$

Thus, the part of an incident wave that is not absorbed by the black hole is reflected back to infinity. The quantities $|T|^2$ and $|R|^2$ are known as the transition and reflection probabilities (or coefficients), respectively.

A second pair of basic solutions (the *UP*- and *DOWN*-modes) can be defined in a similar way. The *UP*-mode corresponds to purely outgoing waves at spatial infinity, and is specified by the asymptotics

$$\hat{u}_\ell^{\rm up}(r_*,\omega) \sim \begin{cases} B_{\rm out}(\omega)\, e^{i\omega r_*} + B_{\rm in}(\omega)\, e^{-i\omega r_*} , & r_* \to -\infty , \\ e^{+i\omega r_*} , & r_* \to +\infty . \end{cases} \tag{4.2.14}$$

The *DOWN*-mode is the complex conjugate of this function. The coefficients in (4.2.14) are not independent of those that defined the *IN*-mode. Again using constant Wronskians, one can show that

$$B_{\rm out}(\omega) = A_{\rm in}(\omega) , \tag{4.2.15}$$

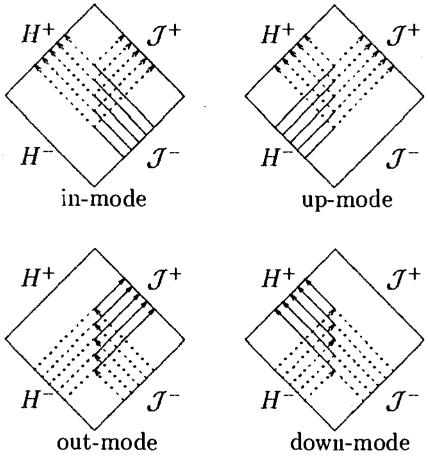


Figure 4.2: *IN-*, *UP-*, *OUT-*, and *DOWN*-modes .

$$B_{\text{in}}(\omega) = -\bar{A}_{\text{out}}(\omega) = -A_{\text{out}}(-\omega)\,. \tag{4.2.16}$$

A bar is used to denote the complex conjugate. Any two of the set introduced above (*IN-*, *UP-*, *OUT-*, and *DOWN*-modes) can be chosen as basic solutions in a given problem. In the following sections we will typically use the combination *IN–UP* and the Wronskian

$$W(IN,UP) \equiv \hat{u}_\ell^{\text{in}} \frac{d\hat{u}_\ell^{\text{up}}}{dr_*} - \hat{u}_\ell^{\text{up}} \frac{d\hat{u}_\ell^{\text{in}}}{dr_*} = 2i\omega A_{\text{in}}(\omega) = 2i\omega B_{\text{out}}(\omega) \tag{4.2.17}$$

will then prove useful.

Let us digress on the physical meaning of these solutions. By combining the radial solutions with $\exp(-i\omega t)$, we get functions describing wave propagation. They have simple physical interpretations. The *DOWN*-solution satisfies the boundary condition that there is no radiation escaping to infinity. This means that exactly the right amount of radiation with just the right phase must emerge from the past horizon H^- to cancel any radiation that might otherwise be scattered back to infinity from a wave originally incoming from past infinity. Thus, in this solution there is radiation coming in from infinity, radiation emerging from H^- to meet it, and radiation going down the black hole at H^+. The amplitudes of the various waves are such that *DOWN*-mode is an acceptable solution to the radial wave equation (4.2.7). The *UP*-mode is defined analogously by the boundary condition that there be no incoming radiation from infinity. In a similar way the *IN*-solution does not contain radiation outgoing from H^-, while the *OUT*-mode has no radiation going down the black hole, at H^+. The situation is presented graphically in Figure 4.2.

One can use the diagrams presented in Figure 4.2 as mnemonic rules for the definition of the basic functions. The regions inside the squares represent the spacetime in the exterior of the eternal version of the black hole. The straight lines at the angle of 45° represent null rays. Two boundaries $\mathcal{J}^+$ and $\mathcal{J}^-$ correspond to asymptotic future and past infinities. The other two boundaries H^+ and H^- are the event horizon and the past horizon, respectively. We shall see in Chapter 5 that this type of diagrams can be obtained by special conformal transformations that bring infinitely removed points of the spacetime to a finite distance. The corresponding *Penrose-Carter conformal diagram* proved to be a very powerful tool for the study of the global structure of a spacetime. Asymptotic values of massless fields in an asymptotically flat physical spacetime are related to the boundary values at the null surfaces $\mathcal{J}^+$ and $\mathcal{J}^-$, representing the so-called future and past null infinities. We shall give the exact definition of an asymptotically flat spacetime and discuss Penrose-Carter conformal diagrams in more detail in the next chapter. In this chapter we shall refer to diagrams presented in Figure 4.2 only in order to recall the boundary conditions used in the definition of different basic solutions.

4.2.3 Weak fields of higher spins

Let us now consider a general massless field with integral spin s in the Schwarzschild spacetime [Thorne (1976)]. For $s = 0$ this is the scalar case that was considered above, $s = 1$ corresponds to electromagnetic waves, and $s = 2$ pertains to gravitational waves. A complete set of gauge-invariant dynamic variables can be found for each field. That is, there exist a set of functions $\Phi^{(s)}$, defined in the exterior of the black hole, such that [Price (1972a,b), Bardeen and Press (1973)]:

1. $\Phi^{(s)}$ and $\partial_t\Phi^{(s)}$ can be fixed at any initial instant of time.
2. Once $\Phi^{(s)}$ and $\partial_t\Phi^{(s)}$ are fixed, their evolution is completely determined by a single wave equation.
3. Once $\Phi^{(s)}$ is known, it is straightforward to calculate all relevant quantities for the field (such as the energy and the angular momentum flux).
4. $\Phi^{(s)}$ can be determined from given parameters of the field.

Thus, knowledge of the behavior of $\Phi^{(s)}$ is equivalent to knowing the evolution of the field of interest. The perturbation problem consequently reduces to the one of finding $\Phi^{(s)}$.

The general approach to this problem is as follows. First, the field is expanded into spherical harmonics (scalar harmonics for $s = 0$, vector harmonics for $s = 1$, tensor harmonics for $s = 2$, and so on). Each spherical harmonic is characterized by its multipole index ℓ: $\ell = 0$ for a monopole, $\ell = 1$ for a dipole, etcetera. Multipoles for which $\ell < s$ do not evolve with time. They correspond to conserved quantities. For obvious reasons, we consider only the non-trivial "radiative" multipoles with $\ell \geq s$ in the following. For gravitational perturbations only multipoles with $\ell \geq 2$ are thus relevant. A gravitational perturbation with $\ell = 0$ describes a change in the black hole mass, and $\ell = 1$ corresponds to a displacement as well as a small increment of angular momentum (rotation of the black hole).

Price (1972a,b) showed that for each radiative multipole (of any spin s) there is a scalar field $\Phi_\ell^{(s)}$ that depends only on r and t. From this scalar quantity one can reconstruct all components of the original field. Each of the functions $\Phi_\ell^{(s)}$ satisfies a wave equation [similar to (4.2.4)] with an effective potential. This potential is a function of ℓ and s (and, of course, r and M as well) and can be written as

$$V_\ell^{(s)}(r) = \left(1 - \frac{2M}{r}\right)\left[\frac{\ell(\ell+1)}{r^2} - \frac{2M(1-s^2)}{r^3}\right] . \tag{4.2.18}$$

Hence, the potentials for fields of different spin are very similar. For this reason, one would expect the evolution of radiative multipoles with different values of s to be quite similar as well. In his study Price found that multipoles corresponding to $\ell \geq s$ are completely radiated away. Thus, only the conserved multipoles remain, and (as Wheeler put it) "the black hole has no hair".

The case of gravitational perturbations is especially interesting. It will be discussed in more detail in the following section. But before moving on to that discussion we refer the reader interested in the analysis of non-classical fields to papers by Hartle (1971, 1972), Unruh (1973, 1974), Martellini and Treves (1977), Iyer and Kumar (1978), Lee (1977) and Chandrasekhar (1976b). The books by Chandrasekhar (1983) and Sibgatullin (1984) also contain relevant discussions.

4.2.4 Gravitational perturbations of a Schwarzschild black hole

Gravitational perturbations ($s = 2$) of a black hole are of special interest because of the ongoing search for gravitational waves (see Section 9.9 for an introductory discussion). If we manage to open the gravitational-wave window to the Universe, it is important that we have some understanding of the physics involved in the generation of such waves. Black holes with their extremely large gravitational field are considered among the most promising sources of the gravitational waves in the Universe [see Thorne (1987, 1994a) for a discussion of the large-scale interferometric gravitational-wave detectors that are presently under construction and possible sources of gravitational waves] . Hence, there has been much work devoted to gravitationally perturbed black holes.

Gravitational perturbations are fundamentally different from scalar and electromagnetic ones. Whereas the latter two correspond to wave fields which evolve in a fixed background geometry, the gravitational case corresponds to minute changes in the metric itself. It therefore seems natural to begin by considering a general perturbation of the Schwarzschild metric. The perturbed metric can be written in the general axisymmetric form [Chandrasekhar (1983)]

$$ds^2 = -e^{2\nu}\, dt^2 + e^{2\psi} \left(d\phi - q_1\, dt - q_2\, dr - q_3\, d\theta\right)^2 + e^{2\mu_2}\, dr^2 + e^{2\mu_3}\, d\theta^2 \ . \quad (4.2.19)$$

In the unperturbed case

$$e^{2\nu} = e^{-2\mu_2} = 1 - \frac{2M}{r} \ , \quad (4.2.20)$$

and

$$e^{\mu_3} = r \ , \quad e^{\psi} = r \sin\theta \ , \quad q_1 = q_2 = q_3 = 0 \ . \quad (4.2.21)$$

When this metric is perturbed by an external agent:

- q_1, q_2, and q_3 become first-order quantities
- Linear increments $\delta\nu, \delta\psi, \delta\mu_2$, and $\delta\mu_3$ must be taken into account

It is easy to see that these two cases correspond to different kinds of perturbations. The first induces dragging of inertial frames (rotation of the black hole), whereas the second set is independent of the sign of ϕ and therefore does not induce rotation. In the following we will refer to the first kind of perturbations as "axial" and the second, as "polar". This terminology was introduced by Chandrasekhar (1983). In the literature the different perturbations are often referred to as *odd-parity* and *even-parity*, respectively. The reason for this nomenclature is that the first set transforms as $(-1)^\ell$ under space inversion. The second set, on the other hand, transforms as $(-1)^{\ell+1}$. Strictly speaking, this means that the first kind of perturbations are of alternating odd/even parity for even/odd multipoles ℓ. Chandrasekhar's (1983) terminology avoids this confusion.

It turns out that the two kinds of perturbations decouple and can consequently be studied separately. In both cases one can derive a wave equation analogous to (4.2.4). The axial equation was first derived by Regge and Wheeler in 1957 and the corresponding effective potential $V_\ell^{(2)}(r)$ is given by (4.2.18) above. The equation governing polar perturbations was first derived by Zerilli (1970b) [see also Edelstein and Vishveshwara (1970)], and the relevant effective potential can be written

$$V_\ell(r) = 2\left(1 - \frac{2M}{r}\right) \frac{n^2(n+1)r^3 + 3n^2Mr^2 + 9nM^2r + 9M^3}{r^3(nr+3M)^2} \ , \quad (4.2.22)$$

where

$$n = \frac{1}{2}(\ell - 1)(\ell + 2) \ . \quad (4.2.23)$$

Even though the axial and the polar potentials have quite different mathematical appearance, they contain much the same physical information. For example, the corresponding reflection and transmission coefficients are identical. The reflected waves differ only in their phases. As was shown by Chandrasekhar (1975) [see also

Chrzanowski (1975), Heading (1977) and Chandrasekhar (1980)], the different equations can be transformed into each other by means of differential operations. They are also connected to the master equation that Bardeen and Press (1973) derived via the Newman-Penrose formalism. In fact, one can show that there exist an infinite number of potentials that contain the same information [Anderson and Price (1991)]. A legitimate question is whether it is possible to further simplify the equations. Anderson and Price conclude that it is not, but the reduction of a general perturbation problem into one of a single wave equation is, by itself, remarkable.

4.2.5 Solutions in the high- and low-frequency limits

Before ending this general description of the equations that govern a perturbed black hole, we will briefly discuss two limiting cases for which approximate solutions can be obtained analytically. These cases correspond to high and low frequencies, respectively. For clarity we will only give the results for scalar perturbations of Schwarzschild black holes, but similar formulas can be derived for other spin fields (and also for Kerr black holes).

In the high frequency limit, when $\omega M \to +\infty$, the Regge-Wheeler equation can be approximated by a confluent hypergeometric equation [see e.g., Liu and Mashhoon (1995)]. Then one can show that the asymptotic amplitudes that are relevant for the construction of the solution $\hat{u}_\ell^{\text{in}}(r_*,\omega)$, cf. (4.2.10), are

$$A_{\text{out}} \approx \frac{\Gamma(1-4i\omega M)(4i\omega M)^{-1/2+4i\omega M}}{\sqrt{\pi}} e^{-4i\omega M} , \tag{4.2.24}$$

$$A_{\text{in}} \approx \frac{i\Gamma(1-4i\omega M)(4i\omega M)^{-1/2}}{\Gamma(1/2-4i\omega M)} . \tag{4.2.25}$$

Given these equations, it is easy to use Stirling's formula and show that the reflection coefficient behaves as

$$|R|^2 \sim e^{-8\pi\omega M} \quad \text{as} \quad \omega M \to +\infty . \tag{4.2.26}$$

This is the anticipated behavior: For very high frequencies a wave is hardly affected by the presence of the curvature potential at all, and the reflection coefficient is exponentially small.

For low frequencies, when $\omega M << 1$, the desired approximations can be obtained, using the approach of Starobinsky and Churilov (1973) [see also Page (1976b)]. They approached the problem for Kerr black holes and arbitrary spin. The overall solution is found by matching two solutions: For large r the differential equation again becomes a confluent hypergeometric equation, and for small r it reduces to a hypergeometric equation. By matching the two, one can infer that (for scalar waves and Schwarzschild black holes)

$$A_{\text{out}} \sim (4i\omega M)^{-\ell-1} \left\{ 1 - \frac{1}{4}\left[\frac{(\ell!)^3}{(2\ell)!(2\ell+1)!}\right]^2 (4\omega M)^{2\ell+2} \right\} , \tag{4.2.27}$$

$$A_{\rm in} \sim (-4i\omega M)^{-\ell-1}\left\{1+\frac{1}{4}\left[\frac{(\ell!)^3}{(2\ell)!(2\ell+1)!}\right]^2(4\omega M)^{2\ell+2}\right\} . \tag{4.2.28}$$

(A similar approximation was recently derived by Poisson and Sasaki (1995).) It follows from this result that the behavior of the reflection coefficient for low-frequency waves is

$$|R|^2 \approx 1-\left[\frac{(\ell!)^3}{(2\ell)!(2\ell+1)!}\right]^2(4\omega M)^{2\ell+2} . \tag{4.2.29}$$

That is, as $\omega M \to 0$, waves are totally reflected by the black hole barrier (the reflection coefficient approaches unity). This is, of course, exactly what one would expect.

4.3 Evolution of Wave Fields Around a Black Hole

4.3.1 Early "numerical relativity"

In the early 1970's perturbation equations like (4.2.4) were used as the basis for numerical investigations of various "physical" scenarios involving black holes. The problems that were studied fall into three main categories:

- Wavepackets scattering by black holes [Vishveshwara (1970a), Press (1971)]
- Small bodies (test particles) falling into – or passing close by – black holes [Davis *et al.* (1971, 1972), Ruffini (1973)]
- Slightly non-spherical gravitational collapse to form a black hole [de la Cruz, Chase, and Israel (1970); Cunningham, Price, and Moncrief (1978, 1979); Gaiser and Wagoner (1980)]

These studies led to some surprising results. The radiation generated when a black hole interacts dynamically with its surroundings is almost completely independent of the perturbing agent. The general features of the emitted waves are always similar, and can be divided into three components:

1. An initial wave burst that contains radiation emitted directly by the source of the perturbation
2. Exponentially damped "ringing" at frequencies that do not depend on the source of the perturbation at all
3. A power-law "tail" that arises because of backscattering by the long-range gravitational field

These features are illustrated in Figure 4.3.

It became clear that a perturbed black hole oscillates at frequencies that are characteristic to the black hole. The spectrum depends only on the three parameters

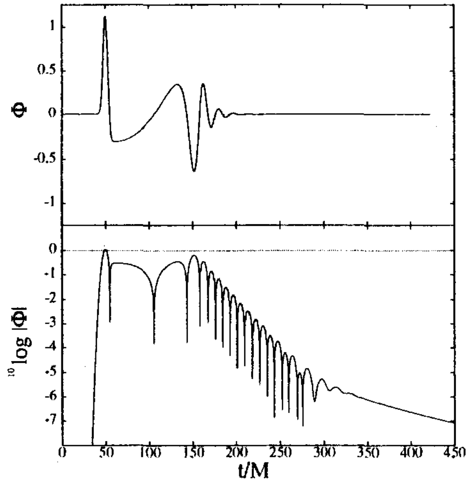


Figure 4.3: The response of a Schwarzschild black hole as a Gaussian wavepacket of scalar waves impinges upon it. The first bump (at $t = 50M$) is the initial Gaussian passing by the observer on its way towards the black hole. Quasinormal-mode ringing clearly dominates the signal after $t \approx 150M$. At very late times (after $t \approx 300M$) the signal is dominated by a power-law fall-off with time.

of the black hole: its mass, electric charge, and angular momentum. The frequencies have complex values because of radiation damping. These ringing modes are now known as the "quasinormal" modes of the black hole. We will discuss these modes in more detail in Section 4.4.

4.3.2 A Green's function analysis

Let us analyze the response of a black hole to an external perturbation in somewhat more detail. As a useful example, we consider the evolution of a scalar massless field from given initial data. That is, we solve (4.2.4) for the case when $u_\ell(r_*, 0)$ and $\partial_t u_\ell(r_*, 0)$ are given. The extension of the discussion below to the case of gravitational perturbations is straightforward. Early studies along these lines were carried out by Detweiler (1978) and Gaiser and Wagoner (1980).

The time evolution of the field $u_\ell(r_*, t)$ from given initial data can be written

$$u_\ell(r_*, t) = \int G(r_*, y, t)\, \partial_t u_\ell(y, 0)\, dy + \int \partial_t G(r_*, y, t)\, u_\ell(y, 0)\, dy \qquad (4.3.1)$$

for $t > 0$. The retarded Green's function is defined by

$$\left[\frac{\partial^2}{\partial r_*^2} - \frac{\partial^2}{\partial t^2} - V_\ell(r)\right] G(r_*, y, t) = \delta(t)\,\delta(r_* - y) \tag{4.3.2}$$

together with the initial condition $G(r_*, y, t) = 0$ for $t < 0$.

One would typically use an integral transform to reduce the problem to an ordinary differential equation. The integral transform

$$\hat{G}(r_*, y, \omega) = \int_{0^-}^{+\infty} G(r_*, y, t)\, e^{i\omega t}\, dt \tag{4.3.3}$$

is well defined as long as $\mathrm{Im}\ \omega \geq 0$. In fact, $\hat{G}(r_*, y, \omega)$ is a holomorphic function of $\omega = \omega_0 + i\omega_1$ for $\omega_1 > 0$. For the analysis to be meaningful, it is also necessary that $u_\ell(r_*, t)$ is pointwise bounded for all values of r_* and t. This is equal to a requirement that the black hole is stable against perturbations. We will discuss the stability of black holes further in Section 4.9.

A change $s = -i\omega$ makes our transform (4.3.3) equal to the standard Laplace-transform as used in the black hole studies of Leaver (1985), Sun and Price (1988), and Nollert and Schmidt (1992). We prefer to use the form (4.3.3) here since it is analogous to the Fourier-transform that has been assumed in most of the existing black hole literature.

Anyway, it follows immediately that

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - V_\ell(r)\right] \hat{G}(r_*, y, \omega) = \delta(r_* - y). \tag{4.3.4}$$

The required Green's function can now be expressed in terms of two linearly independent solutions to the homogeneous equation (4.2.7). One of these solutions is determined by the causal boundary condition that no waves emerge from H^-. We represent this solution by the *IN*-mode discussed previously; see (4.2.10). This solution is square integrable at $r_* \to -\infty$ for ω in the upper half-plane. A second solution, which is square integrable at $r_* \to +\infty$, corresponds to purely outgoing waves at infinity ($\mathcal{J}^+$) and a linear combination of outgoing and ingoing waves at the event horizon. This is clearly the *UP*-mode defined by (4.2.14). In terms of these solutions the Green's function can be written

$$\hat{G}(r_*, y, \omega) = -\frac{1}{2i\omega A_{\mathrm{in}}(\omega)} \begin{cases} \hat{u}_\ell^{\mathrm{in}}(r_*, \omega)\, \hat{u}_\ell^{\mathrm{up}}(y, \omega)\ , & r_* < y\ , \\ \hat{u}_\ell^{\mathrm{in}}(y, \omega)\, \hat{u}_\ell^{\mathrm{up}}(r_*, \omega)\ , & r_* > y\ , \end{cases} \tag{4.3.5}$$

where we have used the Wronskian relation (4.2.17). Finally, the integral transform must be inverted in order to recover the time domain. Then we get the final expression for the Green's function that propagates the initial data:

$$G(r_*, y, t) = \frac{1}{2\pi} \int_{-\infty+ic}^{+\infty+ic} \hat{G}(r_*, y, \omega)\, e^{-i\omega t}\, d\omega\ , \tag{4.3.6}$$

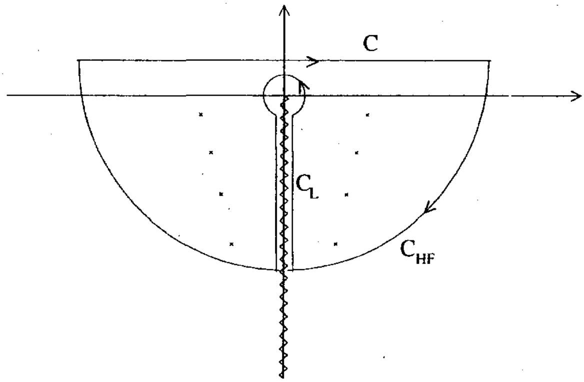


Figure 4.4: Integration contours in the complex ω-plane. The crosses represent the first few quasinormal modes, and the necessary branch cut is taken along the negative imaginary axis.

where c is some positive number (that ensures convergence of the integral).

In order to infer the behavior of the Green's function in different time intervals, it is convenient to deform the contour of integration in (4.3.6) in the complex ω-plane. For this purpose, we need to know the analytic properties of $\hat{u}_\ell^{\text{in}}$ and $\hat{u}_\ell^{\text{up}}$, not only in the upper half-plane where they are a holomorphic functions of ω, but also in the lower half-plane. It turns out that the Wronskian $W(IN, UP) = 2i\omega A_{\text{in}}(\omega)$ has isolated zeros there. This leads to poles of the Green's function $\hat{G}(r_*, y, \omega)$. These singularities correspond directly to the quasinormal modes of the black hole, and their contribution to the radiation can be accounted for by means of the residue theorem [Leaver (1986b)]. It is straightforward to show that the quasinormal modes are symmetrically distributed with respect to the imaginary ω-axis; if ω_n corresponds to $A_{\text{in}} = 0$, then $-\bar{\omega}_n$ must also do so (see Figure 4.4).

In the upper half of the complex ω-plane the solutions to (4.2.7) which are bounded at either end must behave like

$$\begin{aligned} \hat{u}_\ell^{\text{in}}(r_*, \omega) &\sim e^{-i\omega r_*} \quad \text{as } r_* \to -\infty, \\ \hat{u}_\ell^{\text{up}}(r_*, \omega) &\sim e^{+i\omega r_*} \quad \text{as } r_* \to \infty . \end{aligned} \tag{4.3.7}$$

Their analytical continuations into the lower half-plane will show the same behavior. Hence, the time-domain Green's function always satisfies "future outgoing" conditions. This Green's function propagates waves emitted by the source to H^+ and $\mathcal{J}^+$. It is therefore clear that the solutions corresponding to the quasinormal modes are regular both at H^+ and $\mathcal{J}^+$. But it also follows that they will diverge at H^- and $\mathcal{J}^-$.

Careful analysis shows that it is necessary to introduce a branch cut in order to make $\hat{u}_\ell^{\rm up}$ a single-valued function [Ching *et al.* (1995b)]. This cut is customarily placed along the negative imaginary axis, as in Figure 4.4. Given this information, the radiation produced in response to a perturbation of the black hole can be divided into three components, in accordance with the contributions of different parts of the deformed contour in the lower half of the ω-plane:

1. Radiation emitted directly by the source
2. Exponentially damped quasinormal-mode oscillations [contribution of the poles of the Green's function]
3. A power-law tail [contribution of the branch-cut integral]

These conclusions are in perfect agreement with the observations made in the previous section.

It is worth pointing out that the quasinormal modes of the black hole do not form a complete set of dynamical variables for the system. The power-law tail and the transients propagating directly from a source to an observer can not be described by an expansion in terms of quasinormal modes. However, it has recently been shown [Ching *et al.* (1995c,d)] that the modes can be made complete by introducing an infinitesimal change in the potential (for example, letting $V(r) \equiv 0$ for very large r). Such a change typically affects the mode spectrum considerably, but it may nevertheless prove useful. This kind of analysis stems from relevant astrophysical questions. For example, to what extent will the black holes environment, an accretion disk or a host galaxy affect the spectrum of quasinormal modes?

4.4 Quasinormal Modes

If a black hole is to be an astrophysically realistic proposition, solutions to the original wave equation (4.2.4) must be damped with time for each value of r_*. Hence, all acceptable mode solutions [to (4.2.7)] should have Im $\omega_n < 0$. This means that a quasinormal mode is distinguished by solutions to the radial equation that grow exponentially as we approach $r_* = \pm\infty$. At first, this may seem peculiar (and indeed undesirable), but it is easy to see why this happens. Properly defined, the modes correspond to purely outgoing waves reaching H^+ and $\mathcal{J}^+$. For example, at $\mathcal{J}^+$ we expect to have $u_\ell(r_*, t) \sim \exp[-i\omega_n(t - r_*)]$. From this it is clear that once we have used the integral transform (4.3.3), we require solutions that behave as $\hat{u}_\ell \sim \exp(i\omega_n r_*)$, and for Im $\omega_n < 0$ the solution will diverge as $r_* \to +\infty$. In other words: Because a mode solution is expected to be damped with time at any fixed value of r_*, it must diverge as $r_* \to +\infty$ at any fixed value of t. In the complete solution, the apparent divergence is balanced by the fact that it takes a signal an infinite time to reach e.g., $\mathcal{J}^+$.

Anyway, we are clearly not dealing with a standard Sturm-Liouville eigenvalue problem, and the identification of quasinormal modes is somewhat complicated. In order to identify a mode, one must ensure that no contamination of ingoing waves remains at infinity – and that no waves are coming out of the horizon – but these unwanted contributions are exponentially small. In this section we will briefly discuss some of the methods that have been developed to handle this problem, but let us first discuss some simple approximations that lead to surprisingly good results.

4.4.1 Simple approximations

As mentioned previously, the black hole problem is similar to the problem of quantum scattering of a particle by a potential barrier. Hence, the concepts of quantum mechanics might be useful. It is commonly accepted that scattering resonances (which are the quantum analogues to quasinormal modes) arise for energies close to the top of a potential barrier. In the black hole case, this immediately leads to the approximation [Schutz and Will (1985), Ferrari and Mashhoon (1984a,b)]

$$\text{Re}\,\omega_0 \approx \sqrt{V_\ell^{\max}} \approx \frac{1}{3\sqrt{3}\,M}\left(\ell + \frac{1}{2}\right) . \tag{4.4.1}$$

Here we have neglected the field-dependent term in the potential (4.2.18). This approximation for the fundamental mode is poor for low ℓ (the error is something like 30 percent for $\ell = 2$), but it rapidly gets accurate as ℓ increases.

For the imaginary part of the frequency – the lifetime of the resonance – the curvature of the potential contains the relevant information [Schutz and Will (1985)]. One finds that

$$\text{Im}\,\omega_0 \approx -\frac{1}{2}\left|\frac{1}{2V_\ell}\frac{d^2V_\ell}{dr_*^2}\right|^{1/2}_{r=r_{\max}} \approx -\frac{\sqrt{3}}{18M} , \tag{4.4.2}$$

which is accurate to within 10 percent for the fundamental mode.

Interestingly, similar approximations follow from a (seemingly) different approach. Consider a congruence of null rays circling the black hole in the unstable photon orbit at $r = 3M$. The fundamental mode frequency then follows if the beam contains ℓ cycles [Thorne (1978)]. The damping rate of the mode can be inferred from the decay rate of the congruence if the null orbit is slightly perturbed [Mashhoon (1985)].

It is interesting to compare a black hole to other resonant systems in nature. If we define a quality factor in analogy with the standard harmonic oscillator,

$$Q \approx \frac{1}{2}\left|\frac{\text{Re}\,\omega_n}{\text{Im}\,\omega_n}\right| , \tag{4.4.3}$$

the quasinormal-mode approximations given here lead to $Q \approx \ell$. This should be compared to the typical value for an atom: $Q \sim 10^6$. The Schwarzschild black hole is thus a very poor oscillator.

4.4.2 The complete spectrum of black hole modes

During the last twenty years there have been numerous attempts to calculate quasinormal-mode frequencies with acceptable accuracy. Accurate integration of the Regge-Wheeler equation (4.2.7) is difficult since the unwanted solutions – corresponding to, for example, waves coming in from infinity – inevitably become smaller than the numerical uncertainty in the exponentially growing mode solution. Still, by studying the logarithmic derivative of $\hat{u}_\ell$, Chandrasekhar and Detweiler (1975a) computed the first few modes. Ten years later, Leaver (1985) determined very accurate values for the quasinormal-mode frequencies using a three-term recurrence relation together with a numerical solution of continued fractions [for an alternative way of using the recurrence relation, see Majumdar and Panchapakesan (1989)]. Leaver's calculations suggested that an infinite number of modes exist for each value of ℓ. For a given value of ℓ the real part of the mode frequencies approaches a nonzero constant, and at the same time the damping increases in a systematic way.

There have been many attempts to verify Leaver's results. Among the methods employed are:

- WKB-type approximation schemes [Schutz and Will (1985), Iyer and Will (1987), Iyer (1987), Guinn *et al.* (1990), Zaslavskii (1991), Galtsov and Matiukhin (1992), Fröman *et al.* (1992), Andersson and Linnæus (1992), Andersson, Araújo, and Schutz (1993a,b); Araújo, Nicholson and Schutz (1993)]
- Use of the inverted black hole potential [Blome and Mashhoon (1984), Ferrari and Mashhoon (1984a,b)]
- Laplace-transform combined with analytic continuation [Nollert and Schmidt (1992)]
- Numerical integration using complex coordinates [Andersson (1992)]

Most of these schemes give good results for the first few modes, but break down as the damping increases. However, several studies have shown that Leaver's results are reliable [Nollert and Schmidt (1992), Andersson (1992), Andersson and Linnæus (1992)]. Now there are (at least) four independent methods that lead to results that agree perfectly for the first ten or so modes for each value of ℓ. The first five quasinormal-mode frequencies for $\ell = 2, 3, 4$ are given in Table 4.4.2.

Bachelot and Motet-Bachelot (1992) have shown that an infinite number of quasinormal modes should exist for each given ℓ. Their work does not however shed any light on the asymptotic behavior of the mode frequencies. This behavior was elucidated by Nollert (1993). He analyzed the asymptotic behavior of the continued fraction that had been used by Leaver, and found that (for gravitational perturbations)

$$\omega_n M \approx 0.0437 - \frac{i}{4}\left(n + \frac{1}{2}\right) + O\left(\frac{1}{\sqrt{n+1}}\right) \tag{4.4.4}$$

n	$\ell = 2$	$\ell = 3$	$\ell = 4$
0	$0.37367 - 0.08896i$	$0.59944 - 0.09270i$	$0.80918 - 0.09416i$
1	$0.34671 - 0.27391i$	$0.58264 - 0.28130i$	$0.79663 - 0.28433i$
2	$0.30105 - 0.47828i$	$0.55168 - 0.47909i$	$0.77271 - 0.47991i$
3	$0.25150 - 0.70514i$	$0.51196 - 0.69034i$	$0.73984 - 0.68392i$
4	$0.20751 - 0.94684i$	$0.47017 - 0.91565i$	$0.70152 - 0.89824i$

Table 4.1: The first five quasinormal-mode frequencies for the three lowest radiating multipoles ($\ell \geq 2$) of a gravitationally perturbed Schwarzschild black hole. The frequencies are given in units of $[GM/c^3]^{-1} = 2\pi(32312\,\mathrm{Hz}) \times (M/M_\odot)^{-1}$. This means that for a ten solar-mass black hole the fundamental oscillation frequency would be roughly 7.6 kHz.

when the positive mode index n is very large. The behavior is independent of the value of ℓ to leading order, cf. Figure 4.5. This formula has been tested against the semi-analytic method of Andersson and Linnaeus (1992). By comparing the two, one finds that they agree to one part in 10^5 for $n = 10^5$ [Andersson (1993b)]. The same asymptotic behavior for the damping was derived in recent work by Liu (1995). He showed that the damping is the same to leading order for all fields. Moreover, Liu and Mashhoon (1995) have argued that the high overtone spacing is intimately connected with the long-range nature of gravity.

4.4.3 Quasinormal modes for charged and rotating black holes

Electrically charged black holes

Although it is hard to imagine a physically realistic situation where the electric net charge of a black hole is significant, an investigation of the Reissner-Nordström solution does not lack interest. On the contrary, since it provides a more general framework than the Schwarzschild geometry, this case may contribute significantly

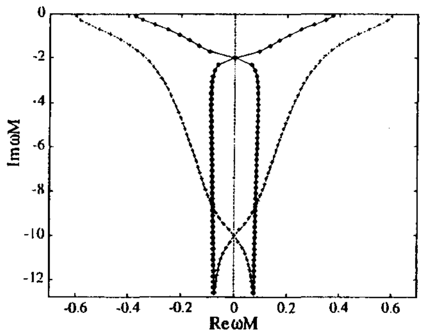


Figure 4.5: The spectrum of quasinormal modes for a Schwarzschild black hole. The first fifty modes for $\ell = 2$ (diamonds) and $\ell = 3$ (crosses) are shown. [The modes were calculated using the method of Andersson and Linnæus (1992)]

to our understanding of more complicated spacetimes.[3]

Perturbation equations relevant for electrically charged, non-rotating black holes were first derived by Zerilli (1974) and Moncrief (1974b,c, 1975). They have also been discussed by Chandrasekhar (1979a) and Bičák (1980b, 1982). It is known that the equations decouple into two second-order differential equations of the form (4.2.7) for axial and polar perturbations. However, whereas electromagnetic and gravitational perturbations can be studied separately in the Schwarzschild limit (when the charge of the black hole vanishes), this is not the case for charged black holes. In a charged environment a variation in the electromagnetic field will induce gravitational radiation and vice versa; see Section 7.3. A quasinormal mode of a Reissner-Nordström black hole therefore generally corresponds to emission of both gravitational and electromagnetic radiation.

The Reissner-Nordström black hole provides the simplest context in which one can study the conversion of gravitational waves into electromagnetic ones (and vice versa) (see Section 7.3). An obvious question regards what fraction of the incident energy is reflected as waves of the other kind. Interestingly, one can show that the conversion factor is the same whether the incident energy is gravitational or electromagnetic [Chandrasekhar (1979a)]. Gunter (1980) considered the problem numerically and showed that the conversion is relevant only in a narrow frequency band: It vanishes

[3]In fact, the highly charged case may give us some information about what to expect in an investigation of a rapidly rotating Kerr black hole. We shall see that the Reissner-Nordström solution for a charged black hole similar to the Kerr metric has an inner (Cauchy) horizon as well as the event horizon. In this respect the effect of the charge is similar to that of the angular momentum in the Kerr case. This is particularly relevant for studies of black holes interior (see Chapter 14), but the existence of the inner horizon also affects the appearance of the perturbation equations for the exterior spacetime.

both as $\omega \to \infty$ and $\omega \to 0$.

Several of the methods used to determine quasinormal-mode frequencies for Schwarzschild black holes have been generalized to the electrically charged case, e.g., the inverted potential approach [Ferrari and Mashhoon (1984a,b)] and the WKB-type methods [Kokkotas and Schutz (1988); Andersson, Araújo, and Schutz (1994)]. To the lowest order of approximation, the WKB studies suggest the following approximate behavior for the slowest damped mode

$$\text{Re}\ \omega_0 \approx \left(\ell + \frac{1}{2}\right) \left[\frac{M}{r_0^3} - \frac{Q^2}{r_0^4}\right]^{1/2}, \tag{4.4.5}$$

$$\text{Im}\ \omega_0 \approx -\frac{1}{2} \left[\frac{M}{r_0^3} - \frac{Q^2}{r_0^4}\right]^{1/2} \left[\frac{3M}{r_0} - \frac{4Q^2}{r_0^2}\right]^{1/2} \tag{4.4.6}$$

in the limit $\ell \gg 1$. Here $Q \leq M$ is the charge of the black hole, and we have defined r_0 as the position where the black hole potential attains its maximum value. This means that $2r_0 = 3M + \sqrt{9M^2 - 8Q^2}$.

These approximations indicate the general behavior of a quasinormal-mode frequency as the charge of the black hole increases. The oscillation frequency increases, and the damping rate reaches a maximum value after which it decreases rapidly. These features have been verified by more accurate numerical work. The most accurate results to date are those obtained by Leaver's (1990) continued fraction approach and Andersson's (1993a) numerical integration scheme.

The present understanding is the following. For modes adhering to the function that corresponds to purely gravitational waves in the Schwarzschild limit, both the real and the imaginary part of the frequency attain extreme values as Q approaches its limiting value ($Q = M$). These extrema are followed by tiny wiggles (for $Q > 0.9M$ or so); see Figure 3 of Andersson (1993a). For the other function, that limits to purely electromagnetic waves as $Q \to 0$, the oscillation frequency does not reach a maximum. It increases monotonically with Q. The peculiar behavior of the mode frequencies close to the limit of extremely charged black hole is not well understood.[4]

Rotating black holes

Whereas it is unlikely that we will ever observe astrophysical black holes with a considerable net charge, most black holes are expected to be rotating. Hence, perturbation studies of the Kerr solution are of great interest. We will discuss the perturbation equations for rotating black holes in some detail in 4.8. The Kerr problem

[4]Recently numerical calculations by Onozawa *et al.* (1996) and Andersson and Onozawa (1996) have shown that the quasinormal frequency trajectories of photons ($s = 1$), gravitino ($s = 3/2$), and gravitons ($s = 2$) with increasing charge meet at the same point in the extreme black hole limit $Q = M$. General explanation of this property by using the ideas of supersymmetry was proposed by Kallosh, Rahmfeld, and Wong (1997).

is considerably more involved than the Schwarzschild one. Nevertheless, one can find quasinormal modes also for rotating black holes. When the black hole has nonzero angular momentum, a, the azimuthal degeneracy is split. For a multipole ℓ there are consequently $2\ell+1$ distinct modes that approach each Schwarzschild mode in the limit $a \to 0$. These modes correspond to different values of m, where $-\ell \leq m \leq \ell$.

Quasinormal modes for Kerr black holes were first calculated by Detweiler (1980a,b Some typical resonance frequencies are shown as functions of the rotation parameter a in Figure 4.6. (Here and later on the figures in this chapter we use a notation $\sigma = \text{Re}\ \omega$ and $\alpha = -\text{Im}\ \omega$.) The results are in many ways similar to those for Reissner-Nordström black holes. Specifically, Detweiler found that

$$\text{for} \quad \ell = m \quad \left\{ \begin{array}{l} \text{Im}\ \omega_n \text{ is almost constant} \\ \text{Re}\ \omega_n \text{ increases monotonically} \end{array} \right\} \text{ as } a \to M \ , \tag{4.4.7}$$

and

$$\text{for} \quad \ell = -m \quad \left\{ \begin{array}{l} \text{Im}\ \omega_n \to 0 \\ \text{Re}\ \omega_n \to -m/2 \end{array} \right\} \text{ as } a \to M \ . \tag{4.4.8}$$

This behavior has been verified by accurate calculations by Leaver (1985). The problem has also been approached approximately [Ferrari and Mashhoon (1984a,b), Seidel and Iyer (1990), Kokkotas (1991)].

The available results suggest that a large number of quasinormal modes become very slowly damped as $a \to M$. Leaver suggests that these modes all coalesce to an undamped one for $a = M$. They all approach the critical frequency $m\,\Omega_H$ for superradiant scattering (see Section 4.8 for a discussion of this effect). Interestingly, the limit $a \to M$ is amenable to analytic methods [Teukolsky and Press (1974), Detweiler (1980b), Sasaki and Nakamura (1990)]. One can show that in the extreme Kerr limit ($a = M$) and for $\ell = m$ there exist an infinite sequence of resonant frequencies with a common limiting point. However, it seems likely that every quasinormal mode of a physically realizable black hole ($a < M$) will have at least some small damping [Leaver (1985), Sasaki, and Nakamura (1990)].

The fact that some quasinormal modes become very long-lived for rapidly rotating black holes could potentially be of great importance for gravitational-wave detection. This is basically because the effective amplitude of a periodic signal buried in noise improves as $\sqrt{n_{\text{cyc}}}$, where n_{cyc} is the number of observed cycles. Hence, it is imperative that the contribution of long-lived modes to the black hole response be investigated in detail. Ferrari and Mashhoon (1984a,b) have addressed this issue within their approximate framework. They argue that the contribution of these modes to the radiation becomes small as $a \to M$ (it vanishes in the limit). On the other hand, Sasaki and Nakamura (1990) determine the contribution to the Green's function (see the following Section) from each mode. They argue that the corresponding flux diverges logarithmically. This could indicate that the extremal Kerr black hole is marginally unstable. It is clear that the behavior of the quasinormal modes as the black hole approaches maximal rotation warrants more detailed investigations.

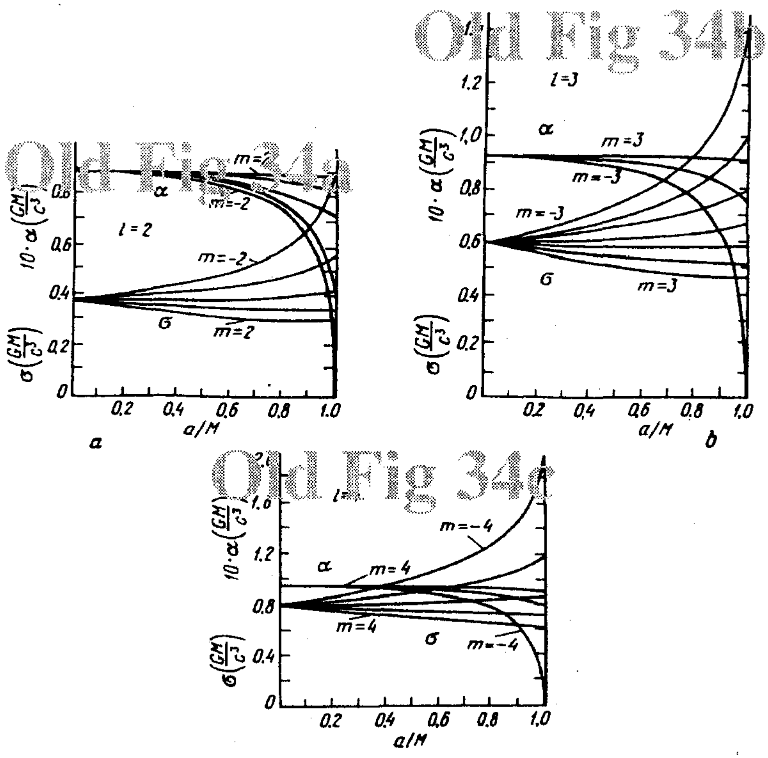


Figure 4.6: Resonance frequencies $\omega = \sigma - i\alpha$ as functions of the parameter a for different l and m [data from Detweiler (1980b)].

Finally, it should be mentioned that the WKB method has been applied to massive scalar perturbations in the Kerr background [Simone and Will (1992)]. This method was also used by Kokkotas (1993) in one of the two available mode studies for black holes which are both electrically charged and rotating; the Kerr-Newman solution. Kerr-Newman quasinormal modes were obtained in the eikonal approximation ($\ell = |m| >> 1$) by Mashhoon (1985).

4.4.4 Quasinormal-mode contribution to the radiation

Let us now return to the initial-value problem considered in Section 4.3, and take a closer look at the contribution from the quasinormal modes to the radiation. It is a straightforward task to continue the integration in (4.3.6) into the lower half of the complex ω-plane and then use the residue theorem to infer the contribution to the integral from each pole [Leaver (1986b)].

Before doing this, let us assume (for clarity of presentation) that the observer is situated sufficiently far away from the black hole that we can use the asymptotic form for $\hat{u}_\ell^{\rm up}$. Then we get

$$G(r_*, y, t) = -\frac{1}{4\pi i}\int_{-\infty+ic}^{+\infty+ic} \frac{\hat{u}_\ell^{\rm in}(y,\omega)}{\omega A_{\rm in}(\omega)} e^{-i\omega(t-r_*)}\, d\omega\,. \tag{4.4.9}$$

The quasinormal-mode contribution is then (using the fact that modes in the third and fourth quadrants are in one-to-one correspondence)

$$G_Q(r_*, y, t) = \mathrm{Re}\left[\sum_{n=0}^{\infty} \hat{u}_\ell^{\rm in}(y,\omega_n)\,\frac{e^{-i\omega_n(t-r_*)}}{\omega_n\alpha_n}\right]\,, \tag{4.4.10}$$

where we sum only over poles in the fourth quadrant of the ω-plane. We have defined α_n according to

$$A_{\rm in}(\omega) \approx (\omega-\omega_n)\alpha_n\,, \tag{4.4.11}$$

close to the nth quasinormal-mode frequency.

In order to use expression (4.4.10), we must be able to compute α_n. Although many methods have been designed to the task of finding the eigenfrequencies, very few attempts to proceed to actual calculations for initial-value problems have been made. The only available results for the residues α_n are those of Leaver (1986b) and Andersson (1995a).

To evaluate (4.4.10), one must also compute the integral over the product of the quasinormal-mode eigenfunctions and the initial data. This is not a trivial – if at all possible – task since the eigenfunction $\hat{u}_\ell^{\rm in}(r_*,\omega_n)$ (say) diverges as $r_* \to \pm\infty$. Nevertheless, it has been shown that (4.4.10) accurately accounts for the excitation of quasinormal modes in cases when the integrand can be evaluated [Leaver (1986b), Sun and Price (1988)]. Such cases include, most importantly, initial data of compact support.

A simple, but surprisingly accurate, approximation suitable for data that is initially localized far away from the black hole is obtained if one assumes that the initial data has considerable support only in the region where $\hat{u}_\ell^{\rm in}(r_*,\omega_n)$ can be replaced by its asymptotic form. Then equation (4.4.10) simplifies to

$$G_Q(r_*, y, t) = \mathrm{Re}\left[\sum_{n=0}^{\infty} \frac{A_{\rm out}(\omega_n)}{\omega_n\alpha_n}\, e^{-i\omega_n(t-r_*-y)}\right]\,. \tag{4.4.12}$$

This form facilitates analytic calculation of the quasinormal-mode excitation for many kinds of initial data. Results obtained in this way for Gaussian initial data are presented in Figure 3 of Andersson (1995a).

4.5 Power-Law Tails

4.5.1 Late-time behavior

Let us now consider the asymptotic behavior of a perturbation as the black hole returns to its unperturbed state as $t \to \infty$. A typical process for which this behavior is relevant is when a nearly spherical star collapses to form a black hole. The question is: Does an initial deformation of the star remain small during the collapse, or will the perturbation grow and halt the collapse in some way?

When a star undergoes gravitational collapse its surface approaches the black hole boundary (*i.e.*, moves towards $r_* = -\infty$) at a physical velocity tending to the speed of light (see Chapter 2). This means that as $t \to \infty$, all processes in the perturbation source must "freeze" according to an observer at rest in the Schwarzschild reference frame (see Chapter 2). Due to time dilation, a perturbation field Φ will take the following form at the surface of the star (for any integer-spin field and any multipole ℓ)

$$\Phi = a + b\exp\left(\frac{r_* - t}{4M}\right), \tag{4.5.1}$$

where a and b are constants [Price (1972a,b)]. The characteristic time-scale of the exponentially decaying term is such that the wave part of Φ will partly be reflected by the potential barrier that surrounds the hole and fall back onto it (return to to $r_* = -\infty$). The remainder will pass through the barrier and escape to infinity ($r_* = \infty$). At the same time, the constant term, a, generates a perturbation of infinite wavelength that will be completely reflected by the black hole barrier and cannot reach an external observer. As a result, only exponentially decaying waves can reach the observer. The curvature of spacetime prevents any information about the final surface field from emerging.

But the damping of the radiation that is seen by a distant observer will not be purely exponential. He will also see waves that have been scattered by the "tail" of the potential barrier (*i.e.*, by the spacetime curvature). Price (1972a) found that backscattering by the asymptotic "tail" of the potential gives rise to a power-law fall-off at late times. Specifically, he inferred the late-time behavior [a clear motivation of this behavior can be found in Thorne (1972)]

$$\Phi \sim t^{-(2\ell+3)}. \tag{4.5.2}$$

Price put his conclusions in the following succinct form: "Anything that can be radiated will be radiated". Consequently, a black hole rids itself from all bumps after it is formed by a non-spherical collapsing star. That a similar result holds for neutrinos was shown by Hartle (1972). That a *power-law tail* dominates the late-time behavior is clear from, for example, the gravitational collapse calculations by Cunningham, Price, and Moncrief (1978, 1979).

In his Green's function analysis of perturbed black holes, Leaver (1986b) associated the power-law tail with the branch-cut integral along the negative imaginary axis in the complex ω-plane. He showed that the behavior (4.5.2) follows from a consideration of the $|\omega M| << 1$ part of that integral. He also indicated that there would be "radiative" tails observable at $\mathcal{J}^+$ and H^+. This result was verified by recent work of Gundlach, Price, and Pullin (1994a).

Interestingly, the late-time behavior is the same also for Reissner-Nordström black holes and fields of different spin. Furthermore, a fully nonlinear study of a collapsing scalar field shows that tails develop even when no black hole is formed [Gundlach, Price, and Pullin (1994b)]. This is strong evidence that the late-time power-law tail depends only on the gravitational field in the far zone and not on the details of the central body (if there is one). Thus, the same power-law tail should be observed from a perturbed black hole, an imploding field that fails to form a black hole, or even a perturbed star.

4.5.2 Analyzing the Green's function

We will now outline a detailed analysis of the tail phenomenon. As already mentioned, the power-law tail can be associated with the branch cut (usually taken along the imaginary ω-axis; see Figure 4.4) in the Green's function (4.3.5) [Leaver (1986b)]. The existence of such a cut depends on the asymptotics of the effective black hole potential. Specifically, a cut is a general feature if $V_\ell(r_*)$ asymptotically tends to zero slower than an exponential [Ching *et al.* (1995a,b)]. In the case of black holes, one can show that [cf. (4.2.18)]

$$V_\ell^{(s)}(r_*) \approx \begin{cases} \dfrac{\ell(\ell+1)}{r_*^2} + \dfrac{4M\ell(\ell+1)}{r_*^3}\log\left(\dfrac{r_*}{2M}\right) , & \ell \neq 0 \ , \ r_* \to +\infty \ ; \\[2ex] \dfrac{2M(1-s^2)}{r_*^3} , & \ell = 0 \ , \ r_* \to +\infty \ ; \\[2ex] \dfrac{\ell(\ell+1)+1-s^2}{(2M)^2} \exp\left(\dfrac{r_*}{2M}\right) , & r_* \to -\infty \ . \end{cases} \tag{4.5.3}$$

Hence, there will be a cut in the Green's function (4.3.5). The cut is associated with the function $\hat{u}_\ell^{\rm up}(r_*,\omega)$, that corresponds to purely outgoing waves at spatial infinity. Because the potential falls off exponentially as $r_* \to -\infty$, there will be no such cut in the function $\hat{u}_\ell^{\rm in}(r_*,\omega)$, which is purely ingoing at the black-hole event horizon.

We must therefore consider the function $\hat{u}_\ell^{\rm up}$ in some detail, especially for $r_* \to \infty$. From (4.5.3) it is clear that it is sufficient to consider an effective potential

$$V_\ell^{(s)}(r_*) \approx \frac{\ell(\ell+1)}{r_*^2} + \hat{V}(r_*) \tag{4.5.4}$$

as long as we restrict the analysis the $\ell \neq 0$. [For $\ell = 0$ the analysis proceeds in a similar way, see Ching *et al.* (1995b)]. The required function $\hat{u}_\ell^{\rm up}$ is thus a solution to

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - \frac{\ell(\ell+1)}{r_*^2}\right] \hat{u}_\ell(r_*,\omega) = \hat{V}(r_*)\,\hat{u}_\ell(r_*,\omega) \tag{4.5.5}$$

that behaves as purely outgoing waves as $r_* \to \infty$. One can now consider the right-hand side as "small" perturbation, and use the Born approximation [Ching *et al.* (1995b)]. This way the contribution from the branch-cut integral to the Green's function can be deduced.

For a Schwarzschild black hole this results in a branch-cut contribution to the Green's function [Leaver (1986b), Ching *et al.* (1995b)]

$$G_B(r_*,y,t) = (-1)^{\ell+1} \frac{(2\ell+2)!}{[(2\ell+1)!!]^2} \frac{4M(r_*y)^{\ell+1}}{t^{2\ell+3}} . \tag{4.5.6}$$

This result – leaving no adjustable parameters – has been shown to agree perfectly with numerical studies.

One final comment is in order here. The above result suggests that for initial data that are static on a Cauchy hypersurface the predicted power-law tail will be $t^{-(2\ell+4)}$. At first sight, this seems to be in contradiction with the result of Gundlach, Price, and Pullin (1994a) who find that "static" initial data lead to a tail $t^{-(2\ell+2)}$. The difference is that in the latter description initial data was introduced on a null surface. It is therefore not surprising that the two results differ, but, unfortunately, there is no apparent way to translate one into the other. The important result is that the two studies predict the same result in the general case, when the data is not static.

4.6 Gravitational Radiation from a Test Particle in the Field of a Black Hole

In general, the equation governing a black hole perturbation is not homogeneous. One must typically also include a source term for the physical situation under consideration. Perhaps the simplest relevant problem is that of a test particle moving in the gravitational field of a black hole. When the mass m of the particle is sufficiently small compared to that of the black hole ($m << M$), this problem can be viewed as a perturbation problem. In this Section we assume that this is the case, and we also neglect any backreaction of the radiation on the motion of the particle. The particle thus moves along a geodesic of the unperturbed spacetime.

4.6.1 Particles plunging into the black hole

The radiation emitted by a test particle of mass m which falls radially into a Schwarzschild black hole was calculated as one of the first astrophysical applications of the

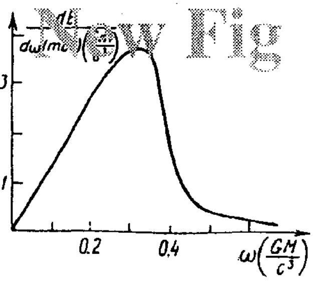


Figure 4.7: The spectrum of quadrupole ($\ell = 2$) gravitational radiation emitted when a particle falls radially into a Schwarzschild black hole. The particle starts off at rest at infinity. The spectrum (averaged over all directions) peaks at the slowest damped quasinormal-mode frequency ($\omega M \approx 0.37$) [data from Davis *et al.* (1971)].

perturbation equations [Davis *et al.* (1971, 1972)]. Typical results obtained for the case when the particle starts off at rest at infinity (i.e., for a parabolic motion[5]) are shown in Figures 4.7 and 4.8. Figure 4.7 shows the spectrum of gravitational radiation (averaged over all directions) as measured by a distant observer. The spectrum is clearly peaked at the slowest damped quadrupole quasinormal mode (at $\omega M \approx 0.37$). The total energy emitted is $\Delta E \approx 0.01 m^2/M$, and roughly 90% of this energy is radiated through the quadrupole modes. In fact, 97% of the total energy goes into quasinormal-mode oscillations. That the radiation is dominated by quasinormal-mode ringing at late times is obvious from Figure 4.8. Here we show the typical form of the resultant gravitational waves (represented by one of the metric perturbations). The particle reaches the event horizon at roughly $t = 0$ in Figure 4.8, and a final burst of radiation directly from the particle is followed by typical mode-ringing. The power-law tail that should dominate at very late times cannot be seen on this scale. Worth noticing is the similarity between this radiation and that resulting from scattering of a wavepacket by the black hole; see Figure 4.4. A reasonable approximation of the initial wave burst is obtained if one assumes that the particle radiates as if it were moving through flat space [Ruffini and Sasaki (1981)].

Qualitatively, these results are hardly altered at all if the particle of mass m is given some initial specific angular momentum $\tilde{L} = L/(2Mm)$ or an initial velocity [Detweiler and Szedenitz (1979), Oohara and Nakamura (1983a,b)]. The energy spectrum in each multipole peaks at the relevant quasinormal-mode frequency. Moreover, one finds that the contribution from $\tilde{m} = \ell$ generally dominates the radiation.[6] An

[5]We recall that the total energy E for a parabolic motion of a test particle is mc^2, and the velocity of the particle at infinity vanishes. Different types of the parabolic motion are specified by their angular momentum L. Sometimes, we shall say that a particle is at rest at infinity. This does not mean that the particle is moving radially since, in the general case, its angular momentum does not vanish.

[6]In order to distinguish between the mass of the particle (m) and its azimuthal quantum number,

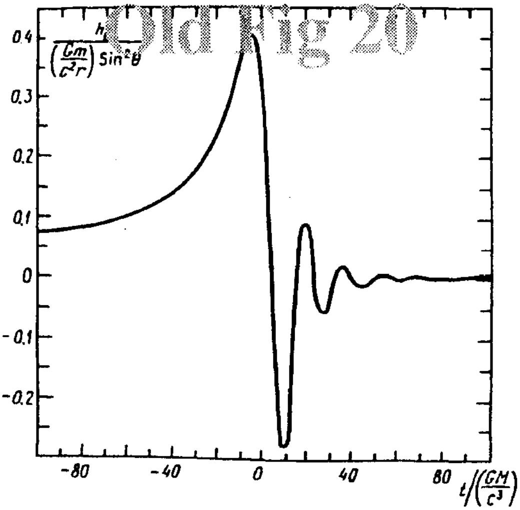


Figure 4.8: Typical quadrupole ($\ell = 2$) gravitational waves radiated when a test particle plunges radially into a non-rotating black hole. The particle starts off at rest at infinity. The direction $\theta = 0$ coincides with the trajectory of the particle. At late times the radiation is clearly dominated by quasinormal-mode ringing [data from Petrich, Shapiro, and Wasserman (1985)]. The sign of h in this and other similar figures depends on the convention; different authors use different definitions, which is of no great significance.

interesting empirical relation for the total energy radiated in each multipole was first noted by Davis *et al.* (1971). It turns out that the emerging energy obeys the relation

$$\Delta E_\ell \approx a e^{-b\ell} m^2/M \,. \tag{4.6.1}$$

(Here m is the mass of the particle, and M is the mass of the black hole.) Thus, the quadrupole radiation always dominates. The coefficients a and b obviously vary with the initial angular momentum of the particle. As $\tilde{L}$ increases, b decreases, and thus the importance of the higher multipoles is enhanced [Detweiler and Szedenitz (1979)]. In Figure 4.9 we show the total radiated energy ΔE and angular momentum ΔL (the angular momentum spectrum is similar to the energy one) as functions of the initial specific angular momentum $\tilde{L}$. The radiated energy can be increased by as much as a factor of 50 if the particle is given initial angular momentum. The apparent divergence at $\tilde{L} = 2$ should not be taken too seriously. That case corresponds to injecting the particle in a circular orbit. Since these calculations neglect the effect of radiation reaction on the particles motion, the particle will stay in a circular orbit forever. Consequently, the total radiated energy must be infinite.

we denote the latter (only in this section) by $\bar{m}$.

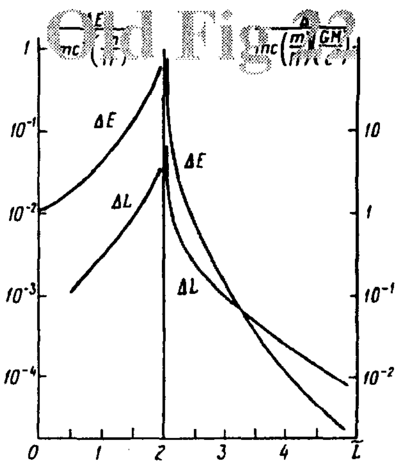


Figure 4.9: The total radiated energy ΔE and angular momentum ΔL shown as functions of the initial angular momentum of the particle. The particle is initially at rest at infinity [data from Oohara and Nakamura (1984)].

Let us also mention an interesting result for Reissner-Nordström black holes. Johnston, Ruffini, and Zerilli (1973) considered an uncharged particle falling into an electrically charged black hole. This problem can be used as indicator of the efficiency of the conversion of gravitational energy into electromagnetic energy that will be a feature of a charged environment. One finds that the gravitationally induced electromagnetic radiation carry a total energy $\Delta E \approx 0.03 m^2 Q^2 / M^3$. Notably, this is of the same order as the energy radiated as gravitational waves if the black hole has a considerable charge.

Particles scattered by the black hole

When the particle (initially at rest at infinity) has initial specific angular momentum $\tilde{L} > 2$, it is not captured by the black hole. Instead it escapes to infinity; that is, it is "scattered" by the black hole. The greater the initial value of $\tilde{L}$ is, the further away from the black hole does the particle reach periastron. Hence, one would expect less radiation to be emitted as $\tilde{L}$ increases (see Figure 4.9). When $\tilde{L}$ becomes very large, the particle never gets close enough to the black hole to induce relativistic effects. Typical waveforms for a particle scattered off a Schwarzschild black hole are shown in Figure 4.10.

We have seen that the bulk of the energy is radiated through the quasinormal modes when the particle falls into the hole. In contrast, the modes are hardly excited at all when the particle is scattered off to infinity. This is true even when periastron is quite close to the black hole. A typical spectrum for a scattered particle

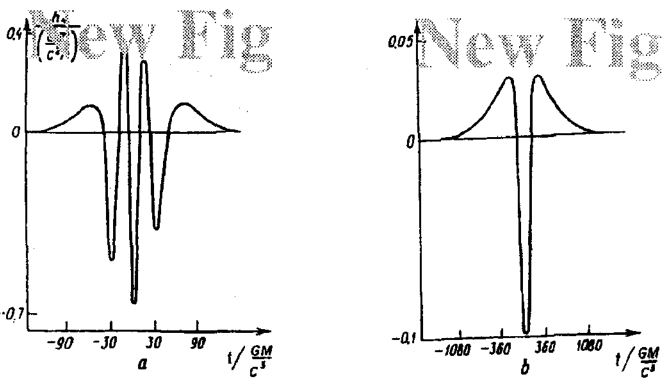


Figure 4.10: Gravitational waves generated when a particle is scattered by a Schwarzschild black hole. The particle starts off at rest at infinity and with initial specific angular momentum (a) $\tilde{L} = 2.005$ and (b) $\tilde{L} = 2.5$. An observer which registers this radiation is in the plane of the orbit in the direction of the periastron of the particle's trajectory. In the first case one can clearly see that the particle orbits the black hole several time before it escapes to infinity. Note the absence of quasinormal-mode ringing in both cases [data from Oohara and Nakamura (1984)].

is shown in Figure 4.11. Here the peak of each multipole spectrum is unrelated to the quasinormal-mode frequency. Instead, the peak frequency depends on the initial angular momentum $\tilde{L}$ of the particle. One can show that the spectrum depends on the angular velocity Ω_p at periastron [Oohara and Nakamura (1983a,b, 1984)]. The position of all peaks in the spectrum are explained in the same way: The source term contains a factor $\cos(\omega t - \tilde{m}\phi)$. Close to periastron it is reasonable to approximate $\phi \approx \Omega_p t$. Then it follows that the spectrum will be peaked at $\omega \approx \tilde{m}\,\Omega_p$, or since the $\tilde{m} = \ell$ term dominates: $\omega_{\max} \approx \ell\,\Omega_p$. The smaller peaks should correspond to other maxima of the cosine in the source term [Oohara and Nakamura (1984)].

This also explains why the quasinormal modes are not excited. For a particle that falls into the black hole the quasinormal modes are excited as the particle passes through the peak of the potential barrier. For a scattered particle, the modes can only be excited by the part of the gravitational radiation emitted by the particle that falls onto the hole. But the frequency of these gravitational waves is typically such that they get reflected off the black-hole potential barrier before getting close to the hole. For example, for $\ell = 2$ one gets $\omega_{\max} M < 0.2$, and such waves will not excite the quasinormal modes considerably. For this very reason, the standard formula [Landau and Lifshitz (1975)] which is valid for particles moving at $v \ll c$ in flat space (and ignores quasinormal modes) provides a reasonable approximation of the radiated energy (we will discuss this in more detail in Section 4.10).

Finally, let us consider the gravitational radiation emitted when the particle is given an initial velocity. The total energy radiated will clearly increase with the

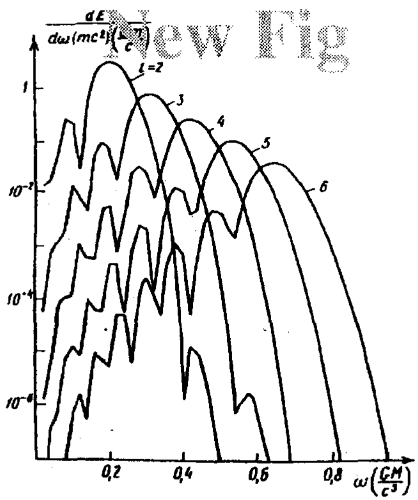


Figure 4.11: Spectrum of gravitational radiation generated by a particle with initial angular momentum $\tilde{L} = 2.005$ and zero velocity at infinity [data from Oohara and Nakamura (1984)].

initial velocity. Furthermore, the relative importance of the higher multipoles is enhanced [Ruffini (1973), Ferrari and Ruffini (1981)]. When an ultra-relativistic particle – that is, a particle that is injected from infinity with a velocity close to the speed of light – scatters off the black hole, the quasinormal modes are excited [Oohara (1984)]. There are two factors that stimulate this excitation. The first one is that the periastron of such a particle may be much closer to the black hole than for a particle that starts from rest at infinity. In fact, periastron may lie inside the peak of the potential barrier. If so, one would expect the particles passage to excite quasinormal-mode ringing. Secondly, an ultra-relativistic particle emits gravitational waves at frequencies that are much higher than that corresponding to the velocity at periastron. This is because of the generated gravitational synchrotron radiation [see, e.g., Doroshkevich *et al.* (1972), Chrzanowski and Misner (1974), Ternov *et al.* (1975), and Breuer (1975)]. Such high-frequency waves can penetrate through the barrier and excite the quasinormal modes.

Radiation from extended bodies

Once one has done a calculation for a test particle moving in the geometry of a black hole, the generalization to the case of "extended" bodies is quite simple. For example, one can study the effect of a "finite-sized" star or a dust shell falling onto the black hole [Sasaki and Nakamura (1981a), Haugan, Shapiro, and Wasserman (1982), Oohara and Nakamura (1983a,b)]. In a way, the latter process approximates a non-spherical gravitational collapse [Shapiro and Wasserman (1982), Petrich, Shapiro, and Wasserman (1985)]. To the lowest order of approximation, each particle follows a geodesic in the background spacetime and is unaffected by the presence of other particles. The total perturbation of the metric simply follows as the sum of the various individual contributions.

The general consensus from these studies is that, for a fixed amount of infalling mass less gravitational radiation is generated than in the case of a single particle. The energy emitted by infalling dust will never exceed that of a particle with the same mass [Shapiro and Wasserman (1982), Petrich, Shapiro and Wasserman (1985)]. For a shell of total mass m one finds that the radiated energy will not be greater than $\Delta E \approx 7.8 \times 10^{-4} m^2/M$ [Sasaki and Nakamura (1981a)]. This is more than a factor of ten smaller than the energy emitted in the test-particle case. The reduction is mainly due to phase cancellations. An extended body essentially acts as a collection of particles and the gravitational waves generated by each such particle interfere destructively with each other.

The results obtained from test-particle calculations are in reasonable agreement with fully relativistic numerical calculations. The waves emitted during a gravitational collapse are strongly dominated by a few quasinormal modes [Stark and Piran (1985), Seidel (1991), Abrahams and Evans (1992)]. The situation is similar for two colliding black holes [Anninos, Hobill *et al.* (1993)]. In fact, if extrapolated to the case of two equal masses, the predictions from test-particle calculations agree surprisingly well with the results of numerical relativity.

4.7 Scattering of Waves by Black Holes

4.7.1 The scattering problem

Much can be learned about a physical system by letting waves of a known character be scattered off it. A considerable effort has gone into studies of waves scattered off black holes. In this section we discuss some of the obtained results. More details can be found in the book by Futterman, Handler, and Matzner (1988) which is devoted entirely to this subject.

The basic scattering problem is simple. One assumes that a plane wave impinges on the black hole. The asymptotic behavior of the resultant field can then be used as a probe of spacetime close to the black hole. The key quantity that tells us what effect the central object has on the original wave is the scattering amplitude $f(\theta)$. If we, for simplicity, focus on massless scalar waves in the Schwarzschild geometry, the scattering amplitude can be extracted from the field at infinity

$$\Phi \sim \Phi_{\text{plane}} + \frac{f(\theta)}{r} e^{i\omega r_*} , \quad r_* \to +\infty . \tag{4.7.1}$$

The scalar field is, of course, described by (4.2.1). The differential cross-section – the intensity scattered into a certain solid angle – follows immediately from

$$\frac{d\sigma}{d\Omega} = |\, f(\theta)\,|^2 . \tag{4.7.2}$$

Here $d\Omega$ is an element of the solid angle ($d\Omega = 2\pi \sin\theta \, d\theta$) .

In order to extract the relevant information, we need to solve two problems. First, we need to define what we mean by a "plane" wave in a black hole spacetime, and then we need to solve (4.2.7) for the field of interest. For a massless scalar field, the first issue has been discussed by Matzner (1968) [see Chrzanowski *et al.* (1976) for a discussion of other fields]. He found that the desired expression for a plane wave at infinity is

$$\Phi_{\rm plane} \sim \frac{1}{\omega r} \sum_{\ell=0}^{\infty} i^\ell (2\ell+1)\, P_\ell(\cos\theta)\, \sin\left(\omega r_* - \frac{\ell\pi}{2}\right) \quad \text{as } r_* \to +\infty . \tag{4.7.3}$$

It should be noted that we retain the standard flat spacetime description by replacing r_* in (4.7.3) by r. The situation is analogous to the well-known one of Coulomb scattering. This is not very surprising. When described in terms of the Schwarzschild coordinate r, the perturbation equation (4.2.7) is a Schrödinger-like equation with a long-range potential (that falls off as $1/r$ for large r).

Now it is straightforward to show [using solution (4.2.10)] that

$$f(\theta) = \frac{1}{2i\omega} \sum_{\ell=0}^{\infty} (2\ell+1) \left[e^{2i\delta_\ell} - 1 \right] P_\ell(\cos\theta) , \tag{4.7.4}$$

where the phase-shifts δ_ℓ are defined as, cf. (4.2.10),

$$e^{2i\delta_\ell} = (-1)^{\ell+1} \frac{A_{\rm out}}{A_{\rm in}} . \tag{4.7.5}$$

The phase-shifts are typically complex valued. The imaginary part accounts for absorption by the black hole.

Before discussing results obtained in this way, we will provide some useful approximations.

4.7.2 Approximate results

Forward focusing

In the classical problem of scattering of a beam of parallel rays off a spherically symmetric center, one has the following expression for the differential cross-section

$$\frac{d\sigma}{d\Omega} = b \frac{db}{d\theta} (\sin\theta)^{-1} . \tag{4.7.6}$$

Here $b = b(\theta)$ is the impact parameter of a ray that is deflected by an angle θ away from the direction of incidence. For a black hole, and in the geometrical optics approximation, the dependence $b = b(\theta)$ can be obtained by solving equations (2.8.5)–(2.8.6). The calculation of the corresponding cross-section is trivial.

An obvious feature of the black hole cross-section concerns forward scattering ($\theta \approx 0$). A small scattering angle corresponds to a wave that passes at a large distance from the black hole (large b or large ℓ in the spherical harmonics expansion). Then the deflection is essentially governed by the Newtonian potential. This means that the beam deviation is $\theta \sim 1/b$. Substituting this relation into (4.7.6), we obtain

$$\frac{d\sigma}{d\Omega} \sim \frac{1}{\theta^4}. \tag{4.7.7}$$

Thus, the cross-section for forward scattering diverges as $\theta \to 0$.

Low frequencies

As already mentioned above, one would expect the scattering of low-frequency waves to be essentially elastic and due to the field far away from the black hole. In fact, if we introduce

$$y = \left(1 - \frac{2M}{r}\right)^{-1/2} \hat{u}\,, \tag{4.7.8}$$

it follows from the Regge-Wheeler equation (4.2.7) that

$$\left[\frac{d^2}{dr^2} + \omega^2 + \frac{4M\omega^2}{r} - \frac{\ell(\ell+1)}{r^2}\right] y = 0 \tag{4.7.9}$$

whenever $M\omega << \ell$. For low frequencies (or large ℓ) it is sufficient to analyze this equation. For obvious reasons, Futterman, Handler, and Matzner (1988) refer to this as "the comparison Newtonian problem". Since (4.7.9) is identical to the radial Coulomb wave equation, one can take over well-known results from quantum mechanics. In particular, the corresponding phase-shifts δ_ℓ are given by

$$e^{2i\delta_\ell} = \frac{\Gamma(\ell+1-2iM\omega)}{\Gamma(\ell+1+2iM\omega)}, \tag{4.7.10}$$

and the scattering amplitude is

$$f(\theta) = M\,\frac{\Gamma(1-2iM\omega)}{\Gamma(1+2iM\omega)} \left(\sin\frac{\theta}{2}\right)^{-2+4iM\omega}. \tag{4.7.11}$$

Note that for small angles θ this approximation leads to a differential cross-section with the anticipated divergence, cf. (4.7.7). The approximation (4.7.11) is valuable, especially, since it provides a large ℓ complement to numerical calculations for the low ℓ partial waves [Handler and Matzner (1980)].

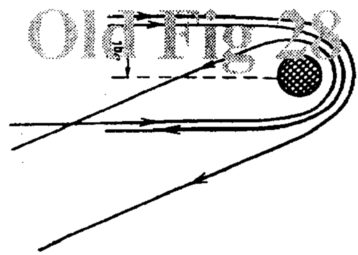


Figure 4.12: Trajectories of rays with impact parameters close to the critical one for glory scattering, b_{gl}. The path difference between rays is small, and the resulting difference in phase produces interference.

Glory effect

Whenever the classical cross-section diverges in either the forward or the backward direction, a diffraction phenomenon called a glory arises. This phenomenon is well known in optics, but it also arises in quantum scattering [Ford and Wheeler (1959a,b)]. For black holes glories arise because a wave can be deflected through an angle greater than π; see Figure 4.12 [Futterman, Handler, and Matzner (1988)]. The glory scattering is associated with the unstable circular photon orbit at $r = 3M$. Its effect is that, for a critical value of the impact parameter, $b_c = 3\sqrt{3}\,M$, we get a logarithmic singularity in the deflection; see Figure 4.13. Darwin (1959) deduced the following approximate expression for the impact parameter close to this singularity

$$b(\Theta) \approx 3\sqrt{3}\,M + 3.48Me^{-\Theta}\,. \tag{4.7.12}$$

Here Θ is the total angle of scattering or a so-called *deflection function.* Whereas the scattering angle θ is restricted to the interval $[\,0, \pi\,]$, the function Θ can attain any value (this is apparent in Figure 4.13). We define b_{gl} as the value of the impact parameter b corresponding to scattering by $\theta = \pi$. For high frequencies $\omega M \gg 1$ one has $b_{gl} \approx 5.35M$.

An approximation of the glory effect has been derived for scattering of waves with arbitrary spin [DeWitt-Morette and Nelson (1984), Zhang and DeWitt-Morette (1984)]. For waves of spin s the resultant formula is

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{glory}} \approx 2\pi\omega b_{gl}^2 \left|\frac{db}{d\theta}\right|_{\pi} J_{2s}^2(\omega b_{gl}\sin\theta)\,, \tag{4.7.13}$$

where $J_{2s}(x)$ is the standard Bessel function. When combined with the Darwin formula (4.7.12) – and using $b_{gl} = b(\pi) \approx 5.35M$ – this provides a good approximation whenever $\omega M >> 1$ and $|\,\theta - \pi\,| << 1$ (see Matzner *et al.* (1985) and also Figure 4.14 here).

From this approximation one can infer that the glory rings are bright at the centre for a scalar field, but dark in the case of fields with nonzero spin. Thus, one expects the cross-section to have a maximum at $\theta = \pi$ for scalar waves, but minima for electromagnetic and gravitational waves.

Orbiting resonances

A particle that is scattered by a potential can sometimes orbit the centre several times before it escapes. This is known as an orbiting resonance [Ford and Wheeler (1959a,b)]. Orbiting typically leads to dips in the cross-section at certain angles. Finding an analytic description of this phenomenon has proved an enormous challenge. Recently, Anninos, DeWitt-Morette *et al.* (1992) have presented an approximation scheme and applied it to scalar wave scattering from black holes.

The scalar wave cross-section due to orbiting can be written as

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{orbiting}} \approx \frac{M^2}{\sin^4\theta/2}\left|1+\frac{iA}{\alpha}\exp(i\beta I_2)\right|^2 . \tag{4.7.14}$$

Here

$$\alpha = \omega b_\theta\sqrt{2\pi\omega(b_c-b_\theta)}\,, \tag{4.7.15}$$

$$\beta = -\omega[b_\theta-(1-\theta)b_c]-\pi/4+\Phi\;, \tag{4.7.16}$$

and

$$I_2 = \omega b_c\cos\theta\left\{\frac{\omega b_c}{\pi}\log\frac{\sin|\pi[\omega(2b_c-b_{\text{gl}})-n]|}{\sin|\pi[\omega b_c-n]|}+i\omega^2 b_c(2b_c-b_{\text{gl}})\right\} . \tag{4.7.17}$$

Here n is to be chosen such that $\omega(b_c-b_{\text{gl}})-n<1$ and $\omega b_c-n<1$. We recall that $b_c = 3\sqrt{3}\,M$, $b_{\text{gl}} = b(\pi)$, and have introduced the shorthand notation $b_\theta = b(\theta)$. The approximation (4.7.14) leaves two parameters (A and Φ) free, but Anninos, DeWitt-Morette *et al.* (1992) obtain good results by adjusting these parameters to fit the overall amplitude and phase of numerically obtained cross-sections. They conclude that the oscillatory features seen away from the forward and backward directions are due to orbiting [see Figure 9 of Anninos, DeWitt-Morette *et al.* (1992)].

4.7.3 Wave scattering

The approximations described above – although instructive and useful – cannot provide a complete description of black hole scattering. First of all, one must proceed beyond geometrical optics to the wave problem, which allows for interference effects and absorption by the black hole. Secondly, it is impossible to obtain approximations that are valid for all scattering angles. Hence, the problem must be approached numerically. Once the phase-shifts δ_ℓ are obtained, for a given wave frequency ω and a large number of partial waves ℓ, all physical quantities of interest follow.

The scattering problem has been studied in a multitude of papers. Scalar fields were considered by Matzner (1968), Sánchez (1976, 1977, 1978a,b, 1997), and recently by Andersson (1995b). The scattering of various massless fields in the low-frequency limit was considered by Matzner and Ryan (1977). Extensive calculations for gravitational waves were made by Matzner and Ryan (1978) and Handler and Matzner (1980).

When extracting information from a set of phase-shifts, the semi-classical scattering description of Ford and Wheeler (1959a,b) [see also Handler and Matzner (1980)] is useful. In this picture the impact parameter is related to ℓ (the angular momentum), via

$$b = \left(\ell + \frac{1}{2}\right)\frac{1}{\omega} . \tag{4.7.18}$$

That is, each partial wave is considered as impinging on the black hole an initial distance b away from the axis.

Most of the physical information can be extracted from the deflection function. It corresponds to the angle by which a certain partial wave is scattered by the black hole, and is related to the real part of the phase-shifts through

$$\Theta(\ell) = 2\frac{d}{d\ell}\,\mathrm{Re}\,\delta_\ell . \tag{4.7.19}$$

(Here it is assumed that ℓ can attain any real value.) An example of this function for scalar wave scattering is shown in Figure 4.13. For large values of the impact parameter, b, one would expect the value of the deflection function to agree with the classic result $\Theta \approx -4M/b$. As can be seen in Figure 4.13, this is, indeed, the case.

Because of the singularity in the deflection function, one should expect interference between waves that are scattered through angles differing by multiples of π; see Figure 4.13. This leads to the glory phenomenon, and it will be most pronounced in the backward direction $\theta \approx \pi$. Figure 4.14 shows a typical cross-section for scattering of monochromatic scalar waves off a Schwarzschild black hole. The scattered wave has a frequency $\omega M = 2$. For large values of the deflection angle θ a beautiful example of glory oscillations can be seen. Figure 4.14 demonstrates that the glory approximation (4.7.13) is in excellent agreement with the partial-wave result for $|\,\theta - \pi\,| << 1$.

A feature of obvious importance in black hole scattering is absorption. For scalar waves, one can show that the total absorption cross-section is [Sánchez (1978a, 1997), Andersson (1995b)]

$$\sigma^{\text{abs}} = \frac{\pi}{\omega^2}\sum_{\ell=0}^{\infty}(2\ell+1)\left[1 - |\,R\,|^2\right] = \frac{\pi}{\omega^2}\sum_{\ell=0}^{\infty}(2\ell+1)\left[1 - e^{-4\,\mathrm{Im}\,\delta_\ell}\right] . \tag{4.7.20}$$

From the low-frequency approximation given in section 4.2 it follows that the quantity in square brackets in (4.7.20) behaves as $\omega^{2\ell+2}$ as $\omega \to 0$. Thus, it is clear that only the $\ell = 0$ partial wave contributes to absorption in this limit. This means that only

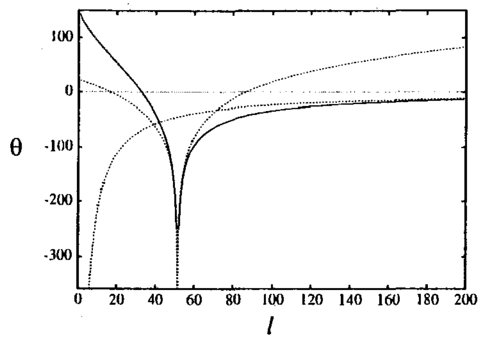


Figure 4.13: The deflection function for a plane scalar wave scattered from a Schwarzschild black hole. The frequency of the scattered wave is $\omega M = 10$. The deflection function as obtained from the phase-shifts (solid line) is compared to two approximations (dashed lines): 1) Einstein deflection and 2) the predictions of Darwin's formula. The first is accurate for large ℓ, and the second is reliable for impact parameters close to that associated with the unstable photon orbit ($\ell \approx 52$ in this case) [data from Andersson (1995b)].

waves that "approach the black hole radially" will be absorbed. Consequently, one would expect the absorption cross-section to become equal to the surface area of the black hole in the low-frequency limit. Indeed, given the approximations in Section 4.2, it is easy to show that the absorption cross-section approaches the anticipated value, $16\pi M^2$ [Unruh (1976c)]. In the high-frequency limit the absorption cross-section approaches the geometrical-optics value $27\pi M^2$. This means that all high-frequency waves that have an impact parameter $b < 3\sqrt{3}\, M$ will be absorbed.

The analysis of electromagnetic and gravitational wave scattering is, in principle, identical to that for scalar waves. But in the case of gravitational waves an additional complication enters. Axial and polar waves scatter differently, and the scattering amplitude consists of a sum of their individual contributions [Chrzanowski *et al.* (1976), Matzner and Ryan (1977, 1978)]. Hence, the cross-section may show features due to interference between these two contributing terms. We will return this issue when we discuss scattering from rotating black holes in Section 4.8.

4.7.4 The complex angular momentum paradigm

Although calculations are straightforward within the standard partial-wave picture of scattering, it suffers from some deficiencies. For example, the partial-wave sum (4.7.4) is formally divergent [Futterman, Handler, and Matzner (1988), Andersson (1995b)]. Moreover, it is not trivial to interpret the results for scattering of gravitational waves. The situation is especially complicated for Kerr black holes; see Section 4.8. These are reasons why alternative descriptions are of interest.

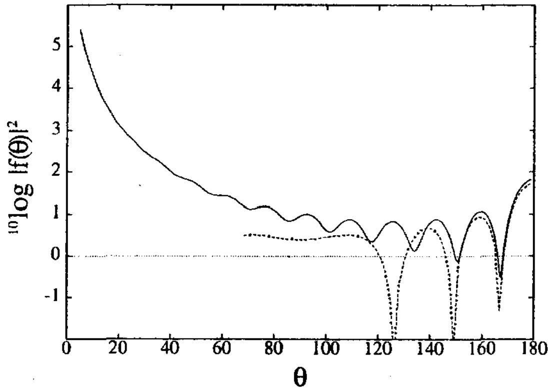


Figure 4.14: The differential cross-section for scalar waves scattered off a Schwarzschild black hole ($\omega M = 2.0$, solid). For comparison we also show the glory approximation (dashed), that agrees well with the partial-wave results [data from Andersson (1995b)].

One such alternative, that makes use of a complex angular momentum, has recently been developed for scalar wave scattering [Andersson and Thylwe (1994), Andersson (1994)]. The generalization of the idea to other fields is straightforward. After a Watson-Sommerfeld transformation, which schematically corresponds to

$$f(\theta) = \sum_{\ell} \to \int d\ell , \tag{4.7.21}$$

the scattering amplitude (4.7.4) splits naturally into two terms. One is a sum over the so-called Regge poles, and the other is a smooth background integral that can be approached by asymptotic methods.

The Regge poles are singularities of $S(\ell) = e^{2i\delta_\ell}$, that is correspond to complex values ℓ_n such that the solution to the Regge-Wheeler equation behaves as purely ingoing waves falling across the event horizon, and at the same time behaves as purely outgoing waves at spatial infinity, cf. (4.7.5). In fact, the Regge poles are complex-angular-momentum analogues to the quasinormal modes of a black hole; see Section 4.4. Many of the methods used to compute quasinormal-mode frequencies could therefore potentially be used also to determine the Regge pole positions ℓ_n. The approximations (4.4.1) and (4.4.2) for quasinormal modes can be inverted to approximate the Regge poles. Approximate values for the first few ℓ_n, and the corresponding residues, in the context of scalar wave scattering from Schwarzschild black holes have recently been obtained [Andersson (1994)].

Figure 4.15 shows a typical result obtained within the complex angular momentum description. The Regge-pole sum is dominated by the first pole for large angles ($\theta \geq 40°$). Each consecutive pole then gives a contribution that is roughly two orders

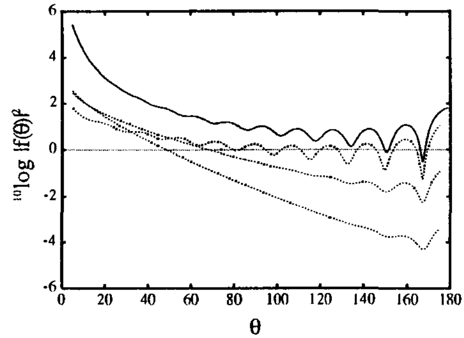


Figure 4.15: The differential cross-section for scalar waves and $\omega M = 2$ as obtained from phase-shifts (solid line) compared to the individual contribution from each of the first three Regge poles (dashed lines). A one-pole approximation is sufficient to describe the black hole glory [data from Andersson (1994)].

of magnitude smaller than that of the preceding pole. In order to obtain a reasonably accurate description of the black hole glory, it is sufficient to include only one Regge pole. In contrast, a large number of partial waves had to be retained in the standard description.

Figure 4.15 illustrates the usefulness of complex angular momenta for black holes. But equally important is the interpretative power of the new representation. One can show that each Regge state should be interpreted as two surface waves that travel around the black hole in opposite directions [Andersson (1994)]. At the same time, the waves decay at a rate corresponding to the inverse of the imaginary part of ℓ_n. Moreover, it is easy to show that the anticipated diffraction oscillations in the backward direction (the black hole glory) will have a period of $\pi/\text{Re}\ (\ell_n + 1/2)$. One can also infer that the real part of each Regge pole is associated with the distance from the black hole at which angular decay occurs, *i.e.*, $\text{Re}\ (\ell_n + 1/2) \approx \omega R_n$. In the case of a Schwarzschild black hole such surface waves should be localized close to the circular unstable photon orbit at $r = 3M$ (or, strictly speaking, the maximum of the Regge-Wheeler potential). This would correspond to $R_n = 3\sqrt{3}\,M \approx 5.196M$. The first Regge pole for various frequencies leads to R_n very close to this value [Andersson (1994)].

4.8 Wave Fields around a Rotating Black Hole

The analysis of weak fields in the geometry of a rotating black hole is in many ways similar to that for non-rotating holes. Thus, perturbations of a Kerr black hole can be studied using the techniques we have already discussed. This is fortunate since the exterior Kerr geometry is believed to be the universal late-time limit reached by gravitational collapse of any rotating body. Since most stars are rotating, it seems likely that black holes will be created with at least some angular momentum. Hence, the Kerr black hole should be of more importance to astrophysics than the Schwarzschild one.

The analysis of Kerr black holes is not completely identical to that for non-rotating holes, however. Some additional features arise because the Kerr metric has only axial symmetry. Carter (1968b) was the first to demonstrate that the variables can be separated for geodesic motion also in the Kerr background. The scalar wave problem was discussed by Brill *et al.* (1972). For some time it was believed that the wave equation would not be separable for fields of nonzero spin. The main reason for this was that the Regge-Wheeler approach – where one considers direct metric perturbations; see Section 4.2.4 – leads to a complicated, and seemingly inseparable, set of coupled equations. Indeed, when separability was demonstrated, it was through the use of an alternative approach. Teukolsky (1972, 1973) used the Newman-Penrose formalism (see Appendix E) to study the problem. From the Newman-Penrose point of view, the Kerr solution is very similar to the Schwarzschild one (Teukolsky's derivation does, in fact, apply to any Petrov type D vacuum metric), and the perturbation equations are completely separable into a set of ordinary differential equations. Amazingly, a single master equation – now known as the *Teukolsky equation* – describes scalar, electromagnetic and gravitational perturbations of a Kerr black hole. This equation also governs the evolution of spin 1/2 (neutrino) fields.

4.8.1 The Teukolsky equation

When no sources are present, the Teukolsky equation takes the form (see Teukolsky (1973) and Appendix G.1)

$$\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\frac{\partial^2\psi}{\partial t^2} + \frac{4Mar}{\Delta}\frac{\partial^2\psi}{\partial t\partial\phi} + \left[\frac{a^2}{\Delta} - \frac{1}{\sin^2\theta}\right]\frac{\partial^2\psi}{\partial\phi^2}$$

$$-\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial\psi}{\partial r}\right) - \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) - 2s\left[\frac{a(r-M)}{\Delta} + \frac{i\cos\theta}{\sin^2\theta}\right]\frac{\partial\psi}{\partial\phi}$$

$$-2s\left[\frac{M(r^2-a^2)}{\Delta} - r - ia\cos\theta\right]\frac{\partial\psi}{\partial t} + (s^2\cot^2\theta - s)\psi = 0\,. \tag{4.8.1}$$

Here, $a \leq M$ is the rotation parameter, connected with the black holes angular momentum J, $a = J/M$; s is the spin-weight of the perturbing field ($s = 0, \pm1, \pm2$), and we have used standard Boyer-Lindquist coordinates (see Section 3.2). We have also defined $\Delta \equiv r^2 - 2Mr + a^2$. The event horizon of the black hole follows from $\Delta = 0$, and thus correspond to $r_+ = M + \sqrt{M^2 - a^2}$. The field quantities that satisfy (4.8.1) are listed in Table 4.2.

s	0	1	-1	2	-2
ψ	Φ	Φ_0	$\rho^{-2}\Phi_2$	Ψ_0	$\rho^{-4}\Psi_4$

Table 4.2: The field quantities (in the Newman-Penrose formalism) that correspond to the function ψ in the Teukolsky equation for various (integer) spin-weights s. The quantity ρ is equal to $-(r - ia\cos\theta)^{-1}$.

The variables in (4.8.1) can be separated by decomposing ψ into modes according to

$$ {}_s\psi_{\ell m} = {}_sR_{\ell m}(r,\omega)\; {}_sZ_{\ell m}(\theta,\phi)\, e^{-i\omega t}\,, \tag{4.8.2}$$

where

$$ {}_sZ_{\ell m}(\theta,\phi) = \frac{1}{\sqrt{2\pi}}\; {}_sS_{\ell m}(\theta)\, e^{im\phi}\,. \tag{4.8.3}$$

Then the equations governing the radial function ${}_sR_{\ell m}$ and the angular function ${}_sS_{\ell m}$ are (suppressing all indices for clarity)

$$ \Delta^{-s}\frac{d}{dr}\left(\Delta^{s+1}\frac{dR}{dr}\right) + \left[\frac{K^2 - 2is(r-M)K}{\Delta} + 4is\omega r - \lambda\right] R = 0 \tag{4.8.4}$$

and

$$ \frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dS}{d\theta}\right) + \left[a^2\omega^2\cos^2\theta - \frac{m^2}{\sin^2\theta} - 2a\omega s\cos\theta \right.$$
$$ \left. - \frac{2ms\cos\theta}{\sin^2\theta} - s^2\cot^2\theta + E - s^2\right] S = 0\,. \tag{4.8.5}$$

Here we have introduced

$$ K \equiv (r^2 + a^2)\omega - am \tag{4.8.6}$$

and

$$ \lambda \equiv E - s(s+1) + a^2\omega^2 - 2am\omega\,. \tag{4.8.7}$$

The solutions to the angular equation are generally called *spin-weighted spheroidal harmonics.* The analytic properties of the functions ${}_sS_{\ell m}$ and the solutions to the radial equation were discussed in some detail by Hartle and Wilkins (1974). Leaver (1986a) derived a power-series expansion for the angular functions and an expansion in terms of Coulomb wave functions for the radial solutions. We will return to the radial functions later. The angular eigenfunctions are discussed in Appendix G.2.

Exactly as for Schwarzschild black holes, one can generate all non-trivial features of a perturbing field from the solution ψ to (4.8.1) [Wald (1973)]. Moreover, Chrzanowski (1975) has shown how the perturbed metric is obtained from Ψ_0 and Ψ_4 (see Appendix G.6 for more details). This completes the perturbation picture for Kerr black holes. The method of separation of variables has been used to analyze the stability of the Kerr metric and to study the scattering of electromagnetic, gravitational, and neutrino fields by Kerr black holes. We will discuss the results of such studies later.

Attempts to separate the variables in the Dirac equation with nonzero mass were unsuccessful until the work of Chandrasekhar (1976b). He suggested a method in which the separation of variables was carried out before the system of equations was decoupled. Page (1976c) extended this approach to the case of Kerr-Newman black holes. Lee (1976), Chitre (1975, 1976), and Bose (1975) treated the interacting electromagnetic and gravitational perturbations in the Kerr-Newman metric. In this case, one generally finds coupled wave equations for, e.g., Φ_2 and Ψ_4. A more complete exposition of the mathematical theory of the propagation of physical fields in the spacetime of a rotating black hole can be found in Chandrasekhar's monograph (1983).

The radial Teukolsky functions

Once a solution to the Teukolsky equation (4.8.1) is known, it is possible to calculate all relevant details of the perturbing field. The Kerr problem thus essentially reduces to an analysis of the radial equation (4.8.4). From that equation, one complication is obvious: The "effective" potential depends directly on the frequency ω. It is also clear that rotation breaks the angular degeneracy that is present for Schwarzschild black holes. Hence, all the $2\ell + 1$ values of m must be considered. Anyway, just as in the Schwarzschild case, it is useful to introduce a tortoise coordinate defined by

$$\frac{d}{dr_*} = \frac{\Delta}{r^2 + a^2} \frac{d}{dr} \tag{4.8.8}$$

and a new variable

$$\chi = (r^2 + a^2)^{1/2} \Delta^{s/2} R \,. \tag{4.8.9}$$

Then the radial equation (4.8.4) takes the form

$$\left[\frac{d^2}{dr_*^2} + \frac{K^2 - 2is(r-M)K + \Delta(4ir\omega s - \lambda)}{(r^2+a^2)^2} - G^2 - \frac{dG}{dr_*} \right] \chi = 0 \,, \tag{4.8.10}$$

where the function G is defined as

$$G = \frac{s(r-M)}{r^2+a^2} + \frac{r\Delta}{(r^2+a^2)^2} \,. \tag{4.8.11}$$

Now, the asymptotic behavior of (4.8.10) implies that the solutions behave as

$$\chi \sim r^{\pm s} e^{\mp i\omega r_*} \quad \text{or} \quad {}_sR_{\ell m} \sim \begin{cases} e^{+i\omega r_*}/r^{2s+1} \\ e^{-i\omega r_*}/r \end{cases} \quad \text{as } r \to \infty \,. \tag{4.8.12}$$

The different power-law fall-offs are in accordance with the so-called *peeling theorem* [Newman and Penrose (1962)].

Close to the event horizon r_+ the behavior is

$$\chi \sim \Delta^{\pm s/2} e^{\pm i\varpi r_*} \quad \text{or} \quad {}_sR_{\ell m} \sim \begin{cases} e^{+i\varpi r_*} \\ \Delta^{-s} e^{-i\varpi r_*} \end{cases} \quad \text{as } r \to r_+ \,. \tag{4.8.13}$$

Here ϖ is defined as $\varpi = \omega - m\Omega_H$, where $\Omega_H = a/(2Mr_+)$ is the angular velocity of the black hole. For rotating black holes the formulation of the physical boundary condition at the horizon requires some care. One must require that a wave be ingoing "in the frames of all physical observers" (who will be dragged around the hole by its rotation). One can show that this means that one should impose ${}_sR_{\ell m} \sim \Delta^{-s} e^{-i\varpi r_*}$ as $r \to r_+$ [Teukolsky (1973)].

The fact that the outgoing and ingoing parts of a solution to (4.8.4) have amplitudes that diverge asymptotically presents a serious obstacle to numerical studies. It would clearly be advantageous if one could transform the original equation into one with a short-range potential. That this may be possible is illustrated by the Schwarzschild case where the equation derived by Bardeen and Press (1973) [which corresponds to the $a = 0$ limit of (4.8.1)] can be transformed into both the Regge-Wheeler and the Zerilli equations [Chandrasekhar (1975)]. Inspired by this example, Chandrasekhar and Detweiler approached the Kerr problem in a series of papers. They found that there are, in principle, an infinite number of alternatives – neither of which one has reason to prefer a priori. Imposing some reasonable constraints – that an analytic solution is preferable to one that involves numerical quadratures, for example – they derive four possible effective potentials for gravitational perturbations of a Kerr black hole [Chandrasekhar and Detweiler (1975b, 1976)]. These potentials are all of short range, and two of them limit to the Regge-Wheeler potential, while two approach the Zerilli potential as $a \to 0$. When the black hole is rotating, the potentials are typically complex valued. It can be shown that all four potentials lead to identical reflection and transmission coefficients. Unfortunately, the potentials derived by Chandrasekhar and Detweiler are rather complicated. Hence, we will not give them explicitly here. Moreover, the new potentials have some added complications. For example, there may be a singularity in the region $r > r_+$ (outside the black hole) when one has superradiance (see below for a discussion of this effect). Yet these equations may be better suited to numerical studies than (4.8.4).

Chandrasekhar (1976a) used a similar approach to derive a short-range potential for an electromagnetic field in the Kerr geometry. As an alternative, Detweiler derived an equation with a manifestly real-valued effective potential for this case [Detweiler (1976)]. The neutrino-problem was discussed by Chandrasekhar and Detweiler (1977).

Another alternative formulation of the Kerr problem was suggested by Sasaki and Nakamura (1981b, 1982a). They were motivated by a difficulty that arises when one considers test-particle motion in a black hole geometry. One encounters divergent integrals, essentially because one must "integrate over the infinite past history of the particle". One way of dealing with these divergences is to integrate by parts and then discard divergent surface terms [Detweiler and Szedenitz (1979)]. Sasaki and Nakamura take a different route and transform the problem into one without these divergences [see also Tashiro and Ezawa (1981)]. The resultant equation has a short-range potential and reduces to the Regge-Wheeler equation in the limit when $a \to 0$. In general, the new equation contains a d/dr_* term.

The Teukolsky-Starobinsky identities

Before concluding the general discussion of the equations describing perturbations of a rotating black hole, we want to point out some useful relations. One reason for doing so is obvious from (4.8.12), which describes the asymptotic behavior of solutions to the radial Teukolsky equation. If we consider waves of spin-weight $s = -2$ as an example, it follows that the Weyl scalar Ψ_4 will consist of an "outgoing-wave" part that falls off as $1/r$ and an "ingoing-wave" part that falls off as $1/r^5$ for large values of r. This result is in accordance with the "peeling theorem", but we know that we should calculate the energy flux from the $1/r$ part of the field. So at first it would seem as if we can only infer the outgoing wave flux from Ψ_4, and this would be inconsistent with the claim that all non-trivial features of a perturbation can be calculated once a single solution to the Teukolsky equation is known [Wald (1979a,b)]. In the given example, we must be able to evaluate also the ingoing-wave flux from Ψ_4, i.e., from the $1/r^5$ part of the field. The resolution to this mystery is more or less obvious. The solutions for $s = -2$ (Ψ_4) and $s = +2$ (Ψ_0) are not independent. One solution is obtainable from the other by means of differential transformations. These relations have been given by Teukolsky and Press (1974) and Starobinsky and Churilov (1973), and are often called the *Teukolsky-Starobinsky identities.* They are discussed in more detail in Chandrasekhar's book (1983). Furthermore, Chandrasekhar (1980) has derived similar identities for arbitrary spin s.

In the case of gravitational perturbations, these relations imply that a solution

$$\Psi_0 = {}_2R_{\ell m}\, {}_2S_{\ell m}\, e^{-i\omega t + im\phi} \tag{4.8.14}$$

corresponds to

$$(r - ia\cos\theta)^4 \Psi_4 = \mathcal{C}\, {}_{-2}R_{\ell m}\, {}_{-2}S_{\ell m}\, e^{-i\omega t + im\phi} \tag{4.8.15}$$

with some constant $\mathcal{C}$.

Given this result, one can make one interesting observation: It is possible to have $\Psi_4 = 0$ and yet have a nonzero Ψ_0. (The reverse situation is also attainable, but to see that, we need to use a relation different from the one above.) This situation corresponds to a so-called "algebraically special" solution, and arises if the constant $\mathcal{C}$ vanishes. In the Schwarzschild limit (when $a = 0$) the constant becomes

$$\mathcal{C} = (\ell-1)\ell(\ell+1)(\ell+2) - 12i\omega M \,. \tag{4.8.16}$$

Thus, algebraically special solutions can be found for a single, purely imaginary, frequency for Schwarzschild black holes. As it turns out, the perturbation equations can be solved analytically for this case [Chandrasekhar (1984)]. For example, algebraically special solutions have also been studied by Couch and Newman (1973).

4.8.2 Superradiant scattering

In general, one would expect an incident wave to partly penetrate the potential barrier that surrounds a black hole and be absorbed, while the remaining parts of the wave scatters to infinity (see the discussion in Section 4.2). Thus, one would expect to find that the amplitude of the scattered waves always is smaller than (or equal to) the amplitude of the incident wave. An analysis of rotating black holes shows that this is not necessarily the case. An impinging wave can be amplified as it scatters off the hole. This phenomenon, called *superradiance*, was first discussed by Zel'dovich (1970, 1971, 1972) and Misner (1972). The additional energy that is radiated to infinity is drawn from the rotational energy of the black hole (see Section 7.1).

To illustrate this phenomenon, let us study the specific case of a scalar field, and extend the discussion of Section 4.2 to the Kerr black hole. A physically acceptable scalar-field solution to (4.8.10) is defined by

$$\chi^{\text{in}} \sim \begin{cases} e^{-i\varpi r_*} & \text{as} \quad r \to r_+ \,, \\ A_{\text{out}}(\omega)\, e^{i\omega r_*} + A_{\text{in}}(\omega)\, e^{-i\omega r_*} & \text{as} \quad r \to \infty \,. \end{cases} \tag{4.8.17}$$

Using this function together with its complex conjugate and the fact that two linearly independent solutions to (4.8.10) must lead to a constant Wronskian, it is easy to show that

$$\left(1 - \frac{m\Omega_H}{\omega}\right) |\, T\,|^2 = 1 - |\, R\,|^2 \ , \tag{4.8.18}$$

where the transmission amplitude T and the reflection amplitude R are defined as in Section 4.2. From this equation it is clear that one will have superradiance ($R > 1$) if

$$\omega < m\Omega_H = \frac{ma}{2Mr_+} \ , \tag{4.8.19}$$

which is only achievable for positive values of m.

Alternatively, one can see that energy can be extracted from the black hole immediately from the boundary condition (4.8.13). If ϖ becomes negative, the solution

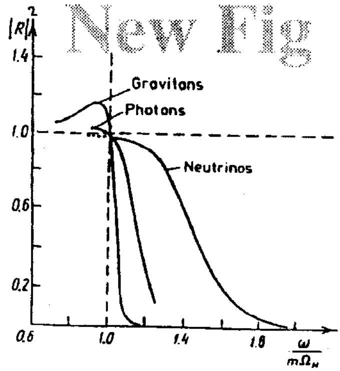


Figure 4.16: Reflection coefficient $|R|^2$ as a function of ω. Superradiance is possible when $\omega - m\Omega_H < 0$ [data from Chandrasekhar and Detweiler (1977)].

which behaves as $e^{-i\varpi r_*}$ and is "ingoing" according to a local observer, will correspond to waves coming out of the hole according to an observer at infinity.

Superradiance has been discussed by several authors [Press and Teukolsky (1972), Starobinsky (1973), Starobinsky and Churilov (1973), Teukolsky and Press (1974)]. One finds that the maximum amplification of an incoming wave is 0.3% for scalar waves, 4.4% for electromagnetic waves, and an impressive 138% for gravitational waves [Teukolsky and Press (1974)]. Figure 4.16 shows the reflection coefficient $|R|^2$ as a function of the frequency of the incident waves for different types. The figure reveals superradiance for gravitational and electromagnetic waves in the range $\omega < m\Omega_H$. No superradiance is seen for neutrinos. This is connected with the Pauli principle of exclusion valid for fermions [Unruh (1973), Chandrasekhar and Detweiler (1977), Martellini and Treves (1977), Chandrasekhar (1979b), and Iyer and Kumar (1978)]. We return to the discussion of this point in Chapter 10, where quantum aspects of scattering of particles by black holes are considered.

4.8.3 Radiation from a test particle moving in the Kerr background

Let us now turn to the radiation generated by a particle moving in the gravitational field of a rotating black hole. The problem is, in all essential respects, identical to that for Schwarzschild black holes [Sasaki and Nakamura (1981b, 1982a)]. Hence, the main conclusions will be the same as in Section 4.6. Of course, some added features will enter because the black hole rotates. We focus our attention on these features here.

The first problem considered in the Kerr background was that of a particle falling along the symmetry axis of the black hole [Sasaki and Nakamura (1981, 1982b),

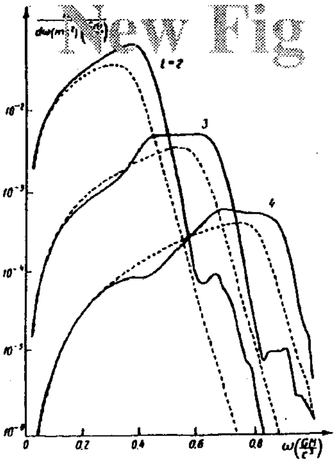


Figure 4.17: The spectrum of gravitational radiation (in different multipoles ℓ) generated by a test particle falling into a Kerr black hole. The particle is initially at rest at infinity, and falls along the symmetry axis of the hole. The case of a rapidly rotating black hole ($a/M = 0.99$, solid lines) is compared to the Schwarzschild case (dashed lines) [data from Sasaki and Nakamura (1982b)].

Nakamura and Sasaki (1981a,b)]. The particle is initially at rest at infinity (parabolic motion). This problem is relatively simple since one need only consider radiation in the $m = 0$ modes. Figure 4.17 shows the spectrum emitted when such a particle falls into a rapidly rotating black hole ($a = 0.99M$). For comparison, the corresponding spectrum for a Schwarzschild black hole ($a = 0$) is also shown. The spectra in Figure 4.17 peak at the corresponding quasinormal-mode frequencies. Compared to the Schwarzschild case the peaks are shifted towards higher frequencies because the oscillation frequency of the slowest damped $m = 0$ quasinormal mode typically increases with increasing a. The smaller peaks at high frequencies in Figure 4.17 probably occur because of coupling between different multipoles [Sasaki and Nakamura (1982b)]. Note that there is a minor peak in the $\ell = 2$ spectrum at the same position as the main peak in the $\ell = 3$ spectrum, etc. This explanation can also be used for the extra bumps in the high ℓ spectra at frequencies lower than that of the main peak.

The total energy radiated in this process generally increases (almost quadratically) with increasing a. When $a = 0.99M$, the energy radiated is 1.65 times that for a non-rotating hole. Moreover, the empirical relation (4.6.1) can be shown to hold also for rotating black holes [Sasaki and Nakamura (1982b)] so that the quadrupole always dominates the radiation.

Particles falling in the equatorial plane

The fact that the results for a particle falling along the symmetry axis of a Kerr black hole are very similar to those for a Schwarzschild hole is perhaps not too surprising. There is no frame-dragging effect if the particle falls onto the black hole along the axis. Most of the features that come into play when the black hole is rotating will be found in the $m \neq 0$ modes. These features should become apparent when one considers particles plunging into the black hole in the equatorial plane ($\theta = \pi/2$). Indeed, for a particle of zero angular momentum that falls into a Kerr black hole the spectrum is different from those in Figure 4.17 in one important respect. There are peaks corresponding to all the different $-\ell \leq m \leq \ell$ quasinormal modes for each ℓ [see Figure 1 of Kojima and Nakamura (1983a)]. This has the effect that the spectrum is extended towards higher frequencies. The total radiated energy is further enhanced, and for $a = 0.99M$ it is 4.27 times that for a Schwarzschild black hole.

In Figure 4.18 we show the total amount of emitted energy for particles with moderate angular momentum which are ultimately captured by the black hole (for a few different values of a). When the black hole is rotating, the curves in Figure 4.18 are clearly asymmetric for positive and negative $\tilde{L}$. They also do not reach a minimum at $\tilde{L} = 0$ as might be expected. Instead, the $a \geq 0.7M$ curves have minima at the negative value of $\tilde{L}$. This can be understood in the following way [Kojima and Nakamura (1983b, 1984a)]. Positive values of $\tilde{L}$ correspond to a particle that co-rotates with the black hole, whereas negative values are for counter-rotation. When a particle that was initially counter-rotating reaches the vicinity of the black hole, it will be slowed down because of frame-dragging. Thus, less gravitational waves are radiated. Similarly, an initially co-rotating particle is speeded up, and the amount of gravitational waves that emerges is increased. When the initial angular momentum is large, it dominates that of the black hole, and the radiated energy increases with $|\tilde{L}|$.

We have seen that quasinormal-mode ringing dominates the emerging radiation, and that most of the energy is radiated through the quadrupole, but what can be said about the relative excitation of modes corresponding to different values of m? For co-rotating particles (positive $\tilde{L}$) the $m = \ell$ mode typically dominates (Kojima and Nakamura 1984a); see Figure 4.19. The situation is markedly different when the particle is counter-rotating. Then the excitation of the $m = -\ell$ mode is the largest. However, this mode is more rapidly damped than the $m = \ell$ mode (see Figure 4.6). This has the effect that the relative importance of the other modes is enhanced (see Figure 4.19).

That the emerging gravitational waves are markedly different for co- and counter-rotating particles is clear from Figure 4.20. Because the $m = \ell$ quasinormal mode is slower damped, the ringing lasts longer for co-rotating particles than for counter-rotating ones (where the $m = -\ell$ mode dominates).

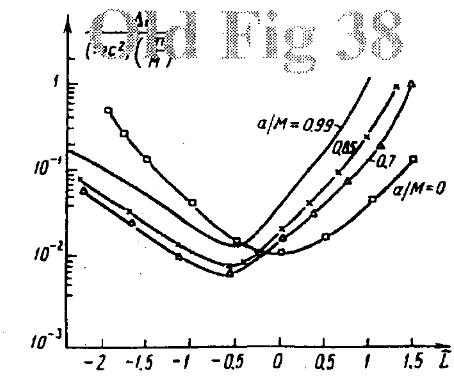


Figure 4.18: Total energy radiated (for $m = 0$) by a particle falling, at rest at infinity, into a Kerr black hole. The particle moves in the equatorial plane of the black hole. The radiated energy is shown as a function of the initial angular momentum of the particle for three different values of a. Also shown is the result for a Schwarzschild black hole [data from Kojima and Nakamura (1983b)].

Particles scattered by a Kerr black hole

Finally, we consider the emission of gravitational waves by a particle that is scattered by a Kerr black hole. The particles motion is in the equatorial plane ($\theta = \pi/2$); its velocity at infinity is zero, but its initial angular momentum $\tilde{L}$ is sufficiently large that it will not be captured by the black hole. This problem was treated by Kojima and Nakamura (1984b).

Recall that when a non-relativistic particle was scattered by a non-rotating black hole, the quasinormal modes were not excited (see Section 4.6). The main reason was that periastron of the particles trajectory lies beyond the potential barrier that surrounds the hole, and the waves emitted by the moving particle have frequencies that are too low to penetrate the barrier. However, if the black hole is rotating, the frequency of the gravitational waves emitted at periastron may be sufficiently close to the quasinormal-mode frequencies that the modes will be excited.

For a scattered particle one can make two general observations. The first one is that periastron lies further away from the hole (for any given a) for counter-rotating particles than for co-rotating ones. Secondly, for a given $|\tilde{L}|$ periastron is further away in the co-rotating case. This means that a co-rotating particle can generally excite the quasinormal modes, whereas a counter-rotating particle can not. Figures 4.21a and 4.21b show typical gravitational waves from a particle scattered off a black hole with $a = 0.99M$. The two waveforms correspond to initial specific angular momenta $\tilde{L} = 2.21$ and $\tilde{L} = 2.6$, respectively. In the first case, the periodic gravitational waves that are due to the particle orbiting the black hole several times are followed by distinguishable quasinormal-mode ringing. No such ringing can be

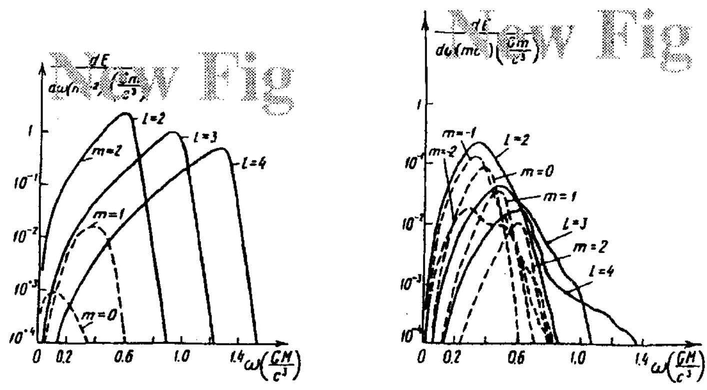


Figure 4.19: Total energy radiated in different multipoles for a particle falling in the equatorial plane of a rotating black hole ($a = 0.85M$). Here the initial specific angular momentum of the particle is (a) $\tilde{L} = 1.3$, *i.e.*, it co-rotates with the black hole. The $m = \ell$ mode clearly dominates the radiation. (b) $\tilde{L} = -2.25$. The particle is thus counter-rotating, and even though the $m = -\ell$ mode dominates the radiation, the importance of the other modes is enhanced [data from Kojima and Nakamura (1983b)].

seen in the second case. The amplitude of the quasinormal modes is small because they were excited by gravitational waves impinging on the hole.

4.8.4 Scattering of waves by Kerr black holes

Scattering of gravitational waves incident along the axis of symmetry of a Kerr black hole was first studied by Matzner and Ryan in 1978, and later reconsidered by Handler and Matzner (1980). It was found, not very surprisingly, that the shape of the scattering cross-section depends on the rotation parameter a. Moreover, certain features evolve as this parameter is changed.

In scattering from Kerr black holes two effects that do not exist for non-rotating holes play a role: superradiance and the polarization of the incident wave. In the plane-wave study of Chrzanowski *et al.* (1976) it is assumed that the impinging wave is circularly polarized. For incidence along the symmetry axis of the black hole one can consequently have either co- or counter-rotating waves. The two cases lead to quite different results [Matzner and Ryan (1978)]. Although the general features of the corresponding cross-sections are similar, they show different structure in the backward direction. This difference seemingly arises because of the phase-difference between axial and polar waves [the two "parities" of gravitational perturbations scatter differently, see Chandrasekhar (1980)].

The available results are best illustrated by an example. Let us consider the case $a = 0.9M$ and various frequencies. Typical results are shown in in Figure 4.22. Consider first the co-rotating case ($\omega > 0$). Two features are clear. First, the glory

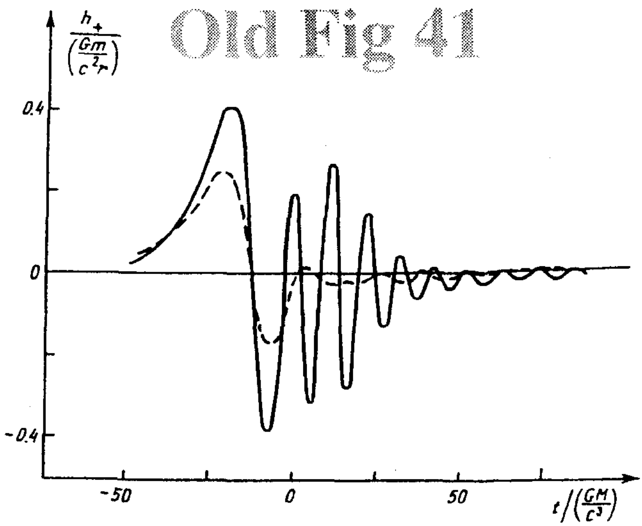


Figure 4.20: Typical quadrupole gravitational waves generated by a particle plunging into a Kerr black hole. The particle moves in the equatorial plane. An observer which registers this radiation is in the plane of the orbit in the direction of the periastron of the particle's trajectory. The two cases are the same as in Figure 4.19a (solid curve) and Figure 4.19b (dashed curve) [data from Kojima and Nakamura (1983b)].

near $\theta = \pi$ is very pronounced for $\omega M = 0.75$ (cf. the discussion of glories in Section 4.7). Secondly, there is a deep dip near $\theta = \pi/2$ for $\omega M = 1.5$. This is presumably due to orbiting [Handler and Matzner (1980)]. Compared to these, the features of the counter-rotating cross-section ($\omega < 0$) are much weaker. As Handler and Matzner have put it: "The glory is considerably less glorious in the counter-rotating case". Nevertheless, one finds that maxima and minima generally occur at the same angles in both cases. In general, all cross-sections reach minima as $\theta \to \pi$, as expected from the glory approximation (4.7.13). It also seems as if the effect of superradiance is to wash out interference minima, cf. Figure 4.23.

It is interesting to estimate the apparent size of the black hole as viewed along the rotation axis, cf. Table 4.3. It is clear that the size of the black hole appears markedly different depending on whether the infalling waves are co- or counter-rotating with the hole. A Kerr black hole appears much smaller to waves that are co-rotating.

4.9 Stability of Black Holes

Maybe the most important question related to the possible existence of black holes in our Universe is whether or not they are stable against perturbations. This question

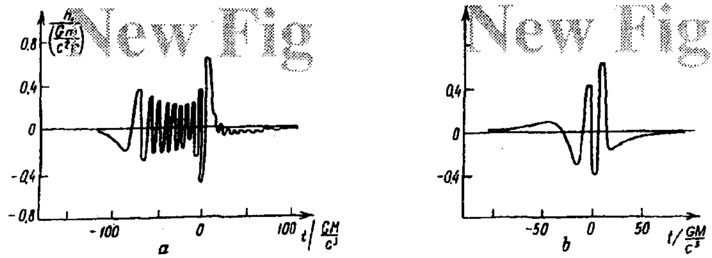


Figure 4.21: Typical gravitational waveforms for a particle scattered off a Kerr black hole with $a = 0.99M$. The particle is initially at rest at infinity and has initial specific angular momentum (a) $\tilde{L} = 2.21$, (b) $\tilde{L} = 2.6$. An observer which registers this radiation is in the plane of the orbit in the direction of the periastron of the particle's trajectory. The figures correspond to the $m = \ell = 2$ mode. This mode dominates the radiation [data from Kojima and Nakamura (1984b)].

ωM	-1.5	-0.75	0.75	1.5
σ^{abs}	$80.3M^2$	$88.7M^2$	$62.5M^2$	$36.5M^2$
b	$5.06M$	$5.31M$	$4.46M$	$3.41M$

Table 4.3: The apparent size b of a Kerr black hole as viewed along the rotation axis. The values are all for $a = 0.9M$ and are estimated from the total absorption cross-section $\sigma^{\text{abs}} \approx \pi b^2$. Positive frequencies co-rotate with the black hole whereas negative ones are counter-rotating. The values should be compared to $b = 5.2M$ for a Schwarzschild black hole [data from Handler and Matzner (1980)].

was the original motivation for studies of perturbed black holes, and it dates back to the seminal work of Regge and Wheeler in 1957. A specific question to be answered can be formulated in the following way. Suppose we set up a perturbation in a compact part of the region external to a black hole on an initial spacelike hypersurface. That is, we specify the value $\Phi(r,\theta,\varphi,t)$ and $\partial_t\Phi(r,\theta,\varphi,t)$ for $t = 0$ (say). To the future of the initial surface the evolution of the perturbation is governed by (4.2.1) or, more generally, (4.8.1). We then ask whether there is an upper bound for $|\Phi|$ at all points in the exterior of the horizon to the future of the initial surface. If not, the perturbation is considered (linearly) unstable.

The first attempts to address this issue ruled out perturbations of the Schwarzschild geometry that grow exponentially with time. A simple argument can be based on the use of physically acceptable boundary conditions [Vishveshwara (1970b)]. Let us consider perturbations of a Schwarzschild black hole and an unstable mode corresponding to a purely imaginary frequency $\omega = i\alpha$; that is, a solution to (4.2.4) with time-dependence $\exp(\alpha t)$. The asymptotic behavior of such solutions must be

$$\Phi \sim e^{\pm\alpha r_*}. \tag{4.9.1}$$

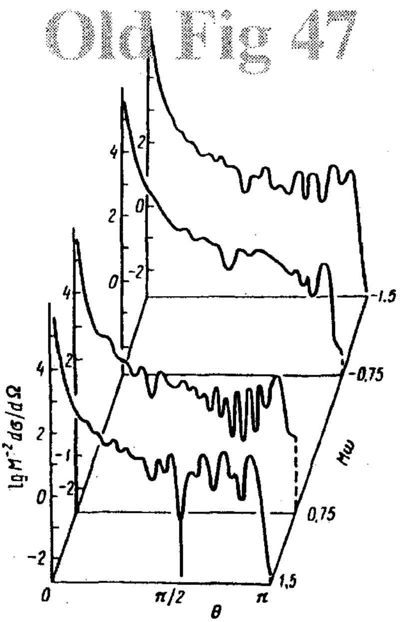


Figure 4.22: Differential cross-sections for scattering of gravitation radiation by a rotating black hole with $a = 0.9M$ for different frequencies ωM [data from Handler and Matzner (1980)].

both as $r_* \to +\infty$ and $r_* \to -\infty$. For regularity at spatial infinity we must clearly choose the lower sign. But from the fact that the effective potential in the Regge-Wheeler equation is positive definite it follows immediately that $\partial^2_{r_*}\Phi$ never becomes negative in $r \in (2M, \infty)$. Hence, the regular solution at spatial infinity cannot be matched to the regular solution at the horizon, and vice versa. A perturbation that grows exponentially with time is therefore physically unacceptable, and the metric should be stable.

This naive argument for mode stability can be put on more rigorous footing by the use of an "energy"-type integral [Detweiler and Ipser (1973), Wald (1979a)]. For Schwarzschild black holes the argument is straightforward. Multiply (4.2.4) by $\partial_t \bar{u}_\ell$ and add the resultant equation to its complex conjugate. Then we have

$$\frac{\partial}{\partial r_*}\left[\frac{\partial \bar{u}_\ell}{\partial t}\frac{\partial u_\ell}{\partial r_*} + \frac{\partial u_\ell}{\partial t}\frac{\partial \bar{u}_\ell}{\partial r_*}\right] = \frac{\partial}{\partial t}\left[\left|\frac{\partial u_\ell}{\partial t}\right|^2 + \left|\frac{\partial u_\ell}{\partial r_*}\right|^2 + V_\ell \left| u_\ell \right|^2\right] .$$

The integral on the left-hand side of this relation over r_* from $-\infty$ to $+\infty$ vanishes;

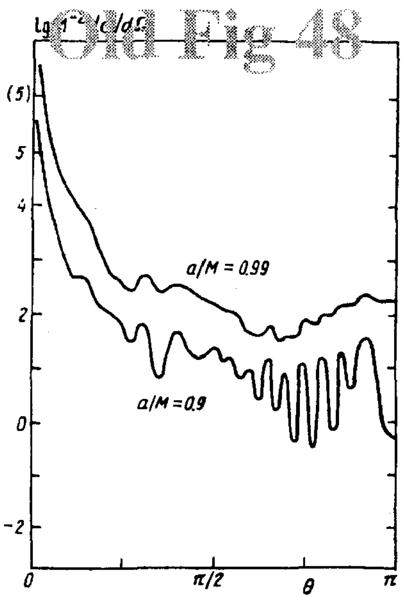


Figure 4.23: Differential cross-sections for scattering of gravitational radiation with $\omega M = 0.75$ by black holes with $a/M = 0.9$ and $a/M = 0.99$. Superradiance is an important effect in the latter case. The curve representing this situation is shifted upwards by unity along the coordinate axis to make the plot clearer. The number 5 in parenthesis on the coordinate axis refers only to this curve [data from Handler and Matzner (1980)].

hence,

$$\int_{-\infty}^{+\infty} \left[\left| \frac{\partial u_\ell}{\partial t} \right|^2 + \left| \frac{\partial u_\ell}{\partial r_*} \right|^2 + V_\ell \, | \, u_\ell \, |^2 \right] dr_* = \text{constant} . \tag{4.9.2}$$

Since V_ℓ is positive definite, this equation bounds $\partial_t u_\ell$ and excludes exponentially growing solutions to (4.2.4). Although more satisfactory than our first naive argument, this demonstration of stability has two deficiencies:

1. Perturbations that grow linearly (or slower) with t are not ruled out.
2. We have only obtained a bound for integrals of u_ℓ. The perturbation may still blow up in an ever narrowing region as $t \to \infty$.

An argument that avoids these two difficulties was provided by Wald (1979a). He used the Sobolev inequality to show that perturbations of the Schwarzschild geometry are bounded at each point in spacetime. This constitutes a complete proof of (linear) stability for Schwarzschild black holes.

An entirely different approach is that of Kay and Wald (1987). They prove that the solution u_ℓ remains pointwise bounded when (4.2.4) is evolved from smooth bounded initial data that has compact support on Cauchy surfaces in the Kruskal extension. This is the best available result for the initial value problem, and goes beyond statements of mode stability.

Unfortunately, the above stability proofs are not straightforwardly generalized to rotating black holes. Many studies have presented unsuccessful searches for instabilities, rather than direct proofs of their nonexistence. A typical example is the work of Press and Teukolsky (1973). They show that mode frequencies do not cross the real frequency axis into the unstable upper half of the complex ω-plane. Similarly, Detweiler and Ipser (1973) postulate as a sufficient condition for stability that no eigenmodes exist in the upper half-plane. Their numerical search for such modes proved unsuccessful, and thus supports stability. Subsequently, Stewart (1975) showed that no unstable modes will exist for finite Cauchy data, and Friedman and Schutz (1974) provided an analytic proof (based on a variational principle) that all axisymmetric modes of a Kerr black hole are stable.

The energy-integral approach is not easily extended to the Kerr case. To see the reason for this, we specifically consider scalar perturbations. For a scalar field the Teukolsky equation becomes, cf. (4.8.1)

$$\frac{\partial}{\partial r}\left[\Delta \frac{\partial \Phi}{\partial r}\right] + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta \frac{\partial\Phi}{\partial\theta}\right] + \left[\frac{1}{\sin^2\theta} - \frac{a^2}{\Delta}\right]\frac{\partial^2\Phi}{\partial\phi^2}$$

$$- \frac{4aMr}{\Delta}\frac{\partial^2\Phi}{\partial t\partial\phi} - \left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\frac{\partial^2\Phi}{\partial t^2} = 0\,. \qquad (4.9.3)$$

Problems with the energy-integral approach enter because of the factor in front of $\partial_\phi^2\Phi$. The quantity $1/\sin^2\theta - a^2/\Delta$ is not positive definite. Specifically, it becomes negative inside the ergosphere. Thus, one could imagine the possibility of an unstable perturbation that is growing in time, with the scalar field energy becoming increasingly negative inside the ergosphere, while it grows to a large positive value outside the ergosphere [Detweiler and Ipser (1973)]. However, Whiting (1989) has showed that it is possible to find a transformation into new dependent variables for which all terms in the integral are positive definite. Thus, one can prove mode stability also for Kerr black holes.

Nevertheless, a complete stability proof for Kerr black holes still eludes us. This is one of few truly outstanding problems that remain in the field of black hole perturbations.

4.10 Gravitational Waves from Binary Systems

During the last few years the demands for accurate modeling of processes generating gravitational waves have increased tremendously. The reason is that construction of large-scale interferometric detectors for such waves is now under way [Thorne

(1994a)], and accurate theoretical predictions are required to ensure detection (see Section 9.9 for an introductory discussion). Numerical calculations are required for the complete study of gravitational radiation from coalescing black holes. This problem provides a fantastic challenge (see Section 7.7). However, interestingly, both the inspiraling phase and the final merger of the two black holes can be modeled using perturbation theory. We will discuss results obtained in this way in this section.

4.10.1 The inspiraling phase

The problem of binary inspiral is usually approached through some kind of post-Newtonian expansion [Will (1994)]. This essentially comprises of an expansion in terms of v, where v is a typical velocity inside the matter source. One need impose no restrictions on (say) the relative masses of the two bodies. Full post-Newtonian calculations are very cumbersome, however. An interesting alternative is provided by black-hole perturbation theory. Let us assume that one of the masses in the binary system is much smaller than the other. This means that $\mu/M << 1$ where the reduced mass μ is practically equal to the smaller mass, and the total mass M is essentially the larger mass. The small mass creates a perturbation in the gravitational field of the larger mass, and the evolution of the binary is governed by the perturbation equations, e.g., the Teukolsky equation (4.8.1). In contrast with the post-Newtonian approach we need impose no restrictions on the orbital velocity v.

We thus assume that the large body is a Schwarzschild black hole of mass M. The small orbiting body is modeled as a point particle of mass μ. For simplicity, but also because eccentric orbits are expected to be circularized by radiation reaction, we can choose the world line of the particle to be a circular geodesic of the Schwarzschild spacetime (see also discussion in Section 2.10). From the standard quadrupole formula [Landau and Lifshitz (1975)] it then follows that the gravitational-wave luminosity approaches

$$\left(\frac{dE}{dt}\right)_N \equiv \frac{32}{5}\frac{\mu^2M^3}{r_0^5} \tag{4.10.1}$$

for large orbital radii r_0.

In the perturbation framework, the binary problem reduces to solving the inhomogeneous Teukolsky equation with a source term appropriate for a point particle in circular orbit around the black hole. This problem can be solved by the Green's function approach discussed for initial value problems in Section 4.3 (the problem is identical to that discussed in Section 4.6). This was first done by Detweiler (1978), who numerically calculated gravitational waveforms and luminosities for rotating holes. When it was realized that very accurate predictions were needed to ensure the successful detection of gravitational waves, the problem was reconsidered [Cutler *et al.* (1993)]. Now the accuracy obtainable through post-Newtonian (or other) approximation schemes was an issue of great importance. Unfortunately, numerical perturbation calculations indicated that the convergence of the post-Newtonian expansion is very poor, and that one will require corrections of rather high order, at least $O(v^6)$ [Cutler, Finn, Poisson, and Sussman (1993)].

The perturbation approach is not only well suited for numerical studies of this problem, however. One can combine the perturbation equations with a "slow-motion approximation" [a low frequency expansion of the solutions to (4.8.1)]. This facilitates direct comparison with the Post-Newtonian expansion. Moreover, as long as $v \ll 1$ (in the weak field regime, where $R_0 \gg 3M$), this problem can be solved by analytic approximations. In the first study along these lines Poisson (1993a) derived the relativistic corrections to (4.10.1) of order v^3.

Since then there have been several, increasingly accurate, studies of the problem. Using an ingenious iterative scheme [Sasaki (1994)], Tagoshi and Sasaki (1994) were able to derive an impressive result for the gravitational wave luminosity,

$$\frac{dE}{dt} = \left(\frac{dE}{dt}\right)_N \left[1 - \frac{1247}{336} v^2 + 4\pi v^3 - \frac{44711}{9072} v^4 - \frac{8191}{672} \pi v^5 \right.$$
$$+ \left(\frac{6643739519}{69854400} - \frac{1712}{105}\gamma + \frac{16}{3}\pi^2 - \frac{3424}{105}\ln 2 - \frac{1712}{105}\ln v\right) v^6$$
$$- \frac{16285}{504}\pi v^7 + \left(-\frac{323105549467}{3178375200} + \frac{232597}{4410}\gamma - \frac{1369}{126}\pi^2\right.$$
$$\left.\left. + \frac{39931}{294}\ln 2 - \frac{47385}{1568}\ln 3 + \frac{232597}{4410}\ln v\right) v^8 + \cdots\right]. \tag{4.10.2}$$

Here the presence of $\ln v$ terms should be noticed, as well as that of the Euler number $\gamma \simeq 0.5772$. The first four terms reproduce the Post-Newtonian result [Blanchet *et al.* (1995)] in the limit $\mu/M \to 0$ (and when spin-orbit and spin-spin interactions are neglected). The expression (4.10.2) also agrees perfectly with the high-precision numerical result of Tagoshi and Nakamura (1994).

Calculations of this type have also been performed for a particle in an eccentric orbit around the black hole [Tanaka *et al.* (1993); Apostolatos *et al.* (1993); Cutler, Kennefick, and Poisson (1994)]. Furthermore, the perturbation approach has been extended to rotating black holes: Poisson (1993b) approached the problem analytically in the slow-rotation limit. Shibata *et al.* (1995) extended these results to rapidly rotating holes. These studies concern circular orbits, but eccentric orbits have been considered also for rotating black holes [Tagoshi (1995), Shibata (1993a,b, 1994)].

An interesting question concerns at what order the effects of the event horizon will enter the calculations. Sasaki (1994) has shown that the existence of the horizon will not affect the outgoing radiation until at very high order: $O(v^{18})$. Meanwhile, black hole absorption will be an effect of order v^8 [Poisson and Sasaki (1995)]. The relative smallness of this effect can be understood in terms of the potential barrier close to the black hole. Most of the low-frequency gravitational waves traveling towards the

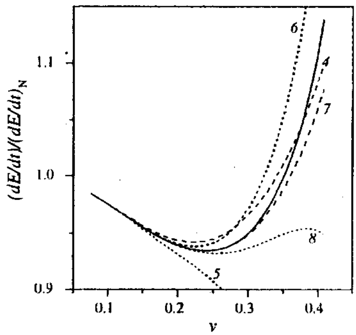


Figure 4.24: The gravitational-wave luminosity (obtained in various orders of the post-Newtonian expansion) as a function of the orbital velocity v (dashed curves). The solid curve represents the "exact" numerical result [data from Poisson (1995)].

hole will be reflected. Consequently, effects associated with the horizon are expected to be of little relevance for gravitational wave observations of inspiraling binaries. But the possible presence of an event horizon will obviously be relevant for the final coalescence of the two bodies.

Given the expression (4.10.2), it is easy to see that the post-Newtonian expansion converges poorly. For example, the second and the third terms become equal in magnitude for a frequency well inside the observational range for the laser-interferometers (roughly 10 Hz - 2 kHz). So higher order terms must certainly be used. This poor convergence is a serious obstacle to the construction of accurate measurement templates [Poisson (1995)]; see Figure 4.24. It is not clear that pushing to higher orders in post-Newtonian theory will be sufficient to produce acceptably accurate theoretical waveforms (the expansion may well be asymptotic). Other ways of dealing with this problem may be required.

4.10.2 Black hole collisions

One obvious way of studying the binary inspiral would be to avoid approximations altogether and perform direct numerical integration of Einstein's equations. However, this is a very difficult problem. Its eventual solution will require state-of-the-art computing combined with clever formulations of the problem. But even though the solution to the full problem of two spiraling black holes (assuming no symmetries) is still remote some progress has been made (see Section 7.7). Recently, Anninos, Hobill *et al.* (1993) studied head-on collisions of two black holes of equal mass.

Interestingly, this and similar problems can be approximated using perturbation theory. Let us assume that the two black holes start off so close together that they are surrounded by a common horizon. Then one can consider the situation as corresponding to a single perturbed black hole. That this "perturbation" is sufficiently "small" to make the approach sensible is not obvious, but as was shown by Price and Pullin (1994) this idea leads to surprisingly accurate results. They re-expressed the initial data as "the Schwarzschild background + something else" and showed that the full problem can be considered as an initial-value problem for the Zerilli equation (see section 4.2). The new initial data can be written [Anninos, Price *et al.* (1995)]

$$u_\ell(r,0) = \frac{2nr^2}{nr+3M}\left\{F\left[\mathcal{G}_\ell - \sqrt{F}\frac{d}{dr}\left(\frac{r\mathcal{G}_\ell}{\sqrt{F}}\right)\right] + (n+1)\mathcal{G}_\ell\right\}, \tag{4.10.3}$$

where

$$n = \frac{(\ell-1)(\ell+2)}{2}, \quad F = 1 - \frac{2M}{r}, \tag{4.10.4}$$

$$\mathcal{G}_\ell = 8\kappa_\ell(\mu_0)\sqrt{\frac{4\pi}{2\ell+1}}\frac{(M/R)^{\ell+1}}{1+M/2R}. \tag{4.10.5}$$

Here we have also used

$$\kappa_\ell(\mu_0) \approx \frac{\zeta(\ell+1)}{|4\ln\mu_0|^{\ell+1}}, \quad R = \frac{1}{4}\left(\sqrt{r}+\sqrt{r-2M}\right)^2. \tag{4.10.6}$$

Here ζ is the Riemann zeta-function, $\mu_0 = L/M$, L is the initial separation of the two holes, and M is the total mass of the spacetime. For $\mu_0 < 1.36$ the two black holes are surrounded by a common horizon. As is apparent from (4.10.3), the parameter μ_0 enters the calculation only through a multiplicative factor. Thus, it affects only the amplitude of the generated gravitational waves, and a single perturbation calculation can be used to describe all initial separations.

We have already discussed the evolution of the perturbation equations from given initial data in Section 4.3. There is essentially nothing new in the evolution that follows from (4.10.3). The waveforms are dominated by quasinormal-mode ringing, and at late times they follow the anticipated power-law tail [Price and Pullin (1994)]. However, the extent to which this approximation agrees with the full numerical simulation of the nonlinear equations is somewhat surprising. The results are compared in Figure 4.25. It is clear that the agreement extends well beyond the small μ_0 region. In fact, the perturbation calculation is reasonably accurate also in cases when the two black holes are not initially surrounded by a single horizon.

Anninos *et al.* (1995) compared the various results in more detail, including the actual waveforms that are generated. Their comparison shows that the perturbation results are robust, and although this technique will never be able to replace full numerical relativity, it provides an important complement. It can obviously be used as a test-bench for numerical simulations, and also predict results and improve our understanding of the physics involved. Moreover, since it is computationally

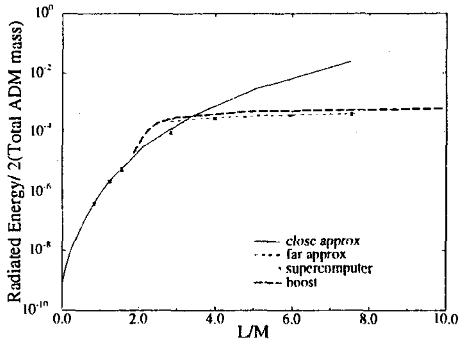


Figure 4.25: Comparing the radiated energy in the close-limit approximation to that of the full numerical simulation [data from Price and Pullin (1994)].

inexpensive, the approximation can be used to search the parameter space for interesting phenomena. Once identified, these phenomena can be studied in more detail numerically.

The close limit approximation for colliding black holes can be extended in several interesting directions. For example, it can be applied to initial data that are only known numerically. As an illustration of this, Abrahams and Cook (1994) considered an extension of the initial data that allows the black holes to have initial linear and angular momentum. They found that the radiation efficiency ($\Delta E/E$) saturated at about 2%. This result seems to suggest that, no matter how large the initial momentum is, it will be impossible to achieve high efficiency in black hole collisions. Further details on how numerical initial data can be incorporated in the calculation have been given by Abrahams and Price (1996a,b).

The close limit approximation shows that perturbation calculations often are much more accurate than one has any reason to expect. In the case of colliding black holes this can be explained by the following argument: The spacetime is only strongly distorted in the region close to the horizon. Because of the existence of the potential barrier (that has a peak at roughly $r = 3M$) in the Zerilli equation, most of this initial perturbation will never escape to spatial infinity. It will be reflected and fall down the hole. Hence, the predicted waveforms in the outer region may be very accurate. But however plausible, this explanation is only based on our intuition. If the perturbation approach is to provide actual predictions, it need be supplemented by some kind of error estimate. The obvious way to obtain such an estimate would be through perturbation calculations to second order. Then the difference between first- and second-order quantities can be used to estimate the achieved accuracy.

It is not at all straightforward to derive such second-order equations. Nevertheless, recent work by Gleiser *et al.* (1995) looks promising. They find that the second-order extension of the Zerilli equation has the form

$$\left[\frac{\partial^3}{\partial t\,\partial r_*^2} - \frac{\partial^3}{\partial t^3} - V(r)\,\frac{\partial}{\partial t}\right] u_\ell^{(2)}(r,t) = S(u_\ell^{(1)})\,. \tag{4.10.7}$$

The second-order quantity $u_\ell^{(2)}(r,t)$ is described by an equation identical to the time-derivative of the first-order equation, cf. (4.2.7), that governs $u_\ell^{(1)}$. The source-term S is a complicated expression quadratic in the first-order perturbations. This result is interesting, and opens up a multitude of possibilities. In the future many of the results discussed in this chapter may be extended to higher orders. That is, the first-order perturbation results can be associated with "error bars". At this time, it is too early to speculate where this may lead, but it is clear that perturbation calculations will continue to play an important role in black hole physics.

Chapter 5

General Properties of Black Holes

5.1 Asymptotically Flat Spacetimes

5.1.1 Asymptotic properties of Minkowski spacetime

So far we have restricted the description of the properties of black holes to an analysis of the Schwarzschild and Kerr metrics and their perturbations. Both Schwarzschild and Kerr spacetimes are stationary and possess certain symmetries. A study of geodesics (corresponding to the motion of test particles and light rays) and wave fields in these metrics enabled us to describe a number of interesting and important features of physical effects in the field of such black holes. It is natural to ask: Are there black holes that differ from those already described? What are their properties? In order to answer these questions, the definition must first of all be extended to the general case of non-stationary spacetime. This generalization is obvious. It is reasonable to define the *black hole* in general as the region of spacetime from which no information-carrying signals are allowed to escape to a distant observer.

However, in order to make this definition rigorous, one must elaborate what class of observers is meant and what is understood by "distant" in geometrically invariant terms. The necessary refinement is easily achieved in the physically important case in which there is no matter and no sources of fields far from the black hole. The greater the distance from the black hole, the smaller the deviations of the spacetime geometry are from flatness. A spacetime with this property is said to be *asymptotically flat.*

The need for rigorous definitions of seemingly clear concepts is obvious in studying black holes because the very existence of these objects modifies the structure of spacetime and its global properties in a fundamental way as compared with flat Minkowski spacetime. For example, the Schwarzschild spacetime contains a singularity, and some geodesics do not extend to infinity. Note that one such geodesic is the circular orbit of a light ray at $r = 1.5\, r_g$; this geodesic lies completely outside the black hole. The formation of black holes, their dynamic interaction, and their merging may produce especially complicated situations. Semi-intuitive, visually plausible arguments are clearly inadequate here.

A rigorous definition of asymptotically flat spaces was suggested by Penrose (1963). The following line of reasoning can be used to arrive at this definition.

Let us consider first the structure of a flat Minkowski spacetime at infinity. To do this, we use an approach typical in geometry: we perform a conformal transformation that brings infinitely remote points to a finite distance. First, we transform the ordinary spherical coordinates (t, r, θ, ϕ) in spacetime M to new coordinates $(\psi, \xi, \theta, \phi)$, using the following relations:

$$t + r = \tan \frac{1}{2}(\psi + \xi), \quad t - r = \tan \frac{1}{2}(\psi - \xi), \tag{5.1.1}$$

$$-\pi \leq \psi - \xi \leq \psi + \xi \leq \pi. \tag{5.1.2}$$

Then the interval ds^2 takes the form

$$ds^2 = \Omega^{-2} d\tilde{s}^2, \quad d\tilde{s}^2 = -d\psi^2 + d\xi^2 + \sin^2\xi\, d\omega^2, \tag{5.1.3}$$

where

$$\Omega = 2\cos\frac{1}{2}(\psi + \xi)\cos\frac{1}{2}(\psi - \xi), \quad d\omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2. \tag{5.1.4}$$

Points at infinity in Minkowski spacetime correspond to $\psi + \xi$ and $\psi - \xi$ assuming the values $\pm\pi$. At these values the metric ds^2 becomes meaningless, but the metric $d\tilde{s}^2$ conformal to ds^2 is regular.[1] When studying the conformal structure on a manifold with a boundary, (5.1.2), we thereby study the conformal structure of Minkowski spacetime, including the infinitely remote points. Recall that it is the conformal structure that is important for studying the general properties of spacetime because it determines the causal properties of the neighborhood of a point, including the properties of null cones.

The $d\tilde{s}^2$ metric is the metric of the four-dimensional cylinder $S^3 \times R^1$ (Figure 5.1a). Inequalities (5.1.2) cut a region corresponding to M on the cylinder; it is hatched in Figure 5.1a. (Of course, we can show only two of the four coordinates; θ and ϕ are omitted.) For the region M cut out of the cylinder, dissected at the point I^0 and developed on a plane, the result is shown in Figure 5.1b. This is the form typically chosen to represent the conformal Minkowski world. This is the so-called *Penrose-Carter conformal diagram* for M. Recall that the left- and right-hand points I^0 coincide, that is, must be "glued together".

All timelike curves in the Penrose-Carter conformal diagrams begin at the point I^- and end at the point I^+, and all spatial sections pass through I^0. Therefore, I^- is said to be the *past timelike infinity*; I^+ is the *future timelike infinity*, and I^0 is the *spatial infinity*.

[1] The metric $d\tilde{s}^2$ has removable coordinate singularities at $\xi = 0$ and $\xi = \pi$.

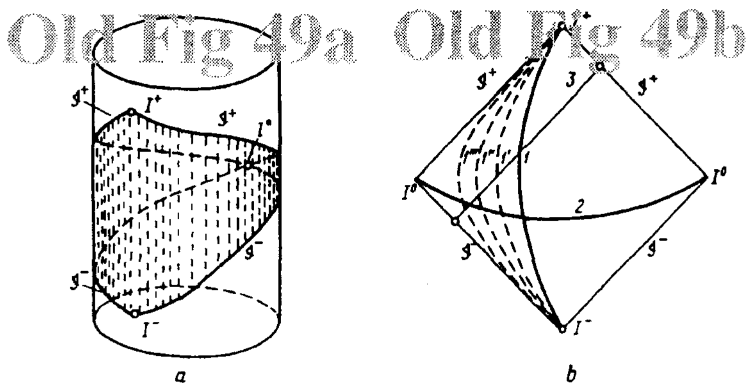


Figure 5.1: (a) Conformal structure of Minkowski spacetime. (b) Penrose-Carter conformal diagram of Minkowski spacetime. The diagram shows timelike (1, 1′, ...), spacelike (2), and null (3) geodesics.

All null geodesics in M begin at the boundary $\mathcal{J}^-$ (at the future null cone of the point I^-) and end at $\mathcal{J}^+$ (at the past null cone of I^+). The boundaries $\mathcal{J}^-$ and $\mathcal{J}^+$ are said to be the *past null infinity* and *future null infinity*, respectively. Clearly, M is bounded by the "infinities" $\mathcal{J}^-$ and $\mathcal{J}^+$ and the points I^-, I^+, I^0.

Penrose-Carter conformal diagrams are especially convenient for studying the global structure of spacetime when the geometry is curved. The coordinates typically used for this purpose represent null rays by straight lines at 45° angles, (e.g., the coordinates ψ and ξ used above possess this property). The causal structure dictated by the arrangement of local null cones is especially clear in such coordinates. Obviously, two-dimensional Penrose-Carter conformal diagrams map the geometry of certain two-dimensional sections of spacetime.

Let us now return to the problem of infinitely distant observers. The world lines of such observers which are at rest at the points r, r', r'', r''' $(r < r' < r'' < r''')$ are shown in Figure 5.1b by the lines 1, 1′, 1″, 1‴, respectively. The greater r, the closer to $\mathcal{J}^-$ and $\mathcal{J}^+$ the corresponding line is. As $r \to \infty$, it tends to $\mathcal{J}^-$ and $\mathcal{J}^+$. It is thus logical to refer to $\mathcal{J}^-$ and $\mathcal{J}^+$ as infinitely distant boundaries of M. Note that the factor Ω in (5.1.3)–(5.1.4), which performs the conformal transformation, vanishes at $\mathcal{J} \equiv \mathcal{J}^- \cup \mathcal{J}^+$, and its gradient $\partial\Omega/\partial x^\mu|_{\mathcal{J}} \neq 0$ is a null vector tangent to the generators of the surface $\mathcal{J}$.

Sometimes it is more convenient to study the regions close to $\mathcal{J}$, using coordinates other than (5.1.1)–(5.1.2). Note that the retarded null coordinate $u = t - r$ can be used to write the interval in Minkowski world:

$$ds^2 = -\,du^2 - 2\,du\,dr + r^2 d\omega^2\,. \tag{5.1.5}$$

The transformation $\rho = r^{-1}$ then gives the following conformal form of the metric

$$ds^2 = \Omega^{-2}\, d\tilde{s}^2, \qquad d\tilde{s}^2 = -\,\rho^2\, du^2 + 2\, du\, d\rho + d\omega^2\,, \tag{5.1.6}$$

where the conformal factor Ω is

$$\Omega = \rho = r^{-1}\,. \tag{5.1.7}$$

In these coordinates the surface $\mathcal{J}^+$ is described by the equation $\rho = 0$. A point on $\mathcal{J}^+$ with coordinates u_0, θ_0, ϕ_0 corresponds to the "end point" $r \to \infty$ of the outgoing null ray $u = u_0$, $\theta = \theta_0$, $\phi = \phi_0$. Likewise, $\mathcal{J}^-$ can be described by replacing u with the advanced null coordinate $v = t + r$.

5.1.2 Definition and properties of asymptotically flat spacetime

Assuming that the properties of asymptotic flat spaces in the neighborhood of "infinity" must be similar to those of Minkowski space, Penrose (1963, 1964, 1965b, 1968) suggested the following definitions. The so-called *asymptotically simple* spacetimes are defined first.

A spacetime M with metric $g_{\mu\nu}$ is said to be *asymptotically simple* if there exists another ("unphysical") space $\widetilde{M}$ with boundary $\partial\widetilde{M} \equiv \mathcal{J}$ and a regular metric $\tilde{g}_{\mu\nu}$ on it such that:

1. $\widetilde{M} \setminus \partial\widetilde{M}$ is conformal to M, and $g_{\mu\nu} = \Omega^{-2}\, \tilde{g}_{\mu\nu}$ in M

2. $\Omega|_M > 0$, $\Omega|_{\partial\widetilde{M}} = 0$, $\Omega_{,\mu}|_{\partial\widetilde{M}} \neq 0$

3. Each null geodesic in M begins and ends on $\partial\widetilde{M}$

We call $\widetilde{M}$ the *conformal Penrose space.* Penrose proved that if the metric $g_{\mu\nu}$ satisfies Einstein's vacuum equations in the neighborhood of $\mathcal{J}$ (or Einstein's equations with the energy-momentum tensor that decreases at infinity sufficiently fast) and the natural conditions of causality and spacetime orientability are satisfied, then an asymptotically simple space has the following properties:

1. The topology of the space M is R^4; its boundary $\mathcal{J}$ is lightlike and consists of two unconnected components $\mathcal{J} = \mathcal{J}^+ \cup \mathcal{J}^-$, each with topology $S^2 \times R^1$.

2. The generators of the surfaces $\mathcal{J}^\pm$ are the null geodesics in $\widetilde{M}$; tangent vectors to these geodesics coincide with $\tilde{g}^{\mu\nu}\, \Omega_{,\mu}|_{\mathcal{J}}$.

3. The curvature tensor in the physical space M decreases as we move along the null geodesic to infinity, and the so-called *peeling off property* holds. Here we do not focus on this property as its detailed analysis can be found in the literature [see, e.g., Sachs (1964), Penrose (1968)].

Property 1 signifies that the global structure of the asymptotically flat space is the same as that of Minkowski space. For example, it has a similar causal structure and does not "allow" black holes. In order to take into consideration the possibility of the existence of localized regions of strong gravitational fields which do not alter the asymptotic properties (as $r \to \infty$) of spacetime, it is sufficient to analyze the class of spaces that can be converted into asymptotically simple spaces by "cutting out" certain inner regions with singularities of some kind (due to the strong gravitational field) and by subsequent smooth "patching" of the resultant "holes". Such spaces are said to be *weakly asymptotically simple.*

To be more rigorous, a space M is said to be weakly asymptotically simple if there exists an asymptotically simple space $\widetilde{M}$ such that for an open subset of $\widetilde{M}$, $K(\partial \widetilde{M} \subset K)$, the space $\widetilde{M} \cap K$ is isometric to a subset of M. One says that a weakly asymptotically simple space is *asymptotically flat* if its metric in the neighborhood of $\mathcal{J}$ satisfies Einstein's vacuum equations (or Einstein's equations with an energy-momentum tensor that decreases sufficiently fast).

5.1.3 Penrose-Carter conformal diagrams

The Schwarzschild spacetime (2.2.1) is asymptotically flat. It is a vacuum solution. In order to relate it with a weakly asymptotically simple space, one can cut out the spacetime region of the interior of the sphere of radius $R > r_g$ which contains the black hole and glue the resulting space with a regular internal solution for a spherical matter distribution which has the same mass as the mass of the black hole. This construction implies that we really have an asymptotically flat spacetime.

It is evident that the structure of the Schwarzschild space infinity is similar to the structure of the infinity in flat space. For this reason, one can expect that there must exist coordinate and conformal transformations similar to (5.1.3)–(5.1.4) that bring infinitely removed points to a finite distance. To construct the corresponding transformations consider the Kruskal metric (2.7.17)

$$ds^2 = -\frac{4r_g^3}{r} e^{-(r/r_g-1)} dV dU + r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{5.1.8}$$

$$-UV = \left(\frac{r}{r_g} - 1\right) \exp\left(\frac{r}{r_g} - 1\right). \tag{5.1.9}$$

Let us introduce new coordinates (ς, η), connected with (U, V) by the relations

$$U = \sinh[\tan(\varsigma)], \quad V = \sinh[\tan(\eta)]. \tag{5.1.10}$$

In these coordinates the infinities $U = \pm\infty$ correspond to the finite values $\varsigma = \pm\pi/2$, and the infinities $V = \pm\infty$ correspond to the finite values $\eta = \pm\pi/2$ (see Figure 5.2). The event horizon H^+ $(U = 0)$ is given by the equation $\varsigma = 0$, while the past horizon

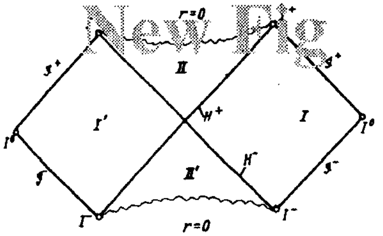


Figure 5.2: Penrose-Carter conformal diagram for the Schwarzschild "eternal" black and white holes.

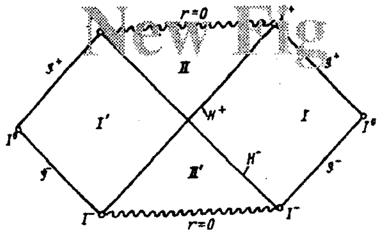


Figure 5.3: "Canonical form" of the Penrose-Carter conformal diagram. On this diagram the lines representing the singularities $r = 0$ are straightened by the additional coordinate transformation given in the text.

H^- $(V = 0)$ is given by $\eta = 0$. Near the horizons one has $U \approx \varsigma$ and $V \approx \eta$. In the coordinates (ς, η) the Kruskal metric (5.1.8) takes the form

$$ds^2 = r^2\, d\tilde{s}^2\,, \tag{5.1.11}$$

$$d\tilde{s}^2 = \frac{4(r/r_g - 1)}{(r/r_g)^3 \tanh[\tan(\varsigma)]\, \tanh[\tan(\eta)]}\, \frac{d\varsigma\, d\eta}{\cos^2\varsigma\, \cos^2\eta} + d\omega^2\,. \tag{5.1.12}$$

The boundary points $(-\pi/2 < \varsigma < 0,\ \eta = \pi/2)$ and $(-\pi/2 < \eta < 0,\ \varsigma = \pi/2)$ correspond to the future null infinities $\mathcal{J}^+$. The boundary points $(0 < \varsigma < \pi/2,\ \eta = -\pi/2)$ and $(0 < \eta < \pi/2,\ \varsigma = -\pi/2)$ correspond to the past null infinities $\mathcal{J}^-$. One can verify that the conformal metric (5.1.12) remains regular at the points of $\mathcal{J}^-$ and $\mathcal{J}^+$. In the coordinates (ς, η) the singularities $r = 0$ are described by the following implicit equation

$$\sinh[\tan(\varsigma)]\, \sinh[\tan(\eta)] = e^{-1}. \tag{5.1.13}$$

Certainly the coordinate and conformal transformations with the required properties are not unique. For example, one can make the transformation $\varsigma \to \tilde{\varsigma} = f(\varsigma)$, $\eta \to \tilde{\eta} = f(\eta)$, where f is a regular monotonic function on the interval $(-\pi/2, \pi/2)$. One can use this ambiguity to straighten the lines representing the singularity $r = 0$. Let $\eta = g(\varsigma)$ be a solution of equation (5.1.13). It evidently possesses the property $\varsigma = g(\eta)$. The coordinate transformation which straightens the singularity is given by the function $f(x) = (1/2)[g(x) - x] + \pi/4$. The corresponding "canonical form" of the Penrose-Carter conformal diagram with the straightened singularity lines is shown in Figure 5.3.

Quite often in the literature, when discussing the causal structure of spacetime, one makes only a coordinate transformation bringing coordinate infinities to a finite "distance", and does not consider a subsequent conformal transformation. For example, one uses the following coordinate transformations [see e.g., Misner,Thorne, and Wheeler (1973), Hawking and Ellis (1973)]

$$U = \tan x\,, \qquad V = \tan y\,. \tag{5.1.14}$$

The complete spacetime after this transformation will again be represented by the domain similar to that shown in Figure 5.3. However, when working in these coordinates, one must be careful because these coordinates are not regular at $\mathcal{J}^{\pm}$.[2]

The Penrose-Carter conformal diagram for a rotating (Kerr) black hole and its charged (Kerr-Newman) generalization are more complicated. They will be discussed in Section 6.6.

5.1.4 Bondi-Metzner-Sachs group of asymptotic symmetries

We return now to the discussion of the main properties of asymptotically flat spacetimes. The decrease of the curvature tensor at infinity (see property 3 at page 153) signifies that the effects due to curvature are negligible in asymptotically flat spaces in the neighborhood of $\mathcal{J}$, and the spacetime itself near $\mathcal{J}$ is nearly flat. Thus, the ordinary law of energy-momentum conservation holds in this region with good accuracy, and the motion of test particles can be treated as approximately uniform and on a straight line. Correspondingly, the group of asymptotic symmetries can be defined in asymptotically flat spaces.

For this purpose, consider first the Poincaré transformations of Minkowski space which in Cartesian coordinates X^μ have the form

$$X^\mu \to X'^\mu = \Lambda^\mu{}_\nu X^\nu + a^\mu\,, \tag{5.1.15}$$

where $\Lambda^\mu{}_\nu$ is the Lorentz transformation matrix, and a^μ is the translation vector corresponding to a shift of the origin of the frame. Now we introduce in Minkowski space retarded (u, r, θ, ϕ) or advanced (v, r, θ, ϕ) coordinates and denote by w either the retarded (u) or advanced (v) time coordinate. Then transformation (5.1.15) corresponds to the following transformation of the coordinates $(w,\ r,\ \theta,\ \phi)$:

$$w' = w'(w, r, \theta, \phi)\,, \qquad r' = r'(w, r, \theta, \phi)\,,$$

$$\theta' = \theta'(w, r, \theta, \phi)\,, \qquad \phi' = \phi'(w, r, \theta, \phi)\,. \tag{5.1.16}$$

As $r \to \infty$, the functions describing this transformation take on a simple form. Thus, the following transformations correspond in this limit to shifts $(\Lambda^\mu{}_\nu = \delta^\mu_\nu)$ in physical spacetime:

$$w' = w + a_0 \pm a_1 \sin\theta\cos\phi \pm a_2 \sin\theta\sin\phi \pm a_3 \cos\theta\,,$$

[2]The authors are grateful to Don Page for discussion of this point.

$$\theta' = \theta, \qquad \phi' = \phi. \tag{5.1.17}$$

In these relations signs + and − must be taken for $w = u$ and $w = v$, respectively. Formulas (5.1.17) describe shifting the surface $\mathcal{J}$ along its generators.

This result can be described more formally in a way that admits a natural generalization to the case of arbitrary asymptotically flat spaces. Let ξ^μ be a Killing vector field corresponding to the symmetry transformation in the physical space M ($\nabla^{(\nu}\xi^{\mu)} = 0$); hence, ξ^μ in the conformal space $\widetilde{M}$ obeys the equation

$$\widetilde{\nabla}^{(\mu}\xi^{\nu)} - \frac{\xi^\alpha \widetilde{\nabla}_\alpha \Omega}{\Omega}\, \tilde{g}^{\mu\nu} = 0, \tag{5.1.18}$$

where $\widetilde{\nabla}^\mu$ is the covariant derivative in the metric $\tilde{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}$.

In the general case, equation (5.1.18) has no non-trivial solutions if the spacetime M does not allow exact isometries. For example, this is true for general type asymptotically flat spaces. However, if the analysis is restricted to the neighborhood of $\mathcal{J}$ and we stipulate that equation (5.1.18) is satisfied only on $\mathcal{J}$, it again has non-trivial solutions. These solutions determine vector fields that generate the transformations of asymptotic symmetries. Note the spectacular fact that the group corresponding to these transformations is independent of the choice of a specific representative of the class of asymptotically flat spaces. This group is known as the *Bondi-Metzner-Sachs group* (the BMS-group). A detailed presentation of its properties and a description of its representations can be found in papers by Sachs (1962), Bondi, van der Burg, and Metzner (1962), Penrose (1964), McCarthy (1972a,b, 1973), McCarthy and Crampin (1973), Volovich *et al.* (1978). Here we only briefly outline the main properties of this group which are important for the presentation that follows.

The BMS-group is infinite-dimensional; hence, it is much wider than the Poincaré group, which exactly preserves the metric of a flat space. This happens because the BMS-group preserves only the asymptotic form of the metric and the gravitational field decreases slowly at infinity. An important property of the BMS-group is that it contains a uniquely identifiable normal four-dimensional subgroup of translations. The action of this subgroup on $\mathcal{J}$ in Minkowski space coincides with that of (5.1.17). In the general case, it is possible to introduce in the neighborhood of $\mathcal{J}$, in asymptotically flat spaces, coordinates in which the transformations of the translation subgroup have the form (5.1.17). Such coordinates are known as *conformal Bondi coordinates* [see Tamburino and Winicour (1966)].

We have described a class of asymptotically flat spaces whose asymptotic behavior is similar to the asymptotic behavior of Minkowski space, and briefly outlined its properties. The concept of an asymptotically distant observer, who moves almost inertially, can be very naturally introduced in this class of spaces. Now we can give a rigorous definition of the black hole. But first, we will briefly treat one further aspect, connected with the scattering of massless fields in asymptotically flat spaces, as it will be useful in subsequent chapters.

5.1.5 Massless fields in asymptotically flat spacetime

The above definition of asymptotically flat space is especially convenient for dealing with the problem of the scattering of massless fields and, in particular, for describing the properties of gravitational radiation. The universal asymptotic behavior ($\propto 1/r$) of these fields in the wave zone allows the use of conformal mapping for transforming the problem of scattering in the physical spacetime into a problem with regular initial data at the past null boundary of $\mathcal{J}^-$ in the conformal Penrose space. It is then found that the regular behavior of the conformally transformed field in the neighborhood of $\mathcal{J}$ implies a specific form of decrease of this field in the asymptotic region.

We will illustrate these arguments with an example of a conformally invariant massless scalar field described by the equation

$$(\Box - \frac{1}{6} R)\,\varphi = 0 \tag{5.1.19}$$

in the asymptotically flat space $\{M, g\}$. Now we use the conformal mapping $g_{\mu\nu} = \Omega^{-2}\,\tilde{g}_{\mu\nu}$ in order to transform $\{M, g\}$ to the conformal Penrose space $\{\widetilde{M}, \tilde{g}, \Omega\}$, supplementing it with a conformal mapping of the field, $\varphi \to \widetilde{\varphi} = \Omega^{-1}\,\varphi$. We refer to the values of the field $\widetilde{\varphi}$ at past null infinity $\mathcal{J}^-$ and future null infinity $\mathcal{J}^+$ as the *images of the field* φ at $\mathcal{J}^-$ and $\mathcal{J}^+$ and denote them by an appropriate capital letter:

$$\widetilde{\varphi}\,|_{\mathcal{J}^\pm} = \Phi^\pm\,. \tag{5.1.20}$$

The field $\widetilde{\varphi}$ in the Penrose space satisfies the equation

$$\left(\widetilde{\Box} - \frac{1}{6}\widetilde{R}\right)\widetilde{\varphi} = 0\,, \tag{5.1.21}$$

where $\widetilde{\Box} = \tilde{g}^{\alpha\beta}\,\widetilde{\nabla}_\alpha\,\widetilde{\nabla}_\beta$, and $\widetilde{R}$ is the scalar curvature of the metric $\tilde{g}_{\alpha\beta}$. Having fixed the image Φ^- of the field $\widetilde{\varphi}$ in the asymptotically simple space, one can find $\widetilde{\varphi}$ by solving equation (5.1.21) with initial data on the regular null surface $\mathcal{J}^-$, whereby Φ^+ is determined. In other words, there is a one-to-one correspondence

$$\Phi^- \leftrightarrow \varphi \leftrightarrow \Phi^+ \tag{5.1.22}$$

between the field φ and its images at $\mathcal{J}^-$ and $\mathcal{J}^+$.

The condition of asymptotic regularity, that is, the condition of the existence of regular images at $\mathcal{J}$, plays the role of the radiation condition, and the problem of classical scattering theory can be formulated as the problem of finding the image of a solution φ at $\mathcal{J}^+$, having a specified image at $\mathcal{J}^-$. Note that the asymptotically regular field in the asymptotic region (close to $\mathcal{J}$) is of the form

$$\varphi \sim \Phi\,\Omega\,. \tag{5.1.23}$$

In Minkowski spaces in (5.1.6) coordinates, the resulting asymptotic behavior in the wave zone is

$$\varphi \sim \frac{\Phi^{+}(u,\theta,\phi)}{r}. \tag{5.1.24}$$

The method presented above is readily generalized to other massless fields [see e.g., Penrose (1965b, 1968), Pirani (1964), Frolov (1979, 1986)].

As the asymptotically flat space has a group of asymptotic symmetries, it is possible to determine characteristics of massless fields such as energy and momentum of the incoming or outgoing flux. Let $\xi^{\nu}_{(a)}$ $(a = 0,1,2,3)$ be the generators of the translation subgroup of the BMS-group; $\xi^{\nu}_{(a)}$ acting on $\mathcal{J}^{\pm}$. The expression for the energy $(a = 0)$ and momentum $(a = 1,2,3)$ of the incoming (outgoing) radiation is

$$P^{a}_{\mathcal{J}^{\pm}} = -\int_{\mathcal{J}^{\pm}} \widetilde{T}_{\mu\nu}\, \xi^{\nu}_{(a)}\, \tilde{g}^{\mu\alpha}\, d\sigma_{\alpha}\,, \tag{5.1.25}$$

where $\widetilde{T}_{\mu\nu} = \Omega^{-2}\, T_{\mu\nu}$ is the metric energy-momentum tensor of the field of interest. It can be readily shown that $P^{a}_{\mathcal{J}^{\pm}}$ for asymptotically regular fields in a flat spacetime coincides with the total energy-momentum of the system provided we define it in terms of vector Killing fields corresponding to translations. In the general case, expression (5.1.25) for asymptotically flat space can be written in terms of massless field images at $\mathcal{J}^{\pm}$. Thus, we have for the scalar field satisfying equation (5.1.19)

$$P^{a}_{\mathcal{J}} = \int_{\mathcal{J}} N_a \left[(\partial_u \Phi)^2 - \frac{1}{6}\, \partial_u^2 (\Phi^2) \right] du\, d\omega\,, \tag{5.1.26}$$

where

$$N_a = (1,\, \sin\theta\cos\phi,\, \sin\theta\sin\phi,\, \cos\theta\,)\,, \qquad d\omega = \sin\theta\, d\theta\, d\phi\,. \tag{5.1.27}$$

Note that for wavepacket-type fields φ with finite energy, the value of $|\,\partial_u \Phi^2\,|$ decreases as $|\,u\,| \to \infty$ so that the second term can be dropped

$$P^{a}_{\mathcal{J}} = \int_{\mathcal{J}} N_a (\partial_u \Phi)^2\, du\, d\omega\,. \tag{5.1.28}$$

Expressions for the energy-momentum of the incoming and outgoing fluxes of other massless fields in terms of field images at $\mathcal{J}^{\pm}$ can be written in a similar manner [see, e.g., Frolov (1986)].

5.2 Event Horizon and its Properties

5.2.1 Event horizon

Now we can give a rigorous definition of the concept of a black hole. A *black hole* in asymptotically flat spacetime is defined as a region such that no causal signal (i.e., a signal propagating at a velocity not greater than that of light) from it can reach $\mathcal{J}^{+}$.

The causal curve describing the propagation of such a signal is either a smooth timelike or a null curve (a vector u^μ tangent to it has the property $u_\mu u^\mu \le 0$) or consists of segments of such curves. Let us define the *causal past* $J^-(Q)$ of a set Q as the set of points with the following property: for each point there is a future-directed casual curve connecting it with one of the points of Q. The set of points visible to a distant observer coincides with $J^-(\mathcal{J}^+)$. The boundary[3] of this set, $\dot{J}^-(\mathcal{J}^+)$, denoted by H^+, is called the *event horizon*. The event horizon constitutes the boundary of the black hole.

Obviously, not just one but several black holes may exist in a bounded region of spacetime; new black holes may be created, and those already existing may interact and merge. In such cases, $\dot{J}^-(\mathcal{J}^+)$ is the union of the boundaries of all black holes. If an asymptotically flat space contains no horizon, then all events taking place in this space can be recorded, after an appropriate time, by a distant observer. If an event horizon appeared, it means that a black hole has been born, and the strongly enhanced gravitational field has produced qualitative changes in the causal structure of spacetime. The enhanced gravitational field prevents signals from escaping so that the observer can never find out about the events inside the black hole unless he chooses to cross the event horizon and fall into it.

Figures 5.4a and 5.4b represent a spherically symmetric black hole produced by the collapse of a spherical star. We already know that this is the simplest type of black hole. Figure 5.4a plots the spacetime of this in Eddington-Finkelstein coordinates. Figure 5.4b is the Penrose-Carter conformal diagram for the corresponding spacetime which can be obtained from the Penrose-Carter conformal diagram of the total spacetime of the eternal black hole, shown in Figure 5.4c, by "cutting" it along line 1 corresponding to the motion of the surface of the collapsing body, and then "gluing" onto the left the part of the Penrose-Carter conformal diagram that describes the metric inside the collapsing body. The last figure shows that the gravitational-radius surface outside the collapsing body coincides with H^+; the region inside H^+ is the black hole. The structure of the infinity of this spacetime (Figure 5.4b) is identical to that of the infinity of Minkowski spacetime. Note also that the region from which light rays cannot escape to infinity (i.e., the black hole) arises not at the moment τ_1 when the star contracts to a size equal to its gravitational radius, but at an earlier moment τ_0. The event horizon H^+ is formed by signals that, having left the center of the star at a moment τ_0, propagate at the speed of light and reach the surface of the star precisely when it collapses to its gravitational radius.

Consider a *Cauchy surface* Σ, that is, a spacelike hypersurface with global extent sufficient to specify a unique evolution by means of initial Cauchy data. The intersection of the event horizon with a Cauchy surface $\Sigma(\tau)$, whose equation is of the type $\tau(x) = \text{const}$, consists in general of a set of closed two-dimensional surfaces $\partial\mathcal{B}_i(\tau)$ $(i = 1, \ldots, n)$ which can be regarded as boundaries of the black holes that exist at

[3]The boundary $\dot{Q}$ of a set Q is defined as $\dot{Q} = \overline{Q} \cap \overline{(M - Q)}$, where M is the complete spacetime, and $\overline{Q}$ is the closure of Q.

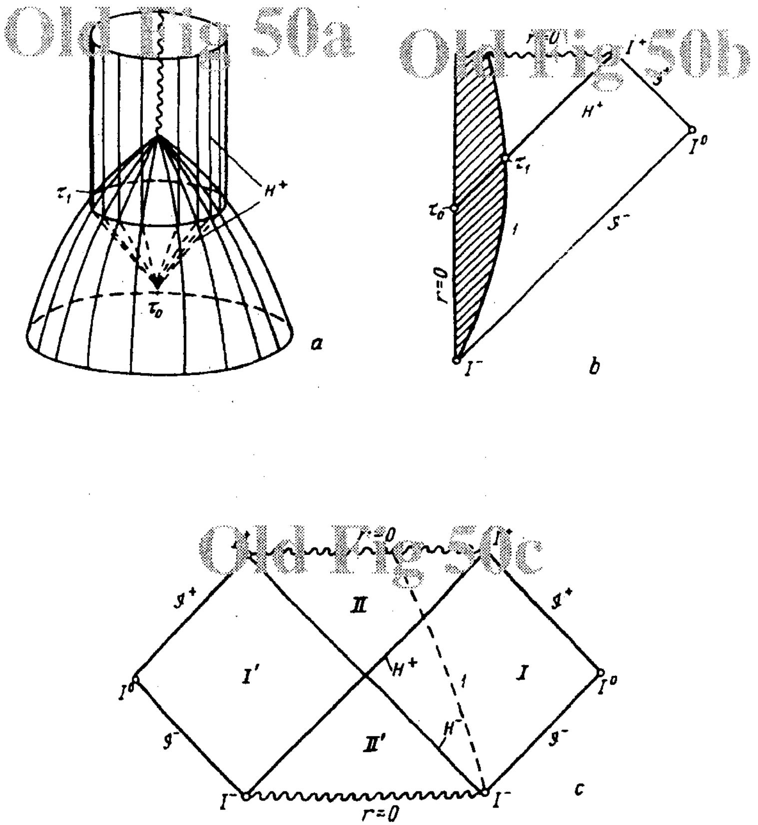


Figure 5.4: (a) Spacetime of a spherically symmetric black hole in Finkelstein coordinates. (b) Penrose-Carter conformal diagram for the spacetime of a spherically symmetric black hole. Line l plots the motion of the surface of the collapsing body. The conformal factor is chosen to have line $r = 0$ (the center of the body) vertical. (c) Penrose-Carter conformal diagram for "eternal" black and white holes. Line 1 is the world line of the test particle.

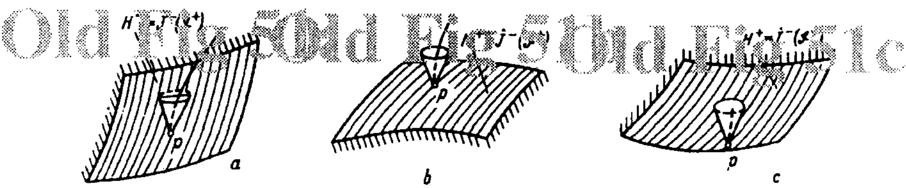


Figure 5.5: The regular part of the event horizon is a null surface. The assumption that the event horizon contains a segment where it is (a) timelike, or (b, c) spacelike produces a contradiction. The region inside the event horizon is hatched.

time τ. The part of $\Sigma(\tau)$ which is bounded by $\partial\mathcal{B}_i(\tau)$ is called the black hole $\mathcal{B}_i(\tau)$ at the given moment of time. The number of black holes can be time-dependent.

As we see from Figure 5.4a, the surface of the event horizon is not smooth. For example, there is an apex point at the moment of time τ_0 when the event horizon is created. Irregular points can also arise on the event horizon after the fall of matter into the hole. The surface of the event horizon beyond these irregular points is lightlike.

It is easy to arrive to a contradiction by assuming that a regular part of the event horizon is either timelike or spacelike. To show this, let us analyze the situations presented in Figure 5.5. First, assume that the surface of the event horizon in a neighborhood of some regular point p of the horizon is timelike (Figure 5.5a). A contradiction is obtained if we consider the behavior of light rays in a small neighborhood of the point p. Because the structure of spacetime is locally the same as that of Minkowski spacetime, the light rays emitted from p form a future directed local null cone. The timelike part of the horizon divides this cone into two parts: the rays go partly into the black hole (to the left of H^+) and partly into the space outside it (to the right of H^+). However, the points of the external space are inside $J^-(\mathcal{J}^+)$; hence, a causal signal can be sent from p and reach $\mathcal{J}^+$. Obviously, this is also possible for points to the left of H^+ that are sufficiently close to p. But this conclusion contradicts the assumption that these points lie inside the black hole.

Assume now that the event horizon has a regular spacelike part. If the set $J^-(\mathcal{J}^+)$ lies to the future of this part of H^+ (Figure 5.5b), then the future-directed light rays emitted at a point p or its small neighborhood inside H^+ can reach $J^-(\mathcal{J}^+)$ and are thus visible to a distant observer. This conclusion is in contradiction with the assumption that the region inside H^+ is a black hole. If the set $J^-(\mathcal{J}^+)$ lies toward the past from the part of H^+ that we consider (Figure 5.5c), then a point p exists in a small neighborhood of this part of H^+ such that the future-directed null cone with its apex at this point crosses H^+. This means that the event p cannot be seen by a distant observer, and p cannot belong to $J^-(\mathcal{J}^+)$, which contradicts the original assumption.

Even though these arguments are not absolutely rigorous, they show why the regular segments of the event horizon are lightlike surfaces.

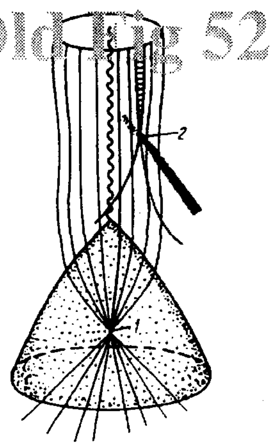


Figure 5.6: Illustration of the Penrose theorem on the structure of the event horizon.

5.2.2 Penrose theorem

A more detailed description of the structure of the event horizon is provided by a theorem proved by Penrose (1968). According to this **theorem**, *the event horizon is formed by null geodesics (generators) that have no end points in the future.*

If we monitor the behavior of any generator in the future, we find that it never leaves the horizon H^+ and never intersects another generator. Two alternatives are possible in the past-directed motion along a chosen generator: either it always lies on H^+, or this light ray leaves H^+ at some point. The null generator enters H^+ at this point having left the region $J^-(\mathcal{J}^+)$ outside the black hole. The entry point of this generator at H^+ is the point of its intersection with other generators (the *caustics*). Only one generator passes through each point of the horizon beyond the caustics.

Figure 5.6 illustrates the Penrose theorem. In this figure, new pencils of rays enter H^+ at caustic 2 when matter falls into the black hole. Caustic 1 corresponds to the point at which the event horizon first appears. Caustics also appear when gravitational radiation is incident on the black hole, or when two or more black holes merge. General properties of such caustics are described in [Friedrich and Stewart (1983)]. (For more details concerning the structure of the event horizon for colliding black holes see Section 7.7.2.)

The situation shown in Figure 5.7, in which an isolated black hole subjected to an external influence breaks into two (or more), is impossible. Indeed, in this process two (or more) points of the black hole surface which are separated at the initial moment τ_0 move closer and intersect at the moment τ_1 when a bifurcation is formed. This means that at least two generators of the event horizon intersect when extended to the future. This conclusion is in contradiction with the Penrose theorem. The

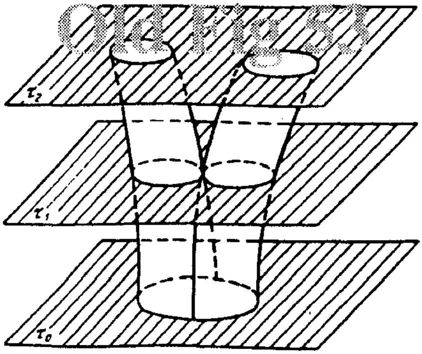


Figure 5.7: A black hole cannot break in two or more black holes. The situation shown in the figure (an isolated black hole breaking in two) is impossible.

proposition that a black hole cannot split or vanish can be proved rigorously. The proof can be found in the monograph by Hawking and Ellis (1973).

5.2.3 Surface topology of black holes

There exist a number of results indicating that the surface topology of black holes must be simple. The corresponding theorems assume asymptotic flatness and a suitable energy condition on the matter field. The formulation of the corresponding energy conditions can be found in the Appendix A.15.

The first of these results was obtained by Hawking (1972a) who proved that in a stationary spacetime the surface topology of a black hole must be a two-sphere (we shall discuss this result and the corresponding assumptions in the Section 6.2). The first theorem concerning the topology of non-stationary black holes was proved by Gannon (1976). This **theorem** claims: *The horizon of the black hole must be either spherical or toroidal, provided the dominant energy condition is satisfied and the horizon is smooth to the future of some slice. The latter condition entails in particular that no new generators enter in the future of that slice.*

An extension and generalization of this result uses the so called *topological censorship*. Consider an asymptotically flat spacetime and denote by γ_0 a timelike curve with past end point at $\mathcal{J}^-$ and future end point in $\mathcal{J}^+$ that lies in a simply connected neighborhood of $\mathcal{J}^- \cup \mathcal{J}^+$. The theorem proved by Friedman, Schleich, and Witt (1993) reads:

Topological censorship theorem. *If an asymptotically flat, globally hyperbolic spacetime satisfies the averaged null energy condition, then every causal curve from* $\mathcal{J}^-$ *to* $\mathcal{J}^+$ *can be continuously deformed to* γ_0.

In other words, in such a space any causal curve connecting $\mathcal{J}^-$ and $\mathcal{J}^+$ may be deformed (in the sense of homotopy) into a similar curve which never enters the region where the gravitational field is strong.

Jacobson and Venkataramani (1995) pointed out that a black hole with toroidal surface topology provides a potential mechanism for violating topological censorship. Specifically, a light ray sent from past null infinity to future null infinity and passing through the hole in the torus would not be deformable to a ray which did not come close to the black hole. Thus, topological censorship implies that a toroidal horizon (if it exists) must close up quickly, before a light ray can pass through. The theorem proved by Jacobson and Venkataramani (1995) claims that the cross-section of the event horizon must be topologically a two-sphere, provided the averaged null energy condition is satisfied and the horizon topology persists for long enough. This theorem strengthens the result obtained by Browdy and Galloway (1995) who showed that the surface geometry of a black hole at a given moment of time must be a two-sphere if no new null generators enter the horizon at later times. Their assumptions include a null energy condition.

Recent computer simulations of the collapse of rotating toroidal configurations of collisionless particles to Kerr black holes demonstrate the formation of a *temporarily toroidal horizon* [Abrahams *et al.* (1994), Hughes *et al.* (1994), Shapiro, Teukolsky, and Winicour (1995)]. In complete agreement with the general results, a temporarily toroidal horizon exists only for short period of time, which is not enough for a null ray to pass through the hole of the torus and escape to infinity. In the computational model, the toroidal black hole has a nonsmooth inner rim. At this rim new generators emerge from crossover points. The toroid becomes a sphere. Just prior to this moment the surface of the black hole is not even a (Hausdorff) manifold [Shapiro, Teukolsky, and Winicour (1995)].

5.3 Ehlers-Sachs Theorem

5.3.1 Light wavefront and light rays

The fact that the event horizon bounding a black hole is a null surface at its regular points immediately leads to a number of important consequences. Hence, it is logical to give a more detailed description of the general properties of null surfaces in curved spacetime before starting the analysis of black hole properties.

Consider a surface Γ which in coordinates x^μ is defined by the equation $v(x^\mu) = 0$. Γ is a null surface if the gradient $v_{,\mu}$ is a null vector on Γ, i.e., $g^{\mu\nu}\, v_{,\mu}\, v_{,\nu}\,|_\Gamma = 0$. In general, this relation does not hold outside Γ. But it can be shown [see, e.g., Courant and Hilbert (1962)] that the freedom of choice of v can be so used that the surface Γ can always be included into the one-parameter family of null surfaces Γ_c defined by the equation $v(x) = c$. We assume, with no loss of generality, that this choice was made so that

$$g^{\mu\nu}\, v_{,\mu} v_{,\nu} = 0\,. \tag{5.3.1}$$

If we define $l^\mu = g^{\mu\nu}\, v_{,\nu}$, then (5.3.1) signifies that l^μ is a null vector that is orthogonal to $v_{,\mu}$, and hence is tangent to the surface Γ_c. Furthermore, making use of the property $v_{,\mu;\nu} = v_{,\nu;\mu}$ and of expression (5.3.1), we find $l^\nu\, l_{\mu;\nu} = l^\nu\, l_{\nu;\mu} = \frac{1}{2}\,(l^\nu\, l_\nu)_{,\mu} = 0$. In other words, the integral curves of the vector field l^μ,

$$\frac{dx^\mu}{dr} = l^\mu\,, \tag{5.3.2}$$

are geodesics, and r is the affine parameter.

If the initial point of integral curve (5.3.2) lies on the null surface Γ, the entire curve belongs to this surface, and Γ itself is formed by a two-dimensional family of null geodesics (*generators*). Let r be the affine parameter along the generators, and y^a $(a = 1, 2)$ be continuous coordinates that parameterize these lines. The solution to equation $v = 0$ can then be written in the following parameterized form: $x^\mu = f^\mu(r, y^a)$. In this parameterization

$$l^\mu = \frac{\partial f^\mu}{\partial r}\,, \qquad l_\mu\,\frac{\partial f^\mu}{\partial y^a} = 0\,. \tag{5.3.3}$$

We also have

$$l^\mu\, l_{\nu;\mu} = 0\,, \qquad l^\mu\, l_\mu = 0\,. \tag{5.3.4}$$

From the physical standpoint, the surface Γ (*characteristic*) describes the propagation of a light wavefront, and its generators (*bicharacteristics*) are light rays perpendicular to the front.

5.3.2 Optical scalars. Ehlers-Sachs theorem

If we pick a narrow beam of light rays, the following experiment will provide information on their behavior. Place an opaque object in the path of the beam (perpendicular to it), and place a screen at some distance from the object also perpendicularly to the beam. Then the **theorem** proved by Ehlers and Sachs [Jordan, Ehlers, and Sachs (1961), Sachs (1961)] states that:

All parts of the shadow reach the screen simultaneously. The size, shape, and orientation of the shadow depend only on the position of the screen and are independent of the velocity of the observer. If the screen is at a short distance δr from the object, the expansion and distortion of the shadow are given by the quantities $\theta\,\delta r$ and $|\,\sigma\,|\,\delta r$, where

$$\theta = \frac{1}{2}\, l^\alpha_{\ ;\alpha}\,, \qquad |\,\sigma\,| = \left(\frac{1}{2}\, l_{\alpha;\beta}\, l^{\alpha;\beta} - \theta^2\right)^{1/2}\,. \tag{5.3.5}$$

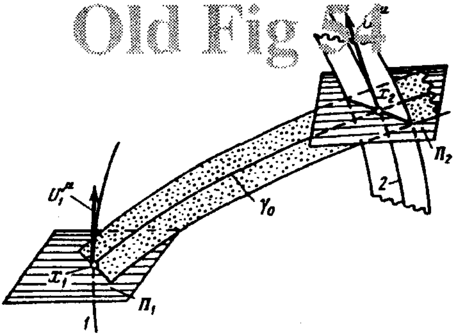


Figure 5.8: Illustration of the Ehlers-Sachs theorem on the propagation of light rays.

Here we reproduce the main steps of the **proof** of this theorem, adapted to the case under consideration. This will enable us to describe in detail a number of important characteristics of null surfaces.[4]

Consider a family (beam) of light rays $x^\mu = f(r, y^a)$, where y^a are parameters enumerating rays ($a = 1, 2$). Denote by l^μ tangent vectors to the rays. Let a light ray γ_0 described by the equation $x^\mu = f^\mu(r, y_0^a)$ intersect the world line 1 of the first observer at a point $x_1^\mu = f^\mu(r_1, y_0^a)$; the four-velocity of the observer at this moment is U_1^μ (Figure 5.8). For the observer, the set Π_1 of the events $x_1^\mu + dx^\mu$ simultaneous with this event x_1 satisfies the condition

$$U_{1\mu}\, dx^\mu = 0\,. \tag{5.3.6}$$

Choose the affine parameter r so as to satisfy the condition $r = r_1$ at the point of intersection of other rays of the beam and Π_1. If we demand that in addition to (5.3.6) the displacements dx^μ satisfy the condition

$$l_\mu\, dx^\mu = 0\,, \tag{5.3.7}$$

then these two conditions jointly define a two-dimensional surface Σ perpendicular (in the reference frame of U_1^μ) to the beam of light rays. Let e_a^α ($a = 1, 2$) be mutually orthogonal unit vectors tangent to Σ and $m^\alpha = 2^{-1/2}(e_1^\alpha + i\, e_2^\alpha)$. This gives

$$m^\alpha\, m_\alpha = \bar{m}^\alpha\, \bar{m}_\alpha = 0\,, \qquad m^\alpha\, \bar{m}_\alpha = 1\,. \tag{5.3.8}$$

Assume now that an object is placed in the path of the beam so that a part of the two-dimensional surface Σ bounded by the curve

$$x^\alpha(\phi) = x_1^\alpha + \zeta_1^\alpha(\phi)\,, \quad \zeta_1^\alpha(\phi) = \zeta(\phi)\, \bar{m}^\alpha + \bar{\zeta}(\phi)\, m^\alpha \tag{5.3.9}$$

[4] For details on congruences of null geodesics in curved spacetime, see, e.g., the reviews by Pirani (1964) and Frolov (1976a) where the relevant references are given.

is opaque to light rays. The rays passing through the boundary of the opaque object form a one-dimensional family parameterized by ϕ. For this family the parameters $y_0^a + \delta y^a(\phi)$ are determined by the equation

$$f^\mu(r_1, y_0^a + \delta y^a(\phi)) = x_1^\alpha + \zeta_1^\alpha(\phi)\,. \tag{5.3.10}$$

A shadow Γ then forms behind the object. Its boundary $\partial\Gamma$ is defined by the equation

$$x^\mu = z^\mu(r,\phi) \equiv f^\mu(r, y_0^a + \delta y^a(\phi))\,. \tag{5.3.11}$$

Let the light ray γ_0 extended to a point $x_2^\mu = f^\mu(r_2, y_0^a)$ intersect there the world line 2 of another observer whose velocity at that moment is U_2^μ (see Figure 5.8). For this observer, the space Π_2 of events simultaneous with x_2^μ is defined by the vectors $dx^\mu = x^\mu - x_2^\mu$ which satisfy the relation $U_{2\mu}\, dx^\mu = 0$. We now use the ambiguity $(r \to r' = A(y^a)(r - r_1) + r_1)$ in choosing the affine parameter in order to arrange that $r = r_2$ for all light rays in the beam at the points of their intersection with Π_2. The relation (5.3.3) implies that a two-dimensional area element $x^\mu = f^\mu(r_2, y_0^a + \delta y^a)$, which describes the position of wavefront at the moment x_2^μ in the reference frame of U_2^μ, is orthogonal to the direction of the light ray whose tangent vector is

$$l_2^\mu = \frac{\partial f^\mu}{\partial r}(r_2, y_0^a)\,. \tag{5.3.12}$$

We have therefore proved that all parts of the shadow simultaneously reach the screen placed perpendicularly to the ray beam.

Consider the boundary $z^\mu(r_2,\phi)$ of the shadow projected on the screen by the opaque object. Denote by $\delta x^\mu(\phi) = z^\mu(r_2,\phi) - x_2^\mu$ a vector connecting the center x_2 of the screen and the boundary. The size, shape, and orientation of the image on the screen are uniquely determined if the scalar products $\delta x^\mu(\phi_1)\cdot\delta x_\mu(\phi_2)$ are known for an arbitrary pair of points on its boundary. It is not difficult to verify that the vectors $\delta x^\mu(\phi)$ are transformed in the transition to a different reference frame $U_2'^\mu$ by the formula $\delta' x^\mu(\phi) = \delta x^\mu(\phi) + a l^\mu$, while the scalar products remain unchanged:

$$\delta x^\mu(\phi_1)\cdot\delta x_\mu(\phi_2) = \delta' x^\mu(\phi_1)\cdot\delta' x_\mu(\phi_2)\,. \tag{5.3.13}$$

We have thus demonstrated that the size, shape, and orientation of the shadow are independent of the observer's velocity.

For the sake of convenience, we take for U_2^μ the vector obtained from U_1^μ by its parallel transport along γ_0 since the characteristics of the image on the screen are independent of the choice of observer. Denote by $m^\alpha(r)$ the vector obtained by the parallel transport of m^α along γ_0 and $m_2^\alpha = m^\alpha(r_2)$. Parallel transport preserves the orthogonality of m^α to the vectors l^μ and U^μ; hence, m_2^α and $\bar{m}_2^\alpha$ span the two-dimensional element of the area of the screen orthogonal to the light beam (in the reference frame U_2^μ). Note that the vectors $\zeta^\alpha(r,\phi) = \zeta(\phi)\,\bar{m}^\alpha(r) + \bar{\zeta}(\phi)\,m^\alpha(r)$ are also parallelly propagated so that

$$l^\beta \zeta^\alpha_{\ ;\beta} = 0\,. \tag{5.3.14}$$

Define $\zeta'^{\alpha}(r,\phi) = z^{\alpha}(r,\phi) - f^{\alpha}(r, y_0^a)$, then for vectors on the screen located on the shadow boundary we have

$$\zeta'^{\alpha}(r_2,\phi) = \zeta'(\phi)\bar{m}_2^{\alpha} + \bar{\zeta}'(\phi) m_2^{\alpha} . \tag{5.3.15}$$

The vectors $\zeta'^{\alpha}(r,\phi)$ are "frozen" into the ray congruence, and hence they are Lie-propagated[5] along γ_0

$$\mathcal{L}_l \zeta'^{\alpha} = l^{\beta}\,\zeta'^{\alpha}_{\ ;\beta} - \zeta'^{\beta}\,l^{\alpha}_{\ ;\beta} = 0 . \tag{5.3.16}$$

Since the vectors ζ'^{α} and ζ^{α} coincide at the initial point r_1, we find for small distances $\delta r = r_2 - r_1$ that

$$\zeta'^{\alpha} = \left(\delta^{\alpha}_{\beta} + l^{\alpha}_{\ ;\beta}\,\delta r\right)\zeta^{\beta} . \tag{5.3.17}$$

Multiplying both sides of this equality by m_{α} and denoting

$$\rho = -\,l_{\alpha;\beta}\,m^{\alpha}\,\bar{m}^{\beta}, \qquad \sigma = -\,l_{\alpha;\beta}\,m^{\alpha}\,m^{\beta}, \tag{5.3.18}$$

we obtain

$$\delta\zeta \equiv \zeta' - \zeta = -(\rho\zeta + \sigma\bar{\zeta})\,\delta r . \tag{5.3.19}$$

The mapping

$$\zeta \to \zeta' = \zeta + \delta\zeta = \zeta(1 - \rho\,\delta r) - \sigma\bar{\zeta}\,\delta r \tag{5.3.20}$$

thus establishes the relation between the shape of the object and that of the shadow.

If the projected object is a circle with a boundary $\zeta(\phi) = \exp(i\phi)$, the boundary of the shadow is given by the expression

$$\zeta' = (1 - \rho\,\delta r)\exp(i\phi) - \sigma\,\delta r\exp(-i\phi) , \tag{5.3.21}$$

which describes an ellipse with semi-axes

$$a_{\pm} = 1 - \left[\frac{1}{2}(\rho + \bar{\rho}) \mp |\,\sigma\,|\right]\delta r , \tag{5.3.22}$$

and area

$$\mathcal{A} = \pi\,a_+\,a_- = \pi[1 - (\rho + \bar{\rho})\,\delta r] . \tag{5.3.23}$$

Thus, the factor $\theta = -\mathrm{Re}\,\rho$ describes the expansion of the linear scale. The shear modulus $|\,\sigma\,|$ gives the degree of deformation of the circle; it is found from the relation $a_+/a_- = 1 + 2|\,\sigma\,|\,\delta r$.

The quantities ρ and $|\,\sigma\,|$ are independent of the choice of vectors $(m^{\alpha}, \bar{m}^{\alpha})$ and characterize the *convergence* and *shear* of light ray congruences. They are called

[5]This relation can be easily verified if one notes that $\zeta^{\alpha}(r,\phi) = [\partial f^{\alpha}(r, y_0^a)/\partial y^a]\,\delta y^a(\phi)$ and uses the symmetry relation $\partial^2 f^{\alpha}/(\partial r\,\partial y^a) = \partial^2 f^{\alpha}/(\partial y^a\,\partial r)$.

optical scalars. It should be emphasized that $\rho\,\delta r$ and $\sigma\,\delta r$ are invariants, while ρ and σ are linearly transformed when the affine parameter changes. It is readily verified that if the null surface in question, $v(x)=0$, is included into a family of null surfaces $v(x)=c$ and $l_\alpha = \partial_\alpha v$, then

$$\rho = -\frac{1}{2}\,l^\alpha_{\ ;\alpha}\,, \qquad |\,\sigma\,|^2 = \frac{1}{2}\,l_{\alpha;\beta}\,l^{\alpha;\beta} - \rho^2\,. \tag{5.3.24}$$

Let us emphasize that since $l_{\alpha;\beta} = v_{;\alpha;\beta} = l_{\beta;\alpha}$, the optical scalar ρ for the congruence of light rays forming the surface Γ is a real quantity: $\rho=\bar\rho$. The relations obtained above complete the proof of the Ehlers-Sachs theorem.

5.3.3 Focusing theorem

If the cross-sectional area of a narrow light beam is denoted by $\delta\mathcal{A}$, its change is described by the following equation implied by (5.3.23) for $\rho=\bar\rho$:

$$\frac{d}{dr}(\delta\mathcal{A})^{1/2} = -\,\rho\,(\delta\mathcal{A})^{1/2}\,. \tag{5.3.25}$$

If $\rho>0$, the cross-sectional area decreases as the affine parameter increases.

The behavior of optical scalars ρ and σ along the light rays is described by a system of first-order ordinary differential equations. The derivation of these equations is based on using the identity

$$l_{\mu;\alpha;\beta} = l_{\mu;\beta;\alpha} + R^\nu_{\ \mu\alpha\beta}\,l_\nu\,. \tag{5.3.26}$$

Multiplying both sides of this identity by $l^\beta\,m^\mu\,\bar m^\alpha$, then choosing m^μ, for the sake of convenience, to be parallel-transported along l^μ, and taking into account the geodesic condition $l^\mu\,l^\alpha_{\ ;\mu}=0$ and the relation $R_{\nu\beta}\,l^\nu\,l^\beta = 2\,R_{\nu\mu\alpha\beta}\,l^\nu\,l^\beta\,m^\mu\,\bar m^\alpha$, we arrive at the following equation:

$$\frac{d\rho}{dr} = \rho^2 + \sigma\,\bar\sigma + \Phi\,, \tag{5.3.27}$$

where $\Phi = (1/2)\,R_{\alpha\beta}\,l^\alpha\,l^\beta$. The equation for σ is derived in a similar manner:

$$\frac{d\sigma}{dr} = \sigma(\rho+\bar\rho) + \Psi\,, \tag{5.3.28}$$

where $\Psi = C_{\alpha\beta\gamma\delta}\,l^\alpha\,m^\beta\,l^\gamma\,m^\delta$, and $C_{\alpha\beta\gamma\delta}$ is the Weyl tensor.[6] It can be easily verified that Ψ does not depend on the particular choice of m^μ provided $\boldsymbol{l}\cdot\boldsymbol{m}=0$. If l^μ are tangent vectors to the null surface, then ρ satisfies the condition $\rho=\bar\rho$.

[6]Equation (5.3.27) is the analogue of the Raychaudhuri equation (A.35) for the congruence of timelike geodesics. Note also that (5.3.27) and (5.3.28) coincide with the first two Newman-Penrose equations of the system (E.12) adapted for the problem under consideration.

If we assume that there is a test beam, formed of light rays, for which the condition $\rho = \sigma = 0$ is satisfied at an initial moment, then part of the curvature Φ (for $\Psi = 0$) acts as a lens free of astigmatism (σ stays equal to zero), while part of the curvature Ψ (for $\Phi = 0$) acts as a purely astigmatic lens.

Relation (5.3.27) makes it possible to prove the following important proposition.

Focusing theorem. *Let $\Phi \geq 0$ and let $\rho \equiv \rho_0 > 0$ at some point $r = r_0$ of the light beam. Then at some finite distance $r - r_0 \leq \rho_0^{-1}$ from this initial point the beam reaches the focal point where the beam cross-sectional area drops to zero.*

To prove this proposition, it is sufficient to make use of the following relation:

$$\frac{d^2}{dr^2}(\delta\mathcal{A})^{1/2} = -\,(\sigma\,\bar{\sigma} + \Phi)(\delta\mathcal{A})^{1/2}. \tag{5.3.29}$$

This formula is obtained from (5.3.25) by differentiation and by using (5.3.27). Since the right-hand side of this equality is non-positive for $r > r_0$, we have $d(\delta\mathcal{A})^{1/2}/dr \leq -\rho_0(\delta\mathcal{A})_0^{1/2}$, where $\mathcal{A}_0$ is the value of $\mathcal{A}$ at r_0. Thus, $(\delta\mathcal{A})^{1/2}$ vanishes at a value of the parameter r which satisfies the inequality $0 < r - r_0 \leq \rho_0^{-1}$.

If the gravitational field is described by Einstein's equations, then the condition $\Phi \geq 0$ is equivalent to the relation $T_{\mu\nu}\, l^\mu\, l^\nu \geq 0$. For example, this condition is satisfied if the energy-momentum tensor describing the matter and field distributions obeys the *weak energy condition* (see Appendix A.15); that is, the energy density $T_{\mu\nu}\, u^\mu\, u^\nu$ in the reference frame of an arbitrary observer ($u_\mu\, u^\mu = -1$) is non-negative. To prove that $\Phi \geq 0$ follows from the weak energy condition, it is sufficient to consider the limiting case of $\alpha(u)\, u^\mu \to l^\mu$.

There are grounds for assuming that the weak energy condition always holds when matter and physical fields are described within the framework of *classical theory*. This means that if one neglects quantum effects, any energy-momentum flux across the null surface has a focusing effect on light rays.

5.4 Hawking's Area Theorem

The Penrose theorem implies that the event horizon is a null surface whose generators never intersect when extended into the future. The caustics at the horizon ($\rho \to -\infty$) can arise as a result of the falling of matter or radiation into a black hole, or upon the collision and merging of black holes, or when external source acts on the black hole. At caustics new bundles of generators enter the horizon. This property together with the focusing theorem, proved in the previous section allows one to derive a number of important propositions on the general properties of black holes.

Let us consider an infinitely narrow pencil of generators of the event horizon. Let the cross-section of this pencil at a point with the affine parameter r be $\delta\mathcal{A}(r)$. Note that in view of Ehlers-Sachs theorem, $\delta\mathcal{A}(r)$ is independent of the special choice of a local observer who measures it, and thus is invariant. Assume that the cross-sectional area begins to decrease at some point r_0 and the energy-momentum tensor describing the matter and the physical fields surrounding the black hole (and possibly falling into it) satisfies the weak energy condition. The focusing theorem then implies that the generators in the pencil must intersect at a finite value of the affine parameter. In order to match this result to the Penrose theorem, we have to conclude that either there is a physical singularity at the horizon and the generators of the horizon reach it before intersecting, or that the assumption is wrong that the cross-sectional area of the pencil of generators might start to decrease. In other words, the weak energy condition and the assumption that there are no singularities that can be encountered by the event horizon, imply that the cross-section area of the pencil of generators of the event horizon cannot start to decrease.

Hawking (197lb, 1972a) proved the theorem that the cross-section of a pencil of generators does not decrease with time even if the condition of no singularities at the event horizon is replaced with the condition that no singularities are visible from $\mathcal{J}^+$. Such singularities (visible from $\mathcal{J}^+$) are said to be *naked.* More rigorously, the condition of the absence of naked singularities is formulated as the condition of the existence of a regular spacelike surface Σ such that all past directed causal curves that come out of $\mathcal{J}^+$ necessarily intersect Σ when extended into the past. The existence of such a surface guarantees that if initial data are fixed on it so as to describe completely the state of particles and fields, then the evolution of the system in the region visible from $\mathcal{J}^+$ is uniquely determined. This is equivalent to stating that there are no singularities visible from $\mathcal{J}^+$. In Hawking's terminology, such spaces are *asymptotically predictable.*

Therefore, on the one hand, if we assume that there are no singularities (either at the event horizon or beyond it), then the cross-section of each pencil of generators of the event horizon is non-decreasing in time. On the other hand, if the event horizon contains caustics where new pencils of generators are formed, the horizon cross-section increases. This implies the sum of areas $S(t)$ of the surfaces of the black holes $\mathcal{B}_i(\tau)$ is a non-decreasing function of "time" τ. (We assume that $l^\mu \tau_{,\mu} < 0$; that is, the cross-section of the event horizon at a later instant τ corresponds to larger values of the affine parameter along each generator.) A similar conclusion that the surface area does not decrease is also valid for an individually taken black hole $\mathcal{B}_i(\tau)$. These results are stated in the theorem proved by Hawking (1971b, 1972a) (Figure 5.9).

There are assumptions used to prove the theorem: (1) naked singularities are absent, and (2) the weak energy condition is satisfied.

The absence of a naked singularity was formulated by Penrose (1969) as a hypothesis known as the *cosmic censorship conjecture*, which states that a naked singularity (i.e., a singularity visible to distant observers) cannot evolve from a regular state of the system under any physically reasonable assumptions concerning the properties of the matter. We shall discuss this conjecture in Section 5.7. As for the weak energy condition, the following has to be emphasized: Although this condition seems to be satisfied in the analysis of a black hole interaction with matter and fields in

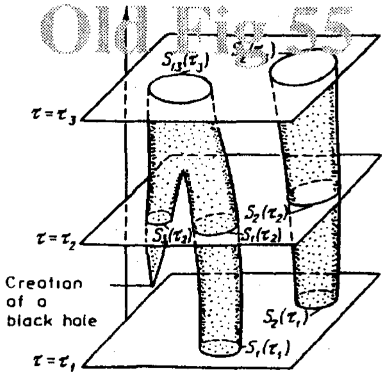


Figure 5.9: Possible processes with black holes (illustration for Hawking's theorem). The planes τ_1, τ_2, τ_3 are spatial sections at the corresponding moments of time; $S_a(\tau_i)$ is the area of the black hole a at the instant of time τ_i. Two black holes can merge into one and a new black hole can emerge. The area of a single black hole does not decrease with time. Hawking's theorem states that the total area of the surface of black holes at a moment τ is a non-decreasing function of time.

the framework of classical theory, it may be violated (and is indeed violated) when quantum effects are taken into account.[7] Therefore, Hawking's black-hole-area theorem is directly applicable only to processes for whose description quantum effects are negligible.

5.5 Trapped Surfaces. Apparent Horizon

5.5.1 "Teleological nature" of the event horizon

The boundary of a black hole is the event horizon H^+. At first glance, the definition of a black hole as a region inside the event horizon is very natural. However, if we consider the processes that are possible during the formation of a black hole or in the course of its subsequent evolution, it becomes clear that this definition actually does not describe quite what it was meant to do.

Indeed, let us imagine that a spherical shell of mass ΔM falls onto a spherical black hole of mass M some time after it was formed (Figure 5.10). It would seem that the boundary of the black hole before the shell reached it was the gravitational radius $r_{g,1} = 2M$, and after the fall the boundary widens and becomes $r_{g,2} = 2(M + \Delta M)$. In fact, the statement that the boundary of the black hole before the fall of ΔM was the $r_{g,1}$ surface is incorrect. Indeed, after the fall of ΔM into the sphere of radius

[7]The possible violation of the Hawking's area theorem in the quantum domain was first pointed out by Markov (1974).

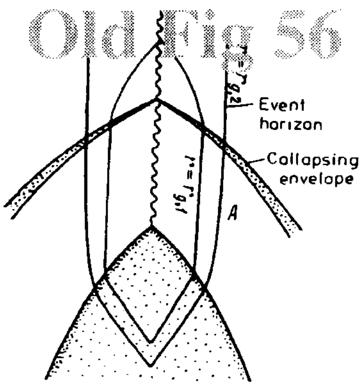


Figure 5.10: Nonstatic spherically symmetric black hole.

$r_{g,1}$, the null geodesics running along $r_{g,1}$ converge and fall into the singularity (see Figure 5.10). These geodesics are not the boundary of the region from which rays do not escape to infinity. This region is wider, it is bounded by the rays A shown in Figure 5.10. Before the fall of the mass ΔM, these rays ran outward from $r_{g,1}$ and somewhat diverged. If ΔM had not fallen in, the rays would have escaped to the infinity $\mathcal{J}^+$. But the falling mass eliminates this divergence and makes the rays go along $r_{g,2}$.

The boundary of the black hole, H^+, is thus determined not only by some specifics of the spacetime at a given moment (say, a strong field in some region) but also by the entire future history (whether a mass ΔM falls or does not fall, etc.). The problem of finding the event horizon H^+ is a problem with final, not initial, conditions. In order to stress this property, one usually refers to it as the *teleological nature of the horizon*, in contrast to the usual causal nature of dynamical systems.[8]

The following situation is even more impressive. Recall the process of creation of a spherical black hole (Figure 5.11a). We know that H^+ (and hence the black hole) is born at the moment τ_0, earlier than the star collapses to the gravitational radius (before τ_1). But imagine that the star explodes at a moment between τ_0 and τ_1, and its matter is thrown out to infinity; that is, no black hole is formed (Figure 5.11b).[9] Of course, it would be wrong to say that the black hole temporarily existed in the time interval from τ_0 to the explosion because there is no horizon H^+ in this example.

The boundary H^+ thus bounds not so much a region with an especially strong gravitational field (although this field is certainly necessary, or H^+ would not appear

[8]Novikov (1997) emphasized that the properties of the event horizon looks causal (not "teleological") if one considers the direction of time from the future into the past. Thus, figuratively speaking, one can say that a black hole "lives" in time flowing in the opposite direction.

[9]Note that after the star has contracted to r_g (after τ_1), the explosion cannot eject the matter out to infinity. Nothing can escape from under $r = r_g$.

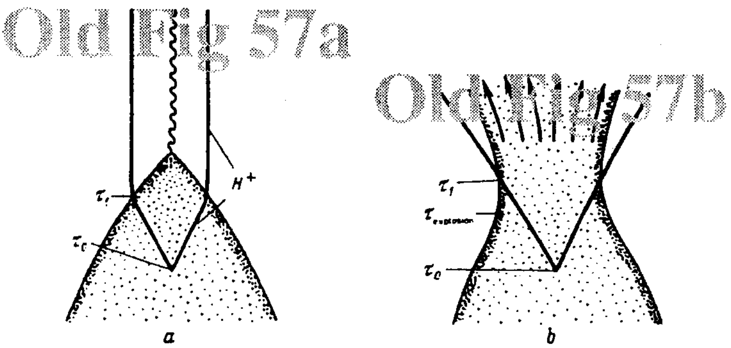


Figure 5.11: The position of the event horizon at a given moment of time depends on the entire subsequent evolution of the system. When a spherical collapse forms a black hole (a), the event horizon appears at a moment τ_0. An explosion of the collapsing star occurring after the moment τ_0 may completely preclude the formation of the event horizon (b).

at all) but rather a region with very specific global properties; namely, no rays escape from this region to infinity. It is this property – the invisibility from infinity, the impossibility for particles and light rays to escape – that justifies the name "black hole" for this region. In addition, the event horizon is formed by null geodesics, for which a number of strong theorems can be formulated (we gave some of them above). This is another reason for choosing H^+ as the definition for the black hole boundary.

5.5.2 Trapped surfaces

There is a question, though: Does the spacetime inside a black hole have specific properties which are qualitatively different from those outside the black hole at a given moment, so that it is possible to say that a black hole does exist at this moment, without analyzing the entire infinite future history of the world? We shall now see that, in general, such specific properties do exist.

As we have already mentioned, the following gedanken experiment can detect the fall of a spherically collapsing body into a black hole. Let the surface of the collapsing body emit a flash at some moment of time. If the body is transparent, then after a short interval of time there are two surfaces, one outside and one inside the body, which correspond to the positions of the outer and inner fronts of the light wave. It is characteristic of the situation inside the black hole that the surface areas of both the inner and the outer wavefronts decrease with time, and the light rays orthogonal to them converge. The gravitational field inside the black hole is so strong that it forces even the outgoing rays that leave the collapsing body to fall to the center (the positions of null cones inside the black hole are shown in Figure 5.12).

In the general case, closed orientable smooth two-dimensional spacelike surfaces, such that both families of null geodesics orthogonal to them converge ($\rho > 0$), are

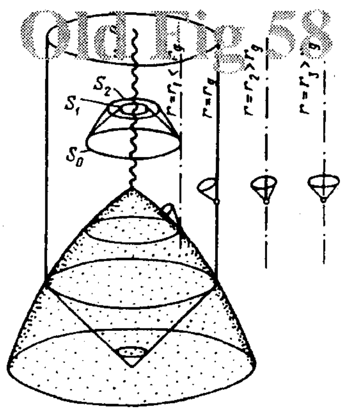


Figure 5.12: Position of local null cones inside and outside the black hole. The areas of the inner (S_1) and outer (S_2) radiation wavefronts emitted normally to a two-dimensional surface S_0 inside the black hole are less than the area of S_0.

called *trapped surfaces*. The presence of a trapped surface indicates that the gravitational field in the region through which this surface passes is very strong. In asymptotically flat spaces (with an asymptotically predictable future[10]), the trapped surface cannot be visible from $\mathcal{J}^+$, unless the weak energy condition is violated [Hawking and Ellis (1973)]. In other words, trapped surfaces lie inside black holes and their existence points to the formation of a black hole.

In the absence of naked singularities the presence of a trapped surface is thus the condition *sufficient* for the existence of a black hole at a given moment of time. But this condition is not *necessary*. The example shown in Figure 5.11a illustrates this. A black hole is formed at the moment τ_0. The star's explosion at τ_t ($\tau_0 < \tau_t < \tau_1$) prevents the formation of a black hole. In this situation a trapped surface cannot exist between τ_0 and τ_1 since the existence of such a trapped surface would make the

[10]In proving rigorous results concerning the properties of black holes, one usually has to make certain assumptions on the global properties of spacetime. These assumptions are reasonable in the physical context and play an important "technical" role in the proofs of theorems since they allow us to ignore various "pathological" situations. Such assumptions are usually listed in detail when the relevant theorems are discussed [see Hawking and Ellis (1973)]. We made an attempt to restrict as much as possible the use of numerous terms meant to describe these properties. Suffice it to mention that a partial Cauchy surface is a spacelike surface intersected by each causal curve not more than once. By definition, *a space with a strongly asymptotically predictable future* is a space with a partial Cauchy surface such that not only the "exterior" of the black hole but also a small neighborhood of the event horizon become predictable by fixing initial data on this surface. We also use the concept of a *regularly predictable space*. This is a space with a strongly asymptotically predictable future and a partial Cauchy surface Σ with the following properties: (1) the intersection of Σ and $J^-(\mathcal{J}^+)$ is contained in $J^+(\mathcal{J}^-)$ and is homeomorphic to R^3 from which an open set with compact closure is cut out, and (2) the surface Σ is simply connected.

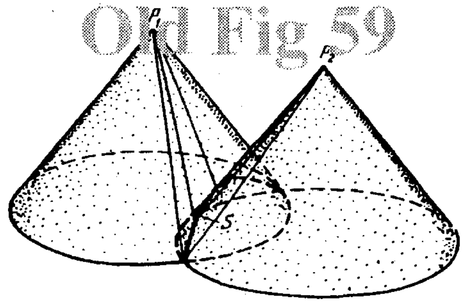


Figure 5.13: An example of a noncompact two-dimensional surface in Minkowski space. Both orthogonally emerging families of light rays are convergent. Such a surface S is not a trapped one.

process impossible. We return to the discussion of this subject later.

The condition that a trapped surface is closed is very important. Such a surface surrounding the gravitating center contracts so "swiftly" that the property described above (convergence) occurs even for light rays propagating out of this surface. If we do not impose the condition that the surface is closed, then all other properties of the trapped surface can be realized without gravitation in Minkowski flat world. Thus, the intersection of two past null cones for two points p_1 and p_2 in Minkowski spacetime (Figure 5.13) gives a two-dimensional spacelike surface S having the required properties, except for the property of being closed.

Monitoring the behavior of the outgoing radiation is found to be sufficient for solving a number of problems connected with the propagation of signals in the field of a black hole. Consequently, the following definition proves to be convenient: An *outer trapped surface* is a compact orientable spacelike surface with the property that the convergence of the outgoing null geodesics orthogonal to this surface is nonnegative ($\rho \geq 0$). This definition assumes that it is possible to determine in an invariant manner which of the two families of light rays orthogonal to the surface is the outgoing one. We next describe how this can be done.

Let the trapped surface S in question arise as a result of the evolution of a system with initial conditions given on the Cauchy surface Σ (Figure 5.14). Assume, for simplicity, that Σ is of topology R^3, and consider an arbitrary congruence of smooth timelike lines (the existence of such congruences can be guaranteed, at least in the future Cauchy region $D^+(\Sigma)$ of the surface Σ [Hawking and Ellis (1973)]).[11] The individual curves of a congruence can be treated as world lines of local observers. If we trace past-directed congruence lines passing through the surface S, the points of intersection of these lines with the Cauchy surface Σ form a closed orientable

[11] The region $D^+(\Sigma)$ of the Cauchy future of set Σ is defined as the set of all points p such that each causal curve nonextendable into the past and passing through p necessarily intersects Σ.

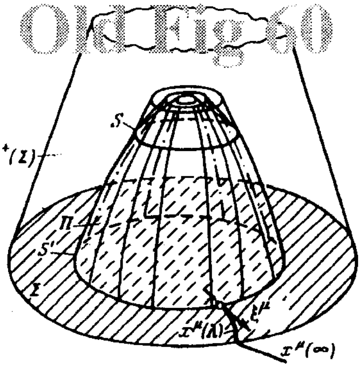


Figure 5.14: Illustration of the definition of an outer trapped surface.

surface S' on Σ. Finding the outward direction on this surface is straightforward. For instance, one can take any smooth curve $x^\mu(\lambda)$ which begins on S' ($x^\mu(0) \in S'$) and reaches the spatial infinity. Then the vector $\xi^\mu = dx^\mu/d\lambda$ for $\lambda = 0$ defines the outward direction at a point $x^\mu(0)$. As the surface Π formed by the congruence passing through S is orientable, the outward direction ξ^μ defined at one point of this surface determines the outward directions at each of its points, including the points of the original surface S. Such a definition is invariant and can be generalized to the case of asymptotically predictable spaces [Hawking and Ellis (1973)].

5.5.3 *R*- and *T*-regions. Apparent horizon

We say that a point p lies in the trapped region (briefly, T_--region) if there exists an outer trapped surface passing through this point. In the important particular case of spherically symmetric spaces, a point p belongs to a trapped region if the condition $(\nabla r \cdot \nabla r)_p \leq 0$ is satisfied.

Figure 5.15a shows a T_--region for the simplest case of the collapse of a spherical dust cloud not followed by a subsequent fall of matter into the black hole. In view of the importance of this concept, the same spacetime is shown again in Figure 5.15b in Lemaître coordinates (see Section 2.7). The metric inside the dust ball can be written in the form

$$ds^2 = -\,d\tau^2 + \left[\frac{3}{2}\,\frac{r_g - \tau}{r_g}\right]^{4/3} \left(dR^2 + R^2\,d\omega^2\right); \tag{5.5.1}$$

the boundary of the ball corresponds to $R = r_g$. The metric outside the ball, in the vacuum, is of the form

$$ds^2 = -\,d\tau^2 + \frac{dR^2}{B} + B^2\,r_g^2\,d\omega^2\,, \tag{5.5.2}$$

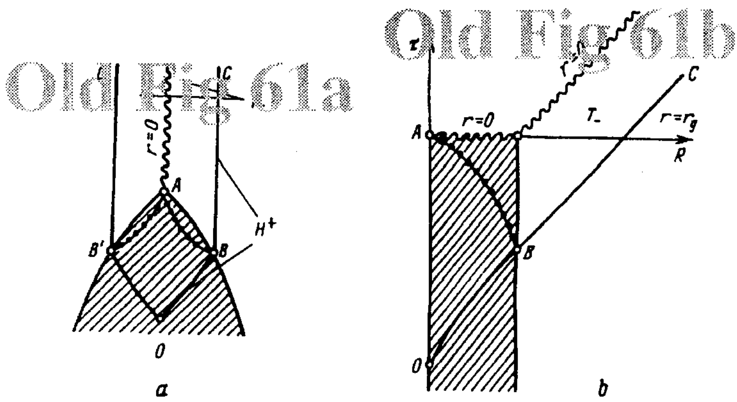


Figure 5.15: (a) Position of the T_--region in the collapse of a spherical dust cloud; (b) the same spacetime in Lemaître coordinates.

where

$$B = \left(\frac{3}{2} \frac{R - \tau}{r_g} \right)^{2/3}.$$

The boundary of the T_--region in matter (line AB) is described by the equation

$$\tau + \frac{2}{3} r_g \left(\frac{R}{r_g} \right)^3 = r_g . \tag{5.5.3}$$

This boundary is spacelike. The boundary of the T_--region outside the matter, in the vacuum, (line BC) is r_g:

$$\frac{3}{2} (R - \tau) = r_g . \tag{5.5.4}$$

This boundary is lightlike.

The outer component $\partial T_i(\tau)$ of the connected part $T_i(\tau)$ of the intersection of the trapped region and a spacelike surface $\tau(x) = \text{const}$ is called the *apparent horizon*.

The apparent horizon is a two-dimensional surface for which the outgoing orthogonal null geodesics have zero divergence ($\rho = 0$). When the weak energy condition is satisfied, the apparent horizon either coincides with the event horizon or lies inside it. In stationary black holes, the apparent horizon and the event horizon coincide (e.g., this is the case for stationary spherically symmetric black holes; see Figure 5.15).

In the general case, the outgoing rays emerging orthogonally to the apparent horizon $\partial T(\tau)$ have zero divergence. In view of equation (5.3.27), the condition $\rho = 0$ still holds along these rays until they intersect the region where $\Phi > 0$ or $\sigma \neq 0$. In this region ρ becomes positive so that light rays leave the apparent horizon

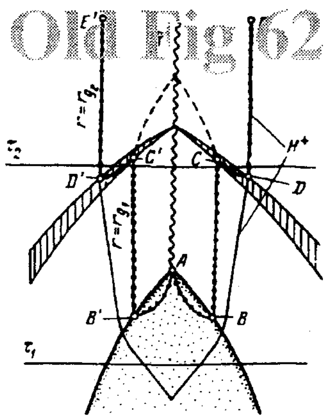


Figure 5.16: Accretion of matter onto a spherical black hole.

and go into the trapped region. In other words, the outer boundary of the T_--region is a null surface in the region where $\sigma = 0$, $\Phi = 0$, but becomes a spacelike region where $\sigma \neq 0$ and (or) $\Phi > 0$ (see Figure 5.15). If the weak energy condition is satisfied, there always exists the event horizon located outside the apparent horizon. Nevertheless, in general, there may be no trapped surfaces inside the event horizon. On the other hand, several connected components of the trapped region can exist inside one black hole.

This situation is illustrated in Figure 5.16. This represents a spherically symmetric black hole that becomes non-stationary for some time because of the matter falling into it. The T_--region lies inside the boundary EDCBAB′C′D′E′ shown by the dotted line. The outer boundary of each connected component of the T_--region in the $\tau =$ const section forms the apparent horizon. There is no T_--region (no trapped surfaces) in the section $\tau = \tau_1 =$ const inside the black hole (inside H^+). This proves that the presence of trapped surfaces in the section $\tau =$ const inside the black hole is not a necessary condition for the existence of the event horizon. The section $\tau = \tau_2 =$ const contains two connected T_--regions; the inner apparent horizon is $r = r_{g,1}$, and the outer one is $r = r_{g,2}$. A similar situation where several connected components of T_--regions are formed may arise when two black holes merge.

The following definition will prove to be useful: A closed orientable smooth two-dimensional spacelike surface is said to be *antitrapped* if both families of null geodesics orthogonal to it are divergent ($\rho < 0$). A point belongs to a T_+-region if there exists an antitrapped surface passing through it. The formation of T_+-regions is typical of white holes. We refer to the region of spacetime that lies beyond both the T_-- and T_+-regions as the R-region.

5.5.4 Dynamics of apparent horizons. Numerical results

The apparent horizon is defined to be the outermost trapped surface. It can be computed on any given Cauchy surface Σ , which gives a slice of spacetime. This makes the apparent horizon extremely convenient for numerical computations. We know that the apparent horizon indicates that a black hole has formed. Moreover, in the study of perturbed black holes it was shown by Anninos, Bernstein *et al.* (1995) that apparent horizons can closely approximate the event horizons even in dynamical spacetimes, and share many dynamical features with them.

Let S be a two-surface which is topologically a two-sphere embedded in Σ, a τ = const spacelike slice of spacetime. Denote by n^μ a future-directed unit vector normal to Σ, and by s^μ the outward-directed spacelike unit vector normal to S in Σ. The surface S is called a *marginally trapped surface* if the null vector $k^\mu = n^\mu + s^\mu$ obeys the condition

$$k^\mu{}_{;\mu} = 0 \tag{5.5.5}$$

everywhere on S.

The apparent horizon is the outermost marginally trapped surface. As we mentioned earlier, it may consist of several disjoint components. Let h_{ij} be the induced 3-metric on Σ, and K_{ij} be its extrinsic curvature. Then equation (5.5.5) can be written as [York (1989)]

$$s^i{}_{:i} + K_{ij}\, s^i\, s^j - h^{ij}\, K_{ij} = 0\,, \tag{5.5.6}$$

where $(\ )_{:i}$ denotes the three-dimensional covariant derivative in the metric h_{ij}. If $F(x^i) = 0$ is the equation defining the coordinate location of the marginally trapped surface S, then the normal vector s^i is proportional to the gradient of F. Thus, equation (5.5.6) is a second-order nonlinear equation for F. The existence, location, and physical properties of the apparent horizon for different initial data sets was discussed by Čadež (1974) and Bishop (1982).

The problem of finding the apparent horizon is greatly simplified under the assumption of axial symmetry. In this case, the coordinate position of the apparent horizon can be fixed by one function of one variable. For example, in cylindrical coordinates (z, ρ, ϕ) the equation of the apparent horizon is $z = z(\rho)$. Relation (5.5.6) in this case is reduced to a second-order nonlinear ordinary differential equation for $z(\rho)$. The evolution of the apparent horizon in an axially symmetric geometry was studied in [Abrahams *et al.* (1992), Bernstein *et al.* (1994a,b), Anninos, Bernstein *et al.* (1994), Brandt and Seidel (1995a,b)]. The analysis of perturbations of single rotating and nonrotating black holes show close similarities to each other [Anninos, Bernstein *et al.* (1994)]. For different types of perturbation the initially highly distorted prolate or oblate apparent horizon geometries quickly relax toward an equilibrium shape. The apparent horizon oscillates at a frequency that is predominantly the $l = 2$ quasinormal-mode frequency of the final black hole. These oscillations damp away in time as the black hole emits gravitational radiation. The importance of these results is connected with the fact that the main characteristics of the evolution of the apparent horizon as well as the accompanying process of the gravitational wave radiation in this highly nonlinear problem coincides to very high accuracy with the predictions of perturbation theory (see Chapter 4). According to this picture, a disturbance in the gravitational field of a black hole generates gravitational waves at the peak of the gravitational scattering potential (near $r = 3M$ for a non-rotating black hole). These waves, emitted at the quasinormal frequency, propagate away from the peak. The waves propagating outward result in gravitational radiation, while the waves propagating inward cause both a shearing and expansion of the apparent horizon. As a result, the apparent horizon oscillates with the quasinormal frequency.

5.6 Theorems on Singularities Inside Black Holes

5.6.1 Singularities in general relativity

We have mentioned in the analysis of spherical collapse that it inevitably produces a singularity, at least in the framework of general relativity. The invariants characterizing the spacetime curvature grow in the course of collapse, and the curvature at the center of the collapsing body becomes infinite after a finite time as measured by a clock fixed to the body. This situation sets in when the boundary of the T_--region crosses the $r = 0$ line. Further continuation of the world lines of particles and light rays that "reached" the newly formed singularity becomes impossible. As a result, the incompleteness of the space due to the cutoff of light rays and world lines at the singularity at a finite value of the affine parameter cannot, in principle, be eliminated.

A situation in which the density of matter and the tidal forces grow infinitely is also possible when the collapse of a dust ball is described in terms of the conventional Newtonian theory of gravitation. The essential feature in the Newtonian case is that pressure and small deviations from spherical symmetry, when taken into account, change the picture drastically: The maximal values of matter density and tidal forces (which in Newtonian theory are analogous to spacetime curvature) become bounded from above. Singularities in Newtonian theory are thus degenerate and unstable in the sense that they arise only in very special situations. Small perturbations are sufficient to eliminate such singularities.

The situation in general relativity is essentially different and the creation of a singularity inside a black hole happens inevitably under very general conditions. This conclusion is supported by a number of rigorous theorems.

If one assumes, on the one hand, that the weak energy condition is satisfied and a trapped surface has formed (in other words, that there is a black hole), the areas of the wavefronts of both outgoing and incoming radiation will decrease. On the other hand, since the velocity of motion of matter does not exceed the speed of light, the gap between these contracting surfaces always contains all the matter that ever got into this region. The matter will be compressed; its density will increase. It is natural to expect that a singularity or some other "unpleasantness" will appear.

What "unpleasantness" could this be? Until now we have interpreted the singularity as infinite curvature of spacetime. This singularity is definitely to be classified as physical because tidal forces become infinitely large near it. Any world line of a particle that runs into this singularity cannot, in principle, be continued beyond it. Here the particle ceases to exist. But this behavior does not exhaust all possible peculiarities of the spacetime that are connected with a singularity. For instance, consider the following situation. Let infinite curvature (a singularity) exist at some point in spacetime. We now cut this singularity, together with a neighborhood, out of spacetime. The remaining manifold has finite curvature only. Can this manifold be considered as free of singularities? Obviously, this conclusion would be incorrect. The point is that the world lines that earlier ran into infinite curvature are now truncated at the boundary of the excised region. This is also a physical irregularity that is covered by the term "singularity".

Normally, the term singularity encompasses not only infinite curvature but also any end-point on the world line of a particle (or a photon) or on a timelike geodesic provided the line cannot, in principle, be continued beyond this point. It should be stressed that the end-point—the singularity—must lie at a finite distance or at a finite value of the affine parameter for a null geodesic. In the most general case, the singularity is thus defined as the incompleteness of world lines in spacetime [for details and classification of spacetime singularities, see Geroch (1968), Schmidt (1971), Geroch *et al.* (1972), Hawking and Ellis (1973), Clarke (1973, 1975, 1976), Geroch *et al.* (1982), Tipler *et al.* (1980)].

5.6.2 Theorems on singularities

Having given these clarifications, we can return to discussing the problem of the inevitability of singularities inside trapped surfaces. The relevant theorem was proved by Penrose (1965a). It states:

Penrose theorem on singularities. *Let the weak energy condition be satisfied, and a trapped surface S exist in a spacetime admitting a noncompact Cauchy surface Σ. Then this spacetime cannot be complete with respect to null geodesics.*

In other words, the theorem states that this space contains at least one light ray that cannot be continued and that ends at a finite value of the affine parameter. Hence, the definition given above implies that a singularity exists.

The idea of the **proof** can be outlined as follows. Consider a set $J^+(S)$ of points that can be connected with S by a causal past-directed curve (Figure 5.17). Local analysis shows that at points where the boundary $\partial J^+(S)$ of this set is non-singular, this boundary is lightlike and consists of segments of null geodesics that are orthogonal to S at the initial points. If the null generators of $\partial J^+(S)$ have end points, these

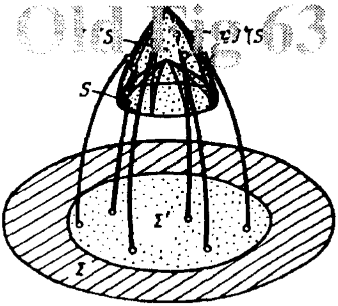


Figure 5.17: Illustration to the proof of the Penrose theorem on the singularity inside a black hole.

points coincide with the singularities of $\partial J^+(S)$ (caustics or intersection points). Now it is possible to prove, using the weak energy condition and the convergence of the generators of $\partial J^+(S)$ on the surface S, that each of the light rays emitted orthogonally to S inevitably arrives at a caustic; furthermore, this occurs when the affine parameter does not exceed ρ_S^{-1}, where ρ_S is the maximum value of ρ on S for both families of outgoing rays. (The existence of ρ_S is guaranteed by the smoothness and compactness of the surface S.) Hence, the boundary $\partial J^+(S)$ is compact because it is formed by a compact system of finite closed segments. It can be proved that $\partial J^+(S)$ is a three-dimensional manifold without edge. Note that the proof of compactness of $\partial J^+(S)$ essentially uses the assumption that the spacetime is complete with respect to null geodesics so that the generators of $\partial J^+(S)$ are not terminated before arriving at a caustic or an intersection point.

The next step of the proof is to establish that the compactness of $\partial J^+(S)$ contradicts the noncompactness of the Cauchy surface Σ, after which it becomes obvious that the assumption on the completeness of spacetime is incompatible with the rest of the hypotheses of the theorem.

The contradiction is established in the following manner. It can be shown that a congruence of timelike curves exists in a spacetime with a Cauchy surface. Since one and only one curve of the congruence passes through each point of space, and since a timelike curve cannot intersect the null surface $\partial J^+(S)$ more than once, this congruence establishes a continuous one-to-one correspondence between $\partial J^+(S)$ and a closed subset Σ' of the surface Σ. Note that Σ' cannot coincide with Σ because by assumption Σ is not compact. Hence, Σ' has a boundary in Σ. However, this is in contradiction with $\partial J^+(S)$ being a manifold without edge. The contradiction thus obtained completes the proof of the Penrose theorem on singularities.

It must be emphasized that the condition of noncompactness of the Cauchy surface Σ was used only in proving that Σ' is not identical to Σ. One could demand instead that at least one timelike curve of the congruence does not intersect $\partial J^+(S)$.

We will reproduce here one additional theorem on singularities (in a sense, this is the strongest one among the theorems of this kind), referring the reader interested in exact formulations to the papers of Penrose (1968, 1979), Hawking and Ellis (1973), Misner, Thorne, and Wheeler (1973), Wald (1984), Tipler *et al.* (1980).

Hawking-Penrose theorem [Hawking and Penrose (1970)]: *The spacetime M necessarily contains incomplete timelike or null geodesics which cannot be continued provided the following conditions are satisfied:*

1. *The spacetime contains no closed timelike curves.*
2. *Arbitrary timelike unit vectors u^μ obey the inequality $R_{\mu\nu} u^\mu u^\nu \geq 0$.*
3. *For each timelike or null geodesic with a tangent vector u^μ, there exists a point in which $u_{[\alpha} R_{\beta]\gamma\delta[\varepsilon} u_{\rho]} u^\gamma u^\delta \neq 0$.*
4. *A trapped surface exists.*

All these conditions appear to be fairly reasonable and general. Condition (1) corresponds to our conventional interpretation of causality.[12] Condition (2) signifies that the energy density ε is non-negative in any physical reference frame and $\varepsilon + 3p \geq 0$. Condition (3) is equivalent to stating that we deal with a spacetime of a general type, without special symmetries. Condition (4) is closely connected, as we have already mentioned, with the existence of a black hole. The Penrose-Hawking theorem guarantees that a singularity arises also in the case when a trapped surface is generated, say, in a closed Universe. Note that in such a case there would be no noncompact Cauchy surface and the Penrose theorem would be inapplicable.

The above theorems on singularities can to some extent be generalized. They permit extension to the case when the energy dominance conditions are met only on average [for details, see Tipler (1978a,b) and Roman (1986, 1988)].

5.7 Cosmic Censorship Conjecture. Critical Behavior in Black Hole Formation

5.7.1 Cosmic censorship conjecture

Penrose's *cosmic censorship conjecture* [Penrose (1969)] plays an important role in black hole physics. It is the statement that a naked singularity (i.e., a singularity

[12]It must be emphasized that closed timelike lines do not contradict the broadly interpreted causality principle. The future cannot be distinguished from the past on a closed time line but this property, by itself, does not produce any contradictions [see Zel'dovich and Novikov (1975), Novikov (1983)]. All events on such a line are "synchronized" with one another. The past cannot be changed once the future is known (as is sometimes said) because all the events along the time line "have already occurred"; they cannot be altered since they are a part of the four-dimensional spacetime. A different wording is possible: It is wrong to speak of the effect of the future on the past because these concepts become meaningless. For more details, see Chapter 16.

visible to distant observers) cannot evolve from a regular initial state of the system under any physically reasonable assumptions concerning the properties of the matter. The cosmic censorship conjecture in this formulation is sometimes referred to as "weak", to distinguish it from the "strong" one. The latter principle, suggested also by Penrose (1978), states that in the general case the singularities produced by gravitational collapse are spacelike so that no observer can see them until he falls into them.[13]

Weak censorship effectively postulates that singularities which might develop in gravitational collapse cannot influence events near future null infinity $\mathcal{J}^+$. One can give the following more accurate formulation of the *weak censorship conjecture* [Wald (1984)]. Let S be a complete spacelike hypersurface on which generic non-singular initial data (induced metric and its time derivative on S) are given. The matter sources are such as to satisfy some energy condition and physically reasonable equations of state. Then the evolution of gravitational collapse from S must be such that the resulting spacetime is strongly asymptotically predictable.

The weak censorship conjecture guarantees that the region outside the black hole is *globally hyperbolic* [Tipler, Clarke, and Ellis (1980)]. If a spacetime fails to obey weak censorship, it must contain a naked singularity, i.e., there exists a future directed causal curve which reaches a distant observer, and in the past it terminates at the singularity.

So far both (weak and strong) cosmic censorship conjectures have not been proved. Difficulties arise as soon as one tries to give a more rigorous definition of what a singularity is.[14] Another difficulty is connected with the fact that the event horizon depends on the whole future behavior of the solution over an infinite period of time, whereas the present theory of quasi-linear hyperbolic equations guarantee the existence and regularity of solutions mainly over a finite interval only [Israel (1984)]. It is also difficult to specify which kind of conditions must be imposed on the stress-energy tensor of matter since there are explicit examples of naked singularities in spacetimes satisfying all the usual energy conditions. In addition, only stable singularities should be considered as really pathological, but there are no sufficiently general criteria for testing stability in general relativity. One of the proposed ideas is that there might be a connection between the strength of the singularity and its stability [Newman (1986b)].

One of the main arguments in favour of the censorship conjecture is the stability of black holes with respect to small perturbations (see Section 4.9). On the other hand, there are a number of results that indicate that at least in its simplest form the censorship conjecture is questionable. Let us consider examples which illustrate the possible creation of naked singularities in situations which might be considered

[13] It is important that the requirement of generality is imposed. Without it the strong censorship conjecture can be easily violated. A timelike singularity inside a charged non-rotating black hole (in Reissner-Nordström geometry) is an evident counter-example.

[14] For the discussion of singularities and references to relevant publications, see Section 5.6.

as physically reasonable.

As the first example, consider a black hole formed by spherically symmetric radiation moving radially with the velocity of light. The corresponding metric is known as the Vaidya (1951, 1953) solution

$$ds^2 = -\left(1 - \frac{2M(v)}{r}\right) dv^2 + 2dv\, dr + r^2\, d\omega^2 . \tag{5.7.1}$$

Here v is the advanced null coordinate. The stress-energy tensor for the null fluid creating this metric is

$$T_{\mu\nu} = \frac{1}{4\pi r^2} \frac{dM}{dv} v_{;\mu} v_{;\nu} . \tag{5.7.2}$$

For $dM/dv \geq 0$ it obeys the weak energy condition. The radiation collapses to the origin $r = 0$, and we assume that $M(v) = 0$ for $v \leq 0$ and $M(v) > 0$ for $v > 0$. The value $dM(0)/dv$ is crucial for the global properties of the solution. In order to investigate these properties, first consider a simplified case when M during the interval $(0, v_0)$ is a linearly growing function of v,

$$M(v) = \lambda v , \tag{5.7.3}$$

where $\lambda > 0$. The equation for outgoing radial null rays in this case can be easily integrated [see e.g., Volovich *et al.* (1976)]. Denote by $x = r/v$, then the radial null rays obey the following relation

$$\frac{|x - x_+|^{x_+/(x_+ - x_-)}}{|x - x_-|^{x_-/(x_+ - x_-)}} = \frac{C}{|v|} , \tag{5.7.4}$$

where $x_\pm = 1/4(1 \pm \sqrt{1 - 16\lambda}\,)$. Different rays are enumerated by a positive parameter C. Near $v = 0$, rays for different C are focused and have the same limiting value $x = x_+$. Besides these radial rays, there exist a family of non-radial rays which are also emitted from $r = 0$, $v = 0$.

The invariant of the square of curvature for the Vaidya metric is

$$R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} = 48 M^2(v)/r^6 , \tag{5.7.5}$$

and hence for $v \geq 0$, there is a curvature singularity at $r = 0$. It is possible to show that for $\lambda \leq 1/16$ the singularity $v = 0$, $r = 0$ is a node of null geodesics, which can reach $\mathcal{J}^+$ [Dwivedi and Joshi (1989, 1991)]. The radial section of the spacetime in (t, r)-coordinates is shown in Figure 5.18.

The naked singularity is strong in the sense that the curvature invariants (for example, the square of curvature (5.7.5)) are divergent near it. A naked singularity with similar properties exists in the Vaidya solutions with an arbitrary mass function $M(v)$ that describes monotonic increase of mass of a black hole from 0 at $v = 0$ to some final value m_0 at v_0 provided $dM/dv|_{v=0} \leq 1/16$ [Dwivedi and Joshi (1989, 1991), Joshi (1993)].

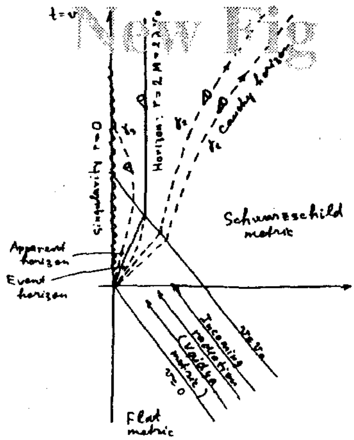


Figure 5.18: A spherical collapse of null fluid with $\lambda < 1/16$. A naked singularity is formed at $v = 0$, $r = 0$. Families of null rays such as γ_1 and γ_2 emitted at the naked singularity escape to infinity. The first family of rays γ_1 leaving the singularity and reaching infinity form the Cauchy horizon.

Infinite curvature is also possible when the contraction of a star is accompanied by radiation that carries away some of its mass provided complete conversion into radiation of the star occurs at the moment when it contracts to a point, and the surface of the star remains outside its gravitational radius during the entire compression phase [an event horizon is not formed in this case, see Steinmüler *et al.* (1975), Lake and Hellaby (1981), Kuroda (1984b)].

Other examples of naked singularities were discussed for the model of spherically symmetric collapsing dust. The corresponding metric (*Tolman solution*) is (see Appendix B.6)

$$ds^2 = -\,d\tau^2 + \frac{r'^2}{1+f}\,dR^2 + r^2 d\omega^2\,, \tag{5.7.6}$$

where $r = r(\tau, R)$ is a solution of the equation

$$\dot{r}^2 = f + \frac{F}{r}\,, \tag{5.7.7}$$

and $f = f(R)$ and $F = F(R)$ are two arbitrary functions, which are specified by initial conditions. The energy density of the dust

$$\epsilon = \frac{F'}{8\pi\, r'\, r^2} \tag{5.7.8}$$

becomes infinite at spacetime points where either $r' = 0$ or $r = 0$. At these points the square of the Ricci tensor is evidently infinite. The first case corresponds to the situation when different layers of the dust, described by different values of the comoving coordinate R intersect one another. This type of singularity is known as a *shell-crossing singularity*. A singularity which develops at $r = 0$ is known as a *shell-focusing singularity*. Both types of singularity are generic for a spherical collapse of dust matter, and with proper choice of initial conditions they are globally naked (i.e., seen from infinity). The nature of shell-crossing singularities appears to be well understood [Yodzis *et al.* (1973, 1974)]. It is shown that they are gravitationally weak [Newman (1986a)], and it is generally believed that such shell-crossing singularities need not be taken seriously as far as the cosmic censorship conjecture is concerned. Shell-focusing singularities are to be taken more seriously [Christodoulou (1984), Newman (1986a) Waugh and Lake (1988)]. They exist and are gravitationally strong for a wide class of Tolman models [Grillo (1991), Dwivedi and Joshi (1992), Joshi (1993)]. These results allow generalization to the case when the pressure does not vanish. For example, Ori and Piran (1990) proved the existence of naked singularities for self-similar spherically symmetric collapse of matter with pressure.

Formulation of the conditions under which cosmic censorship conjecture is valid and a rigorous proof of it remain one of the important unsolved problems in black hole physics [see the review of Israel (1984) and the book by Joshi (1993), the bibliography therein, and also Caderni and Calvani (1979), Calvani (1980), Glass and Harpaz (1981), Yang (1979), Lake (1979), Israel (1985)].

5.7.2 Hoop conjecture

It seems that an understanding of non-spherical collapse might be crucial for the resolution of the problem of naked singularities. Unfortunately, this is a rather complicated problem, and the results are much less complete than in the better understood spherically symmetric case. For this situation quite long ago Thorne (1972) proposed a *hoop conjecture*: ***Black holes with horizons form when and only when a mass M gets compacted into a region whose circumference in every direction is bounded by***

$$C < 4\pi M . \tag{5.7.9}$$

The hoop conjecture is compatible with the occurrence of naked singularities when the collapse is highly non-spherical. For this reason, it might be considered as some kind of alternative which might be used in case cosmic censorship were found not to be valid. If the hoop conjecture is indeed correct, then non-spherical collapse with one or two dimensions larger than the others might lead to a naked singularity. Non-spherical collapse might admit a large variety of singularities; for example, needle-shaped singularities that can result from the collapse of prolate axisymmetric objects. One may suspect that non-spherical types of singularities might be stronger than spherical ones. Shapiro and Teukolsky (1992) have provided evidence in support of this assumption. Namely, they numerically analyzed the behavior of 6000 dust particles, geodesically moving in the average gravitational field of the system, starting with an initially static spheroidal configuration. They found that if the initial spheroid is prolate and sufficiently elongated, then the resulting final configuration at the level of accuracy allowed by their numerical calculations can be interpreted as a naked spindle-type singularity. This result suggests the possibility left open by Thorne's hoop conjecture that the gravitational collapse of a highly deformed configuration will not produce a singularity which is hidden inside a black hole, but a naked singularity.

Although the hoop conjecture is formulated in a very simple manner, the attempts to make it rigorous meet many obstacles. First of all, as was already mentioned, it is virtually impossible to conclude that a black hole is formed by studying the data at a given spacelike surface. This is because the definition of the event horizon involves boundary data at future infinity. For this reason, usually the following modification of the hoop conjecture is considered: an apparent horizon forms when and only when a mass M gets compacted into a region whose circumference in every direction is $C < 4\pi M$.

Another problem is connected with the definition of the mass M which enters (5.7.9). Heuristically, a mass M must be some local characteristic defining the amount of matter located in a compact region under consideration. It is difficult to propose a definition which is both reasonable and covariant. For this reason, when considering an asymptotically flat spacetime, one usually uses the mass defined at spatial infinity (or ADM mass).

As for the definition of the circumference C which enters (5.7.9), several different proposals were made [see e.g., Flanagan (1991), Chiba *et al.* (1994)]. For the axisymmetric case the most satisfactory definition of C is the following [Nakamura, Shapiro, and Teukolsky (1988)]. Let S be an axisymmetric surface with the topology of a two-dimensional sphere surrounding a body. C is defined as the larger of its equatorial and polar circumferences.

Before discussing different results concerning the hoop conjecture, we remark that it is possible to derive the following Penrose-Gibbons *isoperimetric inequality* relating the surface area $\mathcal{A}$ of the apparent horizon and the mass M of the non-rotating body [Penrose (1973), Gibbons (1972)]

$$16\pi M^2 \geq \mathcal{A}. \tag{5.7.10}$$

The proof assumes the validity of the weak energy condition and is based on the following sequence of inequalities

$$\frac{\mathcal{A}}{4\pi(2M)^2} \leq \frac{\mathcal{A}^H}{4\pi(2M)^2} \leq \frac{\mathcal{A}^H}{4\pi(2M_{\text{final}})^2} \leq \frac{\mathcal{A}^H_{\text{final}}}{4\pi(2M_{\text{final}})^2} = 1\,. \tag{5.7.11}$$

The first inequality follows because the apparent horizon always lies inside the event horizon; the second because the mass M_{final} of the Schwarzschild black hole inevitably left at the end of the evolution must not exceed the initial mass; the third because the area of the event horizon can never decrease. The last equality is simply a relation between the mass and the surface area of the Schwarzschild black hole.

For the sphere the circumference C and the surface area $\mathcal{A}$ are related as follows: $C = \sqrt{\pi \mathcal{A}}$. Equation (5.7.10) when applied to this case implies $C \leq 4\pi M$, which reproduces the inequality (5.7.9). The numerical study of time symmetric initial data for different axisymmetric configurations shows that the necessary condition for the creation of the apparent horizon is $C \leq 15.8M$. The quantity staying on the right-hand side of this inequality is greater than $4\pi M$ by the factor 1.25 [Nakamura, Shapiro, and Teukolsky (1988), Chiba *et al.* (1994)]. A very close result ($C \leq 16M$) was obtained by Barrabès *et al.* (1991, 1992) and Tod (1992) for a particular model of collapsing convex massive thin null shells.

A somewhat broader statement similar in spirit to the hoop conjecture is the *event horizon conjecture* of Israel (1984, 1986a,b). In Section 5.5 we mentioned that in the absence of naked singularities (in the asymptotically flat spacetime with asymptotically predictable future) the existence of a trapped surface guarantees the existence of the event horizon provided the weak energy condition is satisfied. Israel's event-horizon conjecture is the requirement that an event horizon with no naked singularities must form whenever the collapse of matter (satisfying appropriate energy conditions) creates a trapped surface.

5.7.3 Critical behavior in black hole formation

Gravitational collapse is usually accompanied by an outflow of matter and outgoing gravitational radiation. Gravitational collapse may produce a black hole. It is also possible that in the collapse no black hole is created. An interesting problem is to study processes at the threshold of black hole formation. Recently the behavior of black holes at the threshold of their formation was investigated [Choptuik (1993)] and a number of interesting general relations characterizing this behavior were established.

Consider a dynamical system describing self-gravitating matter. Let $\mathcal{I}[p]$ denote the one-parameter family of initial states defined above, and let $\mathcal{S}[p]$ be a solution of evolution equations with initial conditions $\mathcal{I}[p]$. We suppose that for a fixed family parameter p is chosen so that for small values of parameter p the gravitational field during evolution is too weak to form a black hole, while for large values of p the black hole is produced. In general, between these two extremes there is a *critical* parameter value, p^*, where black hole formation first occurs. We will refer to the solutions $\mathcal{S}[p < p^*]$ and $\mathcal{S}[p > p^*]$ as subcritical and supercritical, respectively. The quantity $(p - p^*)$ is a natural choice for discussing the phenomenology of the solutions $\mathcal{S}[p]$.

The characteristic properties of the near-critical gravitational collapse were first obtained by Choptuik (1993) for the spherically symmetric collapse of the massless scalar field. He presented convincing numerical evidence that there is *no mass gap* in black hole production; arbitrarily small black holes can be formed in a collapse.

Moreover, the mass of sufficiently small black holes is given by

$$M_{\rm BH} \sim |p - p^*|^\beta\,, \tag{5.7.12}$$

where $\beta \approx 0.37$ is a universal exponent (this relation is referred to as *scaling*).

For marginal data, both supercritical and subcritical, the evolution approaches a certain universal solution $\mathcal{S}[p^*]$ which is the same for all the families of initial data $\mathcal{I}[p]$ (*universality*). This solution, which is unique and corresponds to the field configuration exactly at threshold of black hole formation p^*, is called *critical solution*, and is now commonly referred to as *choptuon*. It is regular and acts as an *intermediate attractor* in the sense that the time evolution first converges onto it, but then eventually diverges from it to either form a black hole or to disperse.

The critical solution for the spherically symmetric gravitational collapse of the massless scalar field has a discrete symmetry: It is periodic in the logarithm of the spacetime scale

$$t' = e^{-\Delta}\, t\,, \quad r' = e^{-\Delta}\, r\,, \tag{5.7.13}$$

$$ds'^2 = e^{-2\Delta}\, ds^2\,, \quad \phi(t', r') = \phi(t, r)\,, \tag{5.7.14}$$

with period $\Delta \approx \ln 30 \approx 3.4$, which is a constant belonging to the choptuon (the moment $t = 0$ captures the formation of black hole). This behavior of the critical solution is referred to as *echoing* because the solution repeats itself at ever-decreasing time and length scale, or discrete self-similarity (DSS).[15]

These features of the near-critical gravitational collapse appear to be quite general. Calculations on gravitational collapse of the massless scalar field using different coordinate systems and numerical algorithms [Garfinkle (1995), Hamadé and Stewart (1996)] confirmed that effects observed by Choptuik are not numerical artifacts.

These results were confirmed and generalized in a number of publications which appeared after the paper by Choptuik. Abrahams and Evans (1993) found a similar phenomenon in the axisymmetric collapse of a gravitational wave with almost the same value of the critical exponent $\beta \approx 0.38$. The corresponding choptuon is also discrete self-similar, but the constant Δ appears to be different $\Delta \approx \ln 1.8 \approx 0.6$. Hirschmann and Eardley (1995a,b) obtained the spherically symmetric solution for coupled Einstein-complex-scalar-field equations which possesses Choptuik-type universal scaling and echoing behavior.

In some cases, the critical solution possesses a stronger symmetry than the discrete self-similarity described above. It is called continuous self-similarity (CSS), or homotheticity, and is defined geometrically as the presence of a vector field ξ such that

$$\mathcal{L}_\xi g_{\mu\nu} = 2 g_{\mu\nu}\,, \tag{5.7.15}$$

[15]On the general geometric definition of discrete self-similarity see e.g., Gundlach (1995, 1997).

where $\mathcal{L}_\xi$ denotes the Lie derivative. The presence of the symmetry allowing to eliminate one of the coordinates from the equations is one of the reasons why it is easier to deal with continuous self-similar solutions; most analytical calculations use continuous self-similarity ansatz. An example of critical behavior with continuous self-similarity was found by Evans and Coleman (1994) in the model of the spherically symmetric collapse of a radiation fluid. The critical exponent β in their case is $\beta \approx 0.36$.

In all the cases when the critical behavior was observed, the generic feature is that the spacetime is asymptotically flat; there is transportation of the energy from a collapsing system to infinity, and the matter content is "massless".

First calculations in different models gave very close values for the mass scaling exponent β. These results indicated that the meaning of the universality might be extended to the independence of the critical exponent of details of the system though initially this meant its independence of initial data. Later calculations for wider class of models did not confirm this conclusion.

Exact analytical solution of the collapse of a thin shell coupled with an outgoing null fluid [Koike and Mishima (1995)] and perturbative analysis of the collapse of a perfect fluid with equation of state $p = \gamma\rho$ (with γ in the range $0 \leq \gamma \leq 0.88$) [Maison (1996)] both have critical exponent β strongly dependent on the parameters of the matter model.

The universality of β among massless scalar field, gravitational waves, and radiation fluid seems to be connected with the fact that these three are *massless* fields, but there is no proof of why it should be so.[16] Observed non-universality goes beyond β varying for different matter models; it affects more fundamental properties of critical solutions. In particular, it was shown by Hirschmann and Eardley (1995b) that in spherically symmetric gravitational collapse of the massless complex scalar field the critical solution is unstable; that is, it has instability other than the obvious black hole one, apparently an oscillatory instability toward the original real choptuon. Especially intriguing is quite recent example of gravitational collapse of a Yang-Mills field [Choptuik, Chmaj, and Bizon (1996)], where there are *two* distinct critical solutions, one discrete self-similarity and allowing black holes of arbitrarily small mass, the other one with a mass gap. The results of study of critical phenomena in gravitational collapse for different models are presented in Table 5.1

A study of the critical behavior in the gravitational collapse requires solutions of complicated nonlinear partial differential equations. For this reason, main results were obtained by using numerical simulations. Evans and Coleman (1994) proposed a perturbation method which can be used to simplify the calculations of the critical exponent. This method was carried out in more details by Koike, Hara and Adachi (1995). Briefly the idea of the method is as follows. The critical solution obtained for $p = p^*$ obeys the self-similarity property, and it usually much easier to find it than

[16]Note, however, that $\beta = 0.387$ for massless complex scalar field [Hirschmann and Eardley (1995b)] which is slightly but noticeably different from $\beta = 0.37$ for the above three fields.

Model	References	β	symmetry
Scalar Fields			
Massless scalar field	1, 2, 3	0.37	DSS
	4, 5, 6	1/2	CSS
	7	1/2	--
	8, 9	0.374	DSS
	10	1	CSS
Complex scalar field	11, 12	0.387	CSS
Charged scalar field	13	0.37	DSS
Other Matter Models			
Gravitational waves	14	0.37	DSS
Radiation fluid	15	0.36	CSS
	16	0.356	CSS
Perfect fluid	17	varies	CSS
Thin shell	18	varies	—
Yang-Mills field	19	0.20	DSS
Other Theories			
Axion-dilaton field	20, 21	0.264	CSS
2D dilaton gravity	22	0.53	—
Nonlinear sigma model	23	varies	both
Brans-Dicke theory	24, 25, 26	varies	both

Table 5.1: Critical behavior in the gravitational collapse. Numbers in References refer to the following publications.

1 – Choptuik (1993), 2 – Garfinkle (1995), 3 – Hamadé and Stewart (1996), 4 – Brady (1994), 5 – Brady (1995), 6 – Oshiro, Nakamura, and Tomimatsu (1994), 7 – Kiem (1994), 8 – Gundlach (1995), 9 – Gundlach (1997), 10 – A. Frolov (1997), 11 – Hirschmann and Eardley (1995a), 12 – Hirschmann and Eardley (1995b), 13 – Hod and Piran (1997), 14 – Abrahams and Evans (1993), 15 – Evans and Coleman (1994), 16 – Koike, Hara, and Adachi (1995), 17 – Maison (1996), 18 – Koike and Mishima (1995), 19 – Choptuik, Chmaj, and Bizon (1996), 20 – Eardley, Hirschmann, and Horne (1995), 21 – Hamadé, Horne, and Stewart (1996), 22 – Peleg, Bose, and Parker (1997), 23 – Hirschmann and Eardley (1997), 24 – Chiba and Soda (1996), 25 – Oliveira (1996), 26 – Liebling and Choptuik (1996)

to solve the problem. The characteristic feature of solutions with initial data close to those of the critical solution is that they first approach the latter, but eventually run away from it; that is, they contain a factor $\exp(\sigma t)$. Evans and Coleman (1994) proposed to use a linear stability analysis for study these run-away solutions. Quite general arguments show that the mass of a black hole which is formed as a result of this instability is proportional to $(p-p^*)^{1/\sigma}$ so the the critical index is $\beta = 1/\sigma$. This method allows both the calculation of β and a test of stability of the critical solution, and has been used in investigation of various matter models [Evans and Coleman (1994), Koike, Hara and Adachi (1995), Maison (1996), A. Frolov (1997)].

The phenomena described above are striking, but as yet there are almost no theoretical explanations for the essence of them, especially for the value of the critical exponent β and the mechanism by which self-similarity and universality are acquired.

5.8 Can One See What Happens "Inside a Black Hole"?

5.8.1 Formulation of the problem

Is it possible for a distant observer to receive information about the interior of a black hole? Strictly speaking, this is forbidden by the very definition of a black hole. What we have in mind in asking this question is the following. Suppose there exists a stationary or static black hole. Can we, by using some device, get information about the region lying inside the apparent horizon?

Certainly it is possible if one is allowed to violate the weak energy condition. For example, if one sends into a black hole some amount of "matter" of negative mass, the surface of the black hole shrinks, and some of the rays which previously were trapped inside the black hole would be able to leave it. If the decrease of the black hole mass during this process is small, then only a very narrow region lying directly inside the horizon of the former black hole becomes visible.

In order to be ahle to get information from regions not close the apparent horizon but deep inside an original black hole, one needs to change drastically the parameters of the black hole or even completely destroy it. A formal solution corresponding to such a destruction can be obtained if one considers a spherically symmetric collapse of negative mass into a black hole. The black hole destruction occurs when the negative mass of the collapsing matter becomes equal to the original mass of the black hole (Figure 5.19). In such a case an external observer can see some region close to the singularity. But even in this case the four-dimensional region of the black hole interior which becomes visible has a four-dimensional spacetime volume of order M^4. It is much smaller than four-volume of the black hole interior, which remains invisible and which is of order M^3T, where T is the time interval between the black hole formation and its destruction (we assume $T \gg M$). The price paid for the possibility of seeing

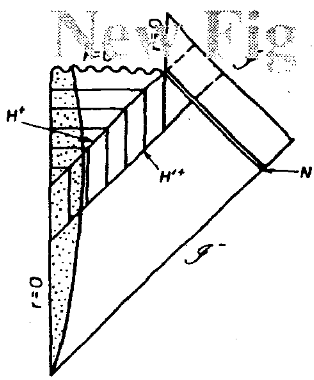


Figure 5.19: Penrose-Carter conformal diagram for the spacetime of a spherical black hole of mass M which is destroyed by the infall of a massive null shell N containing negative mass $M_{\text{shell}} = -M$. The dotted region shows the collapsing matter forming the black hole. H^+ is a true horizon. H'^+ is the null line $r = r_g = 2M$ ("old horizon"). The region lying inside H^+ (shaded by horizontal lines) is invisible from future null infinity $\mathcal{J}^+$. The region lying between H^+ and H'^+ (shaded by vertical lines) becomes visible from $\mathcal{J}^+$ as a result of the negative mass that falls in.

even this small part of the depths of the black hole is its complete destruction.

Does this mean that it is impossible to see what happens inside the apparent horizon without a destructive intervention? We show that such a possibility exists [Frolov and Novikov (1990, 1993a)]. In particular, we discuss a gedanken experiment which demonstrates that traversible wormholes (if only they exist) can be used to get information from the interior of a black hole practically without changing its gravitational field.

5.8.2 Wormholes in the Schwarzschild geometry

Wormholes will be considered in detail later in Chapter 16. Here, we shall describe only those features of traversible wormholes which are required for our gedanken experiment.

In a flat three-dimensional space a wormhole topology arises if one cuts two balls of radius b from the space and identifies their boundaries σ_a ($a = 1, 2$). In this model the jump of the extrinsic curvature at the sphere σ_a means that there is δ-like mass distribution at σ_a. In order to make the geometry smooth, one must assume that there exists a smooth distribution of matter in the handle. A quasi-stationary (or traversible) wormhole is possible only if this matter violates the average weak energy condition (see Chapter 16). The entrances (or mouths) of a wormhole can be located at arbitrarily large distances from one another as measured in the external surrounding space, while the distance between them through the handle can be arbitrarily small. This situation is of the most interest for our purpose. The mouths of a wormhole affected by external forces move like normal physical bodies (in the same manner as black holes). For example, the mouths in an external gravitational field will move along geodesics (free fall). We should also determine the identification of moments of proper times on the boundaries σ_a at some initial moment. The subsequent correspondence of proper time on the boundaries of the balls is uniquely determined and depends on their motion [see Frolov and Novikov (1990)].

Let us suppose now that in the black hole exterior there exists a traversible wormhole. If the mouths of the wormhole are at rest or move so as remain in the exterior region of the black hole, the event horizon is only slightly deformed by the gravitational action of the wormhole. We analyze now what happens when one or both mouths of the wormhole fall into a static (non-rotating) black hole. In what follows we consider only the case when the radii of the mouths of the wormhole and their masses are much smaller than the corresponding parameters of the black hole, and we describe the free fall motion of the mouths by specifying the geodesic lines ($x_a^\mu = x_a^\mu(T)$, $a = 1, 2$), representing their "center of mass" motion. The parameter T is a proper time along the geodesics. For simplicity, we assume that both mouths are freely falling with zero angular momentum and zero velocity at infinity. We also assume that the clocks are synchronized along a closed path passing through the wormhole initially. For such a parabolic motion, the position of a mouth a is described by the equation

$$r_a(T) = \left[r_{0,a}^{3/2} - \frac{3}{2} r_g^{1/2}\, T\right]^{2/3}, \qquad r_g = 2M\,, \tag{5.8.1}$$

where M is the mass of the black hole, and $r_{0,a}$ is the radial coordinate of mouth a at the moment $T = 0$.

For the description of the gravitational field of a black hole we shall use, besides the metric in the standard Schwarzschild coordinates

$$ds^2 = -A\, dt^2 + \frac{dr^2}{A} + r^2\, d\omega^2\,,$$

$$A = 1 - \frac{r_g}{r}, \quad d\omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2, \tag{5.8.2}$$

its form in the Lemaître coordinates (2.4.3)

$$ds^2 = -\,dT^2 + \frac{dR^2}{B} + B^2\, r_g^2\, d\omega^2, \qquad B = \left[\frac{3(R-T)}{2r_g}\right]^{2/3}. \tag{5.8.3}$$

The Lemaître coordinates (T, R) and the Schwarzschild coordinates (t, r) are related by equations (2.4.4) and (2.4.5). The mouths are at rest with respect to the Lemaître coordinates. It is worthwhile to mention that in the case where the mouths are at rest in a static gravitational field, the time gap necessarily grows so that eventually closed timelike curves are formed [Frolov and Novikov (1990]. But the time gap vanishes for free fall of the mouths, even in the case when $R_1 \neq R_2$.

The presence of a wormhole changes the spacetime topology. This means that even in the case when the mouths are at large distances from the black hole and their masses are negligible, the geometry differs from the original geometry of the black hole. When the wormhole's mouths are falling into the black hole, the background geometry is disturbed in addition to this obvious topology change. It is possible to show that when the radius L of the curvature of the background geometry ($L \sim r^{3/2}/r_g^{1/2}$) in the vicinity of the mouths remains much larger than the size b of a mouth and the masses of the mouths are negligibly small, the additional disturbance of the geometry remains small [see Frolov and Novikov (1993a)].

5.8.3 Causal structure of spacetime with a wormhole

If a wormhole connects two distant regions in a flat spacetime, then the light signal sent through one of the mouths into the wormhole re-emerges practically at once from the other mouth. The same is true also when the mouths are moving in an arbitrary curved spacetime. The formally calculated velocity of signal propagation in the external space (the distance between the mouths in the external space divided by the external time difference between the sending of light into one mouth and its re-emergence from the other) can be arbitrary large. The possibility of information exchange via a wormhole between events which are separated by a spacelike interval in the external space may result in changes of the global causal structure of the spacetime in the presence of a wormhole. Especially interesting possibilities arise when a wormhole is placed in a spacetime which originally had non-trivial causal structure. In particular, if one of the mouths of a wormhole is inside the event horizon of the black hole, while the other is in the exterior region, one can obtain information from the "black hole interior". We analyze now how the causal structure of the black hole spacetime is changed.

Consider freely falling mouths (with parabolic velocities). We start by assuming that the size b of the mouths is negligibly small and denote by γ_1 and γ_2 the world lines of the mouths. In the Lemaître coordinates they are described by the equations $R = R_1$ and $R = R_2$, respectively. Points which have the same time coordinate T at γ_1 and γ_2 are to be identified. As we already mentioned, no time gap for the clocks synchronization arises, and no closed timelike curves are created in the spacetime under consideration. If mouth 1 reaches the singularity $r = 0$ at the same moment $T = R_1$ when mouth 2 crosses the gravitational radius $r = r_g$, then $R_2 - R_1 = \frac{2}{3} r_g$. For $R_2 - R_1 < \frac{2}{3} r_g$ mouth 2 crosses the gravitational radius before $T = R_1$, and for $R_2 - R_1 > \frac{2}{3} r_g$ it crosses later than this moment. In the latter case the naked singularity would be seen from $\mathcal{J}^+$. To exclude this in what follows, we consider only the case $R_2 - R_1 < \frac{2}{3} r_g$.

Denote the exterior of a static black hole (i.e., the set of events which can be seen from infinity) in the absence of a wormhole by $P \equiv J^-(\mathcal{J}^+)$. The event horizon $H'^+ = \dot{P}$ in this case coincides with the Schwarzschild surface. In the presence of a wormhole the set of events which can be seen from infinity is modified and becomes $P_\gamma \equiv P \cup J^-(\gamma_1)$. The event horizon will be also modified and becomes $H^+ = \dot{P}_\gamma$.

In order to describe the structure of the modified horizon H^+, we consider the propagation of null rays in the Schwarzschild geometry. We use Eddington-Finkelstein coordinates (v, r, θ, ϕ) in which the line element (5.8.2) takes the form

$$ds^2 = -A\,dv^2 + 2\,dv\,dr + r^2\,d\omega^2\,, \qquad A = 1 - \frac{r_g}{r}\,. \tag{5.8.4}$$

We assume that mouth 1 is moving along the line $\theta = \pi/2, \phi = 0$ in the equatorial plane and consider the past null cone with vertex on the trajectory of mouth 1. For symmetry reasons, it is sufficient to consider null rays (generators of the cone) lying in the equatorial plane $\theta = \pi/2$. The following system of equations describes the rays:

$$\begin{aligned} &-A\,\dot{v} + \dot{r} = \varepsilon\,, \quad r^2\,\dot{\phi} = \tilde{\beta}\,, \\ &-A\,\dot{v}^2 + 2\,\dot{v}\,\dot{r} + r^2\,\dot{\phi}^2 = 0\,. \end{aligned} \tag{5.8.5}$$

The dot in these equations means differentiation with respect to the affine parameter λ. The constants ε and $\tilde{\beta}$ are the integrals of motion. We use the ambiguity in the choice of the affine parameter in order to set $\dot{r} \geq 0$, $\varepsilon = \pm 1$. For this choice we have

$$\begin{aligned} &\dot{r} = \sqrt{1 - (A\,\tilde{\beta}^2/r^2)}\,, \quad \dot{\phi} = \tilde{\beta}/r^2\,, \\ &\dot{v}_\varepsilon = A^{-1}\left[\sqrt{1 - (A\,\tilde{\beta}^2/r^2)} - \varepsilon\right]. \end{aligned} \tag{5.8.6}$$

Here v_ε means that v corresponds to the rays with the chosen parameter ε. Define

$$x = r/r_g\,, \qquad \beta = \tilde{\beta}/r_g\,, \qquad V_\varepsilon = -v_\varepsilon/r_g\,. \tag{5.8.7}$$

Then by eliminating the affine parameter, we get the system of equations

$$\frac{d\phi}{dx} = \frac{\beta}{F}\,, \qquad \frac{dV_\varepsilon}{dx} = \frac{x(F - \varepsilon x^2)}{(1 - x)F}\,, \tag{5.8.8}$$

where

$$F = \sqrt{x^4 + \beta^2(x - x^2)}\,. \tag{5.8.9}$$

We assume that at the moment when mouth 2 crosses the Schwarzschild surface H'^+ mouth 1 is at point p_1 with radial coordinate $r_0 = r_g\,x_0$. We also choose the advanced time coordinate v so that at p_1 one has $v = 0$. The null generators are specified by the parameter $\beta \in (-\infty, \infty)$. For a chosen value of β a null generator is defined by elliptic integrals

$$\phi = \phi(x; x_0, \beta) \equiv \int_{x_0}^{x} \frac{\beta}{F}\, dx\,, \tag{5.8.10}$$

$$V_\varepsilon = V_\varepsilon(x; x_0, \beta) \equiv \int_{x_0}^{x} \frac{x(F - \varepsilon x^2)}{(1-x)F}\, dx\,. \tag{5.8.11}$$

The parameter $\varepsilon = \pm 1$ distinguishes two families of rays. Denote by Q the rays with $\beta = \pm\infty$ which separate the two families. The functions $\phi(x; x_0, \beta)$ and $V_\varepsilon(x; x_0, \beta)$ possess the properties

$$\phi(x; x_0, 0) = 0\,, \tag{5.8.12}$$

$$\phi(x; x_0, \pm\infty) = \pm 2\left(\arcsin\sqrt{x} - \arcsin\sqrt{x_0}\right)\,, \tag{5.8.13}$$

$$V_\varepsilon(x; x_0, \pm\infty) = \frac{1}{2} V_{-1}(x; x_0, 0) = \ln\left(\frac{1 - x_0}{1 - x}\right) - (x - x_0)\,. \tag{5.8.14}$$

By using these relations, it is easy to obtain the equations

$$x = \sin^2(\psi_0 \pm \phi/2)\,, \tag{5.8.15}$$

$$V = 2\ln\left[\frac{\cos\psi_0}{\cos(\psi_0 \pm \phi/2)}\right] \pm \sin(2\psi_0 \pm \phi/2)\sin(\phi/2) \tag{5.8.16}$$

which describe the rays Q. Here $\sin\psi_0 = (x_0)^{1/2}$.

It is easy to see that only the rays of the family $\varepsilon = 1$ when traced back can cross the Schwarzschild surface $r = r_g$. The parametric equation $\phi = \phi(V)$ of the line Γ_1 where these rays cross the Schwarzschild surface is

$$\phi = \phi(1; x_0, \beta)\,, \qquad V = V_1(1; x_0, \beta)\,. \tag{5.8.17}$$

In particular, these relations show that for large values of V the line Γ_1 asymptotically tends to the null generators of H'^+ defined by the equation

$$\phi_\pm \equiv \phi(1; x_0, \pm\infty) = \pm[\pi - 2\arcsin\sqrt{x_0}]\,. \tag{5.8.18}$$

By using (5.8.10), we get

$$\frac{\partial\phi(1; x_0, \beta)}{\partial\beta} = \int_{x_0}^{1} \frac{x^4\, dx}{F^3} > 0\,, \tag{5.8.19}$$

and hence ϕ at the curve Γ_1 grows monotonically with the parameter β from ϕ_- till ϕ_+. In the limit when x_0 tends to 0, the asymptotic equation of Γ_1 takes the form $\phi = \pm\pi$.

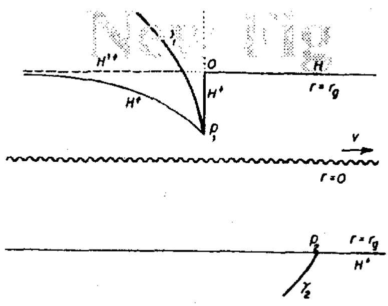


Figure 5.20: The structure of the $(\theta = \pi/2)$-section of the event horizon H^+ in the case where a wormhole falls into a black hole (the angular coordinate ϕ is not shown). Mouth 2 reaches the Schwarzschild surface at point p_2 before mouth 1 reaches the singularity $r = 0$. Dotted lines represent rays which enter the caustic from the exterior before becoming null generators of the horizon.

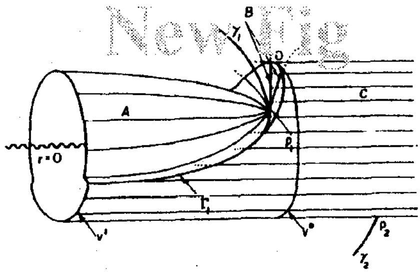


Figure 5.21: The same spacetime as in Figure 5.20 with the angular coordinate ϕ added. The sections V^0 represents the section $V = 0$ of the event horizon which passes through the point p_1, and V^1 represents the black hole surface long before the fall of a wormhole.

Figure 5.20 illustrates the structure of the event horizon H^+ in the case when a wormhole falls into a black hole. Mouth 2 reaches the Schwarzschild surface at point p_2 before mouth 1 reaches the singularity $r = 0$ (event p_1). To make this and following pictures more transparent, we assume that mouth 2 is moving along the line $\phi = \pi$ (i.e., from the side of the black hole diametrically opposite to mouth 1). Events p_1 and p_2 are simultaneous in the Lemaître frame. That is why they have the same value of the time T in the Lemaître coordinates but different values of the advanced time v. The region between H^+ and H'^+ in Figure 5.20 becomes visible from infinity through the wormhole.

The spacetime picture (in the section $\theta = \pi/2$) for the same case is shown in Figure 5.21 . The event horizon is formed by three families (A, B, C) of null generators. The family A is formed by the rays with $\varepsilon = 1$ and $\beta \in (-\infty, \infty)$. The family B is formed by rays with $\varepsilon = -1$ and $\beta \in (-\infty, \infty)$. The family C is formed by the rays

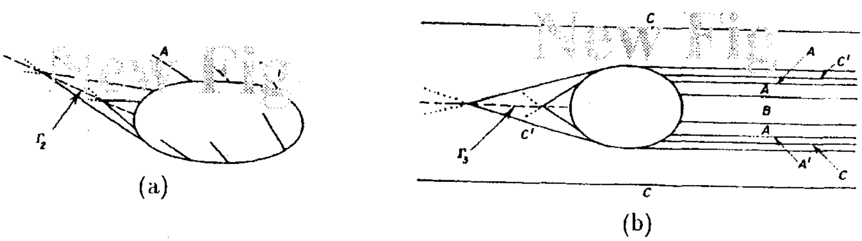


Figure 5.22: The pictures schematically show the local structure of null generators in the vicinity of mouth 1 (a) and mouth 2 (b). Γ_2 shown in (a) is a caustic. In (b) a part of the Schwarzschild surface is shown, and the rays of the subfamily C'' to the left of the mouth and the caustic Γ_3 are shown on the projection on this surface.

propagating along the Schwarzschild surface. The line Γ_1, which is the intersection of families B and C, is a caustic on the Schwarzschild surface . The null generators of B and C when traced back in time leave the horizon at Γ_1 and enter P_γ.

Up to now the size b of the mouths was supposed to be negligibly small. Now we take into account the finiteness of the value of b. First of all, we consider rays propagating in the wormhole geometry. The rays can be separated into four classes. Rays which pass near a mouth at significant distances and practically do not scatter belong to the first class. Closer rays which deflect appreciably and may even make a few circles before moving away in the same space belong to the second class. The analogous rays which move nearer to the mouth and enter through the handle into the second space with appreciable deflection belong to class 3. The latter two classes are formed by rays which approach the mouth almost tangentially. And finally, there are also rays which enter the mouth at a significant angle to its surface and exit into the second space. These rays belong to class 4.

The finiteness of the mouths' sizes requires a more detailed description of the horizon. To be concrete, we assume an identification of the mouths points such that the hemisphere of mouth 1 nearest to the black hole is identified with the nearest hemisphere of mouth 2. (For other identifications the picture of the continuation of the rays through the wormhole will obviously be different. Nevertheless, the main properties of the horizon remain the same.) One can demonstrate that in addition to families A, B, and C of the horizon generators shown in Figure 5.21 there are two additional families of the horizon generators A' and C'. They are shown in Figure 5.22. The generators of these families enter the horizon at caustics.

Let us trace back in time the generators of the horizon belonging to different families. The B-generators enter mouth 2 (see Figure 5.22b), exit from mouth 1 (see Figure 5.22a), and leave the horizon at the caustic Γ_1 (see Figure 5.21). The A-generators enter mouth 2, exit from mouth 1, and go to past infinity (see Figure 5.21). The A'-generators enter mouth 2 almost tangentially, exit from mouth 1 and leave the horizon at the caustic Γ_2 (Figure 5.22a). The C'-generators after scattering on

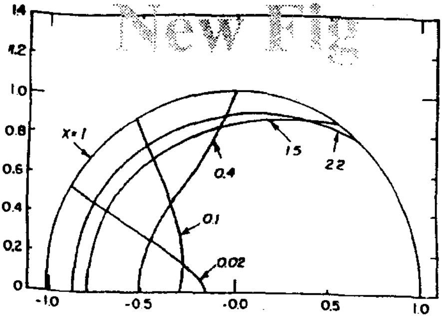


Figure 5.23: The sections of the event horizon by null surfaces V = const for $r_0 = 0.1r_g$ and the values of V =0.02, 0.1, 0.4, 1.5, and 2.2.

mouth 2 leave the horizon at caustic Γ_3 (Figure 5.22b). Those C-generators which cross caustic Γ_1 leave the horizon at these points, while the other C-generators (which never cross Γ_1) go to past infinity.

The area of the horizon surface $\mathcal{A}$ considered as a function of v decreases until it reaches a minimal value $\mathcal{A}_{\min}$ at $v = v_{min}$. After this it returns to the value $\mathcal{A}_0 = 4\pi r_g^2$. The value $\mathcal{A}_{\min}$ becomes smaller when r_0 tends to zero. Numerical calculations show that the absolute minimum (for point-like mouths) is reached at $v_{\min} \simeq -0.5r_g$ and $\mathcal{A}_0 - \mathcal{A}_{\min} \simeq \mathcal{A}_0/8$.

5.8.4 Energy and information extraction

The possibility of observing through a wormhole a spacetime region inside the Schwarzschild surface allows one to get information about these regions and extract bodies and energy from there. This means that one can use a wormhole as a device not only to "look inside a black hole", but also to return objects back safely into the exterior space after their traveling into the black hole interior.

In order to avoid possible confusions, we emphasize that the spacetime region lying inside the new modified horizon described above cannot be seen from infinity and, according to the formal definition, is a black hole. The infall of a wormhole results in temporary shrinkage of the horizon, while the gravitational field outside the mouths during this process remains practically the same. This means that the infall of a wormhole does not change the gravitational field of a black hole, and hence a black hole considered as a physical object remains the same. By looking through a wormhole, one can see what happens inside this physical object (i.e., in the region lying inside the Schwarzschild surface, or apparent horizon).

It is quite interesting that one can also use an infalling wormhole to extract energy by using the enormous gravitational field in the interior of the "former" black hole. For example, consider the following process. A particle of mass m falls freely in the gravitational field of a black hole and meets mouth 1 of a wormhole at a point p inside the gravitational radius . After this the particle goes through the wormhole and re-appears from mouth 2 at the point $\tilde{p}$ outside the black hole. It can be shown (see Frolov and Novikov (1993a)) that the energy $\tilde{E}$ of the outgoing particle can be much larger than its initial energy E. For example, if the point p is close to $r = 0$, then

$$\tilde{E} \simeq \sqrt{m^2 + (J/r_0)^2}\,, \tag{5.8.20}$$

where r_0 is the value of r at point p, and J is the angular momentum of the particle.

Strictly speaking, this process is not the extraction of energy from the black hole. When a particle appears from mouth 2 with energy $\tilde{E}$ the mass of the black hole does not change at all because no energy passes through its horizon. It is the mass of mouth 2 which decreases by the value $\tilde{E}$ in this process.[17] Thus, a black hole plays the role of a "catalyst" for the extraction of energy from a wormhole.

We have shown that, by using a wormhole, one can drastically change the structure of the event horizon. Nevertheless, it should be stressed that no fundamental theorems concerning black hole physics are violated. The generators of the modified horizon are null geodesics which have no end points in the future and which enter the horizon from the black hole exterior at the caustic points.[18] Hence, the Penrose theorem is satisfied. The decrease of the area of the event horizon, which looks "paradoxical", does not mean a violation of the Hawking's area theorem because the matter in the handle of a falling traversible wormhole necessarily violates the weak energy condition.

But what is interesting that this drastic (not small!) change of the event horizon is achieved by wormholes which can be arbitrarily small and contain arbitrarily small amounts of matter violating the weak energy condition. The study of the black hole interior practically does not change its gravitational field. By using one wormhole, one can extract information from the interior of the black hole during a period of time of order r_g. After this the black hole "returns" to its "initial state". In this sense the above-described procedure of studying the black hole interior by means of wormholes could be called "non-demolition". It is this non-demolishing property of the process that distinguishes it from other possibilities of getting information from the inside of the Schwarzschild surface discussed at the beginning of this section. Moreover, one can increase the visible part of the black hole interior by using many wormholes falling along different radial directions and one can increase the time during which the interior is visible by continuing this operation of throwing in new wormholes long enough.

[17] We should remember that the mass of the black hole even increases when the mouth 1 falls down into it as well as when a particle m later crosses the horizon.

[18] Besides the caustics connected with the formation of a black hole, in our case there are also new ones which are connected with the fall of a wormhole.

Therefore, we have found that the event horizon is in some sense very unstable because by using an arbitrarily small amount of matter violating the weak energy condition, it can be deformed drastically.

Chapter 6

Stationary Black Holes

6.1 "Black Holes Have No Hair"

As a result of the lightlike nature of the event horizon enclosing the black hole, the horizon acts as a one-way membrane. Particles and radiation can cross the event horizon from outside and enter the black hole, but escape is forbidden for both particles and radiation. Hence, processes involving black holes are essentially irreversible.

A black hole, if left alone, eventually reaches a stationary state. The following standard arguments lead to this conclusion. Let collapse produce a black hole, with a nonequilibrium configuration of particles and fields in its exterior. This configuration will immediately begin to re-arrange itself. The re-arrangement is accompanied by the radiation of energy to infinity and into the black hole. As the fields and particles outside the black hole originally had finite energies, and since nothing compensates for energy radiated away or absorbed into the black hole, it is to be expected that this process dies down, and the black hole becomes stationary; that is, the geometry of spacetime around the black hole will gradually differ less and less from that of a stationary space possessing the Killing vector field $\xi^{\mu}_{(t)}$. If the black hole is produced in a collapse slightly deviating from spherical symmetry, this conclusion is supported by the results described in Chapter 4. It is also confirmed by the theorem on the stability of stationary black holes with respect to small perturbations [see Section 4.9 and also Price (1972a, b), Wald (1979a, 1980)] and by the property [proved by Chandrasekhar and Detweiler (1975a), Detweiler (1977, 1979), and Chandrasekhar (1983)] that the gravitational perturbations in the spacetime of a stationary black hole cannot have eigenmodes with purely real frequency (zero damping) (see also Chapter 4).

It is not difficult to verify that the equilibrium of a given physical field in the neighborhood of the surface of a stationary black hole implies stringent constraints on the admissible configurations of the field [Israel (1971)]. For the sake of simplification, consider a non-rotating black hole described by the Schwarzschild metric

$$ds^2 = -F\,dt^2 + \frac{dr^2}{F} + r^2\,d\omega^2, \quad F = 1 - \frac{2M}{r}. \tag{6.1.1}$$

Let $T^{\mu\nu}$ be the energy-momentum tensor describing the physical field or medium in the neighborhood of this black hole. This system can be in equilibrium if the "weight" of the field or medium in each element of volume is exactly balanced out by the "buoyancy" due to the action of the components of the energy-momentum tensor (describing the tension) on the surface enclosing the chosen volume. In the local limit, this peculiar version of the "Archimedean principle" reduces to the equation (a conservation law)

$$T^{\nu}_{\ \mu;\nu} = 0 \tag{6.1.2}$$

supplemented with conditions

$$\partial_t T^{\mu\nu} = 0 \quad \text{(staticity)}, \qquad T^{rt} = T^{t\theta} = T^{t\phi} = 0 \quad \text{(zero fluxes)}. \tag{6.1.3}$$

Under these conditions, equation (6.1.2) holds identically if $\mu \neq r$; if $\mu = r$, it yields the relation

$$\frac{1}{2}\frac{\partial_r F}{F}(T^t_t - T^r_r) = \frac{1}{r^2}\,\partial_r(r^2\,T^r_r) + \\ + \frac{1}{\sin\theta}\,\partial_\theta(\sin\theta\,T^\theta_r) + \partial_\phi(T^\phi_r) - \frac{1}{r}\,(T^\theta_\theta + T^\phi_\phi). \tag{6.1.4}$$

Consider the tensor T^μ_ν, which satisfies conditions (6.1.3) and is regular on H^+. It is readily shown (e.g., by converting the coordinates to the Eddington-Finkelstein coordinates (v, r, θ, ϕ) with $v = t + r_*$, which are regular at the horizon H^+) that its components in the coordinates (t, r, θ, ϕ) as well as the expression $F^{-1}(T^t_t - T^r_r)$ are finite at $r = 2M$.[1] Otherwise, the configuration cannot be static. What happens if this condition is violated? The component T^t_r is then inevitably nonzero, producing a flux of field energy across the horizon: Energy flows until the field re-arranges and reaches a sustainable equilibrium. The characteristic time of this process is of order $t \sim r_g/c$.

Wheeler summarized the results of a large number of papers devoted to the final states of the black holes and formulated a conjecture that in its evolution to the stationary state, an isolated black hole sheds through radiation all those characteristics that radiation can remove. Since for a spin-s massless boson field the radiation is connected with changing the multipole moment l of the system (provided $l \geq s$), Wheeler conjectured that a stationary black hole resulting from a collapse of neutral

[1] In a spherically symmetric spacetime the condition $T^t_t = T^r_r$ is necessary and sufficient for the generalized Birkhoff's theorem to hold (see Appendix B.3). Obviously, this condition holds e.g., in the Reissner-Nordström and de Sitter spaces.

matter which interacts only gravitationally ($s = 2$) is described by a metric having only two free parameters: the mass ($l = 0$) and the angular momentum ($l = 1$). This conclusion was also reached by Doroshkevich *et al.* (1965) and Novikov (1969). If the collapsing matter was electrically charged, the stationary metric that finally emerges is described unambiguously by fixing three parameters: M, J, and Q (the electric charge).[2] The absence of the (non-monopole) magnetic field of a black hole was proved by Ginzburg (1964) and Ginzburg and Ozernoi (1964).

The above arguments show that Wheeler's conjecture is equivalent to the following proposition: Regardless of the specific details of the collapse or the structure and properties of the collapsing body, the resulting stationary black hole is described by a geometry specified by the parameters M, J, and Q. Wheeler expressed this property of stationary black holes in the following metaphorical form: "Black holes have no hair" [Ruffini and Wheeler (1971a)].

Wheeler focused on mass, angular momentum, and electric charge because all these quantities subject to a Gauss law and can be measured by a distant observer. One could define "hair" as any other configuration of gravitational and/or electromagnetic field associated with a black hole which for its description requires additional parameters (for example, multiple momentum). In this form the *no hair conjecture* is equivalent to the uniqueness theorem; that is, the statement that the only stationary black hole solution of Einstein-Maxwell equations is given by the Kerr-Newman metric or its special cases. We discuss the validity of this theorem in this chapter (see also a book by Heusler (1996a)].

It should be stressed that quite often the no hair conjecture is interpreted in a more general way. Namely, one defines "hair" as any field (not of gravitational or electromagnetic nature) associated with a black hole (for example, see [Bekenstein (1996)]). There are fields that do obey this conjecture (we discuss examples in Section 6.7). On the other, hand there are now many known black hole solutions with gauge fields (different from Maxwell's field), dilaton scalar field, and other fields. Such black holes naturally arise in theories that unify gravity with other interactions and we discuss them in Chapter 13. Thus, strictly speaking, the no hair conjecture in its widest interpretation certainly is not correct.

6.2 General Properties of Stationary Black Holes

6.2.1 Stationary spacetime with a black hole

Ample reasons exist for assuming that in the absence of external perturbations and with quantum effects ignored, the final state of any isolated black hole is stationary: hence, it is natural to begin the description of the properties of these final states with

[2] If magnetic monopoles exist, and the collapsing system possesses a magnetic charge, the value of this charge must be introduced into the description of the black hole as a fourth parameter.

an analysis of stationary black holes.

The stationarity of the spacetime signifies that one can introduce in it coordinates such that the coefficients of the metric are independent of the "time" coordinate. In "more geometric" terms, this means that the spacetime admits a one-parameter group of motions (isometries) whose generators are $\xi^\mu \partial_\mu$, where ξ^μ is the Killing vector field satisfying the equation

$$\xi_{(\mu;\nu)} = 0 . \tag{6.2.1}$$

As we want the spacetime not to change under "time-translation", it is logical to demand that the vector $\boldsymbol{\xi}$ be timelike ($\boldsymbol{\xi}\cdot\boldsymbol{\xi} < 0$) at least in some region. However, it is impossible to guarantee in the general case that $\boldsymbol{\xi}\cdot\boldsymbol{\xi}$ has the same sign throughout the entire spacetime. An asymptotically flat space will be said to be stationary if it admits a Killing vector field ξ^μ which is timelike in the neighborhood of $\mathscr{J}^+$ and $\mathscr{J}^-$.

The following two assumptions must be made in order to prove the basic propositions concerning general properties of stationary black holes:

1. The spacetime is regularly predictable.
2. The spacetime is either empty or contains fields described by hyperbolic equations and satisfying the energy dominance condition that for arbitrary timelike vectors ξ_1^μ and ξ_2^μ the energy-momentum tensor $T^{\mu\nu}$ of the field satisfies the inequality $T^{\mu\nu}\xi_{1\mu}\xi_{2\nu} \geq 0$.

Assumption 1, concerning the general causal structure of the spacetime and thoroughly discussed in the preceding chapter (see footnote to page 176), is largely technical. The energy dominance condition (which implies, e.g., the weak energy condition) signifies that an arbitrary observer finds the local energy to be non-negative and the local energy flux to be non-spacelike. Assumption 2 definitely holds for the electromagnetic field [for additional details, see Hawking and Ellis (1973)]. Throughout this chapter we assume that the constraints set above are satisfied.

In a stationary spacetime, the black hole area is independent of time. As a result, the convergence of light rays generating the event horizon is identically zero, and hence the apparent horizon coincides with the event horizon. Using relations (5.3.25), (5.3.27), and (5.3.28), it is readily ascertained that the weak energy condition ($\Phi \geq 0$) (which is certainly valid under the above assumed condition 2) and the constancy of the black hole area imply that the quantities σ, Φ, and Ψ vanish at the horizon surface

$$\sigma|_{H^+} = 0 , \qquad \Phi|_{H^+} = 0 , \qquad \Psi|_{H^+} = 0 . \tag{6.2.2}$$

The last two equalities can be interpreted as implying the absence of the fluxes of matter and physical fields ($\Phi = 0$) and of gravitational radiation ($\Psi = 0$) across the event horizon.

At each given instant of time, τ, each connected component of the horizon $\partial B(\tau)$ is compact and simply connected in the stationary spacetime (as it is in the general case). Furthermore, Hawking (1972a) has shown that in the stationary case the topology of the surface of any black hole is that of the two-dimensional sphere S^2. Black hole surface topologies distinct from S^2 are possible if the dominant energy condition is violated [Geroch and Hartle (1982)]. In principle, it cannot be excluded that a stationary spacetime contains several connected components $\partial B(\tau)$ and, correspondingly, several black holes "at rest". Such equilibrium is only possible if the gravitational attraction is compensated for by electromagnetic repulsion (or by repulsion due to forces of a different nature). For example, if there are several non-rotating black holes with masses m_i and charges Q_i satisfying the relation $m_i = Q_i/\sqrt{G}$, a system of black holes is in equilibrium [Hartle and Hawking (1972a), Ohta and Kimura (1982)].

In what follows we treat the case of a single stationary black hole, and restrict the analysis to the region of spacetime exterior to this black hole. In the general case, the entire spacetime of the stationary black hole may contain, together with the event horizon H^+ also the past event horizon $H^- = \dot{J}^+(\mathcal{J}^-)$ (this is readily verified for an eternal Schwarzschild black hole whose Penrose-Carter conformal diagram is shown in Figures 5.2 and 5.3). The region $J^+(\mathcal{J}^-) \cap J^-(\mathcal{J}^+)$ of spacetime outside H^- and H^+ is said to be the ***exterior region of a black hole***. A property inherent to events occurring in the exterior region is that causal curves can be traced which connect these events with both $\mathcal{J}^-$ and $\mathcal{J}^+$. It can be proved that in a stationary space the Killing vector field $\boldsymbol{\xi}$ is nonzero everywhere in the exterior region and on the part $H^+ \cap J^+(\mathcal{J}^-)$ of the event horizon [see, e.g., Hawking and Ellis (1973)].

6.2.2 Static black holes

For a more detailed description of the properties of stationary spaces, it is convenient to introduce the following differential invariant ω^α related to the Killing vector field ξ^μ by the formula

$$\omega^\alpha = \xi_{\mu;\nu}\, \xi_\lambda\, e^{\mu\nu\lambda\alpha}, \tag{6.2.3}$$

where $e^{\mu\nu\lambda\alpha}$ is a completely antisymmetric tensor. A stationary space is said to be *static* if $\omega^\alpha = 0$. By Frobenius theorem, the necessary and sufficient condition for the vector field ξ^μ to be orthogonal to a surface is that ω^α vanishes. In other words, if $\omega^\alpha = 0$, two scalar functions, α and t, can be found such that

$$\xi_\mu = \alpha\, t_{,\mu}\,. \tag{6.2.4}$$

In the region where $\xi_\mu \neq 0$, t can be used as one of the coordinates (the time coordinate); we supplement it with three other coordinates x^i. The coordinates x^i are conveniently chosen as follows. Let us fix an arbitrary surface $t = \text{const}$, define the coordinates x^i on it, extend these coordinates to the entire spacetime, demanding that they be constant along the integral curves of ξ^μ. In such coordinates, the metric of a static space is

$$ds^2 = -V^2\,dt^2 + h_{ij}\,dx^i dx^j. \tag{6.2.5}$$

It is readily shown, using the Killing equation (6.2.1), that $\partial_t h_{ij} = 0$, and that one can make use of the arbitrariness $t \to t' = f(t)$ in order to satisfy the equalities

$$\alpha = -V^2 = \xi_\mu \cdot \xi^\mu\,, \qquad \partial_t V = 0\,, \qquad \xi^\mu\,\partial_\mu = \partial_t\,. \tag{6.2.6}$$

The still remaining ambiguity in the choice of coordinates corresponds to the transformations

$$t \to t' = t + t_0\,, \qquad x^i \to x'^i = f^i(x^j)\,. \tag{6.2.7}$$

Note that since V and h_{ij} are independent of t, metric (6.2.5) is also invariant under time reversal $t \to -t$. The converse is also true; namely, any stationary metric possessing an additional time reversal symmetry $t \to -t$ is static.

An important property of static black holes is that the Killing vector field ξ^μ is timelike throughout their exterior region, while on the part $H^+ \cap J^+(\mathcal{J}^-)$ of the event horizon, which bounds the exterior region, ξ^μ is nonzero, lightlike, and directed along the generators of H^+. This last property is easily proved. Namely, from the equality $\xi_{[\mu;\nu}\,\xi_{\alpha]} = 0$ which follows from the condition $\omega^\alpha = 0$, and from relation (6.2.1) we obtain

$$2\,\xi_{\alpha;[\mu}\,\xi_{\nu]} = -\,\xi_{\mu;\nu}\,\xi_\alpha\,. \tag{6.2.8}$$

This equality immediately yields

$$(V^2)_{;[\mu}\,\xi_{\nu]} = -\,V^2\,\xi_{\mu;\nu}\,, \tag{6.2.9}$$

and hence the surface $V^2 = 0$ is lightlike because the direction of the normal to it, $(V^2)_{;\mu}$, coincides on this surface with that of the null vector ξ_μ. Since $(V^2)_{;\mu}$ and ξ_μ are parallel, we have

$$\frac{1}{2}\,(V^2)^{;\mu} \equiv \xi^{\mu;\nu}\,\xi_\nu = \kappa\,\xi^\mu\,, \tag{6.2.10}$$

and hence ξ^μ is a vector tangent to the null geodesic (to a generator of the surface $V^2 = 0$). These null geodesics, on the one hand, do not emerge on $\mathcal{J}^+$ because they always stay on the surface where $\xi_\mu\,\xi^\mu = 0$, while on $\mathcal{J}^+$, $\xi_\mu\,\xi^\mu = -1$. On the other hand, the divergence of the null generators on the surface $V = 0$ is zero. Hence, this surface is the outer trapped surface, and, at the same time, it is the event horizon because spacetime is stationary [Vishveshwara (1968)]. In the Section 6.3 these results will be generalized to the case of stationary (but not necessary static) black holes.

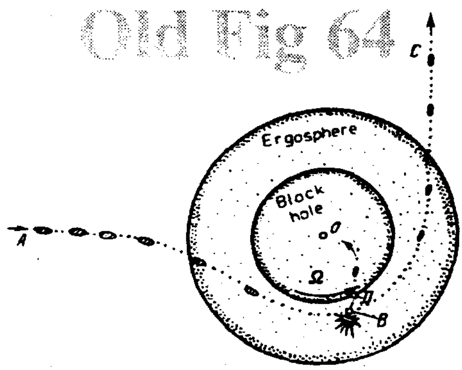


Figure 6.1: Penrose process. A body falling from a certain distance (position A) enters the ergosphere of a rotating black hole and explodes at a point B close to the black hole surface into two fragments. One fragment is absorbed by the black hole at a point D (the "explosion" parameters are chosen so as to give negative energy to this fragment). The other fragment is ejected from the ergosphere (point C) having an energy greater than the energy of the falling body.

6.2.3 Penrose process

If a stationary black hole is not static, the Killing vector field ξ^μ inevitably becomes spacelike in a part of the exterior region [Hawking and Ellis (1973)]. The region of the black hole exterior where $\xi^2 > 0$ is called the *ergosphere*.

The formation of the ergosphere outside a stationary nonstatic black hole has a number of important physical corollaries. They will be discussed in more detail in the chapters to follow. Here we consider one of them. Recall that by Noether's theorem, the symmetries of spacetime imply the laws of conservation of certain physical quantities. Thus, uniformity with respect to time implies the conservation of energy. For a particle moving in a stationary spacetime with a Killing vector field ξ^μ, this conserved quantity (energy ε) is written in the form (see Appendix A.8)

$$\varepsilon = -p^\mu \xi_\mu , \tag{6.2.11}$$

where p^μ is the four-momentum of a particle. For particles beyond the ergosphere, $\varepsilon \geq 0$ because p^μ is a future-directed timelike or null vector. However, some particles or light rays in the ergosphere may satisfy the reversed inequality $\varepsilon < 0$. Obviously, such particles can escape from the ergosphere only by dropping into the black hole. This is impossible unless the ergosphere has common points with the event horizon.

Because of the existence of states with negative energy ε in the ergosphere, the following mechanism of energy extraction from stationary black holes becomes possible [Penrose (1969)]. Imagine that a particle with momentum p_0^μ enters the ergosphere and there decays into a pair of particles with momenta p_1^μ and p_2^μ ($p_0^\mu = p_1^\mu + p_2^\mu$) so

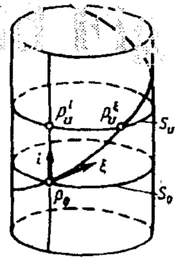


Figure 6.2: A stationary rotating black hole is axially symmetric. Illustration of the proof of Hawking's theorem.

that $\varepsilon_2 = -p_2^\mu \xi_\mu < 0$, and the particle with momentum p_1^μ escapes from the ergosphere (see Figure 6.1). The energy of the escaping particle, $\varepsilon_1 = -p_1^\mu \xi_\mu = \varepsilon_0 - \varepsilon_2$, is then greater than the energy of the incident particle, $\varepsilon_0 = -p_0^\mu \xi_\mu$; hence, energy can indeed be extracted in this process.

An essential difference between the properties of black holes with $\omega^\mu = 0$ and those with $\omega^\mu \neq 0$ is that, in a certain sense, nonstatic stationary black holes are rotating. The appearance of negative energies of particles moving in the field of a rotating black hole can be explained if we take into account the additional gravitational interaction of the angular momentum of this particle and the angular momentum of the rotating black hole. The energy imparted to the particle ejected in the Penrose process is drawn from the rotational energy of the black hole.

6.2.4 Stationary nonstatic black holes are axisymmetric

Having made these remarks, we can return to discussing the general properties of stationary black holes and consider the mutual arrangement of the ergosphere and the event horizon. In principle, it is possible that the ergosphere does not intersect the horizon, lying completely outside it. However, this situation appears to be unstable [Hawking and Ellis (1973)]. Therefore, we assume that the ergosphere does intersect the horizon in the stationary spacetime describing the final state of a black hole (for discussion of this point, see Carter (1987)). This means that the event horizon comprises points at which the Killing vector field ξ^μ is spacelike. Let us show that in this case the black hole is necessarily axially symmetric.

Let S_0 be the surface of a black hole at some moment of time. As we have mentioned earlier, the topology of a stationary black hole surface is that of a two-dimensional sphere, S^2. We denote by S_v those sections of the event horizon that are produced by translating the points of S_0 along the integral curves $x^\mu(v)$ of the field ξ^μ $(dx^\mu/dv = \xi^\mu)$ by an amount corresponding to the parameter v (Figure 6.2). Let a point $p_0 \in S_0$ transform thereby into a point $p_v^\xi \in S_v$. Denote by p_v^l the point of intersection of S_v and the generator of the event horizon passing through p_0. Since

l^μ (a tangent vector to a generator of the horizon) is null, while ξ^μ is spacelike, these vectors are not parallel, and the mapping $p_v^l \to p_v^\xi$ is a non-trivial transformation of S_v into itself. As a result of vanishing of the convergence ρ and shear σ at the event horizon of a stationary black hole, the spacing between any two points p_v^l and q_v^l on S_v is equal to the spacing between the corresponding points p_0 and q_0 on S_0. On the other hand, ξ^μ is the Killing vector field so that this spacing is also equal to the distance between p_v^ξ and q_v^ξ. Hence, the transformation $p_v^l \to p_v^\xi$ is a symmetry transformation that maps S_v into itself. The topology of S_v being that of the sphere S^2, all points of S_v, except two (the "poles"), are moved by the described isometry group along closed circular orbits. In other words, the surface of a stationary nonstatic black hole is axially symmetric.

If the metric describing a stationary black hole is analytic,[3] then the axial symmetry of the event horizon implies the axial symmetry of the entire spacetime. This result is formulated in the following

Theorem [Hawking (1972a)]. *Let the ergosphere in a stationary nonstatic space intersect the event horizon at $H^+ \cap J^+(\mathcal{J}^-)$. Then there exists a one-parameter cyclic group of isometries whose generators commute with ξ^μ and whose orbits are spacelike in the vicinity of $\mathcal{J}^+$ and $\mathcal{J}^-$.*

This theorem remains valid when the metric is non-analytic in isolated regions outside the horizon.

This theorem has the following very important consequence. Consider a rotating black hole surrounded by non-axially-symmetric distribution of matter. Since the gravitational field of the system is not axially symmetric, the black hole in the final state cannot have an angular momentum. This means that there exists a non-Newtonian gravitational interaction of a black hole with the surrounding matter. As a result of this interaction the black hole gradually transfers its angular momentum to the matter [Ipser (1971), Hawking (1972a), Press (1972), Hawking and Hartle (1972)].

To summarize the results outlined in this section: The final state of an isolated black hole is thus described by a stationary metric. The black hole either does not rotate and its metric is static, or it rotates and the spacetime possesses additional axial symmetry. In Sections 6.4 and 6.5 we shall prove the so-called *uniqueness theorems*, showing that stationary (both static and nonstatic) isolated black holes have relatively simple structure. Namely, we analyze the stationary solutions of the Einstein-Maxwell equations and prove that all such solutions that describe a stationary black hole reduce to the Kerr-Newman metric (6.5.29). In the case of no rotation, the solution reduces to the Reissner-Nordström solution. But before the discussion of the uniqueness theorems let us consider properties of the horizon of stationary black holes in more detail.

---

[3]On the analyticity of stationary axisymmetric asymptotically flat solutions of Einstein's equations, see Müller zum Hagen (1970).

6.3 Killing Horizon

6.3.1 Definition and properties of the Killing horizon

In a stationary spacetime with a black hole it is possible to introduce the notion of a *Killing horizon.* Let $\boldsymbol{\xi}$ be a Killing vector, then a Killing horizon is a null surface, H, to which the Killing vector field $\boldsymbol{\xi}$ is normal. This implies that ξ^μ is tangent to the null generators of H, and hence $\xi^2 \equiv \xi_\mu \xi^\mu = 0$ on H. Note that $\xi^2 = 0$ is a necessary but not a sufficient condition for the Killing horizon. For example, the equation $\xi_{(t)}^2 = 0$ in the Kerr spacetime defines the static limit surface which in the presence of rotation is timelike everywhere outside the poles.

Logically, the notion of Killing horizon is completely independent of the notion of event horizon, but there is a close connection between them. This connection is established by the following.

Theorem (see Hawking and Ellis (1973), proposition 9.3.6). *Let a stationary, asymptotically flat spacetime contain a black hole and be a solution of Einstein equations with matter satisfying suitable hyperbolic equations. Then the event horizon H^+ is a Killing horizon.*

In this section we describe some geometrical properties of Killing horizons [Boyer (1969), Wald (1984, 1992)]. First of all, we obtain some useful relations valid for an arbitrary Killing vector field. The Killing equation

$$\xi_{(\alpha;\beta)} = 0 \tag{6.3.1}$$

can be used to verify that any Killing vector ξ^α obeys the relation

$$\xi_{[\alpha}\, \xi^2{}_{;\beta]} = -3\, \xi^\mu\, \xi_{[\alpha}\, \xi_{\beta;\mu]} + \xi^\mu\, \xi_\mu\, \xi_{\alpha;\beta}\,. \tag{6.3.2}$$

Contracting this equation with $\xi^{\alpha;\beta}$, one gets

$$\frac{1}{2}\,(\xi^2)_{;\alpha}\,(\xi^2)^{;\alpha} = -3\,\xi^{[\alpha}\,\xi^{\beta;\gamma]}\,\xi_{[\alpha}\,\xi_{\beta;\gamma]} + \xi^\alpha\,\xi_\alpha\,\xi^{\beta;\gamma}\,\xi_{\beta;\gamma}\,. \tag{6.3.3}$$

We call a *Killing trajectory* (or *orbit*) the integral line $x^\mu = x^\mu(v)$ of the Killing vector field:

$$\frac{dx^\mu}{dv} = \xi^\mu(x)\,. \tag{6.3.4}$$

The Killing equation (6.3.1) implies

$$(\xi^2)_{,\alpha}\,\xi^\alpha \equiv 2\,\xi_{\beta;\alpha}\,\xi^\beta\,\xi^\alpha = 0\,. \tag{6.3.5}$$

Hence, ξ^2 is constant on any Killing trajectory; thus, one-dimensional Killing trajectories can be classified as spacelike, timelike, or null according to the sign or vanishing of ξ^2.

The following ***propagation equations*** for $(\xi^2)_{,\alpha}$ and ξ^α, respectively, are valid on a Killing trajectory [Boyer (1969)]

$$\frac{D(\xi^2{}_{,\alpha})}{dv} = \xi^2{}_{;\alpha\beta}\,\xi^\beta = (\xi^2{}_{,\beta}\,\xi^\beta)_{,\alpha} - \xi^2{}_{,\beta}\,\xi^\beta{}_{;\alpha} = -\,\xi^2{}_{,\beta}\,\xi^\beta{}_{;\alpha}\,, \tag{6.3.6}$$

$$\frac{D\xi^\alpha}{dv} = \xi^\alpha{}_{;\beta}\,\xi^\beta = -\frac{1}{2}(\xi^2)^{,\alpha}\,. \tag{6.3.7}$$

Relation (6.3.6) shows that $\xi^2{}_{,\alpha}$ vanishes either at all points or at no point of a Killing trajectory. (At a fixed point of the Killing trajectories, $\xi^2{}_{;\alpha} = 0$.) The equation (6.3.7) shows that a Killing trajectory is geodesic if and only if on it

$$\xi_{[\alpha}\,\xi^2{}_{,\beta]} = 0\,. \tag{6.3.8}$$

According to the definition, the Killing vector $\boldsymbol{\xi}$ on the surface H of a Killing horizon is normal to H. This implies that on H $\boldsymbol{\xi}$ obeys the condition[4]

$$\xi_{[\alpha}\,\xi_{\beta;\gamma]} \overset{H}{=} 0\,. \tag{6.3.9}$$

(Here and later we use the notation $\overset{H}{=}$ to indicate that the corresponding equality holds only on H.) Since $\xi^2 \overset{H}{=} 0$, the relations (6.3.2) and (6.3.9) imply

$$\xi_{[\alpha}\,\xi^2{}_{;\beta]} \overset{H}{=} 0\,. \tag{6.3.10}$$

Equation (6.3.10) shows that condition (6.3.8) is satisfied on the Killing horizon, and hence Killing trajectories lying on the horizon are geodesics. This property can be also proved if one takes into account that the Killing horizon by definition is a null surface, and hence its generators are null geodesics. On the other hand, ξ^μ is tangent on H to the generators so that Killing trajectories at the Killing horizon coincide with its generators.

6.3.2 Surface gravity

Relation (6.3.10) shows that

$$\xi^2{}_{,\alpha} \overset{H}{=} -2\kappa\,\xi_\alpha\,. \tag{6.3.11}$$

The quantity κ defined on H is called the *surface gravity*. Invariance of (6.3.11) under the isometries generated by $\boldsymbol{\xi}$ implies that κ is constant along each null generator,

[4] In the vicinity of a regular point of H one can choose coordinates (φ, x^i) so that $\varphi = 0$ is the equation of the surface H. In the vicinity of H one has $\xi_\beta = P\varphi_{;\beta} + Q\zeta_\beta$, where $P(0, x^i) \neq 0$, $Q(0, x^i) = 0$, and ζ_β is regular at H. This relation implies at H : $\xi_{[\beta;\gamma]} = P_{[;\gamma}\,\varphi_{\beta]} + \varphi_{[;\gamma}\,\zeta_{\beta]}(\partial Q/\partial\varphi)|_{(0,x^i)}$. By multiplying this expression by ξ_α with further antisymmetrization, we get the required relation.

i.e., $\xi^\alpha \nabla_\alpha \kappa = 0$. In order to show this, it is sufficient to apply the operator $\xi^\alpha \nabla_\alpha$ to the equation (6.3.11) and use the relations (6.3.6) and (6.3.7).

Equation (6.3.11) can be rewritten as

$$\xi^\nu \xi^\mu{}_{;\nu} \overset{H}{=} \kappa \xi^\mu , \tag{6.3.12}$$

which shows that κ measures the failure of the "Killing parameter time", v, to coincide with "affine parameter time", V, on H; namely, one has

$$V = e^{\kappa v} . \tag{6.3.13}$$

We recall that in the general case there is an ambiguity in the choice of the affine parameter $V \to V' = aV + b$ (with a and b constant). We fix the ambiguity by choosing $V = 1$ for $v = 0$ and $V = 0$ for $v = -\infty$. On the Killing horizon H one has

$$\xi^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial v} = \kappa V \frac{\partial}{\partial V} . \tag{6.3.14}$$

If $\kappa \neq 0$ and the null generators of H are geodesically complete,[5] then along each generator there exists a point where $\xi^\mu = 0$. The set of these points form a spacelike two-dimensional surface of bifurcation of the Killing horizon which we denote by S.

The surface gravity κ admits the following useful representation

$$\kappa^2 \overset{H}{=} -\frac{1}{2} \xi_{\alpha;\beta} \xi^{\alpha;\beta} . \tag{6.3.15}$$

To show this, we rewrite equation (6.3.9) in the form

$$\xi_\gamma \xi_{\alpha;\beta} \overset{H}{=} -2 \xi_{\gamma];\alpha} \xi_{\beta]} . \tag{6.3.16}$$

Contracting this relation with $\xi^{\alpha;\beta}$ and using (6.3.12), we get

$$\xi_\gamma \xi_{\alpha;\beta} \xi^{\alpha;\beta} \overset{H}{=} -2\kappa^2 \xi_\gamma , \tag{6.3.17}$$

which (for $\xi^\gamma \neq 0$) gives (6.3.15).

The surface gravity calculated for the Kerr-Newman black hole is

$$\kappa = \frac{r_+ - M}{r_+^2 + a^2} , \tag{6.3.18}$$

where $r_+ = M + \sqrt{M^2 - a^2 - Q^2}$.

[5] We recall that a geodesic is called *complete* if it has a representation in terms of an affine parameter V, which is defined for $-\infty < V < \infty$.

6.3.3 Constancy of surface gravity on H

The surface gravity κ is constant along null generators of the Killing horizon. The remarkable property of Killing horizons is that κ also does not vary from one generator to another, and κ is globally constant on H.

In the case where the generators of the Killing horizon are geodesically complete, one can prove that κ does not vary from generator to generator (and hence is globally constant) by using the following simple arguments (see e.g., Wald (1992)). If $\kappa \neq 0$ along a generator of the Killing horizon, then (at least locally in a neighborhood of this generator) there exists a two-dimensional bifurcation surface, S, on which $\boldsymbol{\xi} = 0$. The derivative of equation (6.3.15) on the bifurcation surface S in a direction s^μ tangent to S gives

$$\kappa\, s^\alpha\, \kappa_{;\alpha} = -\frac{1}{2}\, s^\gamma\, \xi_{\beta;\alpha\gamma}\, \xi^{\beta;\alpha} = \frac{1}{2}\, s^\gamma R_{\alpha\beta\gamma}{}^{\delta}\, \xi_\delta\, \xi^{\beta;\alpha} = 0\,. \tag{6.3.19}$$

Here we use relation (A.30) for the Killing vector and that ξ^α vanishes on S. The surface gravity κ is constant along the generators of the horizon, and hence the value of κ on the generator coincides with its value at the point of intersection of the generator with the bifurcation surface. Thus, if $\kappa \neq 0$ on any generator, then its value cannot vary from one generator to another. Note that this proof does not use the Einstein equations.

If one does not assume that the generators of H are geodesically complete,[6] then one can also prove that κ is constant on H [Bardeen *et al.* (1973)]. This proof, presented below, requires the use of the Einstein equations with matter obeying the energy dominance condition. Following Bardeen *et al.* (1973), it is convenient to use complex null tetrads. We augment $l^\mu \stackrel{H}{=} \xi^\mu$ to a complex null tetrad by choosing complex null vectors m^μ and $\bar{m}^\mu$ tangent to the horizon surface and normalized by the relation $m^\mu\, \bar{m}^\mu = 1$, and a real null vector n^μ orthogonal to m^μ and $\bar{m}^\mu$ and normalized by the condition $l^\mu\, n_\mu = -1$.

We assume that the null tetrad is continued outside the horizon, preserving the normalization conditions. Such a continuation is evidently non-unique, but the corresponding ambiguity is not important for our considerations. Using the complex null tetrad $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$, we can write κ in the following form:

$$\kappa \stackrel{H}{=} -\,l_{\nu;\mu}\, n^\nu\, l^\mu\,. \tag{6.3.20}$$

Here and later in the proof the following important property of the covariant derivatives along H (such as $l^\mu\nabla_\mu$, $m^\mu\nabla_\mu$, and $\bar{m}^\mu\nabla_\mu$) is used. Such covariant derivatives when applied to any tensor field depend only on the value of this tensor field on H.

[6]In a realistic situation a black hole is formed as a result of gravitational collapse, and a spacetime metric is time-dependent. The metric becomes stationary only some time after the collapse. If we are considering a region of stationarity for this problem, the generators of H being restricted to this region are incomplete.

Using this property and the relation $\xi^\mu \stackrel{H}{=} l^\mu$, we can rewrite equation (6.3.11) in the form

$$\kappa l_\nu \stackrel{H}{=} l_{\nu;\mu} l^\mu . \tag{6.3.21}$$

The equation (6.3.20) is obtained by multiplying this equation by n^ν.

Using (6.3.20), we obtain

$$m^\alpha \kappa_{;\alpha} \stackrel{H}{=} - l_{\nu;\mu;\alpha} n^\nu l^\mu m^\alpha - l_{\nu;\mu} n^\nu{}_{;\alpha} l^\mu m^\alpha - l_{\nu;\mu} n^\nu l^\mu{}_{;\alpha} m^\alpha . \tag{6.3.22}$$

The first of the terms on the right-hand side contains covariant derivatives only along H, and hence it depends only on the value of l^μ on H. This allows one to use the fact that l^μ on this surface coincides with the Killing vector field ξ^μ. Thus, we can use relation (A.30) for the Killing vector field to recast the first term on the right-hand side of (6.3.22) in the form $- R_{\alpha\beta\gamma\delta} l^\alpha m^\beta l^\gamma n^\delta$. Using (6.3.21) and the relation $l_\nu n^\nu{}_{;\alpha} = - n^\nu l_{\nu;\alpha}$, we rewrite the second term on the right-hand side of (6.3.22) in the form $\kappa l_{\nu;\alpha} n^\nu m^\alpha$. Let us now show that this term and the last one on the right-hand side of (6.3.22) cancel. First, we remark that the conditions of normalization of the null tetrad yield the relation

$$g^{\mu\beta} = - n^\beta l^\mu - l^\beta n^\mu + m^\beta \bar{m}^\mu + \bar{m}^\beta m^\mu . \tag{6.3.23}$$

Making use of this relation and the conditions of no shear or expansion of the event horizon surface,

$$\rho = - l_{\alpha;\beta} m^\alpha \bar{m}^\beta \stackrel{H}{=} 0 , \qquad \sigma = - l_{\alpha;\beta} m^\alpha m^\beta \stackrel{H}{=} 0 , \tag{6.3.24}$$

we rewrite the last term on the right-hand side of (6.3.22) in the form

$$\begin{aligned} - l_{\nu;\mu} n^\nu l^\mu{}_{;\alpha} m^\alpha &\stackrel{H}{=} - l_{\nu;\mu} n^\nu g^{\mu\beta} l_{\beta;\alpha} m^\alpha \\ &\stackrel{H}{=} l_{\nu;\mu} n^\nu l^\mu n^\beta l_{\beta;\alpha} m^\alpha \stackrel{H}{=} - \kappa l_{\nu;\alpha} n^\nu m^\alpha . \end{aligned} \tag{6.3.25}$$

This expression differs from the second term in (6.3.22) only in sign, and the two cancel out. As a result, we have

$$m^\alpha \kappa_{;\alpha} \stackrel{H}{=} - R_{\alpha\beta\gamma\delta} l^\alpha m^\beta l^\gamma n^\delta . \tag{6.3.26}$$

Note that on the horizon surface

$$l_{\alpha;\beta} m^\alpha{}_{;\gamma} \bar{m}^\beta m^\gamma \stackrel{H}{=} l_{\alpha;\beta} m^\alpha \bar{m}^\beta{}_{;\gamma} m^\gamma \stackrel{H}{=} 0 . \tag{6.3.27}$$

This can be confirmed using (6.3.23), (6.3.24), and the normalization conditions for the tetrad vectors. Therefore, in view of (A.30) and (6.3.24), we find

$$0 \stackrel{H}{=} - \rho_{;\gamma} m^\gamma \stackrel{H}{=} (l_{\alpha;\beta} m^\alpha \bar{m}^\beta)_{;\gamma} m^\gamma \stackrel{H}{=} R_{\epsilon\alpha\beta\gamma} l^\epsilon m^\alpha \bar{m}^\beta m^\gamma$$

$$= - R_{\alpha\beta}\, l^\alpha\, m^\beta + R_{\alpha\beta\gamma\delta}\, l^\alpha\, m^\beta\, l^\gamma\, n^\delta\,. \tag{6.3.28}$$

This relation and Einstein's equation $R_{\alpha\beta}\, l^\alpha\, m^\beta = 8\pi\, T_{\alpha\beta}\, l^\alpha\, m^\beta$ allow the rewriting of (6.3.26) in the form

$$m^\alpha\, \kappa_{;\alpha} \stackrel{H}{=} - 8\pi\, T_{\alpha\beta}\, l^\alpha\, m^\beta\,. \tag{6.3.29}$$

To complete the proof of the constancy of κ, we assume that the energy-momentum tensor $T_{\alpha\beta}$ satisfies the *energy dominance condition*; that is, $T_{\alpha\beta}\, l^\alpha$ is a non-spacelike vector for the null vector l^α (see Appendix A.15). If this condition is satisfied, the vector $T_{\alpha\beta}\, l^\alpha$ has to be null on the event horizon (the case of a timelike vector is excluded because at the horizon $T_{\alpha\beta}\, l^\alpha\, l^\beta \equiv \Phi \stackrel{H}{=} 0$; see (6.2.2)). Hence, $T_{\alpha\beta}\, l^\alpha\, m^\beta \stackrel{H}{=} 0$, and relation (6.3.29) proves that κ is constant at the horizon.

Let us discuss the physical interpretation of κ. Consider a stationary observer moving near a black hole such that the world line of this observer coincides with a Killing trajectory.[7] The velocity four-vector of this Killing observer is

$$u^\mu = \frac{\xi^\mu}{|\, \xi_\alpha\, \xi^\alpha\, |^{1/2}}\,. \tag{6.3.30}$$

Obviously, the motion of this observer is non-geodesic, his four-acceleration a^μ [cf. (A.51)] being equal to

$$w^\mu = - \frac{\xi^\mu_{;\alpha}\, \xi^\alpha}{\xi_\beta\, \xi^\beta}\,. \tag{6.3.31}$$

We use now the relation (6.3.3) which we rewrite in the form

$$3\, \xi^{[\alpha}\, \xi^{\beta;\gamma]}\, \xi_{[\alpha}\, \xi_{\beta;\gamma]} = \xi^\alpha\, \xi_\alpha\, \xi^{\beta;\gamma}\, \xi_{\beta;\gamma} - \frac{1}{2}\, (\xi^2)_{;\alpha}\, (\xi^2)^{;\alpha}\,. \tag{6.3.32}$$

Since $\xi_{[\alpha}\, \xi_{\beta;\gamma]} \stackrel{H}{=} 0$, the gradient of the left-hand side vanishes on the horizon. On the other hand, for $\kappa \neq 0$ the gradient of $\xi_\mu\, \xi^\mu$ does not vanish on H. That is why according to the l'Hospital rule the left-hand side of the equation (6.3.32) divided by ξ^2 must approach zero on the horizon [Wald (1984)]. Using this property and (6.3.15), after dividing both sides of the equation (6.3.32) by ξ^2 and taking the limit $\xi^2 \to 0$, one obtains

$$\kappa = \lim_{H}\ (\alpha\, w)\,, \tag{6.3.33}$$

[7]In the general case when a stationary black hole is rotating, the spacetime is axisymmetric and there exists more than one Killing vector. In this case the Killing horizon (which according to the Hawking's theorem coincides with the event horizon) is defined with respect to a special linear combination of the Killing vectors. It is this linear combination, which we denoted here by ξ^μ, that is used in the considerations. A stationary observer, located near the event horizon and moving along such a Killing trajectory, rotates at an angular velocity equal to that of the black hole.

Here we define $w = |w_\mu w^\mu|^{1/2}$ and $\alpha = |\xi_\beta \xi^\beta|^{1/2}$, and $\lim_H$ indicates taking the limit in which the point where the expression αw is calculated tends to the Killing horizon H.

For a non-rotating black hole, α is the redshift factor ($\alpha = \sqrt{-g_{tt}}$). Imagine that a body of mass m suspended on a weightless rigid string is at rest near the event horizon. In this case the string acts on the body with a force f^μ such that $f = |f_\mu f^\mu|^{1/2} = m\,w$. It can be shown that it is sufficient to apply the force $f_0 = m\,\alpha\,w$ to the other (distant) end of the string (see Section 2.2). The quantity $\alpha\,w$ can be interpreted as the acceleration of a body at rest in the vicinity of the black hole, as measured in the reference frame of a distant observer. In other words, the surface gravity κ characterizes the limiting "strength" of the gravitational field on the black hole surface, evaluated by a distant observer. If the black hole is rotating, κ has a similar meaning, the only difference being that the body rotates at the angular velocity of the black hole.

6.3.4 Angular velocity

For a Kerr-Newman black hole the angular velocity $\Omega^H = a/(r_+^2 + a^2)$ (just like the surface gravity κ) is constant on the event horizon. This property is still preserved if the black hole is surrounded by matter provided the spacetime geometry remains stationary. If a spacetime with a black hole is stationary (but not static), then according to the Hawking's theorem mentioned at the end of Section 6.2, it must be axisymmetric. We show that there is a natural generalization of the notion of the black hole angular velocity to this case and demonstrate that the angular velocity is constant on the horizon.

In an axisymmetric stationary spacetime, in addition to the Killing vector field $\xi^\mu_{(t)}$ which is normalized at infinity by the condition $\xi^\mu_{(t)} \cdot \xi_{(t)\mu} = -1$, there is also a spacelike Killing vector field $\xi^\mu_{(\phi)}$ which corresponds to the symmetry of space with respect to rotation. This field commutes with $\xi^\mu_{(t)}$ and has closed integral curves. The field $\xi^\mu_{(\phi)}$ is nonzero everywhere in the exterior region and at the horizon, except on the rotation axis, where $\xi^\mu_{(\phi)} = 0$. If we define $X = \xi^\mu_{(\phi)}\,\xi_{(\phi)\mu}$, the condition of regularity of spacetime at the rotation axis is satisfied when

$$\left.\frac{X^{;\alpha} X_{;\alpha}}{4X}\right|_{X=0} = 1\,. \tag{6.3.34}$$

This condition implies that at the rotation axis the spacetime is locally flat (i.e., there is no conical singularity).

The vector fields $\xi^\mu_{(t)}$ and $\xi^\mu_{(\phi)}$ possessing the properties described above, including the normalization (6.3.34), are uniquely defined in a stationary axially symmetric asymptotically flat space.

In such a space, we can introduce coordinates t, ϕ, x^X $(X = 1, 2)$ in which the metric takes the form

$$ds^2 = -V\,dt^2 + 2\,W\,d\phi\,dt + X\,d\phi^2 + 2\,g_{tX}\,dx^X\,dt$$
$$+ 2\,g_{\phi X}\,dx^X\,d\phi + \gamma_{XY}\,dx^X\,dx^Y\,, \tag{6.3.35}$$

and the functions V, X, W, g_{tX}, $g_{\phi X}$, and γ_{XY} are independent of t and ϕ. In these coordinates $\xi^\alpha_{(t)} = \delta^\alpha_t$, $\xi^\alpha_{(\phi)} = \delta^\alpha_\phi$, and

$$X = \boldsymbol{\xi}_{(\phi)} \cdot \boldsymbol{\xi}_{(\phi)}\,, \qquad W = \boldsymbol{\xi}_{(t)} \cdot \boldsymbol{\xi}_{(\phi)}\,, \qquad V = -\,\boldsymbol{\xi}_{(t)} \cdot \boldsymbol{\xi}_{(t)}\,. \tag{6.3.36}$$

Metric (6.3.35) is said to satisfy the *circularity condition* if coordinate transformations can be found that preserve the form (6.3.35) but have the coefficients g_{tX} and $g_{\phi X}$ vanish. In this case the two-dimensional surfaces (t = const, ϕ = const) are orthogonal to two-dimensional surfaces formed by the integral curves of the fields $\xi^\mu_{(t)}$ and $\xi^\mu_{(\phi)}$. The necessary and sufficient condition for a metric to satisfy the circularity condition is that the following relations are satisfied [see, e.g., Kramer *et al.* (1980)]:

$$e^{\alpha\beta\gamma\delta}\,\xi_{(\phi)\alpha}\,\xi_{(t)\beta}\,\xi_{(t)\gamma;\delta} = 0\,, \qquad e^{\alpha\beta\gamma\delta}\,\xi_{(t)\alpha}\,\xi_{(\phi)\beta}\,\xi_{(\phi)\gamma;\delta} = 0\,. \tag{6.3.37}$$

It can be shown [Kundt and Trümper (1966), Carter (1973a)] that these relations hold if and only if the Ricci tensor $R_{\alpha\beta}$ satisfies the conditions

$$\xi^\delta_{(t)}\,R_{\delta[\alpha}\,\xi_{(t)\beta}\,\xi_{(\phi)\gamma]} = 0\,, \qquad \xi^\delta_{(\phi)}\,R_{\delta[\alpha}\,\xi_{(\phi)\beta}\,\xi_{(t)\gamma]} = 0\,. \tag{6.3.38}$$

Obviously, these conditions are met for the vacuum solutions to Einstein's equations. It is readily verified that they also hold outside sources in electrovac spaces [Carter (1969)]. Therefore, the circularity condition is satisfied in the case of interest here (stationary axially symmetric solutions of the Einstein-Maxwell equations).

Define a bivector $\rho_{\mu\nu}$:

$$\rho_{\mu\nu} = 2\,\xi_{(t)[\mu}\,\xi_{(\phi)\nu]}\,. \tag{6.3.39}$$

Carter (1973a) demonstrated that if the circularity condition is met, the event horizon of an arbitrary stationary axially symmetric black hole coincides with the set of points at which the bivector $\rho_{\mu\nu}$ becomes null

$$\rho^2 \equiv -\frac{1}{2}\,\rho_{\mu\nu}\,\rho^{\mu\nu} = W^2 + VX = 0\,, \tag{6.3.40}$$

and the vector l^μ tangent to the event horizon coincides in direction with the null vector that lies in the two-dimensional plane generated by the vectors $\xi_{(t)}$ and $\xi_{(\phi)}$. Choosing the normalization l^μ in an appropriate manner, we find

$$l^\mu \stackrel{H}{=} \xi^\mu_{(t)} + \Omega^H\,\xi^\mu_{(\phi)}\,. \tag{6.3.41}$$

The quantity Ω^H which enters this relation is the angular velocity of the black hole. Since $l^2 \stackrel{H}{=} 0$, one has

$$\Omega^H \stackrel{H}{=} \frac{V}{W} \stackrel{H}{=} -\frac{W}{X}, \tag{6.3.42}$$

and $l_\mu \xi^\mu_{(t)} = l_\mu \xi^\mu_{(\phi)} = 0$.

As follows from spacetime symmetry, the angular velocity Ω^H of the black hole cannot be a function of either time t or angle ϕ:

$$\xi^\mu_{(t)} \Omega^H_{,\mu} = \xi^\mu_{(\phi)} \Omega^H_{,\mu} = 0. \tag{6.3.43}$$

Moreover, Ω^H (like κ and Φ) is also independent of the "latitude" of a point on the black hole surface; that is, it is constant everywhere on the horizon. For proof it is sufficient to remark that the event horizon of a stationary black hole coincides with the Killing horizon, and hence there exists a Killing vector field ξ^μ which at the horizon is tangent to its generators

$$\xi^\mu \stackrel{H}{=} \lambda \left[\xi^\mu_{(t)} + \Omega^H \xi^\mu_{(\phi)}\right]. \tag{6.3.44}$$

Using that ξ^μ, $\xi^\mu_{(t)}$, and $\xi^\mu_{(\phi)}$ obey the Killing equation, one derives from (6.3.44)

$$\Omega^H_{(,\nu} \xi_{(\phi)\mu)} = -\frac{1}{\lambda^2} \lambda_{(,\nu} \xi_{\mu)}.$$

Let s^ν be a vector tangent to H; then by multiplying this relation by $s^\nu \xi^\mu_{(\phi)}$ and using (6.3.43), one has

$$\xi^2_{(\phi)} s^\nu \Omega^H_{,\nu} \stackrel{H}{=} 0. \tag{6.3.45}$$

This proves that Ω^H is constant everywhere outside the axis of symmetry. Since Ω^H is continuous, it remains constant also at the axis of symmetry.

It is also possible to prove that Ω^H = const without using the fact that the event horizon coincides with the Killing horizon. Using (6.3.42), one has $X\,\Omega^H = -W$. If this relation is differentiated with respect to x^α and then the commutativity $[\boldsymbol{\xi}_{(t)}, \boldsymbol{\xi}_{(\phi)}] = 0$ is used, it is not difficult to obtain the following equality:

$$X^2 \Omega^H_{,\alpha} = 2\left(\xi^\beta_{(\phi)} W - \xi^\beta_{(t)} X\right) \xi_{(\phi)\beta;\alpha}. \tag{6.3.46}$$

Now we multiply both sides of this equality by $\rho_{\gamma\delta}$ and antisymmetrize it with respect to the indices α, γ, and δ. If we also take into account the relation

$$\xi_{(\phi)\alpha;[\beta} \rho_{\gamma\delta]} = \xi_{(\phi)\alpha} \xi_{(\phi)[\beta;\gamma} \xi_{(t)\delta]} - \xi_{(t)\alpha} \xi_{(\phi)[\beta;\gamma} \xi_{(\phi)\delta]}, \tag{6.3.47}$$

implied by circularity condition (6.3.37), and recall that the invariant $W^2 + XV \equiv \rho^2$ vanishes on the horizon, it can be shown that the right-hand side of expression (6.3.46) equals zero, so that

$$X^2\, \Omega^H_{[,\alpha}\, \rho_{\gamma\delta]} = 0\,. \tag{6.3.48}$$

Taken off the symmetry axis ($X \neq 0$), this condition implies that $\Omega^H_{,\alpha}$ lies in the two-dimensional plane spanned by the vectors $\xi_{(t)}$ and $\xi_{(\phi)}$, and relation (6.3.43) shows that $\Omega^H_{,\alpha} = 0$. The constancy of Ω^H on the event horizon is thereby proved. (At the poles, Ω^H is found from continuity.)

6.3.5 Electric field potential

In describing physical effects in the field of a charged black hole, the relevant expressions include, in addition to κ and Ω^H, another invariant quantity Φ^H; that is, the electric field potential on the black hole surface. This quantity, $\Phi^H = Qr_+/(r_+^2 + a^2)$, is constant on the horizon of a Kerr-Newman black hole. Let us show that this result is general and valid for any stationary (not necessarily isolated) black hole.

Let ξ^μ be a Killing field in a spacetime with a Killing horizon and $F_{\mu\nu}$ be the electromagnetic field tensor satisfying Maxwell's equations

$$F^{\mu\nu}{}_{;\nu} = 4\pi j^\mu\,, \qquad F_{[\mu\nu,\alpha]} = 0 \tag{6.3.49}$$

and obeying the symmetry condition

$$\mathcal{L}_\xi F_{\mu\nu} = 0\,. \tag{6.3.50}$$

Then it is readily shown that the vector $E_\mu = -\,F_{\mu\nu}\,\xi^\nu$ satisfies the condition

$$E_{[\mu;\nu]} = 0 \tag{6.3.51}$$

and therefore is the gradient of a function Φ

$$E_\mu = \Phi_{,\mu}\,. \tag{6.3.52}$$

Let us show that if A_μ is the vector potential of the field $F_{\mu\nu}$ satisfying the symmetry condition

$$\mathcal{L}_\xi A^\mu \equiv \xi^\alpha A^\mu{}_{;\alpha} - A^\alpha\, \xi^\mu{}_{;\alpha} = 0\,, \tag{6.3.53}$$

we can choose for Φ the quantity

$$\Phi = -\,A_\alpha\, \xi^\alpha\,. \tag{6.3.54}$$

Indeed, differentiating (6.3.54) and then using (6.3.53), we obtain

$$\Phi_{,\mu} = -\,A_{\alpha;\mu}\,\xi^\alpha - A_\alpha\,\xi^\alpha{}_{;\mu} = -\,A_{\alpha;\mu}\,\xi^\alpha + A_{\mu;\alpha}\,\xi^\alpha = E_\mu\,. \tag{6.3.55}$$

We also specify that the potential is chosen in such a manner that A_μ vanishes at infinity. The value Φ^H of the corresponding quantity Φ at the Killing horizon is called the *electric potential* of the black hole. We can show that Φ^H is constant on the horizon. To do this, we remark that the condition

$$T_{\mu\nu}\, l^\mu\, l^\nu \overset{H}{=} 0\,, \tag{6.3.56}$$

which is implied by (6.2.2) to hold on the surface of an arbitrary stationary black hole ($l^\mu \overset{H}{=} \xi^\mu$), is equivalent for the electromagnetic field to the relation

$$\left(F_\mu{}^\alpha\, F_{\nu\alpha} - \frac{1}{4}\, g_{\mu\nu}\, F_{\alpha\beta}\, F^{\alpha\beta} \right) l^\mu\, l^\nu \equiv E^\alpha\, E_\alpha \overset{H}{=} 0\,; \tag{6.3.57}$$

that is,

$$\Phi_{,\mu} = E_\mu \overset{H}{=} 4\pi\sigma\, l_\mu \tag{6.3.58}$$

on the horizon surface.[8] Hence, the following equality holds for any vector s^μ tangent to the event horizon:

$$s^\mu\, \Phi^H{}_{,\mu} \overset{H}{=} 0\,, \tag{6.3.59}$$

which means that the electric potential Φ^H at the event horizon is constant.[9]

6.4 Uniqueness Theorem for Static Black Holes

Let us discuss now the uniqueness theorems for stationary black holes. We begin with the uniqueness theorem for static black holes proved by Israel (1967) which was historically the first in the series of these results.

Consider static solutions of Einstein's vacuum equations. We choose the coordinates in the static spacetime as outlined in Section 6.2.2 and write the static metric in the form (6.2.5):

$$ds^2 = -\,V^2\, dt^2 + h_{ij}\, dx^i\, dx^j\,, \quad (i,j = 1,2,3)\,, \tag{6.4.1}$$

$$V = V(x^1, x^2, x^3)\,, \qquad h_{ij} = h_{ij}(x^1, x^2, x^3)\,.$$

Denote by ${}^{(3)}R_{ij}$ the Ricci tensor of three-dimensional space Σ described by the equation t = const and having the metric h_{ij}. Since the extrinsic curvature K_{ij} of Σ vanishes, the Gauss-Codazzi equations (A.76) and (A.78) imply

$${}^{(3)}R_{ij} = R_{ij} - E_{ij}, \qquad {}^{(3)}R = R - 2R_0^0\,,$$

[8] The quantity σ coincides with the "surface charge density" in the membrane paradigm [cf. formula (8.4.4)].

[9] When a charged rotating black hole is described in the framework of a five-dimensional Kaluza-Klein theory of gravitation, the quantities Ω^H and Φ^H enter the expressions in a similar manner, and their properties are to a certain extent similar [Bleyer *et al.* (1987)].

where ${}^{(3)}R = h^{ij}\,{}^{(3)}R_{ij}$, $R = g^{\mu\nu} R_{\mu\nu}$, and $E_{\alpha\beta} = R_{\mu\alpha\nu\beta}\,\xi^\mu \xi^\nu/\xi^2$. By using relation (A.30), one can verify that $E_{ij} = -V_{;ij}/V$. [( )$_{;i}$ denotes the covariant derivative in the metric h_{ij}.] These relations imply that in the static spacetime (6.4.1) Einstein's vacuum equations are equivalent to the following equations:

$$h^{ij} V_{;ij} = 0\,, \tag{6.4.2}$$

$$V_{;ij} - {}^{(3)}R_{ij}\, V = 0\,. \tag{6.4.3}$$

Assume that the spacetime with metric (6.4.1) obeys the following conditions:

1. It is asymptotically flat.
2. It has an event horizon.
3. It has no singularities on or outside the event horizon.

Considered in detail, these assumptions imply that:

1. The space Σ is asymptotically Euclidean; that is, coordinates x^i can be chosen such that

$$h_{ij} = \delta_{ij} + O(r^{-1})\,, \quad \partial_k h_{ij} = O(r^{-2})\,, \quad V = 1 - M/r + \eta\,, \tag{6.4.4}$$

$$M = \text{const}\,, \quad \eta = O(r^{-2})\,, \qquad \partial_i \eta = O(r^{-3})\,, \quad \partial_i\partial_j \eta = O(r^{-4})$$

as $r \equiv (\delta_{ij}\, x^i\, x^j)^{1/2} \to \infty$.

2. The function V has zeros on Σ, the set $V(x^i) = 0$ being a connected regular smooth surface. Rigorously speaking, the points at which $V = 0$, are not covered by the coordinates (t, x^1, x^2, x^3) because metric (6.4.1) has a coordinate singularity. The assumption that there exists a regular event horizon implies that a transition to new coordinates makes it possible to extend the metric to that part of the spacetime which contains the event horizon. The surface $V = 0$ can be treated as the boundary of Σ that arises as a result of the passage to the limit $V = \delta =$ const as $\delta \to +0$. The function V satisfies the elliptic equation $V_{;i}{}^{;i} = 0$, and hence is harmonic. Since $V = 1$ as $r \to \infty$, the values of V are positive and smaller than 1 at finite values of r outside the horizon [on the corresponding property of harmonic functions, see, e.g., Jano and Bochner (1953)].

3. The invariant $\mathcal{R}^2 = R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta}$ constructed from the four-dimensional tensor of curvature is finite everywhere on Σ (for $0 \le V < 1$).

The uniqueness theorem for static black holes: *Any static solution of Einstein's vacuum equations satisfying conditions (1)-(3) is spherically symmetric and coincides with the Schwarzschild metric.*

Israel (1967) proved this theorem under the following additional condition: (4) the equipotential surfaces $V = \text{const} > 0$ are regular simply connected two-dimensional closed surfaces. Later it was proved [see Müller zum Hagen, Robinson, and Seifert (1973), Robinson (1977)] that this condition which implies, among other things, that $V_{,\alpha} \neq 0$ everywhere in the interval $0 \leq V < 1$, follows from conditions (1)–(3).

The main steps of proving Israel's theorem are as follows. The function V ($V_{,\alpha} \neq 0$) is chosen as a coordinate. The remaining two coordinates θ^2 and θ^3 on the surfaces $V = \text{const}$ are chosen in such a manner that curves of $\theta^2, \theta^2 = \text{const}$ be orthogonal to the surfaces $V = \text{const}$. In these coordinates, metric (6.4.1) is written in the form

$$ds^2 = -V^2\, dt^2 + \rho^2\, dV^2 + b_{XY}\, d\theta^X\, d\theta^Y\,, \tag{6.4.5}$$

where $X, Y = 2, 3$; ρ and b_{XY} are functions of V, θ^2, and θ^3, and equation (6.4.2) takes the form

$$\partial_V \left(\frac{\sqrt{b}}{\rho}\right) = 0\,, \qquad b = \det(b_{XY})\,. \tag{6.4.6}$$

Now we define K_{XY} by the relation[10]

$$K_{XY} = \frac{1}{2\rho}\,\frac{\partial b_{XY}}{\partial V}\,. \tag{6.4.7}$$

For the trace of this tensor, $K = b^{XY} K_{XY}$, we can derive the expression $K = \rho^{-1}\,\partial\left(\ln\sqrt{b}\right)/\partial V$; in view of (6.4.6), this gives

$$\frac{\partial \rho}{\partial V} = \rho^2\, K\,. \tag{6.4.8}$$

It can be shown that equation (6.4.3) is equivalent to the following set of equalities

$$\left(\frac{\partial}{\partial V} + \frac{1}{V}\right) K_X{}^Y = -\rho_{|X}{}^{|Y} + \frac{1}{2}\,\rho\; {}^{(2)}\!R\, \delta_X^Y - \rho\, K\, K_X{}^Y\,, \tag{6.4.9}$$

$${}^{(2)}\!R = -K_{XY}\, K^{XY} + K^2 + \frac{2K}{\rho V}\,, \tag{6.4.10}$$

$$\partial_X \rho = \rho^2\, V\left(\partial_X K - K_X{}^Y{}_{|Y}\right)\,, \tag{6.4.11}$$

[10]This two-dimensional tensor is proportional to the extrinsic curvature of the surface $V = \text{const}$, and differs from the latter only in sign. Note also that the sign convention in the definition of the Ricci tensor chosen in Israel (1967) paper differs from the sign convention of this book.

where ${}^{(2)}R$ is the scalar curvature of the two-dimensional surface $V = \text{const}$, and $(\ )_{|X}$ stands for the covariant derivative in the metric b_{XY}. Equations (6.4.6), (6.4.7), and (6.4.9) make it possible to find how the unknown functions ρ, b_{XY}, and K_{XY} depend on V, and equations (6.4.10) and (6.4.11) act as constraints: If they are satisfied for one value of V, the other equations imply that they are satisfied for all V.

The next step is to find the conditions imposed on the unknown functions by assumption (3). First, we rewrite the invariant $\mathcal{R}^2 = R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta}$ in terms of ρ, b_{XY}, and K_{XY}:

$$\frac{1}{8}\,\mathcal{R}^2 = (V\rho)^{-2}\left(K_{XY}\,K^{XY} + \frac{2\,\rho_{|X}\,\rho^{|X}}{\rho^2} + K^2\right). \tag{6.4.12}$$

Equation (6.4.6) implies that $\sqrt{b} = c(x^Y)\rho$; hence, the regularity of the surface $V = 0$ implies that $\rho(V = 0, \theta^2, \theta^3) \neq 0$. The regularity of $\mathcal{R}^2$ for $V = 0$ now gives

$$K_{XY}(V = 0, \theta^2, \theta^3) = 0\,, \qquad \rho(V = 0, \theta^2, \theta^3) = \rho_0 = \text{const}\,,$$

$$\lim_{V\to 0}(K/V) = \frac{1}{2}\,\rho_0\,{}^{(2)}R(V = 0, \theta^2, \theta^3)\,. \tag{6.4.13}$$

If the black hole area $\int_{V=0}\sqrt{b}\,d\theta^2\,d\theta^3$ is denoted by $\mathcal{A}_0$, then by integrating (6.4.6) in V from 0 to 1, and taking into account the boundary conditions (6.4.4) and (6.4.13), we find

$$\mathcal{A}_0 = 4\pi M \rho_0\,. \tag{6.4.14}$$

This equality implies that M is always positive.

The following relations can be obtained from equations (6.4.6) and (6.4.8)–(6.4.10):

$$\frac{\partial}{\partial V}\left(\frac{\sqrt{b}\,K}{\sqrt{\rho}\,V}\right) = -\,\frac{2\sqrt{b}}{V}\left[{}^{(2)}\Delta(\rho^{1/2}) + \rho^{-3/2}\left(\frac{1}{2}\,\rho_{|X}\,\rho^{|X} + \Psi_{XY}\,\Psi^{XY}\right)\right], \tag{6.4.15}$$

$$\frac{\partial}{\partial V}\left[\frac{\sqrt{b}}{\rho}\left(KV + \frac{4}{\rho}\right)\right] = -\sqrt{bV}\,\Big[{}^{(2)}\Delta(\ln\rho) + \rho^{-2}\left(\rho_{|X}\,\rho^{|X} + 2\,\Psi_{XY}\,\Psi^{XY}\right) + {}^{(2)}R\Big], \tag{6.4.16}$$

where

$${}^{(2)}\Delta \equiv (\ )_{|X}{}^{|X} \quad \text{and} \quad \Psi_{XY} = \left(K_{XY} - \frac{1}{2}\,K\,b_{XY}\right)\rho\,. \tag{6.4.17}$$

The last step of the proof consists of integrating relations (6.4.15) and (6.4.16) in V from 0 to 1. Using the boundary conditions (6.4.4) and (6.4.13), the identity

$$\int_{V=\text{const.}} {}^{(2)}\Delta f\,\sqrt{b}\,d\theta^2 d\theta^3 = 0\,, \tag{6.4.18}$$

valid for an arbitrary smooth function f, and the Gauss-Bonnet theorem,

$$\int_{V=\text{const}} {}^{(2)}R\,\sqrt{b}\,d\theta^2 d\theta^3 = 8\pi, \tag{6.4.19}$$

we arrive at the inequalities

$$\rho_0 \geq 4M, \qquad \mathcal{A}_0 \geq \pi\rho_0^2, \tag{6.4.20}$$

which become equalities if and only if

$$\Psi_{XY} = 0, \qquad \rho_{|X} = 0 \tag{6.4.21}$$

everywhere on Σ.

Comparing (6.4.20) and (6.4.14), we easily find that these two relations are not in contradiction only if (6.4.20) contains only equality signs, and hence (6.4.21) holds. These relations show that the relevant vacuum solution to Einstein's equations is spherically symmetric, and hence it coincides with the Schwarzschild solution in accordance with Birkhoff's theorem (Birkhoff (1923); see also Appendix B.3).

A similar uniqueness theorem holds if the condition that Einstein's vacuum equations are valid is dropped, and these are replaced with a system of source-free Einstein-Maxwell equations. In this situation, the black hole may be charged. The corresponding unique solution is spherically symmetric and coincides with the Reissner-Nordström metric [Israel (1968), Müller zum Hagen, Robinson, and Seifert (1973, 1974), and Robinson (1977)]. Some open gaps have been closed more recently such as the extension of certain vacuum results to the electrovac case [Simon (1985)], the exclusion of multiple black hole solutions [Simon (1984), Bunting and Masood-ul-Alam (1987), Ruback (1988), Masood-ul-Alam (1992)], and the proof of the staticity theorem for the Einstein-Maxwell equations [Sudarsky and Wald (1992)].

6.5 Uniqueness Theorem for Stationary Black Holes

Now we will discuss the properties of solutions of the Einstein-Maxwell equations that describe stationary axially symmetric black holes. As we already mentioned in Section 6.3.4, the electrovac solutions outside the sources satisfy the circularity condition so that the metric can be written as follows:

$$ds^2 = -V\,dt^2 + 2\,W d\phi\,dt + X\,d\phi^2 + d\gamma^2\,, \tag{6.5.1}$$

where

$$d\gamma^2 = \gamma_{XY}(x^Z)\,dx^X\,dx^Y\,, \quad (X, Y, \ldots = 1, 2)\,.$$

Carter (1969) proved that if the causality condition is met (there are no closed timelike lines), the quantities

$$\rho^2 \equiv -(1/2)\,\rho_\mu\,\rho^{\mu\nu} = VX + W^2 \tag{6.5.2}$$

and X are positive everywhere in the exterior region, except on the rotation axis, where $X = W = 0$, and on the event horizon bounding the exterior region, where ρ^2 vanishes. For a static black hole, $W = 0$ and the equation of the event horizon becomes $V = 0$.

If the Einstein or Einstein-Maxwell equations are satisfied, the function ρ is harmonic:

$$\frac{1}{\sqrt{\gamma}}\,\partial_X\left(\sqrt{\gamma}\,\gamma^{XY}\,\partial_Y\rho\right) = 0\,. \tag{6.5.3}$$

Any two-dimensional metric being conformally flat, we can write $d\gamma^2$ in the form

$$d\gamma^2 = \widetilde{U}(x^1, x^2)\left[(dx^1)^2 + (dx^2)^2\right]\,. \tag{6.5.4}$$

However, it is more convenient for the description of the properties of a metric in the neighborhood of the horizon to introduce new coordinates λ and μ which in the asymptotically distant region ($V \to \infty$) are related to the conventional spherical coordinates r and θ by the formulas

$$\lambda \approx r - M\,, \qquad \mu \approx \cos\theta\,. \tag{6.5.5}$$

Here M is the black hole mass measured by an asymptotically distant observer and $d\gamma^2$ written in these coordinates is

$$d\gamma^2 = U(\lambda, \mu)\,d\gamma_0^2\,, \tag{6.5.6}$$

$$d\gamma_0^2 = \frac{d\lambda^2}{\lambda^2 - C^2} + \frac{d\mu^2}{1 - \mu^2}\,. \tag{6.5.7}$$

Carter (1971) showed that the coordinates t, λ, μ, and ϕ cover the entire exterior region of a stationary black hole (except for the rotation axis where these coordinates obviously have a singularity). The angle coordinate ϕ is periodic (with a period of 2π). The time coordinate t varies from $-\infty$ to $+\infty$. The coordinate μ varies from -1 to +1, and its boundary values are attained on the "north" and "south" polar axes. The coordinate λ is defined in the region $\lambda > C > 0$. The value $\lambda = C$ corresponds to the event horizon, and for asymptotically distant points, $\lambda \to \infty$. In these coordinates, ρ obtained as a regular solution of (6.5.3) is

$$\rho^2 = (\lambda^2 - C^2)(1 - \mu^2)\,. \tag{6.5.8}$$

The electromagnetic field $F_{\mu\nu}$ outside the sources is written in the form

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,, \qquad A_\mu\,dx^\mu = \Phi\,dt + B\,d\phi\,. \tag{6.5.9}$$

The quantities V, X, W, U, Φ and B are functions of λ and μ.

Now we will outline the main steps in the proof of the uniqueness theorem for axially symmetric stationary black holes. These steps are as follows.

1. The solution of the Einstein-Maxwell equations can be reduced, using the method developed by Ernst (1968a,b) [see also Kramer *et al.* (1980)], to solving a system of two second-order elliptic equations for two complex functions of the variables λ and μ (the *Ernst potentials*). It is then found that the resulting equations coincide with the equations of motion for a specific Lagrangian.

2. An analysis is carried out of the conditions imposed on the coefficients of the metric (6.5.1), (6.5.6)-(6.5.7) and on the components of electromagnetic field (6.5.9). These conditions stem from the constraint of regularity of spacetime in the neighborhood of the event horizon and on the rotation axis, and also from the assumption that the space is asymptotically flat. These conditions are subsequently reformulated in an equivalent manner as boundary conditions for Ernst potentials at the singular points $\lambda = C$, $\lambda = \infty$, $|\mu| = 1$.

3. A differential condition relating two arbitrary solutions is obtained, using the invariance of the Lagrangian introduced for the problem involved. This condition is used to prove that any two solutions satisfying the derived boundary conditions with fixed values of the arbitrary constants are identical.

4. It is shown that the familiar Kerr-Newman solution, describing a charged rotating black hole, satisfies the boundary conditions mentioned above and contains the required number of arbitrary constants. This procedure establishes that this family of solutions exhausts the set of solutions describing stationary axially symmetric black holes.

The following remark serves as the starting point of the program outlined here. Assume that the functions X, W, Φ, and B corresponding to a certain axially symmetric stationary asymptotically flat solution of the Einstein-Maxwell equations are known. Then the function V for this solution is found from (6.5.8), and the function U can be uniquely determined by solving an equation implied by the complete Einstein-Maxwell system [Kramer *et al.* (1980)].

Then a change of variables Φ, W to new variables E, Y is carried out via the following relations:

$$
\begin{aligned}
E_{,\mu} &= (X\,\Phi_{,\lambda} - WB_{,\lambda})/(1-\mu^2)\,,\\
E_{,\lambda} &= -\,(X\,\Phi_{,\mu} - WB_{,\mu})/(\lambda^2 - C^2)\,,\\
Y_{,\mu} &= (X\,W_{,\lambda} - WX_{,\lambda})/(1-\mu^2) + 2(BE_{,\mu} - EB_{,\mu})\,,\\
Y_{,\lambda} &= -\,(X\,W_{,\mu} - WX_{,\mu})/(\lambda^2 - C^2) + 2(BE_{,\lambda} - EB_{,\lambda})\,.
\end{aligned}
\tag{6.5.10}
$$

Here $()_{,\mu}$ and $()_{,\lambda}$ denote partial derivatives with respect to the coordinates μ and λ, respectively. It can be shown that the original Einstein-Maxwell system of equations provides the consistency conditions for system (6.5.10) and yields four nonlinear partial differential equations for four unknown functions (Ernst potentials) X, Y, E, B which can be obtained by varying the following "action" functional:

$$S \equiv \int \sqrt{\gamma_0}\, d\lambda\, d\mu\, \mathcal{L}, \tag{6.5.11}$$

where the "Lagrangian" $\mathcal{L}$ is

$$\mathcal{L} = \frac{1}{2X^2}\left[(\nabla X)^2 + [\nabla Y + 2(E\nabla B - B\nabla E)]^2\right] + \frac{2}{X}\left[(\nabla E)^2 + (\nabla B)^2\right]. \tag{6.5.12}$$

All operations of contraction and index raising are carried out here using the two-dimensional metric $d\gamma_0^2$ (6.5.7). In the absence of an electromagnetic field, it is sufficient for solving Einstein's vacuum equations to set $E = B = 0$: the "Lagrangian" $\mathcal{L}$ then takes the form

$$\mathcal{L} = \frac{(\nabla X)^2 + (\nabla Y)^2}{2X^2}. \tag{6.5.13}$$

Carter (1971, 1973a) proved that the boundary conditions which uniquely determine the solution X, Y, E, B follow from the following assumptions: (a) spacetime is asymptotically flat; (b) spacetime is regular everywhere in the exterior region, including the symmetry axis; (c) the event horizon is a regular surface; that is, it contains no physical singularities. In our case, these assumptions take the following form:

(a) E, B, Y, and $\lambda^{-2}X$ are regular functions of λ^{-1} and μ in the asymptotically distant region (as $\lambda \to \infty$). Their asymptotic behavior is

$$E = Q\mu + O(\lambda^{-1}), \qquad B = P\mu + O(\lambda^{-1}),$$

$$Y = 2J\mu(3-\mu^2) + O(\lambda^{-1}), \qquad \lambda^{-2}X = (1-\mu^2)[1 + O(\lambda^{-1})], \tag{6.5.14}$$

where J, Q, and P are constants that play the roles of angular momentum, electric charge, and magnetic monopole charge of the black hole.

(b) E, B, X, and Y on the symmetry axis (as $\mu \to \pm 1$) are regular functions of μ and λ; furthermore, the following conditions are satisfied:

$$E_{,\mu} = O(1), \qquad E_{,\lambda} = O(1-\mu^2), \qquad Y_{,\lambda} = O(1-\mu^2),$$

$$Y_{,\mu} + 2(EB_{,\mu} - BE_{,\mu}) = O(1-\mu^2),$$

$$B_{,\mu} = O(1), \qquad B_{,\lambda} = O(1-\mu^2),$$

$$X = O(1-\mu^2), \qquad X^{-1}X_{,\mu} = 1 + O(1-\mu^2). \tag{6.5.15}$$

(c) E, B, X, and Y on the event horizon (as $\lambda \to C$) are regular functions of μ and λ, and the following conditions hold

$$E_{,\mu} = O(1), \qquad E_{,\lambda} = O(1), \qquad B_{,\mu} = O(1), \qquad B_{,\lambda} = O(1),$$

$$Y_{,\mu} = O(1), \qquad Y_{,\lambda} = O(1), \qquad X = O(1). \tag{6.5.16}$$

In the case of no electromagnetic field, setting $E = B = 0$ transforms the above conditions into the boundary conditions for problem (6.5.13).

The next (main) step of the proof is to establish a differential identity that relates two arbitrary stationary axially symmetric solutions. In deriving this identity, we follow Mazur (1982). The proof essentially employs the invariance of action (6.5.11) (6.5.12) under the group SU(1,2) of transformations of field variables.[11] In order to establish this invariance, it is convenient to exchange the variables X, Y, E, B for new variables ξ, η via the relations

$$-X + iY - E^2 - B^2 = \frac{\xi - 1}{\xi + 1}, \qquad E + iB = \frac{\eta}{\xi + 1}. \tag{6.5.17}$$

In these variables the Lagrangian density (6.5.12) is rewritten in the form

$$\mathcal{L} = 2(1 - \xi\bar{\xi} - \eta\bar{\eta})^{-2} \times \left[(1 - \eta\bar{\eta})\nabla\xi\,\nabla\bar{\xi} + (1 - \xi\bar{\xi})\,\nabla\eta\,\nabla\bar{\eta} + \xi\bar{\eta}\,\nabla\eta\,\nabla\bar{\xi} + \eta\bar{\xi}\,\nabla\xi\,\nabla\bar{\eta}\right], \tag{6.5.18}$$

and the condition of positiveness of X is equivalent to the inequality

$$\xi\bar{\xi} + \eta\bar{\eta} < 1. \tag{6.5.19}$$

Denote by Φ the following non-degenerate matrix constructed from ξ and η:

$$\Phi = (1 - \xi\bar{\xi} - \eta\bar{\eta})^{-1} \begin{pmatrix} 1 + \xi\bar{\xi} + \eta\bar{\eta} & 2\bar{\xi} & 2\bar{\eta} \\ 2\xi & 1 + \xi\bar{\xi} - \eta\bar{\eta} & 2\xi\bar{\eta} \\ 2\eta & 2\eta\bar{\xi} & 1 + \eta\bar{\eta} - \xi\bar{\xi} \end{pmatrix}. \tag{6.5.20}$$

Let

$$j_Y = \nabla_Y \Phi \cdot \Phi^{-1}, \tag{6.5.21}$$

where $\nabla_Y \Phi$ is a matrix obtained from Φ by term-by-term differentiation of its components. A simple procedure verifies that Lagrangian density (6.5.18) admits the following equivalent re-arrangement

$$\mathcal{L} = \frac{1}{4}\,\mathrm{Tr}(j_Y\, j^Y). \tag{6.5.22}$$

---

[11] Action (6.5.11)–(6.5.12) is a particular case of action of the type

$$S[Z^A] = \int d^n x\,\sqrt{\gamma}\,\gamma^{ab}\,\partial_a Z^A\,\partial_b Z^A\,G_{AB},$$

where $a, b = 1, \ldots, n$; $A, B = 1, \ldots, N$; $\gamma^{ab} = \gamma^{ab}(x)$; $G_{AB} = G_{AB}(Z)$. The extremum $Z^A(x)$ of this action is a harmonic mapping [Misner (1978)]. The invariance mentioned above signifies that in the case considered here, the unphysical space (Z^A are its coordinates, and $G_{AB}(Z)$ is its metric) is homogeneous. Bunting (1983) suggested a different proof of the uniqueness theorem. His proof does not make use of this symmetry but is based on the fact that the curvature tensor of this space is constant in sign. For the exposition of this proof and its possible generalization, see Carter (1985).

Here Tr denotes the operation of finding the trace of the matrix, and operations with the index Y are carried out, using the metric γ_0^{XY}.

Let U be a pseudo-unitary matrix satisfying the condition

$$U^+ \eta U = \eta , \quad \eta = \mathrm{diag}(-1,1,1) , \quad \det U = 1 . \tag{6.5.23}$$

Then the matrix

$$\tilde{\Phi} = U \Phi U^{-1} \tag{6.5.24}$$

is of the same form as (6.5.20) for the transformed variables $\tilde{\xi}$ and $\tilde{\eta}$. If the transformation matrix U is independent of x^Y, the Lagrangian density (6.5.22) obviously remains invariant under transformations (6.5.24). In other words, action (6.5.11), (6.5.18) is invariant under the group $SU(1,2)$ of nonlinear transformations $(\xi, \eta) \to (\tilde{\xi}, \tilde{\eta})$ generated by linear representation (6.5.24). In accordance with Noether's first theorem, this invariance implies conservation laws. In our case, they are equivalent to the relation

$$\nabla_X (\rho j^X) = 0 , \tag{6.5.25}$$

which holds for the solutions Φ of the field equations.

Consider now two arbitrary fields Φ_1 and Φ_2 of type (6.5.20), and construct a matrix $\Phi = \Phi_1 \Phi_2^{-1}$. It is then possible to verify the following differential identity :

$$\mathrm{Tr}\left\{\Phi \left[\nabla_X \left(\rho j_2^X\right) - \nabla_X \left(\rho j_1^X\right)\right]\right\} + \nabla_X \left[\rho \nabla^X (\mathrm{Tr}\Phi)\right]$$
$$= \rho \,\mathrm{Tr}\left\{\Phi \left[j_{1X}\, j_1^X + j_{2X}\, j_2^X - 2 j_{2X}\, j_1^X\right]\right\} , \tag{6.5.26}$$

where

$$j_i^X = \nabla^X \Phi_i \, \Phi_i^{-1} . \tag{6.5.27}$$

Identity (6.5.26) serves to complete the proof of the uniqueness theorem. Let (X_1, Y_1, E_1, B_1) and (X_2, Y_2, E_2, B_2) (or Φ_1 and Φ_2) be the solutions describing two stationary axially symmetric black holes and satisfying the regularity conditions (6.5.14)-(6.5.16). Then the first term on the left-hand side of (6.5.26) identically vanishes, and the second term vanishes if expression (6.5.26) is integrated over the exterior region $\lambda > C$, $-1 \leq \mu < 1$, and boundary conditions (6.5.14)-(6.5.16) are taken into account. On the other hand, it can be shown [Mazur (1982, 1984)] that the expression on the right-hand side of expression (6.5.26) is non-negative. Since the integral of the left-hand side over the region $\lambda > C$, $-1 \leq \mu < 1$ vanishes, the right-hand side of (6.5.26) must vanish identically on the solutions Φ_1 and Φ_2. It is then proved that vanishing of the right-hand side of (6.5.26) implies, when boundary conditions (6.5.14)-(6.5.16) are taken into account, that

$$\Phi_1 = \Phi_2 , \tag{6.5.28}$$

which implies that only one solution of the field equations corresponding to the action (6.5.11) with (6.5.18) satisfies the prescribed boundary conditions. It is proved thereby that any stationary axially symmetric black hole is uniquely defined by fixing the values of four arbitrary parameters: C, J, Q, and P.

To complete the proof, consider the following stationary axially symmetric solution of Einstein-Maxwell equations

$$ds^2 = -\frac{\Delta}{\Sigma}(dt - a\sin^2\theta\, d\phi)^2 + \frac{\sin^2\theta}{\Sigma}\left[a\, dt - (r^2 + a^2)\, d\phi\right]^2 + \Sigma\left(\frac{dr^2}{\Delta} + d\theta^2\right), \tag{6.5.29}$$

$$A_\mu\, dx^\mu = -\rho^{-2}\{Qr(dt - a\sin^2\theta\, d\phi) + P\cos\theta\,[a\, dt - (r^2 + a^2)\, d\phi]\}, \tag{6.5.30}$$

where

$$\Delta = r^2 - 2Mr + a^2 + Q^2 + P^2, \quad \Sigma = r^2 + a^2\cos^2\theta. \tag{6.5.31}$$

This is a generalization of the Kerr-Newman metric (see Appendix D.1) which includes an additional parameter P. The structure of the electromagnetic field determined by this parameter indicates that P must be identified with the monopole magnetic charge.

It can be shown that the solution (6.5.29)–(6.5.30) satisfies boundary conditions (6.5.14)–(6.5.16) and contains exactly four arbitrary parameters: M, a, Q, and P (related to the parameters J and C by the formulas $J = Ma$ and $C = (M^2 - a^2 - Q^2 - P^2)^{1/2}$, respectively). Hence, this solution is the most general one, describing an isolated stationary axially symmetric black hole in the Einstein-Maxwell theory.

It is usually assumed that a black hole has no monopole magnetic charge ($P = 0$). Solution (6.5.29)–(6.5.30) then transforms into the Kerr-Newman metric (D.1)–(D.3).

This proof of the uniqueness theorem is greatly simplified if the black hole has zero electric and magnetic charges [Carter (1976)]. This limiting case is obtained by setting $\eta = E = B = 0$ and by denoting by Φ, instead of matrix (6.5.20), the 2 x 2 matrix obtained from (6.5.20) by deleting the last row and the last column. Identity (6.5.26) then transforms into the identity found by Robinson (1975) in proving the uniqueness theorem for noncharged stationary axially symmetric black holes.

6.6 Analytic Continuation of the Kerr-Newman Metric Inside the Event Horizon

The stationary metric of a rotating uncharged black hole outside the event horizon was discussed in Section 3.4. We have explained there why the Kerr metric extended into the region within the event horizon cannot describe the spacetime inside the black hole. Of course, the same arguments hold in the general case of charged rotating hole described by the Kerr-Newman metric (see Section 3.7).[12]

Nevertheless, in this section we consider a formal continuation of the Kerr-Newman metric to the region within the event horizon. The following reasons justify this analysis.

First, the structure of this continuation proved to be quite unexpected. An analysis showed that the total spacetime in general relativity may be topologically very complicated. Hypotheses were advanced that travel between different spaces can be possible if such a total solution is used provided there exist structures similar to those described by the total Kerr-Newman solution. In reality, the reliability of such hypotheses became very problematic after this solution had been proved to be unstable inside the event horizon.

Second, in order to prove the instability of the Kerr-Newman solution inside the black hole, one necessarily begins by giving the solution itself and then proving its instability. The properties of the analytic continuation of the Kerr-Newman solution within the event horizon are discussed below in this section, while instability is proved in Chapter 14.

The global structure of the Kerr-Newman spacetime is analyzed, in principle, just as in the case of the Schwarzschild metric. An additional difficulty arises because of the absence of spherical symmetry. We assume that $M^2 \geq Q^2 + a^2$; otherwise, the solution does not describe a black hole. First of all, recall that in Boyer-Lindquist coordinates, the event horizon lies at $r = r_+ = M + (M^2 - a^2 - Q^2)^{1/2}$ (see (3.6.1) and Appendix D.1). Metric (3.6.1) has a singularity here. In fact, the singularity is a coordinate one. This can be established by transforming to Kerr coordinates (see Appendix D.7). All curvature invariants are finite at $r = r_+$, and the spacetime has no singularities.

When a metric is considered inside a black hole ($r < r_+$), one has to remember that the coordinates (t, r, θ, ϕ) need not have a simple meaning of the temporal and spherical spatial coordinates, as they had at infinity in the external space. We have already encountered such behavior when analyzing the Schwarzschild metric (see Section 2.4), where the variable r became the time coordinate for $r < r_g$. The physical meaning of coordinates in the Kerr-Newman metric is even more complex. The coordinate grid is produced by coordinate lines "traced" in the curved four-dimensional manifold; their physical meaning can be found at each point by considering their orientation with respect to the null cone.

If $r < r_+$, metric (3.6.1) also has singularities at

$$r_- \equiv M - (M^2 - a^2 - Q^2)^{1/2} \tag{6.6.1}$$

and at

$$\Sigma \equiv r^2 + a^2 \cos^2\theta = 0\,, \tag{6.6.2}$$

[12]Even though we have assumed in this section that the black hole has no magnetic charge, the results presented here are readily generalized to the case when the magnetic charge does not vanish.

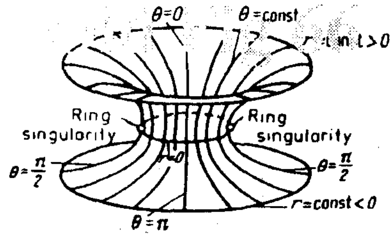


Figure 6.3: Qualitative structure of the section $t = \text{const}$, $\phi = \text{const}$ close to $r = 0$.

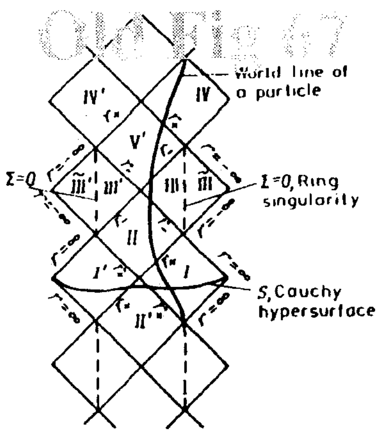


Figure 6.4: Penrose-Carter conformal diagram for the total Kerr-Newman spacetime.

i.e., at the points where $r = 0$ and $\theta = \pi/2$. The singularity (6.6.1) is a coordinate one, like $r = r_+$.

Singularity (6.6.2) is a true singularity where the spacetime curvature is infinite. Figure 6.3 shows the qualitative structure of the spacetime section $t = \text{const}$, $\phi = \text{const}$ at $r = 0$. The true singularity in the section $t = \text{const}$ is a "ring" $r = 0$, $\theta = \pi/2$ in the equatorial plane. If we follow the line $\theta = \text{const} \neq \pi/2$ (follow it in the mathematical sense), no singularities are met on the way along the line. The spacetime is regular at $r = 0$, and we can enter the region where $r < 0$. The spacetime continues up to $r = -\infty$. However, one cannot assume that the section shown in Figure 6.3 is spacelike. As we see from (3.6.1), the coefficient of $d\phi^2$ becomes negative at sufficiently small r and θ close to $\pi/2$; hence, ϕ becomes a timelike coordinate. But ϕ is a cyclic variable with a period of 2π.[13] This means that under these conditions, the section contains *closed timelike curves* (lines that lie along the singular ring and close to it).

The conformal Penrose-Carter diagram of Figure 6.4 shows the complete structure of the true analytic continuation of the Kerr-Newman spacetime.[14] A similar diagram

[13]For metric (3.6.1) to be asymptotically flat at infinity, $r \to \infty$, the variable ϕ must vary from 0 to 2π, and θ, from 0 to π.

[14]The structure of the maximal analytic continuation for an extremal black hole is somewhat different [see Carter (1966a), Hawking and Ellis (1973)]. The maximal analytical continuation for the Reissner-Nordström metric was obtained by Graves and Brill (1960). The maximal analytic continuation for the Kerr spacetime was given by Carter (1968b). The general method of constructing a maximal analytic continuation for stationary metrics with horizons was presented by Walker (1970). See also O'Neill (1994)

for the Schwarzschild spacetime contains four distinct regions (see Figure 5.4c): a white hole, two exterior regions that are asymptotically flat at infinity, and a black hole. The diagram for the Kerr-Newman solution contains an infinite number of regions. Regions I and I′ correspond to similar asymptotically flat regions of the Schwarzschild black hole. Region II′ corresponds to a white hole, and region II, to a black hole. These regions are not bounded by spacelike true singularities, as they are in the case of the Schwarzschild solution. Region II connects to regions III and III′ across two distinct boundaries $r = r_-$. Each of these regions has a ring singularity like the one discussed above; in each of these regions it is possible to go into the region where $r < 0$ (regions $\widetilde{\text{III}}$ and $\widetilde{\text{III}}'$) up to $r \to -\infty$ bypassing the singularity. As $r \to -\infty$, the spaces $\widetilde{\text{III}}$ and $\widetilde{\text{III}}'$ become asymptotically flat. Ring singularities $\Sigma = 0$ manifest themselves in these spaces as "naked singularities" of negative mass.

Regions III and III′ are connected across the boundaries r_- to region V′ which is a white hole completely identical in its properties to region II′. Region V′, in its turn, is connected across the boundaries r_+ to regions IV and IV′ that are completely identical in their properties to I and I′, and so on (to infinity).

The timelike line of a particle that fell into the black hole (region II) from the exterior region I continues to the intersection with one of the boundaries $r = r_-$. Only motions to progressively smaller r are possible in the region II. Having crossed $r = r_-$, a particle enters either the region III′ or III.[15] Here both the motions with decreasing r (down to $r \to -\infty$) and those with increasing r are possible. In the latter case, a particle crosses the boundary $r = r_-$, enters the region V′ (where only motion with increasing r is possible), and crosses one of the boundaries $r = r_+$, thus emerging in the region IV′ and IV. Thus, the particle whose world line is shown in Figure 6.4 can leave "our" exterior space I and enter another, identical space IV.

Note that the topological structure shown in Figure 6.4 is preserved in the case of a charged black hole ($Q \neq 0$) even if the hole is not rotating ($a = 0$) (provided $Q^2 < M^2$). The only difference is that in this case the singularity $\Sigma = 0$ changes from a ring singularity (in the $t = 0$ section) to a point-like one. Now it becomes impossible to bypass this singularity and enter the region $r < 0$. In this case regions $\widetilde{\text{III}}$ and $\widetilde{\text{III}}'$ are absent although it is again possible to pass from I to IV along a timelike world line.

The possibility of such "journeys" generated a number of exotic hypotheses about the outcome of a real gravitational collapse [Novikov (1966a,b, 1970), De la Cruz and Israel (1967), Bardeen (1968)]. However, we have already mentioned that, owing to the instability of the Kerr-Newman solution inside the black hole, the diagram of Figure 6.4 has hardly anything to do with reality.

The boundaries $r = r_-$ of region II are known as the Cauchy horizons. This designation reflects the following fact. If we trace a spacelike Cauchy hypersurfaces

[15] It is possible to get into the region V′ directly from II across the intersection of the boundaries r_-.

S in the entire spaces I and I′ (and possibly, through parts of regions II′ and II) as shown in Figure 6.4, and fix on this surface the Cauchy data for any fields or particles, then these data define the evolution of the fields and the motion of the particles only up to the boundaries $r = r_-$. Field evolution and particle motion in the regions III and III′ can be influenced by sources inside these spaces; these sources are defined indepedently of the data on S. In other words, the spacetime is *unpredictable beyond the Cauchy horizon.*

The following factor constitutes an important property of the Cauchy horizon. Figure 6.4 shows that the later a light signal from the region I reaches II, the closer its world line passes to the boundary r_-. As a result, the world lines of all signals going into the black hole as $t \to \infty$ get "accumulated" close to r_-. The factor that causes the instability of the Kerr-Newman solution inside a black hole with respect to small perturbations is this concentration of signals along $r = r_-$ (see Chapter 14).

6.7 Generalization of the Uniqueness Theorems

Black holes are solutions of highly nonlinear gravity equations that are characterized by non-trivial topology and causal structure. As compared with the field theory in the given flat spacetime background, one might expect greater diversity of black hole solutions and an additional complexity of their structure. In fact, the reverse is observed, however paradoxical it may seem. In the presence of a black hole the class of possible solutions describing their final stationary configurations narrows down dramatically, and its complete description becomes possible under certain constraints. The physical reason for this phenomenon is that the gravitational field is universal and acts on any matter possessing energy-momentum. When a black hole is formed, gravitation gets so enhanced that extremely rigid conditions have to be satisfied in its vicinity for physical fields to be in equilibrium; in fact, these conditions are equivalent to eliminating from the field configuration all degrees of freedom capable of propagating. Hence, the general picture is considerably simplified. The uniqueness theorems proved in this chapter give grounds for this conclusion.

The uniqueness theorems for the black holes solutions of the Einstein-Maxwell theory can be generalized and extended to some other fields. In this section, we discuss briefly these results.

Search for new no hair theorems has quite a long history. First works were trying to prove that barion, lepton, and similar charges cannot be specified for a black hole. Hartle (1971, 1972) showed that a lepton charge absorbed into a black hole cannot be detected by measuring the forces created by this charge due to the exchange of a neutrino-antineutrino pair. Similar results on non-measurability of the lepton, barion and other quantum numbers of a black hole were proved by Teitelboim (1972a,b,c).

Bekenstein (1972a,b,c) proved that a stationary black hole cannot have hair in the form of massive classic fields. He used the following method. Let there be a number of physical fields in addition to the gravitational field. We denote this set by φ_A and assume that the collective index A enumerates all components of the fields. Let these fields be described by the action

$$W[\varphi_A] = \int \mathcal{L}(\varphi_A, \varphi_{A,\mu})\, \sqrt{-g}\, d^4x. \tag{6.7.1}$$

The equations implied by the action are

$$\frac{1}{\sqrt{-g}}\left[\sqrt{-g}\,\frac{\partial \mathcal{L}}{\partial \varphi_{A,\mu}}\right]_{,\mu} - \frac{\partial \mathcal{L}}{\partial \varphi_A} = 0\,. \tag{6.7.2}$$

Multiply this equation by φ_A, take the sum over all A, and integrate over the entire exterior region of the black hole. Denoting

$$b^\mu = \sum_A \varphi_A\,\frac{\partial \mathcal{L}}{\partial \varphi_{A,\mu}} \tag{6.7.3}$$

and making use of *Stokes' theorem*, we obtain

$$-\int b^\mu\, d\sigma_\mu + \sum_A \int \left(\varphi_{A,\mu}\,\frac{\partial \mathcal{L}}{\partial \varphi_{A,\mu}} + \varphi_A\,\frac{\partial \mathcal{L}}{\partial \varphi_A}\right) \sqrt{-g}\, d^4x = 0\,. \tag{6.7.4}$$

Here $d\sigma_\mu$ is an element of the boundary hypersurface, and integration is carried out over the entire boundary of the external region of the black hole; that is, the event horizon, spatial infinity, and the null and the timelike infinities of past and future.

The next relevant fact is that b^μ in the stationary case for massive fields and scalar massless field falls off sufficiently rapidly at spatial infinity, while the quantity $b^\mu\, d\sigma_\mu$ vanishes at the horizon and at infinite time so that the first integral in (6.7.4) vanishes. If the integrand of the second integral is positive definite, its vanishing implies that the corresponding fields φ_A also vanish, which proves the sought result on the absence of "hair" of this field.

Making use of this method, Bekenstein (1972a,b,c) was able to prove that there cannot be a static black hole with a regular scalar, vector or tensor massive fields outside it that are described by linear equations without sources. A similar result is valid for stationary axially symmetric black holes provided the metric is assumed to satisfy the circularity condition [Bekenstein (1972c)]. Unfortunately, the proof of the validity of this condition in the general case was not found. This is an obstacle to carrying out the complete proof at the same level of rigor as the proof of the uniqueness theorem for electrovac black holes.

It can be shown [Bekenstein (1972b,c), Chase (1970)] that a black hole cannot have "hair" due to a scalar massless field φ described by the equation $(\Box - \xi R)\varphi = 0$ and vanishing at infinity provided the value of φ at the horizon is finite. A similar result is also proved in the Brans-Dicke scalar-tensor theory [Hawking (1972b)]. If the condition of finiteness of φ is dropped, it is possible to construct a solution that describes an extremal non-rotating black hole with a scalar massless field [Bocharova *et al.* (1970), Bekenstein (1975)]. This solution was found to be unstable [Bronnikov and Kireev (1978)].

The no hair theorem can be proved for a wider class of scalar fields interacting with a static black hole. Namely, consider the action

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,[R - F_{\mu\nu}F^{\mu\nu} - \mathcal{E}(\Phi,\varphi)]\,. \tag{6.7.5}$$

Here $\Phi = (\nabla\varphi)^2$, and $\mathcal{E}$ is a function obeying conditions $\partial\mathcal{E}/\partial\Phi > 0$ and $\mathcal{E} > 0$. (These conditions guarantee that the energy of the static scalar field is positive.) Then the only static spherically symmetric black hole solution is the Reissner-Nordström one with no scalar field [Bekenstein (1995), Mayo and Bekenstein (1996)]. The theorem generalizes earlier results by Heusler (1992, 1995) and Sudarsky (1995). It admits generalization to many scalar fields [Bekenstein (1995)].

Starting with action (6.7.5) and using a conformal transformation with a conformal factor depending on the scalar field, one can generate an action containing the non-minimal interaction of the scalar field with the curvature. By using this approach Mayo and Bekenstein (1996) were able to generalize their result to the theories with non-minimal coupling. Namely, consider the action

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,[R - F_{\mu\nu}F^{\mu\nu}]$$
$$-\frac{1}{2}\int d^4x\,\sqrt{-g}\,\left[\,|\,D_\mu\varphi\,|^2 + \xi\,R|\,\varphi\,|^2 + V(\,|\,\varphi\,|^2)\right], \tag{6.7.6}$$

where $|\,\varphi\,|^2 = \varphi\varphi^*$, and $D_\mu = \partial_\mu - ieA_\mu$. Then for any regular potential V and for any ξ, the stationary black hole solutions of the action (6.7.6) can have no charged hair, and are restricted to the Reissner-Nordström family. For positive definite potential and for $\xi \le 0$ or $\xi \ge 1/2$, the black hole solutions do not have uncharged hair (with φ real and $e = 0$). For review of these and other results on the no hair theorems for a scalar field interacting with gravitational and electromagnetic fields see [Bekenstein (1996)].

The uniqueness theorem can be proved for gravity interacting with the special class of nonlinear scalar fields known as the ***nonlinear sigma model.*** The corresponding action is of the form

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,\left[R - G_{AB}(\varphi(x))\,g^{\mu\nu}(x)\,\partial_\mu\varphi^A\,\partial_\nu\varphi^B\right]. \tag{6.7.7}$$

In this action $G_{AB}(\varphi)$ is a metric on a Riemannian manifold, and φ^A are coordinates on it. The **theorem** proved by Heusler (1996b) runs as follows: ***The only stationary and axisymmetric, asymptotically flat black hole solution for the action (6.7.7) with regular event horizon consists of the Kerr metric and constant field φ.***

As we already mentioned, the no hair conjecture is not valid for more general classes of theories. For example, it is violated if one includes non-Abelian gauge fields. The existence of non-Abelian gauge field hair of black holes was proved in the early 1990's. Now a variety of black hole solutions is known that exhibit different types of hair. Some of these black holes are characterized by additional new charges that determine characteristics of the hair. Such hair is called *primary hair*. For other black holes the number of parameters specifying the solution is not increased, and hair is uniquely defined by the "old" parameters (*secondary hair*). We shall discuss these solutions later in Chapter 13 in the more general context of black holes in unified theories.

Chapter 7

Physical Effects in the Gravitational Field of a Black Hole

7.1 Extraction of Energy from a Black Hole

7.1.1 Irreducible mass

In this chapter we continue the discussion of the effects of interaction between classical particles and fields and black holes.[1] Let us begin by analyzing the efficiency of energy extraction from black holes. Recall that although a black hole is by definition a region of no escape for material bodies and light rays, situations are possible in which energy can be extracted via certain physical processes. We will see later that this energy is extracted from a field connected to the black hole and surrounding it. Thus, extraction is possible if the black hole rotates or is charged. We already discussed two examples of such processes; namely, the Penrose process (Section 6.2.3) and superradiance (Section 4.8.2). In this section, we establish some general restrictions on the possible efficiency of this kind of process.

Consider the efficiency of the Penrose process (see Figure 6.1). Let $\varepsilon_i = -p_i^\mu \xi_{(t)\mu}$ be the energy and $j_i = p_i^\mu \xi_{(\phi)\mu}$ the angular momentum of a particle of type i with momentum p_i^μ, which moves in the gravitational field of a Kerr black hole ($i = 0$ denotes an infalling particle decaying in the ergosphere; $i = 1$, a particle ejected to infinity, and $i = 2$, a particle absorbed into the black hole). Note now that at the event horizon the vector

$$l^\mu = \xi_{(t)}^\mu + \Omega^H \xi_{(\phi)}^\mu , \tag{7.1.1}$$

where Ω^H is the angular velocity of the black hole, is null and tangent to the generators of the horizon. Since p_i^μ is a timelike vector and l^μ is future directed, we

[1] Those features of physical processes in the field of a black hole for which quantum phenomena are important are discussed in Chapters 10 and 11.

have

$$0 \geq l^{\mu} p_{2\mu} = -\varepsilon_2 + \Omega^H j_2 . \tag{7.1.2}$$

Therefore, a particle sinking into the black hole obeys the inequality

$$j_2 \leq \varepsilon_2/\Omega^H . \tag{7.1.3}$$

Thus, if the outgoing particle has a higher energy than the infalling one ($\varepsilon_1 - \varepsilon_0 = -\varepsilon_2 > 0$), a similar relation holds for their angular momenta:

$$j_1 - j_0 = -j_2 \geq -\varepsilon_2/\Omega^H \geq 0 . \tag{7.1.4}$$

The parameters M and J of the black hole change when a particle is absorbed,

$$\delta M = \varepsilon_2 , \quad \delta J = j_2 , \tag{7.1.5}$$

with condition (7.1.3) signifying that

$$\delta M \geq \Omega^H \delta J . \tag{7.1.6}$$

The physical processes, resulting in changes δM and δJ of the black hole parameters such that they obey the relation

$$\delta M/\Omega^H - \delta J = 0 , \tag{7.1.7}$$

are said to be *reversible.*

The differential equation (7.1.7) relating the changes in M and J in a reversible process can be integrated [Christodoulou (1970)]. Note that the total differential of the function

$$\tilde{A} = M^2 + \sqrt{M^4 - J^2} \tag{7.1.8}$$

can be written in the form

$$\delta \tilde{A} = \frac{J}{\sqrt{M^4 - J^2}} \left(\frac{\delta M}{\Omega^H} - \delta J \right) . \tag{7.1.9}$$

Here

$$\Omega^H = \frac{a}{r_+^2 + a^2} = \frac{a}{2Mr_+} = \frac{J}{2(M^2 + \sqrt{M^4 - J^2}\,)M} \tag{7.1.10}$$

is the angular velocity of the black hole.

Relations (7.1.6) and (7.1.9) show that the following inequality holds for the above-discussed processes involving a particle falling into a black hole:

$$\delta \tilde{A} \geq 0 ; \tag{7.1.11}$$

equality holds if and only if the process is reversible. The quantity

$$M_{\text{ir}} = (\tilde{A}/2)^{1/2} \tag{7.1.12}$$

is known as the *irreducible mass* of the black hole [Christodoulou (1970)]. Equations (7.1.8) and (7.1.12) give

$$M^2 = M_{\text{ir}}^2 + \frac{J^2}{4M_{\text{ir}}^2} \geq M_{\text{ir}}^2 . \tag{7.1.13}$$

This relation implies that the Penrose process cannot make the initial mass M less than M_{ir}, and hence the maximal possible energy gain in this process is $\Delta\mathcal{E} = \Delta M$, where

$$\Delta M = M_0 - M_{\text{ir}}(M_0, J_0) , \tag{7.1.14}$$

and M_0 and J_0 are the initial mass and angular momentum of the black hole; $M_{\text{ir}}(M_0, J_0)$ is the corresponding to irreducible mass.

Simple arguments show that for a given initial mass M_0, the maximal value of ΔM,

$$\Delta M_{\text{max}} = \left(1 - 1/\sqrt{2}\right) M_0 \approx 0.29\, M_0 \tag{7.1.15}$$

is reached for an extremal black hole with $J_0 = M_0^2$.

It is easily shown that the quantity $\tilde{A}$ differs from the area $\mathcal{A}$ of the Kerr black hole only by a numerical coefficient:

$$\mathcal{A} \equiv 4\pi(r_+^2 + a^2) = 8\pi\tilde{A} . \tag{7.1.16}$$

In view of this, condition (7.1.11), which signifies non-decreasing black hole area in the processes analyzed above, is in fact a particular case of the general Hawking's area theorem (Section 5.4).

Hawking's area theorem makes it possible to draw a number of general conclusions concerning processes involving black holes. First of all, inequality (7.1.6) can easily be extended to charged black holes and to processes involving charged particles. This is done simply by making use of expression (7.1.16), where r_+ for a charged black hole is

$$r_+ = M + \sqrt{M^2 - a^2 - Q^2} . \tag{7.1.17}$$

The condition $\delta\mathcal{A} \geq 0$ then yields

$$\delta M \geq \Omega^H \delta J + \Phi^H \delta Q , \tag{7.1.18}$$

where δJ and δQ are the changes in the angular momentum and electric charge of the black hole, and

$$\Phi^H = Q\, r_+/(r_+^2 + a^2) \tag{7.1.19}$$

is its electric potential.

Processes for which equality is satisfied in (7.1.18), which generalizes (7.1.7), are also said to be reversible. The property common to all reversible processes is that the area of black holes does not increase.

It should be emphasized that δJ in (7.1.18) is the total change in the angular momentum of a black hole. Whether this change is caused by the angular momentum of the incoming particle (corresponding to its orbital motion), or by its inherent angular momentum (spin), is irrelevant. The generalized inequality (7.1.18) applied to the latter case serves to show, among other things, that an additional gravitational spin-spin interaction exists between the spin of the particle and the angular momentum of a rotating black hole [Hawking (1972a), Wald (1972), Bekenstein (1973b)].

As an illustration, consider the simplest case of a particle of spin s, charge e, and energy ε falling on a black hole and moving exactly along the symmetry axis. If this particle drops into the black hole, we make use of the conservation laws and obtain

$$\delta Q = e\,, \quad \delta J = \sigma s\,, \quad \delta M \leq \varepsilon\,. \tag{7.1.20}$$

Here $\sigma = 1$ if the spin points in the direction of rotation of the black hole, and $\sigma = -1$ otherwise. The inequality in the last of relations (7.1.20) is possible because energy may be partly radiated away. Relations (7.1.18) and (7.1.20) show that a particle with nonzero spin can fall into a black hole only if its energy ε exceeds the quantity $\sigma s\,\Omega^H + e\,\Phi^H$. The second term, $e\,\Phi^H$, carries the meaning of standard electrostatic energy of repulsion. The first term describes repulsion ($\sigma = 1$) or attraction ($\sigma = -1$) due to the spin-spin interaction in gravitation theory. This interaction exists between any two rotating bodies; for a detailed derivation of the expression for this force, and the description of the analogy of the gravitational spin-spin interaction and electromagnetic interaction between magnetic dipoles, see Wald (1972).

7.1.2 Black hole as an amplifier

Since the motion of particles in the geometrical optics is directly related to the propagation of wave packets it is logical to expect that there exists an analogue of the Penrose process for wave scattering on black holes. Really, we already saw in Chapter 4 that under certain conditions a wave incident on a rotating black hole may also be enhanced. Let us derive (using Hawking's area theorem) the general conditions under which it takes place.

The Kerr-Newman metric describing the geometry of charged black holes is stationary and axially symmetric; hence, expansion in eigenfunctions of the operators $\xi^\mu_{(t)}\,\partial_\mu \equiv \partial_t$, and $\xi^\mu_{(\phi)}\,\partial_\mu \equiv \partial_\phi$ is convenient for describing wave propagation in the background of this metric. Consider the behavior of the field mode φ_A with quantum numbers ω, m. The temporal and angular dependence of the mode is

$$\varphi_A \sim f_A(r,\theta)\,\exp(-\,i\omega t + im\phi)\,. \tag{7.1.21}$$

The field φ_A can describe scalar, electromagnetic, or gravitational waves[2] (or other boson fields whose quanta may have, for instance, mass μ and charge e). Far from the black hole, solution (7.1.21) describes the ensemble of quanta, each having energy $\hbar\omega$, a ϕ-component of angular momentum $\hbar m$, and possibly an electric charge e. Hence, the ratio of the flux of the ϕ-component of angular momentum and of electric charge across a large sphere surrounding the black hole to the flux of energy across this sphere are equal to m/ω and $e/\hbar\omega$, respectively. (This can be proved rigorously by using explicit expressions for the energy momentum tensor and current corresponding to the field φ_A.) Making use of the energy and angular momentum conservation laws which reflect the symmetry of the problem at hand, and electric charge conservation, one can show that the interaction of the wave φ_A with the black hole changes the black hole mass, angular momentum, and charge by δM, δJ and δQ related as follows

$$\delta J = \frac{m}{\omega}\,\delta M, \quad \delta Q = \frac{e}{\hbar\omega}\,\delta M. \tag{7.1.22}$$

Using now the inequality (7.1.18) implied by Hawking's area theorem, we obtain

$$\delta M\left(1 - \frac{m\,\Omega^H}{\omega} - \frac{e\,\Phi^H}{\hbar\omega}\right) \geq 0\,. \tag{7.1.23}$$

Thus, any scattering of the modes which satisfies the condition

$$\hbar\omega < \hbar m\,\Omega^H + e\,\Phi^H \tag{7.1.24}$$

results in a reduction of the mass of black hole. If this condition is satisfied, the scattered wave possesses an energy greater than the incident wave; that is, the incident wave is amplified [Zel'dovich (1971, 1972), Misner (1972), Starobinsky (1973), Starobinsky and Churilov (1973), Unruh (1974)]. This effect, known as *superradiance*, was already discussed in Section 4.8.2.

The possibility of wave amplification by rotating black holes was first noticed by Zel'dovich (1971, 1972), who employed an analogy of such black holes to rotating absorbing bodies. Indeed, consider in ordinary flat space a cylindrical wave incident on a cylinder of radius R, rotating at an angular velocity Ω around an axis coinciding with the z axis. The corresponding solution φ_A has the form

$$\varphi_A = f_A(\rho)\,\exp(-i\omega t + im\phi)\,. \tag{7.1.25}$$

On the cylinder surface ($\rho = R$) this field corresponds to a perturbation propagating at a phase velocity $d\phi/dt = \omega/m$. If the velocity ΩR at which matter moves on the surface of a dielectric or conducting cylinder is greater than the linear velocity $R\omega/m$

---

[2]Electromagnetic and gravitational waves propagating close to a charged black hole may transform into each other (for details, see Section 7.3). This mixing does not change the general condition of wave amplification but calls for a more careful analysis.

at which the phase of the incident wave moves on the cylinder surface, absorption of the wave is replaced by amplification. The corresponding condition is

$$\omega < \Omega m . \qquad (7.1.26)$$

This effect is related to the so-called *anomalous Doppler effect* [3].

Note that the amplification condition (7.1.24) is universal and independent of the spin of the field. For particles with nonzero spin, the azimuthal quantum number m corresponds to the quantum number of total (orbital and spin) angular momentum. By contrast, the wave amplification factor depends essentially on the spin of the field. The maximum increase in the energy of electromagnetic waves is only 4.4 %, but for gravitational waves it is 138 % [Starobinsky and Churilov (1973)].

Under certain conditions, this amplification is possible also for the gravitational radiation of a particle moving close to a rotating black hole. If the particle gained the same amount of energy that it emitted to infinity, it could orbit the black hole without ever falling on it: This particle would be a sort of catalyst for extracting energy from a black hole. Such orbits are called "*floating*" [Misner (1972), Press and Teukolsky (1972)]. Unfortunately, Detweiler (1978) was able to show that regardless of the angular velocity values of the black hole and test particle, this process cannot occur.

The following curious effect is connected with superradiance [Damour *et al.* (1976), Ternov *et al.* (1978), Zouros and Eardley (1979), Detweiler (1980a), Gal'tsov (1986)]. Let a wave packet of a massive scalar field revolve on a circular orbit around a rotating black hole, and let the energy on this orbit be such that massive particles in this packet cannot be radiated to infinity. However, these particles can move across the event horizon. If the frequency of quanta falling into the black hole satisfies the condition of superradiance, the fall is accompanied by a more intensive outward radiation. The particles of this radiation have the same quantum numbers as the particles in the packet; hence, they cannot escape to infinity and have to accumulate close to the orbit of the packet. Finally, this results in an instability. Ternov *et al.* (1978) and Detweiler (1980a) proved that this instability holds for a scalar field of

[3] Let we have a transparent medium with the refraction coefficient $n(\omega)$ and an oscillator or atom moving in it with a constant velocity v. We assume that this system (oscillator or atom) can radiate at frequency ω_0 (as measured in its own reference frame). Then the frequency ω of the radiation from the moving system measured by an observer which is at rest with respect to the medium is

$$\omega(\theta) = \frac{\omega_0\sqrt{1-v^2}}{1 - v\,n(\omega)\cos\theta} .$$

For the emission outside the *Cherenkov cone* ($v\,n(\omega)\cos\theta < 1$) this formula describes an ordinary (normal) Doppler effect. For a superluminal motion ($v > n^{-1}$) and for the direction inside the Cherenkov cone ($v\,n(\omega)\cos\theta > 1$), the radiation with positive ω is possible only when $\omega_0 < 0$; that is, when the system emitting photons becomes excited as a result of emission. This effect is known as the *anomalous Doppler effect* [Ginzburg and Frank (1947)]. A detailed discussion of the effect of superradiance for moving and rotating bodies can be found in [Bolotovsky and Stolyarov (1980)].

mass μ such that $\mu M/m_{\rm Pl}^2 \ll 1$; the characteristic time of evolution of the instability is

$$\tau = 24\,(a/M)^{-1}(m_{\rm Pl}^2/\mu M)^8(m_{\rm Pl}/\mu)\,t_{\rm Pl}\,, \tag{7.1.27}$$

where $m_{\rm Pl}$ and $t_{\rm Pl}$ are the Planck mass ($\sim 10^{-5}$g) and Planck time ($\sim 10^{-44}$ s), respectively. A more detailed discussion of the behavior of massive fields in the spacetime of a rotating black hole can found in Gal'tsov (1986) and Gal'tsov *et al.* (1988). No such instability develops with massless fields [Detweiler and Ipser (1973), Press and Teukolsky (1973), Teukolsky and Press (1974)].

It should be mentioned that even though the processes of energy loss by the black hole described earlier in this section (Penrose process and superradiance) are very important in principle for black hole physics, they could hardly produce appreciable observable effects in actual astrophysical conditions [Mashhoon (1973), Wald (1974c), Kovetz and Piran (1975a,b)]. Analogues of the Penrose process, in which the decay of a particle is replaced by a collision of two particles, may have more attractive implications. The collision produces two new particles one of which is ejected to infinity [Piran *et al.* (1975)]. One version of this effect is the Compton scattering of a freely falling photon on an electron with a large angular momentum, moving in the ergosphere [Piran and Shaham (1977)].

7.2 Electromagnetic Field of a Test Charge

7.2.1 Electric field of a point-like charge in the black hole exterior

In this section we discuss the structure of an electromagnetic field created by a test charge in the spacetime of a black hole. This is the simplest of the problems connected with the electrodynamics of black holes, which is the main subject of the next chapter. We have singled out the problem of the field of electric charges at rest near a black hole in a separate section because several aspects of this problem are of independent importance.

Before considering the formal aspects of the problem we make some general remarks. In the absence of gravitational fields the electric field of a charge is radially directed and the flux of the field through any closed two-dimensional surface is $4\pi q$, where q is the electric charge. If we bring the charge into the gravitational field of a black hole its field will be distorted. Besides the general deformation induced by the "weight" of the electromagnetic field, there will also be a more specific effect: Some of the field lines enter the black hole. It is important to note that until the charge crosses the horizon the total flux of the electric field through the surface of the horizon remains unchanged. This means that besides the "ingoing" field lines there are an equal number of "outgoing" lines. In other words, the electric field of a

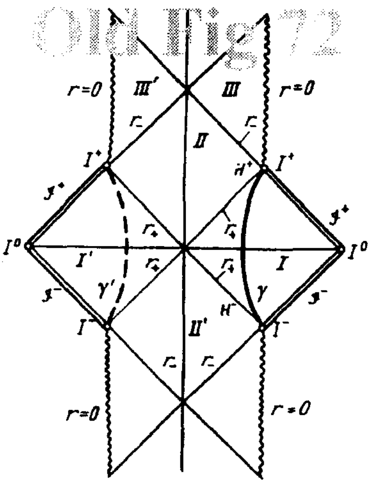


Figure 7.1: Penrose-Carter conformal diagram for the total spacetime of the Reissner-Nordström black hole.

charge near a black hole to some extent resembles the field near a conducting body.[4] Since the field lines cannot end at the horizon, they penetrate into the black hole, and hence there is a non-vanishing electromagnetic field in its interior. An observer falling into the black hole will be moving with respect to the electromagnetic field. Such an observer will register electromagnetic radiation. This effect is closely related to the effect of radiation of a uniformly accelerated charge. Another important aspect of the problem is that due to the nonlocality of the electric field and its deformation under the action of gravity its contribution to the self-energy of a charged particle depends on the position of the particle. As a result of this effect, there is an additional force (proportional to q^2) acting on a charged particle in the field of a black hole.

We begin by considering the electric field in the black hole exterior created by a point-like charge at rest near the horizon. Let a point-like test electric charge q be at rest outside a charged eternal black hole. The black hole metric is fixed and the effect of the field of the test charge on the metric is neglected.[5] The Reissner-Nordström

[4]In the next chapter we shall see that this analogy allows a quite rigorous formulation.

[5]Note that this formulation of the problem is, to a certain extent, idealized. If the field of the test charge on the background of a charged black hole is treated as perturbation, the equation for this perturbation would contain additional terms describing the perturbation of the gravitational field (see Section 7.3). Following Linet and Lèaute (1976), we drop these additional terms. This operation does not change the general conclusion on the specific features of the field inside the black hole. These additional terms disappear if a black hole is uncharged, $Q = 0$. In the present consideration we also ignore the instability of the charged black hole interior (see Chapter 14).

metric, describing the gravitational field of the charged black hole, is

$$ds^2 = -F(r)\,dt^2 + F^{-1}(r)\,dr^2 + r^2\,d\omega^2,$$

$$F = 1 - 2M/r + Q^2/r^2, \quad d\omega^2 = d\theta^2 + \sin^2\theta\,d\phi^2, \tag{7.2.1}$$

where M is the mass, and Q is the electric charge of the black hole.

Figure 7.1 shows the Penrose-Carter conformal diagram for the eternal black hole in question. The coordinates (t, r) cover the region I shown in this figure. The world line of the charge is denoted by γ.

The electric field A_μ created by the test charge at rest at a point (r_0, θ_0, ϕ_0) was obtained by Linet (1976) and Lèaute and Linet (1976) who found and corrected a minor inaccuracy in the formula derived earlier (for $Q \equiv 0$) by Copson (1928). This solution is

$$A_\mu(x) = -q\,\delta^0_\mu\,\frac{1}{r\,r_0}\left(M + \frac{\Pi}{R}\right), \tag{7.2.2}$$

where

$$R^2 = (r-M)^2 + (r_0-M)^2 - 2(r-M)(r_0-M)\lambda - (M^2-Q^2)(1-\lambda^2), \tag{7.2.3}$$

$$\Pi = (r-M)(r_0-M) - (M^2-Q^2)\lambda, \tag{7.2.4}$$

$$\lambda = \cos\theta\,\cos\theta_0 + \sin\theta\,\sin\theta_0\,\cos(\phi-\phi_0)\,. \tag{7.2.5}$$

The electric field $E^i = -g^{ij}A_{0,j}$ for the solution (7.2.2) has non-vanishing r and θ components. One can also characterize the electric field by its components $E^{\hat{i}}$ in the physical (orthonormal) frame $(\mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\theta}}, \mathbf{e}_{\hat{\phi}})$. We discuss the picture of the field lines in Section 8.5 (see Figure 8.2). Here we just mention that the details of the picture of field lines near the horizon depends whether one uses E^i or $E^{\hat{i}}$. It will be shown that the field lines for $E^{\hat{i}}$ always enter a static black hole orthogonally to the surface of the horizon. At a great distance from the black hole (at $r \to \infty$), the electric potential A_0 takes the standard monopole form $A_0 \approx -q/r$.

7.2.2 Electromagnetic field in the black hole interior

Solution (7.2.2) can be extended from region I to the regions located inside the black hole, using the method of analytic continuation. For this purpose, it is convenient to introduce Kruskal coordinates U, V $(-\infty < U, V < \infty)$ which are regular in regions I, II, I′, II′, while in region I they are related to the coordinates (r, t) by the formulas

$$U = -\exp\left[-\frac{r_+ - r_-}{2\,r_+^2}\,(t - r^*)\right], \quad V = \exp\left[\frac{r_+ - r_-}{2\,r_+^2}\,(t + r^*)\right],$$

$$r^* = r + \frac{r_+^2}{r_+ - r_-} \ln \left| \frac{r}{r_+} - 1 \right| - \frac{r_-^2}{r_+ - r_-} \ln \left| \frac{r}{r_-} - 1 \right| ,$$

$$r_\pm \equiv M \pm \sqrt{M^2 - Q^2} . \tag{7.2.6}$$

The expression for the metric in these coordinates can be found in Appendix D.7 (formulas (D.79)–(D.81)).

In Kruskal coordinates (U, V), the nonzero components of the analytically continued electromagnetic potential A_μ and the field strength $F_{\mu\nu}$ are

$$A \equiv A_\mu \, dx^\mu = - \frac{q}{r\, r_0} \left(M + \frac{\Pi}{R} \right) \frac{r_+^2}{r_+ - r_-} \left(\frac{dV}{V} - \frac{dU}{U} \right), \tag{7.2.7}$$

$$F^{UV} = \frac{q}{B r^2 r_0} \left\{ M + \frac{\Pi}{R} - \frac{r}{R^3} [r_0 - M - (r - M)\cos\theta](r_0^2 - 2M r_0 + Q^2) \right\},$$

$$F^{V\theta} = \frac{q}{2\, r\, r_0} \frac{\sin\theta}{R^3} \frac{r_+ - r_-}{r_+^2} \left(r_0^2 - 2M r_0 + Q^2 \right) V , \tag{7.2.8}$$

$$F^{U\theta} = - \frac{q}{2\, r\, r_0} \frac{\sin\theta}{R^3} \frac{r_+ - r_-}{r_+^2} \left(r_0^2 - 2M r_0 + Q^2 \right) U .$$

In these formulas, r is assumed to be a function of U and V, defined by the relation

$$UV = \left(1 - \frac{r}{r_+} \right) \exp \left[\left(r - \frac{r_-^2}{r_+ - r_-} \ln \left| \frac{r}{r_-} - 1 \right| \right) \left(\frac{r_+ - r_-}{r_+^2} \right) \right] . \tag{7.2.9}$$

It should be noted that the above solution obtained by the analytic continuation does not solve our problem. The reason is that we are looking for a solution which has an electric charge q as its only source. The above solution does not possess this property. It is not difficult to see that this solution is invariant under the transformation $U \to -U$, $V \to -V$ which maps region I onto I′. Hence, it has, in addition to a singularity corresponding to the world line γ of the charge, a singularity on the line γ' corresponding to an additional charge $-e$ in region I′. As a result, expressions (7.2.7)–(7.2.8) do not give the solution of the formulated problem of finding the field produced by a single charge.

The sought solution can be obtained if we take into account that it would be natural to choose zero field in regions I′ and II′ that lie outside the region of influence of the test charges.[6] This solution can be written in the following form [Zel'nikov and Frolov (1982), Demianski and Novikov (1982)]:

$$\widetilde{F}_{\mu\nu} = F_{\mu\nu}^{\rm reg} + F_{\mu\nu}^{\rm sing} , \quad F_{\mu\nu}^{\rm reg} = F_{\mu\nu}\, \theta(V) , \quad F_{\mu\nu}^{\rm sing} = \Psi_{\mu\nu}\, \delta(V) , \tag{7.2.10}$$

[6] A discussion of the boundary conditions for the field of a test charge in the spacetime of an eternal black hole can be found in Demianski and Novikov (1982).

where $F_{\mu\nu}$ is solution (7.2.8). The singular term $F^{\rm sing}_{\mu\nu}$ ensures that homogeneous field equations are satisfied on the surface $V = 0$. Substituting (7.2.10) into Maxwell's equations $F^{\mu\nu}{}_{;\nu} = 0$, we obtain

$$\Psi^{\mu V} = 0\,, \qquad \frac{1}{\sqrt{-g}}\,\partial_\nu(\sqrt{-g}\,\Psi^{\mu\nu}) = -\,F^{\mu V}|_{V=0}\,. \tag{7.2.11}$$

The solution of these equations, bounded on H^-, takes the form

$$\Psi_{\mu\nu} = -\,2q\,\delta^V_{[\mu}\,\delta^\theta_{\nu]}\sin\theta\,\frac{\sqrt{M^2-Q^2}}{r_0}\,\frac{M+\sqrt{M^2-Q^2}}{r_0-M-\sqrt{M^2-Q^2}\,\cos\theta}\,. \tag{7.2.12}$$

The analytical continuation readily yields A_μ in regions III and III$'$. In general, it is not possible to extend the solution in a single-valued manner beyond the Cauchy horizon because in this region the solution depends on conditions that must be additionally specified. In the limit of $Q = 0$, the obtained solution describes the fields produced by point-like sources in the spacetime of an eternal Schwarzschild black hole.

It is also possible to obtain a similar solution for a massless scalar field φ obeying the equation

$$\Box\varphi = 0\,. \tag{7.2.13}$$

The scalar field created by a point-like charge g rested at the point $p_0 = (r_0, \theta_0, \phi_0)$ is

$$\varphi = g\sqrt{F(p_0)}\,\frac{1}{R}\,, \tag{7.2.14}$$

where R is defined by equation (7.2.3). The solution vanishing in regions I' and II' is

$$\tilde{\varphi} = \varphi^{\rm reg} + \varphi^{\rm sing}\,, \qquad \varphi^{\rm reg} = \varphi\,\theta(V)\,, \qquad \varphi^{\rm sing} = \Psi\,\delta(V)\,. \tag{7.2.15}$$

The singular term $\varphi^{\rm sing}$ ensures that the homogeneous field equation is satisfied on the surface $V = 0$. Substituting (7.2.15) into equation (7.2.13) yields

$$(1-\lambda^2)\,\frac{\partial^2\Psi}{\partial\lambda^2} - 2\lambda\,\frac{\partial\Psi}{\partial\lambda} - \frac{r_+ - r_-}{r_+}\,\Psi = 0\,, \tag{7.2.16}$$

$$\frac{\partial\Psi}{\partial U} = 0\,, \qquad \lambda = \cos\theta\,.$$

If the black hole is not extremal ($Q < M$), the only solution (7.2.16) bounded on H^- is $\Psi = 0$. If the black hole is extremal ($Q = M$), another solution is $\Psi = \text{const}$ although the value of $\Psi|_{H^-}$ is not determined by the external scalar field, and thus has no relation to the charge g.

A δ-type singularity in the solutions obtained above for the electromagnetic and scalar massless field appears because the problem was so formulated that the charge was assumed to be permanently at rest in the neighborhood of the black hole. A similar singularity appears also in the complete solution describing a massive vector field produced by a source outside the black hole [Frolov (1978a, 1986)].[7]

Note that in a realistic situation when a black hole arises as the result of gravitational collapse the past horizon H^- simply does not exist and the electromagnetic field (as well as other fields) is regular. In the absence of incoming radiation it unambiguously determined by specifying the world line of the charge. But even if we consider an eternal black hole the singularities at $V = 0$ do not arise if a charge was not permanently at rest at a fixed radius but has been inserted close to the black hole at some finite moment of time.

7.2.3 Comparison with the field of a uniformly accelerated charge

The complete solution derived above for the field of a test charge in a black hole spacetime is closely related to the solution for the field of a uniformly accelerated charge in flat spacetime [Zel'nikov and Frolov (1982)]. In order to find this relation, we notice when the mass M of a black hole becomes large, then the effect of curvature in the region of spacetime of finite size $L \ll M$ close to the event horizon can be neglected. Hence, the gravitational field in this region can be regarded as uniform to a high degree of accuracy. The transition to the limit of uniform field is carried out as follows. Introduced coordinates $(\hat{U}, \hat{V}, X, Y)$ related to Kruskal coordinates (U, V, θ, ϕ) by the formulas

$$\hat{U} = 4M\,U\,, \quad \hat{V} = 4M\,V\,,$$

$$X^2 + Y^2 = 16\,M^2 \tan^2(\theta/2)\,, \quad Y/X = \tan\phi\,. \tag{7.2.17}$$

The Kruskal metric (2.7.17) written in these coordinates becomes

$$ds^2 = -\frac{2M}{r}\,e^{1-r/2M}\,d\hat{U}\,d\hat{V} + \left(\frac{r}{2M}\right)^2 \frac{dX^2 + dY^2}{1 + \dfrac{X^2+Y^2}{16M^2}}\,, \tag{7.2.18}$$

where r is related to $\hat{U}$ and $\hat{V}$ by the expression

$$\hat{U}\,\hat{V} = 16\,M^2\left(1 - \frac{r}{2M}\right)\exp\left(\frac{r}{2M} - 1\right). \tag{7.2.19}$$

[7]On the behavior of a massive vector field in the neighborhood of a Schwarzschild black hole, see also Gal'tsov *et al.* (1984b).

If M now tends to infinity with the values of the coordinates $\hat{U}$, $\hat{V}$, X, Y fixed, equation (7.2.18) transforms into the metric of flat space,

$$\begin{aligned} ds^2 &= -\, d\hat{U}\, d\hat{V} + dX^2 + dY^2 = -\, dT^2 + dZ^2 + dX^2 + dY^2 \\ &= -\, dT^2 + dZ^2 + d\rho^2 + \rho^2\, d\phi^2 \,. \end{aligned} \tag{7.2.20}$$

Here

$$\begin{aligned} T &= \frac{1}{2}\,(\hat{U} + \hat{V})\,, \quad Z = -\,\frac{1}{2}\,(\hat{U} - \hat{V})\,, \\ \rho^2 &= X^2 + Y^2\,, \quad \tan\phi = \frac{Y}{X}\,. \end{aligned} \tag{7.2.21}$$

The equation $r = r_0$ of motion of a test charge then takes the form

$$\hat{U}_0\, \hat{V}_0 \equiv -\, Z_0^2 + T^2 = -\, w^{-2}, \quad X_0 = Y_0 = 0\,, \tag{7.2.22}$$

where w is the modulus of the four-acceleration of the charge. In the limit $M \to \infty$, the horizon surfaces $H^{\pm}$ turn into null hyperplanes described by the equation $\hat{U}\,\hat{V} = 0$ (the "horizons" of the Rindler space (Rindler (1966), see also Appendix C). The invariant distance to the horizon for a particle with four-acceleration w then tends to a finite value w^{-1}.

In the limiting transition described above, expressions (7.2.10) and (7.2.15) take the form ($Q = 0$)

$$\begin{aligned} \widetilde{F}_{\hat{U}\hat{V}} &= \frac{2q}{w^2}\, \frac{(\rho^2 + \xi + w^{-2})}{S^3}\, \theta(\hat{V})\,, \quad \widetilde{F}_{\hat{U}\rho} = -\, \frac{4q\rho\, \hat{V}}{\omega^2\, S^3}\, \theta(\hat{V})\,, \\ \widetilde{F}_{\hat{V}\rho} &= \frac{4q\rho\, \hat{U}}{w^2\, S^3}\, \theta(\hat{V}) - \frac{2q\rho}{\rho^2 + w^{-2}}\, \delta(\hat{V})\,, \end{aligned} \tag{7.2.23}$$

$$\widetilde{\varphi} = \frac{q}{w}\, \frac{2\,\theta(\hat{V})}{S}\,, \tag{7.2.24}$$

where

$$S = \left[(\rho^2 - \xi + w^{-2})^2 + 4\,\xi\, w^{-2}\right]^{1/2}, \quad \xi = \hat{U}\,\hat{V}. \tag{7.2.25}$$

Expressions (7.2.23) for $\widetilde{F}_{\mu\nu}$ can be derived from the following four-potential:

$$\begin{aligned} \tilde{A}_\mu\, dx^\mu = &-\, \frac{q}{2}\, \theta(\hat{V}) \left(\frac{\rho^2 - \xi + w^{-2}}{S} - 1 \right) \left(\frac{d\hat{V}}{\hat{V}} - \frac{d\hat{U}}{\hat{U}} \right) - \\ &-\, 2q\, \theta(\hat{V})\, \frac{\rho\, d\rho}{\rho^2 + w^{-2}}\,. \end{aligned} \tag{7.2.26}$$

One can easily show that formulas (7.2.23) coincide with the expression for the field of a uniformly accelerated electric charge [see, e.g., Boulware (1980)]; the term proportional to $\delta(\hat{V})$ in (7.2.23) then reproduces correctly the term with a singularity at $\hat{V} = 0$ introduced by Bondi and Gold (1955).

7.2.4 The shift in the self-energy

In this section we deal with the change of self-energy of a charged particle placed in a stationary gravitational field. The effect is as follows. The total mass of a charged particle is composed of its "mechanical" mass localized at the point where the charge is located, and the "electromagnetic" mass spread over the region where electromagnetic field is nonzero. When a charged particle is placed in a non-uniform gravitational field, this field affects the "local" and "distributed" masses in a different manner, causing a "deformation" of the electric field of the charge; the self-energy is also altered. The shift in self-energy depends on the position occupied by the particle. That is why forces acting in a gravitational field on particles with identical total inertial mass differ when one particle is charged while another is neutral.

The effect of the gravitational field on the self-energy of an electric charge was first analyzed by Fermi (1921). Fermi considered the case of an electric charge at rest in a uniform gravitational field and showed that the electromagnetic interaction causing a change in the inertial mass of a particle changes its gravitational mass by the same amount, in complete agreement with the equivalence principle. Arguments based on the equivalence principle do not work for the system as a whole in a non-uniform gravitational field. In general, one should expect the relationship between the self-energy of a charged particle and the change in the gravitational mass to be more complicated.

This is indeed the case for a particle in the field of a black hole. The necessary corrections (in the approximation $GM/c^2 \ll 1$) were found by Vilenkin (1979a). It was also shown in this paper that the inhomogeneity of the field leads to an additional force that pulls the charge away from the black hole. Still earlier, Unruh (1976a) proved that a similar force acts on a test charge placed inside a thin hollow massive spherical shell. Smith and Will (1980) and Frolov and Zel'nikov (1980) noticed that the shift in the self-energy of an electric charge in the field of a Schwarzschild black hole can be calculated exactly, and carried out the calculation of the additional repulsive force. This result was later generalized to the case of Reissner-Nordström black holes [Zel'nikov and Frolov (1982)], Kerr black holes [Lèaute and Linet (1982)], and Kerr-Newman black holes [Lohiya (1982)].

To calculate the shift in the self-energy of a charge in the field of a black hole we assume that a classical particle; that is, a system of bound electric charges, is at rest on the symmetry axis in a stationary gravitational field; the equilibrium is maintained by an appropriate external force. Denote by $\xi^\mu_{(t)}$ a Killing vector field that is timelike at infinity where it is normalized by the condition $\xi^\mu_{(t)}\,\xi_{(t)\mu} = -1$. We also denote by $\alpha = |\,\xi^\mu_{(t)}\,\xi_{(t)\mu}\,|^{1/2}$ the corresponding redshift factor. The energy of this system is

$$E = -\int_\Sigma T_{\mu\nu}\,\xi^\mu_{(t)}\,d\sigma^\nu, \tag{7.2.27}$$

where $T_{\mu\nu}$ is the total energy-momentum tensor of the system, and $d\sigma_\mu = \delta^0_\mu\,\sqrt{-g}\,d^3x$

is an element of volume of the hypersurface $x^0 = \text{const}$ [see (A.81)].

To specify the problem, we choose a model of the charged particle: a stiff nonconducting thin sphere of mass m_0 and radius ε, with a charge q uniformly distributed over its surface. In view of the final transition to the limit of point-like particle, we assume ε to be much less than the characteristic size of inhomogeneities of the gravitational and external electromagnetic fields; the terms $O(\varepsilon)$ in the final answer will be neglected.

The total energy E of a particle is made up of:

1. The part of energy $E_0 = \alpha\, m_0$ which is connected with the "mechanical" mass m_0 of the particle
2. The self-energy E_{self} or the energy of the self-action of the particle's charge
3. The energy E_{ext} of interaction between the particle and the external field
4. The energy E_{int} of the additional interaction stabilizing the charged particle

The additional interaction is required to satisfy the Laue theorem. If the equilibrium radius of an uncharged particle is denoted by ε_0, then $E_{\text{int}}(\varepsilon) = E_{\text{int}}(\varepsilon_0) + \frac{1}{2} K(\varepsilon - \varepsilon_0)^2$ and the changes (due to the introduction of the charge into the field) in the equilibrium size, $\Delta\varepsilon = \varepsilon - \varepsilon_0$, and in energy, $\Delta E = E_{\text{int}}(\varepsilon) - E_{\text{int}}(\varepsilon_0)$, can be made arbitrarily small by choosing a sufficiently high effective stiffness K. In what follows, the increments $\Delta\varepsilon$ and ΔE are neglected; that is, the stiffness K is assumed to be sufficiently high. Having included the constant quantity $E_{\text{int}}(\varepsilon_0)$ into E_0, we can write the expression for the total energy of a particle as an expansion in powers of the charge q of the particle:

$$E = E_0 + E_{\text{ext}} + E_{\text{self}}\,. \tag{7.2.28}$$

In the case of no external field and far from gravitating bodies, formula (7.2.28) reduces to the expression

$$E = m = m_0 + q^2/2\varepsilon. \tag{7.2.29}$$

The difference between m and m_0 arises because of the field created by the charge. If an uncharged particle is introduced into a static gravitational field, the work extracted in this operation reduces the particle's energy to $E_0 = \alpha\, m$. If the particle is charged, then part of the extracted work goes into changing the field around the particle. As a result, in general the sum $E_0 + E_{\text{self}}$ is not equal $\alpha\, m$.

The expression for E_{self} can be recast as

$$E_{\text{self}} = \int_\Sigma \Pi^\beta\, d\sigma_\beta\,, \tag{7.2.30}$$

where

$$-\,\Pi^\beta = T^\beta_\alpha\, \xi^\alpha_{(t)} \equiv \frac{1}{4\pi}\left(F_{\alpha\mu}\, F^{\beta\mu} - \frac{1}{4}\,\delta^\beta_\alpha\, F_{\mu\nu}\, F^{\mu\nu}\right)\xi^\alpha_{(t)}\,, \tag{7.2.31}$$

and $F_{\alpha\beta}$ is the strength of the field produced by the current

$$j^\mu(x) = \delta^\mu_0 \frac{e}{4\pi\varepsilon^2\sqrt{-g(x)}} \delta(l(\mathbf{x},\mathbf{x}_0) - \varepsilon). \tag{7.2.32}$$

Expression (7.2.32) describes the charge distribution of the particle. Here $l(\mathbf{x},\mathbf{x}_0)$ is the invariant distance between a point $(t,\mathbf{x})$ and the center $(t,\mathbf{x}_0)$ of the charged particle along the geodesic connecting these points, and ε is the invariant size of the particle. The integration of (7.2.30) is carried out over a spacelike surface Σ that intersects the horizon H^+. Note that the integral (7.2.30) over the part of Σ within the event horizon is the energy of the field inside the black hole. Its contribution has already been taken into account in the definition of the total mass of the black hole. Therefore, when we are interested in calculating the energy shift in the field of a black hole with known parameters, we assume these parameters to be fixed, and hence integrate (7.2.30) over that part of Σ which lies outside the black hole.[8]

Using Maxwell's equations

$$F^{\mu\nu}{}_{;\nu} = 4\pi j^\mu, \tag{7.2.33}$$

we can transform expression (7.2.31) to the form

$$\Pi^\beta = B^{\beta\nu}{}_{;\nu} + \frac{1}{2}\xi^\beta A^\alpha j_\alpha - j^\beta \xi^\alpha A_\alpha, \tag{7.2.34}$$

where

$$B^{\beta\nu} = \frac{1}{8\pi}\left(\xi^\nu A_\alpha F^{\alpha\beta} - \xi^\beta A_\alpha F^{\alpha\nu} - 2\xi^\alpha A_\alpha F^{\nu\beta}\right). \tag{7.2.35}$$

Using Stokes' theorem (A.87), we can reduce the integral of the expression $B^{\beta\nu}{}_{;\nu}$ to the sum of the integral over the black hole surface and the integral over a surface at infinity. The second of these integrals vanishes because the field falls off rapidly enough at infinity. It is readily shown that for a particle on the symmetry axis, the first integral (over the black hole surface) also vanishes. Finally, taking into account that ξ^α and j^α are parallel, we find

$$E_{\text{self}} = -\frac{1}{2}\int_\Sigma A_\alpha j^\alpha \xi^\beta d\sigma_\beta. \tag{7.2.36}$$

To obtain an explicit analytical expression for E_{self}, we can make use of the expression for the potential A_α produced by a point-like charge on the symmetry axis in the Kerr spacetime [it was found by Linet (1977a)]. This gives the following expression for E_{self}:

$$E_{\text{self}} = \frac{q^2}{2\varepsilon}\,|\,g_{tt}\,|^{1/2} + \frac{q^2}{2}\frac{M}{r_0^2 + a^2}. \tag{7.2.37}$$

[8]Expressions (7.2.30)–(7.2.31) for the self-energy of a charge, together with the above condition imposed on the choice of integration domain Σ, are in complete agreement with general expression for the change in the mass of a system, containing the black hole, due to a change in the parameters of the entire system which will be obtained in Section 12.2. It is interesting to note that the integral over the part of the surface Σ beneath the horizon is identically zero for the Schwarzschild black hole.

If the force necessary to fix this particle at the point r_0 is compared with that required to fix a neutral particle of mass $m = m_0 + q^2/(2\varepsilon)$, their difference Δf^μ proves to be

$$\Delta f = |\,\Delta f^\mu\,\Delta f_\mu\,|^{1/2} = q^2\,\frac{M r_0}{(r_0^2 + a^2)^2}\,. \tag{7.2.38}$$

This excess force, Δf^μ, applied to the charged particle points along the symmetry axis away from the black hole.

If a charged particle (with a scalar, electric, or gravitational charge) is at rest in the vicinity of a black hole outside its symmetry axis, it is subject to an additional force [Gal'tsov (1982)]. This force is proportional to the angular momentum of the black hole and to the squared charge of the particle. It arises as a reaction to the tidal force exerted by the particle on the black hole and tending to slow down its rotation. This force vanishes if the particle rotates at the same angular velocity that the black hole has or if it is on the symmetry axis.

7.3 Mutual Transformation of Electromagnetic and Gravitational Waves

7.3.1 Geometrical optics approximation

The effect of mutual transformation of electromagnetic and gravitational waves in an external electric field is a well-known corollary of the nonlinearity of Einstein-Maxwell equations [for a detailed discussion of this effect and the relevant bibliography, see, e.g., Sibgatullin (1984)]. This section will briefly describe this effect for the particular case of the propagation of photons and gravitons in the field of a charged black hole [Sibgatullin (1973, 1974, 1984), Sibgatullin and Alekseev (1974), Gerlach (1974, 1975)].

Let there be a metric $g_{\mu\nu}$ and an electromagnetic field A_μ, and let $h_{\mu\nu} = \delta g_{\mu\nu}$ and $a_\mu = \delta A_\mu$ be small perturbations against this background. If assume that both sets, $\{g_{\mu\nu} + h_{\mu\nu}, A_\mu + a_\mu\}$ and $\{g_{\mu\nu}, A_\mu\}$ are solutions of Einstein-Maxwell equations, then we get the following linearized system of equations for the perturbations:

$$k_{\mu\nu;\lambda}{}^{;\lambda} - k^\lambda{}_{\mu;\nu;\lambda} - k^\lambda{}_{\nu;\mu;\lambda} - \frac{1}{2}\,g_{\mu\nu}\,k_{;\lambda}{}^{;\lambda} - 2k^{\alpha\beta}F_{\mu\alpha}\,F_{\nu\beta} - \frac{1}{2}\,k_{\mu\nu}\,F_{\alpha\beta}\,F^{\alpha\beta}$$

$$+\,g_{\mu\nu}\,k^{\alpha\gamma}\,F_{\alpha\beta}\,F_\gamma{}^\beta + k\,T_{\mu\nu} + 2F_\mu{}^\alpha\,f_{\nu\alpha} + 2F_\nu{}^\alpha\,f_{\mu\alpha} - g_{\mu\nu}\,F^{\alpha\beta}f_{\alpha\beta} = 0\,, \tag{7.3.1}$$

$$f^{\alpha\beta}{}_{;\beta} - k_{\beta\rho}\,F^{\alpha\rho;\beta} - k_\mu{}^\alpha{}_{;\beta}\,F^{\mu\beta} - k_{\mu\beta}{}^{;\beta}\,F^{\alpha\mu} + \frac{1}{2}\,k_{,\beta}\,F^{\alpha\beta} = 0\,. \tag{7.3.2}$$

Here

$$k_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\,h\,g_{\mu\nu}\,, \quad h = h_\alpha{}^\alpha\,, \quad k = k_\alpha{}^\alpha\,,$$

$$f_{\mu\nu} = a_{\nu,\mu} - a_{\mu,\nu}, \quad F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}. \tag{7.3.3}$$

Indices are lowered and raised, and covariant differentiation is carried out using the metric $g_{\mu\nu}$; $T_{\mu\nu}$ is the energy-momentum tensor of the field $F_{\mu\nu}$.

This system is invariant with respect to the gauge transformations

$$h_{\mu\nu} \to h_{\mu\nu} - \xi_{\mu;\nu} - \xi_{\nu;\mu},$$

$$a_\mu \to a_\mu + \lambda_{;\mu} - \xi^\alpha A_{\mu;\alpha} - \xi^\alpha_{\ ;\mu} A_\alpha. \tag{7.3.4}$$

The following gauge-fixing conditions conveniently eliminate the gauge freedom:

$$k^{\alpha\beta}_{\ \ ;\beta} = 0, \quad a^\alpha_{\ ;\alpha} = 0. \tag{7.3.5}$$

We will consider the effect of the interaction between electromagnetic and gravitational perturbations due to those terms of (7.3.1) that contain the electric field f and those terms of (7.3.2) that contain the gravitational perturbation k. We assume that the wavelength λ of electromagnetic and gravitational waves is much shorter than the characteristic size L of inhomogeneities in the background fields $g_{\mu\nu}$ and A_μ and use the *geometrical optics approximation*.[9] In this approximation we can write the perturbations $k_{\mu\nu}$ and a_μ in the form

$$a_\mu = \mathrm{Re}\left[(\alpha_\mu + \varepsilon\beta_\mu + \cdots)\,\mathrm{e}^{iS/\varepsilon}\right], \tag{7.3.6}$$

$$k_{\mu\nu} = \mathrm{Re}\left[(\kappa_{\mu\nu} + \varepsilon\pi_{\mu\nu} + \cdots)\,\mathrm{e}^{iS/\varepsilon}\right], \tag{7.3.7}$$

where ε is a parameter characterizing the degree of smallness of a term with respect to a dimensionless parameter λ/L. The phase functions S in (7.3.6)–(7.3.7) are chosen to be identical. This can be done by redefining the preexponential factors if phases differ by a factor of order ε. Otherwise, if the difference between the phases S_a and S_k in the expressions for a_μ and k_μ is not small [$S_a - S_k = O(\varepsilon^0)$], the terms responsible for mixing enter with a high-frequency factor $\exp[i(S_a - S_k)/\varepsilon]$, and there is no mixing in the lowest order in ε.

If we define $l_\alpha = S_{,\alpha}$, substitute (7.3.6)–(7.3.7) into equations (7.3.1)–(7.3.2) and into the gauge conditions (7.3.5), and then set the terms of order ε^{-2} and ε^{-1} equal to zero, we arrive at the eikonal equation

$$l_\alpha l^\alpha = 0 \tag{7.3.8}$$

and the equations

$$l^\mu \kappa_{\mu\nu} = 0, \quad l^\mu \alpha_\mu = 0, \tag{7.3.9}$$

[9] On the application of the geometrical optics approach to the propagation of high-frequency gravitational waves, see [Isaacson (1968a,b)]. A detailed presentation of this approach for electromagnetic and gravitational perturbations can be found in [Misner, Thorne, and Wheeler (1973)].

$$l^{\beta}{}_{;\beta}\,\alpha^{\mu} + 2\,l^{\beta}\,\alpha^{\mu}{}_{;\beta} = N_1{}^{\mu}, \tag{7.3.10}$$

$$l^{\beta}{}_{;\beta}\,\kappa_{\mu\nu} + 2\,l^{\beta}\,\kappa_{\mu\nu;\beta} = N_{2\mu\nu}\,, \tag{7.3.11}$$

$$N_1{}^{\mu} = l_{\beta}\left(\frac{1}{2}\,F^{\mu\beta}\,\kappa^{\alpha}{}_{\alpha} - F_{\gamma}{}^{\mu}\,\kappa^{\gamma\alpha}\right), \tag{7.3.12}$$

$$N_{2\mu\nu} = -4\left[F_{\mu}{}^{\alpha}\,l_{[\nu}\,\alpha_{\alpha]} + F_{\nu}{}^{\alpha}\,l_{[\mu}\,\alpha_{\alpha]} + \frac{1}{2}\,g_{\mu\nu}\,F^{\alpha\beta}\,\alpha_{\alpha}\,l_{\beta}\right]. \tag{7.3.13}$$

Condition (7.3.8) shows that the constant phase surface $S = \text{const}$ is a null surface. The integral lines $x^{\mu} = x^{\mu}(\lambda)$ defined by the equation

$$\frac{dx^{\mu}}{d\lambda} = l^{\mu}(x) \tag{7.3.14}$$

and lying on this surface are therefore null geodesics, and λ is an affine parameter.

Now augment the vector field l^{μ} to a complex null tetrad (l^{μ}, n^{μ}, m^{μ}, $\bar{m}^{\mu}$), and demand that the vectors of this tetrad be normalized by the conditions

$$l^{\mu}\,n_{\nu} = -1, \qquad m^{\mu}\,\bar{m}_{\mu} = 1 \tag{7.3.15}$$

(the remaining scalar products vanish), while the tetrads themselves be covariantly constant along the integral curves l^{μ}:

$$l^{\mu}\,m^{\nu}{}_{;\mu} = l^{\mu}\,n^{\nu}{}_{;\mu} = 0\,. \tag{7.3.16}$$

The orthogonality condition $l^{\mu}\,m_{\mu} = 0$ implies that the vectors m_{μ} and $\bar{m}_{\mu}$ are tangent to the surface $S = \text{const}$. It can be shown that the arbitrariness remaining in the gauge transformations (7.3.4) (they preserve the additional conditions (7.3.9)) can be used to rephrase the expressions for α_{μ} and $k_{\mu\nu}$ in the form

$$\alpha_{\mu} = A\,\bar{m}_{\mu} + \bar{A}\,m_{\mu}\,,$$

$$\frac{1}{2}\,\kappa_{\mu\nu} = H\,\bar{m}_{\mu}\,\bar{m}_{\nu} + \bar{H}\,m_{\mu}\,m_{\nu} \tag{7.3.17}$$

[see, e.g., Sibgatullin (1984)]. Multiplying (7.3.10) by m^{μ} and (7.3.11) by $m^{\mu}\,m^{\nu}$ and defining $\Phi_0 = F^{\alpha\beta}\,l_{\alpha}\,m_{\beta}$ and $\rho = -\frac{1}{2}\,l^{\beta}{}_{;\beta}$, we arrive at the following system of equations:

$$\frac{dA}{d\lambda} - \rho A = \bar{\Phi}_0\,H\,, \tag{7.3.18}$$

$$\frac{dH}{d\lambda} - \rho H = -\Phi_0\,A\,. \tag{7.3.19}$$

These equations imply the relation

$$\left[l^\mu \left(|A|^2 + |H|^2 \right)\right]_{;\mu} = 0\,, \tag{7.3.20}$$

which can be interpreted as the law of conservation of the total number of photons and gravitons.

The system of equations (7.3.18)–(7.3.19) can be slightly simplified by transforming from the field variables A and H to the variables

$$\tilde{A} = zA\,, \quad \widetilde{H} = zH, \tag{7.3.21}$$

where

$$z = z(\lambda) \equiv z_0 \exp\left[-\int_{\lambda_0}^{\lambda} \rho\, d\lambda\right]. \tag{7.3.22}$$

Equations (7.3.18)–(7.3.19) rewritten in terms of these variables take the form

$$\frac{d\tilde{A}}{d\lambda} = \bar{\Phi}_0\, \widetilde{H}, \quad \frac{d\widetilde{H}}{d\lambda} = -\,\Phi_0\, \tilde{A}\,. \tag{7.3.23}$$

If $\Phi_0 = \bar{\Phi}_0$, this system reduces to a second-order equation

$$\frac{d^2\tilde{A}}{dx^2} + \tilde{A} = 0\,, \tag{7.3.24}$$

where $x = \int \Phi_0\, d\lambda$. This equation shows that the amplitudes of both electromagnetic and gravitational fields undergo oscillations due to the mutual transformations of photons and gravitons. The period of these oscillations, $\Delta\lambda$, is found from the condition [Sibgatullin (1974)]

$$2\pi = \int_{\lambda}^{\lambda+\Delta\lambda} \Phi_0\, d\lambda\,. \tag{7.3.25}$$

7.3.2 Mutual transformation of photons and gravitons

The arguments given above are directly extended to the case of high-frequency electromagnetic and gravitational waves propagating in the field of a charged black hole. The eikonal equation (7.3.8),

$$g^{\mu\nu}\, S_{,\mu}\, S_{,\nu} = 0\,, \tag{7.3.26}$$

in the Reissner-Nordström metric (7.2.1) admits a complete integral

$$S = t \pm R(r) \pm \Psi(\theta) + m\,\phi\,, \tag{7.3.27}$$

where

$$R(r) = \int F^{-1} \sqrt{1 - b^2 F/r^2}\, dr\,,$$

$$\Psi(\theta) = \int \sqrt{b^2 - m^2/\sin^2\theta}\, d\theta\,, \qquad F = 1 - 2M/r + Q^2/r^2. \tag{7.3.28}$$

Light rays forming the surface $S = \text{const}$ are parameterized by arbitrary constants b and m denoting the impact parameter and angular momentum, and are described by the equations

$$S = \text{const}\,, \qquad \frac{\partial S}{\partial b} = \text{const}\,, \qquad \frac{\partial S}{\partial m} = \text{const}\,. \tag{7.3.29}$$

For a given congruence of light rays, the affine parameter λ is related to r by the formula $d\lambda = dr\,(1 - Fb^2r^{-2})^{-1/2}$; the complex null tetrad can be chosen to make Φ_0 a real quantity having the form

$$\Phi_0 = bQ/r^3. \tag{7.3.30}$$

Equation (7.3.25) for the oscillation period now takes the form

$$2\pi = Qb \int \frac{dr}{r^3\sqrt{1 - Fb^2r^{-2}}}\,. \tag{7.3.31}$$

If a high-frequency electromagnetic wave with amplitude A_{in} and impact parameter b is incident on a charged black hole, the result of its passage close to the black hole (unless the wave is captured) is the creation of outgoing electromagnetic and gravitational waves with amplitudes A_{out} and H_{out} [Sibgatullin (1974, 1984)]:

$$A_{out} = A_{in}\cos\left[2Qb\int_{r_0}^{\infty} \frac{dr}{r^3\sqrt{1 - Fb^2r^{-2}}}\right],$$

$$H_{out} = A_{in}\sin\left[2Qb\int_{r_0}^{\infty} \frac{dr}{r^3\sqrt{1 - Fb^2r^{-2}}}\right], \tag{7.3.32}$$

where r_0 is the minimal value of r for a light ray with a given impact parameter b. This value coincides with the maximal root of the equation

$$F(r) = r^2/b^2. \tag{7.3.33}$$

For $b = b_{cr}$, where

$$b_{cr}^2 = 4M^2\left[x + \frac{5}{2} + \sqrt{1 + 8x} + (8x)^{-1}(\sqrt{1 + 8x} - 1)\right],$$

$$x = 1 - Q^2/M^2, \tag{7.3.34}$$

this root equals

$$r_0 = r_{cr} = \frac{M}{2}(3 + \sqrt{1 + 8x}) \tag{7.3.35}$$

and becomes multiple. The integrals in (7.3.32) then diverge. The corresponding impact parameter corresponds to an unstable closed circular orbit. The integrals in (7.3.32) diverge because the conditions of validity of the geometrical optics approximation are violated. A finite answer is obtained when the wave properties of light and gravitational radiation are taken into account. For example, for an extremal ($Q = M$) black hole it is found that when $|b - b_{cr}| \ll O(\omega^{-1})$ the total intensity of the outgoing electromagnetic and gravitational waves adds up to a considerable fraction of the intensity of the incoming electromagnetic radiation. That is, the photon-to-graviton conversion rate becomes of order unity. The remaining energy is absorbed by the black hole.

If the black hole is rotating, the above-described effect of mutual transformation of photons and gravitons is accompanied by an additional rotation of their planes of polarization [Sibgatullin (1984)]. Note that the effect of mutual transformation can be realized only in the neighborhood of a charged black hole. Although we have already mentioned that in realistic astrophysical conditions a black hole is unlikely to have an appreciable charge, certain processes which produce a nonzero electric charge on a black hole are nevertheless possible. One of these processes, caused by an external magnetic field acting on a black hole, will be described in the next chapter. Another possible process suggested by Shvartsman (1971) stems for the difference in the effect of radiative pressure on the electrons and ions of the matter accreting onto a black hole.

7.4 Interaction of Cosmic Strings with a Black Hole

7.4.1 Gravitational capture of cosmic strings by a black hole

Cosmic strings are topologically stable one-dimensional objects which are predicted by grand unified theories. Cosmic strings (as well as other topological defects) may appear during a phase transition in the early Universe. (Detailed discussion of cosmic strings and other topological defects can be found in the book by Shellard and Vilenkin (1994).) The characteristic thickness of a cosmic string is $\rho_0 \sim l_{Pl}(m_{GUT}/m_{Pl})$, while its mass μ per unit length is $\mu \sim (m_{Pl}/l_{Pl})(m_{GUT}/m_{Pl})^2$. Here m_{GUT} is the characteristic mass defining the symmetry breaking in the grand unified theory which is responsible for the cosmic string formation, and m_{Pl} and l_{Pl} are the Planck mass and length, respectively.

For cosmic strings of astrophysical interest the parameter ρ_0 is much less than any other parameters such as the length of the string or the size of inhomogeneity of the gravitational field so that one can idealize the motion of the string as a two-dimensional world sheet $x^\mu = x^\mu(\tau, \sigma)$ which describes the evolution in time of the line representing the string. It can be shown [see e.g., Shellard and Vilenkin (1994)]

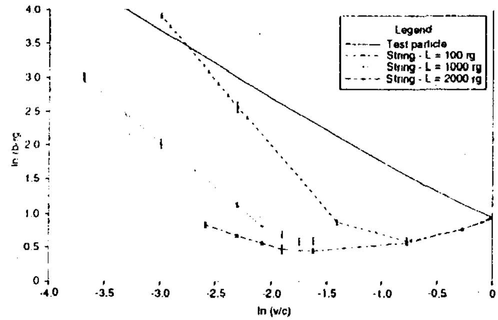


Figure 7.2: The capture impact parameter b as the function of the initial velocity v of the cosmic string.

that the equations of motion of a cosmic string in this approximation can be obtained from the *Nambu-Goto action*

$$I = -\mu \int d^2\varsigma \sqrt{-\det\left[g_{\mu\nu}\frac{\partial x^\mu}{\partial\varsigma^a}\frac{\partial x^\nu}{\partial\varsigma^b}\right]}\,. \tag{7.4.1}$$

Here $g_{\mu\nu}$ is an external metric and ς^a stands for the world-sheet coordinates ($a, b = 0, 1$, $\varsigma^0 = \tau$, $\varsigma^1 = \sigma$). A two-dimensional world sheet which provides an extremum to the Nambu-Goto action is a *minimal surface*.

We consider the interaction of a cosmic string with a black hole. For this problem the gravitational field created by the string can be neglected because the dimensionless parameter which characterizes the strength of the field, $\mu^* = G\mu/c^2$, is negligibly small (for GUT one has $\mu^* \sim 10^{-6}$). It means that we can consider the string as a test object moving in the background of a black hole metric. We also assume that the length of the string L is much greater than the gravitational radius and use the infinite-string approximation.

When a cosmic string is moving in the gravitational fields of a black hole and passes near the horizon, the string can be captured by the black hole. This process resembles the gravitational capture of particles by black holes, discussed in Section 2.9 and Section 3.4.5.

The gravitational capture of cosmic string by a black hole was studies by De Villiers and Frolov (1997). They considered a straight string moving in the gravitational field of a non-rotating black hole with the initial velocity v and used calculations on a supercomputer to define the capture impact parameter b as the function of v. Their results are presented in Figure 7.2. In order to solve the problem numerically, a long finite string with heavy particles of mass m at its ends was considered at first. In the limit $m \to \infty$ the end points move along geodesics. Dashed, dotted, and dashed and dotted lines represent the results for strings of different length L. Calculations show that for a fixed value of the initial velocity v starting with some length $L(v)$ the impact parameter b practically does not depend on L. For comparison, the solid line give the capture impact parameter for a test particle with initial velocity v. Due to the tension, the string must come closer to the black hole than a particle in order to be captured. (Earlier numerical results obtained by Lonsdale and Moss (1988) does not satisfy this properties, and it seems that they contain a mistake.)

After a capture the string remains attached to the black hole. The action of the string on the black hole will finally change the velocity of the black hole. But for a massive black holes and small parameter μ^* the characteristic time T_c of this process is very large. We restrict ourselves by considering time scales much smaller than T_c so that this effect can be neglected.

7.4.2 Stationary string near a rotating black hole

Stationary configurations of cosmic strings, and especially stationary configurations of trapped cosmic strings, are of particular interest. One can expect that these configurations might play the role of final equilibrium configurations. In this section we discuss such configurations following the paper by Frolov *et al.* (1989).

Consider a stationary spacetime with the Killing vector $\xi^\mu_{(t)}$ normalized to unity at asymptotic infinity. A two-dimensional world sheet Σ represent a stationary string if the Killing vector $\xi^\mu_{(t)}$ is tangent to Σ, and Σ is a solution of the Nambu-Goto equations. The metric of a stationary spacetime can be written as [compare (A.57)]

$$ds^2 = -F(\,dt + g_i\,dx^i)^2 + F^{-1}\,h_{ij}\,dx^i\,dx^j, \tag{7.4.2}$$

where $F = -\,\xi^2_{(t)}$, g_i, and h_{ij} are independent of time. Let us choose $\tau = t$, then the string configuration is specified by giving the functions $x^i = x^i(\sigma)$ and Nambu-Goto action becomes

$$I = -E\,\Delta t, \tag{7.4.3}$$

where

$$E = \mu \int d\sigma \sqrt{h_{ij}\,\frac{dx^i}{d\sigma}\,\frac{dx^j}{d\sigma}} \tag{7.4.4}$$

is the energy of the string. The equilibrium stationary configurations correspond to a minimum of the energy, and hence they are geodesics in a three-dimensional space with the metric

$$dl^2 = h_{ij}\,dv^i\,dx^j\,. \tag{7.4.5}$$

In a static spacetime, the meaning of the expression for the energy of the string is transparent. The proper mass of a piece of the string is proportional to its proper length. The integrand in (7.4.4) is nothing but a redshifted proper length.

For the Kerr-Newman geometry one has[10]

$$F = \frac{\Delta - a^2 \sin^2\theta}{r^2 + a^2\cos^2\theta}, \tag{7.4.6}$$

$$g_i = \delta_i^\phi\, g_\phi, \quad g_\phi = a\sin^2\theta\,\frac{2Mr - Q^2}{\Delta - a^2\sin^2\theta}. \tag{7.4.7}$$

The metric h_{ij} is diagonal with

$$h_{rr} = \frac{\Delta - a^2\sin^2\theta}{\Delta}, \quad h_{\theta\theta} = \Delta - a^2\sin^2\theta, \quad h_{\phi\phi} = \Delta\sin^2\theta. \tag{7.4.8}$$

The analysis of stationary string configurations in the Kerr-Newman spacetime is thus reduced to the problem of finding geodesics in a three-dimensional space with metric h_{ij}. This problem is similar in many respects to the study of the motion of test particles in the Kerr-Newman geometry. In particular the geodesic equation in the space dl^2 can be integrated.

7.4.3 Separation of variables

The first integrals of the equations describing stationary string configurations in the Kerr-Newman geometry can be easily obtained by using the Hamilton-Jacobi method. The Hamilton-Jacobi equations read

$$\frac{\partial S}{\partial\sigma} + h^{ij}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j} = 0, \tag{7.4.9}$$

where the parameter σ is a proper distance along the string in metric h_{ij}. For metric (7.4.8), equation (7.4.9) allows separation of variables [Frolov *et al.* (1989)]

$$S = -\frac{1}{2}\mu^2\sigma + L\phi + S_r(r) + S_\theta(\theta). \tag{7.4.10}$$

Substitution of (7.4.10) into (7.4.9) gives

$$\left(\frac{dS_r}{dr}\right)^2 = \mathcal{R}, \tag{7.4.11}$$

$$\left(\frac{dS_\theta}{d\theta}\right)^2 = \Theta, \tag{7.4.12}$$

where

$$\mathcal{R} = \frac{a^2L^2}{\Delta^2} - \frac{Q^2}{\Delta} + \mu^2, \tag{7.4.13}$$

[10]The corresponding formulas can be found in Appendix D.5.1. Note that in this section we use the metric h_{ij}, which differs from the one chosen in the appendix by an additional factor F.

$$\Theta = -\frac{L^2}{\sin^2\theta} - \mu^2 a^2 \sin^2\theta + Q^2. \tag{7.4.14}$$

A non-negative separation constant Q^2 is related to the Killing tensor ξ_{ij}

$$\xi^i_j = \mathrm{diag}(a^2\sin^2\theta, \Delta, \Delta + a^2\sin^2\theta) \tag{7.4.15}$$

by

$$Q^2 = \xi^{ij}\, p_i\, p_j\,, \quad p_i = \frac{\partial S}{\partial x^i}\,. \tag{7.4.16}$$

The existence of the Killing tensor (7.4.15) for the metric (7.4.8) is a property which is inherited from the original Kerr-Newman metric. The discussion of this mechanism and the separation of variables for a stationary string in a curved spacetime can be found in [Carter and Frolov (1989); Carter, Frolov, and Heinrich (1991)].

The geodesics of the metric h_{ij} which represent stationary string configurations follow from

$$\frac{\partial S}{\partial Q} = \mathrm{const}\,, \quad \frac{\partial S}{\partial L} = \mathrm{const}\,, \quad \frac{\partial S}{\partial \mu^2} = \mathrm{const}\,. \tag{7.4.17}$$

These relations in explicit form are

$$\int^\theta \frac{d\theta}{\sqrt{\Theta}} = \int^r \frac{dr}{\Delta\sqrt{\mathcal{R}}}\,, \tag{7.4.18}$$

$$\phi = L\left(\int^\theta \frac{d\theta}{\sin^2\theta\,\sqrt{\Theta}} - a^2\int^r \frac{dr}{\Delta^2\sqrt{\mathcal{R}}}\right), \tag{7.4.19}$$

$$\sigma = -a^2\int^\theta \frac{\sin^2\theta\, d\theta}{\sin^2\theta\cdot\sqrt{\Theta}} + \int^r \frac{dr}{\sqrt{\mathcal{R}}}\,. \tag{7.4.20}$$

For more detailed discussion of stationary string configurations see Frolov *et al.* (1989). It should be emphasized that metric (7.4.8) becomes singular on the static limit surface. For this reason the above described approach is directly applicable only in the regions lying outside the ergosphere. In the next section we demonstrate that there is a special class of these solutions which describe a string crossing the ergosphere and entering the black hole.

7.4.4 Uniqueness theorem

We describe now a special class of stationary string configurations which play an important role in the Kerr-Newman spacetime. For this purpose consider a *principal null geodesic* $\gamma_\pm$ (see Appendix D.6). We use signs $\pm$ to distinguish outgoing (+) and incoming (−) rays. Denote by l^μ a tangent vector to γ and by $\Sigma_\pm$ a timelike surface which is formed by the trajectories of the Killing vector field $\xi^{\mu}_{(l)}$ passing through $\gamma_{\pm}$. We call $\Sigma_{\pm}$ a *principal Killing surface*. Denote by n^{μ}_{R} $(R = 2,3)$ two unit vectors normal to $\Sigma_{\pm}$; then it is easy to show that the trace $\Omega_{RA}{}^{A}$ of the second fundamental form Ω_{RAB} is

$$\begin{aligned}\Omega_{RA}{}^{A} &= g_{\mu\nu}\, G^{AB}\, n^{\mu}_{R}\, x^{\gamma}_{,A}\, \nabla_{\gamma}\, x^{\nu}_{,B} \\ &= g_{\mu\nu}\, n^{\mu}_{R} \left[\frac{2}{(\xi\cdot l)}\, l^{\gamma}\, \nabla_{\gamma}\, \xi^{\nu} + \frac{F}{(\xi\cdot l)^{2}}\, l^{\gamma}\, \nabla_{\gamma}\, l^{\nu} \right].\end{aligned} \tag{7.4.21}$$

An important property of the Kerr-Newman geometry is that the principal null vector l^{μ} is an eigenvector of the antisymmetric tensor $\xi_{(t)\mu;\nu}$. This property and the geodesic equation for l^{μ} imply that $\Omega_{RA}{}^{A} = 0$, and hence the principal Killing surface is a minimal surface [Frolov and Larsen (1995); Frolov, Hendy, and Larsen (1996)]. These minimal surfaces $\Sigma_{\pm}$ are solutions of the Nambu-Goto equations which correspond to the special case when the parameters L and Q obey the relation

$$Q^{2} = \pm 2\, a L, \qquad |\, L\,| \leq a\,. \tag{7.4.22}$$

Such a string lies on the surface $\theta = \theta_0 = \text{const}$, where $\theta_0 = \arcsin(\,|\,L/a\,|^{1/2})$ and its equation is

$$\frac{d\phi}{dr} = \pm\, \frac{a}{\Delta}\,. \tag{7.4.23}$$

The following **uniqueness theorem** for stationary cosmic strings was proved by Frolov, Hendy, and Larsen (1996):

The only stationary timelike minimal two-surface which crosses the static limit surface S_{st} and is regular in its vicinity is a principal Killing surface.

This result implies in particular that a stationary string can cross the static limit surface only along very special directions. For a non-rotating black hole the static limit coincides with the event horizon, and a stationary string enters the horizon orthogonally. This result resembles the no hair theorems discussed in Chapter 6. It has a similar physical origin. Namely, in order for the string to be stationary as a whole, the resulting force acting on a string element must be directed along the string. Near the static limit surface where the gravitational force grows infinitely this condition leaves little freedom for the orientation of the string elements. It is also worth mentioning that the forces acting along the string can "stretch" it, but due to the relativistically invariant nature of the string equation of state its tension remains the same.

7.5 Black Hole in an External Field

7.5.1 Black hole in an external field

A black hole subjected to an external force behaves to a certain extent like a compact elastic body. Some specifics of the "response" of a black hole to this force are mostly caused by the unambiguous dependence of the black hole size on its mass while the gravitational self-action of this mass is extremely high.

Before passing on to a detailed description of black hole behavior in an external field, let us dwell on an aspect that often leads to misunderstanding. Imagine a situation of an isolated (e.g., Schwarzschild) black hole at rest and a distant observer who at some moment of time switches on an external field in order to find the effect on the black hole. For instance, it can be a plane light wave directed at the hole. In general, the pressure of such a wave on an ordinary body results in motion of the body (via the effects of light absorption and scattering). On the other hand, if we monitor the propagation of the light wavefront in the Schwarzschild metric, we discover that an infinitely long time (by the clock of the distant observer) is required for the wavefront to reach the gravitational radius. A detailed calculation shows [e.g., see Hanni (1977)] that the wavefront bends around the black hole after which the wave propagates further. After some time, an outgoing component appears, indicating scattering, and a new wavefront forms around the black hole; it corresponds to the radiation falling into it. The question is: How can a black hole "feel" the pressure of radiation and start moving in a finite time (by the clock of distant observer) if for this observer the radiation never reaches the event horizon?[11]

To answer this question, it is instructive to analyze a similar situation that arises when a light wave is scattered by a body of radius r_0, composed of matter whose refractive index n increases continuously inward from 1 at the boundary (at $r = r_0$) to infinity on some surface inside this body (at $r = r_1$). As in the case of a black hole, the process of light propagation to the surface $r = r_1$ may take an infinitely long time. However, the body starts moving before this process is completed. It can be shown that once the wave front has reached the surface $r = r_0$, the energy-momentum flux across this surface becomes nonzero in the general case, and the force due to this flux causes the motion of the body as a whole.

Likewise, if a black hole is enclosed in an imaginary sphere of radius $r_0 \geq r_g$, a light wave produces a nonzero energy-momentum flux across this surface and the entire region within r_0 (the black hole and "part of the space" surrounding it) starts moving with respect to the external observer. The energy of an ordinary light wave in a volume of order r_g^3 is much smaller than $r_g c^4/G$ so that this wave produces a negligible effect on the black hole metric. As a result, in the immediate vicinity of the black hole all phenomena unfold as if there were no radiation (in the reference

[11]Obviously, an observer falling into the hole will find that it takes the light wavefront a finite time (by his clock) to cross the event horizon.

frame fixed to the black hole). On the other hand, a distant observer will be able to detect the motion of this reference frame with respect to the observer's.

7.5.2 Perturbation theory

Although the general problem of the motion of a black hole in an external field does not admit an analytical solution, a detailed description of the motion is possible in terms of a sort of perturbation theory in the particular case of a black hole not interacting too strongly with its surroundings. For example, this description can be given for the motion in an external gravitational field if the characteristic size of the black hole, determined by its mass M, is much less than the characteristic scale L of inhomogeneities of the gravitational field in which the black hole moves.[12] In this case the external gravitational field slightly changes the metric in the neighborhood of the black hole. Hence, the metric $g_{\mu\nu}$ in the region outside the black hole, determined by the characteristic scale M, can be written in the form

$$g_{\mu\nu} = g^{[0]}_{\mu\nu} + \varepsilon\, g^{[1]}_{\mu\nu} + \varepsilon^2\, g^{[2]}_{\mu\nu} + \cdots\,, \tag{7.5.1}$$

where $g^{[0]}_{\mu\nu}$ is the metric of the unperturbed black hole (the Kerr metric), and $\varepsilon = M/L$. Likewise, the effect of the black hole on the external metric on the scale of order L can be considered small; it can be described as corrections, with the external metric written in the form

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + \varepsilon\, g^{(1)}_{\mu\nu} + \varepsilon^2\, g^{(2)}_{\mu\nu} + \cdots\,. \tag{7.5.2}$$

The assumption of the absence in the black hole neighborhood of matter whose fall can result in a sufficiently fast modification of its parameters, and the condition of weakness of a deformation of an external field by a black hole imply that expansions (7.5.1) and (7.5.2) have a common domain of applicability. In other words, both expansions hold simultaneously in an interval of distances from the black hole that is determined by a characteristic scale $l \sim \varepsilon^\alpha M$ $(0 < \alpha < 1,\ M \ll l \ll L)$. The expansions themselves are found unambiguously by comparing (matching) them in this common region.

The method of matching the asymptotic expansions for analyzing the motion of black holes in external fields and the interaction between the black holes was developed by D'Eath (1975a,b, 1978, 1979). The following results were obtained by this method in the problem of interest here [Demianski and Grishchuk (1974), D'Eath (1975a,b, 1979), Kates (1980, 1981), Damour (1983), Thorne and Hartle (1985)]. The term $g^{[1]}_{\mu\nu}$ in expansion (7.5.1) vanishes. The mass M, momentum P, and angular momentum J of a black hole can be determined from the parameters

[12] In general, the inhomogeneities of the gravitational field are characterized both by the curvature radius L_1 of spacetime and by the scales of the space (L_2) and time (L_3/c) over which this curvature changes appreciably. The smallest of these quantities is denoted by L: $L = \min(L_1, L_2, L_3)$.

of the unperturbed metric $g^{[0]}_{\mu\nu}$ even though the spacetime as a whole need not be asymptotically flat. In order of magnitude, the uncertainties in these parameters due to the terms $\varepsilon^2\, g^{[2]}_{\mu\nu}$ are

$$\Delta M \sim \varepsilon^2 M, \quad \Delta P \sim \varepsilon^2 M, \quad \Delta J \sim \varepsilon^2 M. \tag{7.5.3}$$

The external space metric $g^{(0)}_{\mu\nu}$ is everywhere regular, but the corrections to it, formally extended to the entire space, reveal singularities corresponding to the motion of the black hole on the timelike world lines γ. Coordinates (t, x^i) can be introduced in the neighborhood of γ such that the metric takes the form

$$\begin{aligned} g^{(0)}_{00} &= -1 - E_{ij}\, x^i x^j + \cdots, \\ g^{(0)}_{0i} &= \frac{2}{3}\, \varepsilon_{ijk}\, B^j{}_l\, x^k x^l + \cdots, \\ g^{(0)}_{ij} &= \delta_{ij}\,(1 - E_{kl}\, x^k x^l) + \cdots. \end{aligned} \tag{7.5.4}$$

Here

$$E_{ij} = R_{i0j0} \quad \text{and} \quad B_{ij} = \frac{1}{2}\, \varepsilon_{ikl}\, R^{kl}{}_{j0} \tag{7.5.5}$$

are the "electric" and "magnetic" components of the curvature tensor. A change of the part of perturbation $g_{\mu\nu} - g^{(0)}_{\mu\nu}$ which is regular on the trajectory of the black hole leads to a small uncertainty in the determination of E_{ij} and B_{ij}.

The trajectory of the black hole in the lowest approximation in ε is a geodesic, with the spin of the black hole parallelly transported along this line [D'Eath (1975a)]. The corrections describing the deviations in the motion of the black hole from a geodesic and from the Fermi-transport of spin are given by the following equations [see Thorne and Hartle (1985)]

$$\frac{dP^i}{dt} = -\, B^i{}_j\, J^j, \tag{7.5.6}$$

$$\frac{dJ^i}{dt} = -\, \varepsilon^i{}_{jk}\, E^k{}_l\, \frac{1}{M}\, J^j J^l. \tag{7.5.7}$$

The changes in the momentum and the spin of the black hole that follow from these equations greatly exceed the uncertainties ΔP^i and ΔJ^i of (7.5.3) caused by changes in $E^i{}_j$ and $B^i{}_j$. In principle, therefore, these effects of non-geodesic motion of a rotating black hole and of the precession of its angular momentum are observable.

It should be emphasized that equations (7.5.6) and (7.5.7) coincide in form with the equations of motion of rotating test particles in an external gravitational field. It is important that when the extremely strong gravitational self-action of black holes is taken into account, the form of the equations is not changed. To a distant observer, a black hole continues to move in the external field in the same manner as a small test particle did.

Likewise, it can be shown that when an external electric field E acts on a black hole of charge Q and mass M, the hole acquires an acceleration $a = QE/M$. An exact solution of the Einstein-Maxwell equations describing the motion of a charged black hole in a uniform electric field was obtained by Ernst (1976b) (see also [Bičák (1980)]). The corresponding metric is given in Section 14.4.4 where we shall discuss properties of the Ernst solution in detail.

7.5.3 Deformed black holes

An external field causes a deformation of a black hole. We will briefly outline the changes induced in the properties of a black hole when it is "inserted" into a gravitational field produced by a distribution of matter. If the black field is non-rotating, and the gravitational field is axially symmetric, this problem admits a fairly complete solution [Israel and Khan (1964), Doroshkevich *et al.* (1965), Mysak and Szekeres (1966), Israel (1973), Geroch and Hartle (1982)]. A generalization to the case of a rotating black hole can be found in Tomimatsu (1984).

Recall, first of all, that a static axially symmetric vacuum gravitational field is described by the *Weyl metric*

$$ds^2 = -\,e^{2U}\,dt^2 + e^{-2U}\left[e^{2V}(d\rho^2 + dz^2) + \rho^2\,d\phi^2\right], \tag{7.5.8}$$

where U and V are functions of ρ and z and satisfy the equations

$$U_{,\rho\rho} + \frac{1}{\rho}U_{,\rho} + U_{,zz} = 0, \tag{7.5.9}$$

$$V_{,\rho} = \rho\,(U^2_{,\rho} - U^2_{,z}), \quad V_{,z} = 2\rho\,U_{,\rho}\,U_{,z} \tag{7.5.10}$$

[see, e.g., Kramer *et al.* (1980)]. It is quickly verified that system (7.5.10) is integrable if condition (7.5.9) is satisfied. The vacuum metric (7.5.8) is uniquely determined by choosing a solution of equation (7.5.9) because a solution of system (7.5.10) for a given function $U(\rho, z)$ is easily found by quadrature. Thus, the Schwarzschild metric in these coordinates corresponds to the following solution

$$U = U_S \equiv \frac{1}{2}\ln\left(\frac{\lambda-1}{\lambda+1}\right), \quad V = V_S \equiv \frac{1}{2}\ln\left(\frac{\lambda^2-1}{\lambda^2-\mu^2}\right), \tag{7.5.11}$$

where

$$\lambda = \frac{R_+ + R_-}{2M}, \quad \mu = \frac{R_+ - R_-}{2M}, \quad R_\pm = \sqrt{\rho^2 + (z \pm M)^2}; \tag{7.5.12}$$

the event horizon H is now defined by the condition

$$\rho = 0, \quad -M \le z \le M. \tag{7.5.13}$$

In view of the uniqueness theorem (see Section 6.4), this solution is unique for a static black hole in vacuum in an asymptotically flat spacetime. Any other static *vacuum* solution having a regular horizon cannot be asymptotically flat.

Strictly speaking, therefore, a perturbed black hole is necessarily described by a nonvacuum solution of Einstein's equations. In the simplest case, we can assume that the matter creating the external gravitational field is located at some distance from the black hole. Then the gravitational field in the vicinity of its horizon is described by the vacuum Weyl metric (7.5.8). The exact solution of such complete problems can be found only in some very particular cases of matter distribution. Nevertheless, an analysis of the properties of vacuum Weyl solution close to a regular horizon yields detailed information on the possible effect of external influence on the black hole surface.[13]

The solution describing a perturbed black hole can be written in the following form [Geroch and Hartle (1982)]:

$$U = U_S + \hat{U}, \quad V = V_S + \hat{V}, \tag{7.5.14}$$

where U_S and V_S are given by (7.5.11), and $\hat{U}$ is the solution of the homogeneous equation (7.5.9) which satisfies the regularity condition on the segment $\rho = 0$, $-M \leq z \leq M$ and in a certain neighborhood of this segment; $\hat{U}$ takes on identical values at the two ends of the segment:

$$\hat{U}(\rho = 0, z = -M) = \hat{U}(\rho = 0, z = M) \equiv u\,. \tag{7.5.15}$$

The value of $\hat{V}$ is uniquely determined from equations (7.5.10) provided $V = 0$ on the parts of the axis $\rho = 0$ that lie outside the matter-occupied region. Condition (7.5.15), reflecting the absence of nodal singularities, guarantees that the net force applied by the external field to the black hole as a whole vanishes; hence, a regular static solution is possible. Using the second equation in (7.5.10), one can prove that the following relation holds at the event horizon:

$$\hat{V} = 2\hat{U} - 2u\,. \tag{7.5.16}$$

Using the coordinate transformation

$$\rho = \mathrm{e}^{u}\, r\sqrt{1 - \frac{2M_0}{r}}\,\sin\theta_0\,, \quad z = \mathrm{e}^{u}\,(r - 2M_0)\cos\theta\,, \tag{7.5.17}$$

$$M_0 = M\,\mathrm{e}^{-u}\,, \tag{7.5.18}$$

it is possible to recast metric (7.5.8) to the form

$$ds^2 = -\,\mathrm{e}^{-2\hat{U}}\left(1 - \frac{2M_0}{r}\right)dt^2 + \mathrm{e}^{2\hat{V}-2\hat{U}+2u}$$

[13] It is of interest to note that even though the Weyl metric describing a distorted axially symmetric black hole belongs to the Petrov type I, its Petrov type on the event horizon is D [Papadopoulos and Xanthopoulos (1984)].

$$\times\left[\left(1-\frac{2M_0}{r}\right)^{-1}dr^2+r^2(d\theta^2+e^{-2\hat{V}}\sin^2\theta\,d\phi^2)\right]. \tag{7.5.19}$$

In these coordinates the event horizon is described by the equation $r = 2M_0$, and the two-dimensional metric on its surface becomes, after (7.5.16) is taken into account,

$$ds^2 = 4M_0^2\left(e^{2\hat{U}-2u}\,d\theta^2+e^{-2\hat{U}+2u}\sin^2\theta\,d\phi^2\right). \tag{7.5.20}$$

It is easily found that the horizon surface is a sphere deformed in an axially symmetric manner[14] whose area is

$$A = 16\pi M_0^2. \tag{7.5.21}$$

To a certain extent, the action of the external gravitational field on a black hole, with $\exp\hat{U}$ playing the role of potential, is analogous to the effect produced on an elastic massive body. In order to verify this statement, consider the case where the values of $\hat{U}$ are higher at the poles that at the equator. In this situation any test body at rest near a pole is subject to a force that tends to move the body toward the equator. Correspondingly, the horizon surface is found to be flattened at the poles, as follows from (7.5.20).

The surface gravity κ is constant over the horizon surface:

$$\kappa = e^u/4M_0. \tag{7.5.22}$$

The difference between the value of mass M observed at infinity and the value M_0 determining the area of the event horizon and playing the role of irreducible mass (see Section 7.1) is explained in the following manner. Consider the process of "inserting" a Schwarzschild black hole (obviously, with $M = M_0$) into a given external gravitational field. If this process is carried out slowly enough, the area A of the black hole surface remains constant; hence, M_0 is preserved in this process. On the other hand, the quantity $M = e^u M_0$ undergoes a change when the gravitational potential e^u of the external field changes at the point where the black hole is located. The difference $M - M_0 = M_0(e^u - 1)$ is the work done by the external gravitational field when the black hole is "inserted" into it. It can be shown [Geroch and Hartle (1982)] that if the strong energy condition is satisfied, the potential u assumes only non-positive values.

As an example illustrating these general arguments, consider the explicit expression for the metric that describes a black hole in an external quadrupole field [Doroshkevich *et al.* (1965)]:

$$\hat{U} = \frac{1}{4}q(3\lambda^2-1)(3\mu^2-1), \tag{7.5.23}$$

[14] It should be mentioned that vacuum axisymmetric static metrics are possible, describing the spacetime close to a horizon which has the topology of torus [see, e.g., Geroch and Hartle (1982), Xanthopoulos (1983), and references cited in these papers]. However, the Hawking's theorem implies (see Section 5.2) that if the energy dominance condition is satisfied, the external space cannot then be regular and at the same time asymptotically flat.

$$\hat{V} = -3q\,\lambda(1-\mu^2) - \frac{9}{16}q^2(\lambda^2-1)(1-\mu^2)(9\mu^2\lambda^2-\lambda^2-\mu^2+1)\,, \qquad (7.5.24)$$

where q is a parameter characterizing the quadrupole moment of the system which produces the external gravitational field. At the horizon we have $\lambda = 1$, $u = z/M$ so that we immediately obtain that in the case under consideration $u = q$. The Gaussian curvature of the event horizon is

$$K = \frac{e^q}{4M_0^2}\left(1 + 3q - 12q\mu^2 - 9q^2\mu^2 + 9q^2\mu^4\right). \qquad (7.5.25)$$

A constant external quadrupole field described by the solution (7.5.23)–(7.5.24) can be created by distant masses at rest. This solution also describes approximately the effect produced on a black hole by distant free masses whose velocities due to their mutual attraction are initially small and whose field is nearly static.

We have already indicated that the main difficulty in deriving exact complete axially symmetric stationary solutions, describing the behavior of a black hole in a gravitational field, lies in finding the solution in the region with matter. The cases of a black hole in a uniform electric field [Ernst (1976b)] and in a uniform magnetic field [Ernst (1976a), Ernst and Wild (1976), Gal'tsov and Petukhov (1978), Gal'tsov (1980), Wild and Kerns (1980), Aliev *et al.* (1980), Wild *et al.* (1981), Krori *et al.* (1983, 1984), Dadlich (1983), Dhuraudhar and Dadlich (1984a,b)] are important exceptions.

7.6 Interaction Between Black Holes

7.6.1 Interaction of two non-relativistic black holes

The method of matching of asymptotic expansions (see above) also serves to analyze the interaction between two black holes. If the distance between the black holes greatly exceeds their gravitational radii and the holes move with respect to each other at a velocity much lower than the speed of light, the corresponding equations of motion of such interacting black holes were obtained by D'Eath (1975b, 1979) [see also Thorne and Hartle (1985)]. The gravitational field in the neighborhood of each black hole is described by a perturbed Kerr metric. The metric far from the black holes is found using the post-Newtonian approximation to the required order of accuracy. Matching these expansions yields the following system of equations of motion for one of the black holes and of the precession of its angular momentum in the field of the other black hole:

$$m_1\frac{d^2\mathbf{x}_1}{dt^2} = \mathbf{F}_1^{(1)} + \mathbf{F}_1^{(2)} + O(\varepsilon^4)\,, \qquad (7.6.1)$$

$$\frac{d\mathbf{J}_1}{dt} = \left[\left(\boldsymbol{\Omega}_1^{(1)} + \boldsymbol{\Omega}_1^{(2)} + \boldsymbol{\Omega}^{(3)}\right)\times\mathbf{J}_1\right]. \qquad (7.6.2)$$

Here we denote by $\mathbf{x}_i$ the position and by $\mathbf{v}_i$ the velocity of the ith black hole having a mass M_i and an angular momentum $\mathbf{J}_i$; $\mathbf{r}_{21} = \mathbf{x}_2 - \mathbf{x}_1$, $\mathbf{v}_{21} = \mathbf{v}_2 - \mathbf{v}_1$ are the position and velocity of the second black hole with respect to the first; $r = |\mathbf{r}_{21}|$, and $\mathbf{n} = \mathbf{r}_{21}/r$. Quantities

$$J_i = |\mathbf{J}_i| \quad \text{and} \quad \mathbf{j}_i = \mathbf{J}_i/J_i \tag{7.6.3}$$

are the magnitude and the unit vector in the direction of angular momentum of the ith black hole. The smallness parameter ε is equal to the ratio of the larger of the two gravitational radii to the characteristic distance between black holes. The value of the force $\mathbf{F}_1^{(1)}$ corresponding to the geodesic law of motion of one body in the gravitational field created by the second body [it was found by Einstein, Infeld, and Hoffman (1938)] is

$$\mathbf{F}_1^{(1)} = \frac{M_1 M_2}{r^2}\left\{\mathbf{n}\left(1 - \frac{4M_2 + 5M_1}{r} + \mathbf{v}_1^2 + 2\mathbf{v}_2^2 - 4\mathbf{v}_1\mathbf{v}_2 - \frac{3}{2}(\mathbf{v}_2\mathbf{n})^2\right)\right.$$
$$\left. - \mathbf{v}_{21}[\mathbf{n}(3\mathbf{v}_2 - 4\mathbf{v}_1)]\right\}. \tag{7.6.4}$$

The term $\mathbf{F}_1^{(2)}$ in (7.6.1),

$$\mathbf{F}_1^{(2)} = \frac{M_1 J_2}{r^3}\{6\mathbf{n}([\mathbf{j}_2 \times \mathbf{n}]\mathbf{v}_{12}) + 4[\mathbf{j}_2 \times \mathbf{v}_{12}] - 6[\mathbf{j}_2 \times \mathbf{n}](\mathbf{v}_{12}\mathbf{n})\}$$
$$+ \frac{M_2 J_1}{r^3}\{6\mathbf{n}([\mathbf{j}_1 \times \mathbf{n}]\mathbf{v}_{12}) + 3[\mathbf{j}_1 \times \mathbf{v}_{12}] - 3[\mathbf{j}_1 \times \mathbf{n}](\mathbf{v}_{12}\mathbf{n})\}, \tag{7.6.5}$$

describes the additional force due to spin-orbital interaction. The term $O(\varepsilon^4)$ in the same equation corresponds to the spin-spin interaction and to the interaction of the quadrupole moment of the black hole with the curvature: Both are of order ε^4.

Equation (7.6.2) describes the precession of the angular momentum of a black hole with respect to a comoving orthogonal reference frame that does not rotate with respect to an infinitely distant observer. The gravimagnetic, $\mathbf{\Omega}_1^{(1)}$, and the geodesic, $\mathbf{\Omega}_1^{(2)}$, components of the angular velocity of this precession and the component $\mathbf{\Omega}_1^{(3)}$ due to the coupling of the quadrupole moment of the black hole to curvature are, respectively,

$$\mathbf{\Omega}_1^{(1)} = \frac{1}{r^3}[-\mathbf{J}_2 + 3\mathbf{n}(\mathbf{n}\mathbf{J}_2)],$$
$$\mathbf{\Omega}_1^{(2)} = \frac{M_2}{r^2}\left[\left(\frac{3}{2}\mathbf{v}_1 - 2\mathbf{v}_2\right) \times \mathbf{n}\right], \tag{7.6.6}$$
$$\mathbf{\Omega}_1^{(3)} = \frac{3M_2}{M_1 r^3}\mathbf{n}(\mathbf{n}\mathbf{J}_1).$$

In the limiting case of $M_1 \ll M_2$, these formulas coincide with the equations of motion of a rotating test particle in the field of a massive rotating body. A detailed description of the solution of this last problem can be found in the monograph of Misner, Thorne, and Wheeler (1973) which also gives references to numerous original publications.

For more details on the problem of motion of black holes in a given field and on interaction between them, see Thorne, Price, and Macdonald (1986).

7.6.2 Interaction of relativistically moving black holes

In the opposite case of the relative velocity v of two black holes being close to the speed of light, an expansion in the small parameter γ^{-1} can be used $[\gamma = (1-v^2)^{-1/2}]$. This method was employed by D'Eath (1975b, 1979) to solve the problem of the scattering of two ultra-relativistic black holes moving on parallel, oppositely directed courses. The calculation started with the expression for the metric of an isolated non-rotating black hole moving at a constant velocity, in the limit $v \to c$. This metric can be obtained by transformations from the Schwarzschild metric written in isotropic coordinates:

$$ds^2 = -\left(\frac{1 - M/2r}{1 + M/2r}\right)^2 dt^2 + \left(1 + \frac{M}{2r}\right)^4 (dx^2 + dy^2 + dz^2), \tag{7.6.7}$$

where $r^2 = x^2 + y^2 + z^2$, and M is the black hole mass. Applying the Lorentz boost transformation

$$\begin{aligned} \bar{t} &= (1-v^2)^{-1/2}(t + vx), \quad \bar{y} = y, \\ \bar{x} &= (1-v^2)^{-1/2}(x + vt), \quad \bar{z} = z, \end{aligned} \tag{7.6.8}$$

we rewrite (7.6.7) in the form

$$\begin{aligned} ds^2 &= (1+A)^4 \left(-d\bar{t}^2 + d\bar{x}^2 + d\bar{y}^2 + d\bar{z}^2\right) \\ &+ \left[\left(\frac{1-A}{1+A}\right)^2 - (1+A)^4\right] \frac{(d\bar{t} - v\,d\bar{x})^2}{(1-v^2)}, \end{aligned} \tag{7.6.9}$$

where

$$\begin{aligned} A &= \frac{M}{2r} = \frac{\mu(1-v^2)}{2[(\bar{x} - v\bar{t}\,)^2 + (1-v^2)(\bar{y}^2 + \bar{z}^2)]^{1/2}}, \\ \mu &= \gamma M = M/\sqrt{1-v^2}. \end{aligned} \tag{7.6.10}$$

Taking the limit $v \to 1$ in (7.6.9) at a constant value of energy μ and denoting $\bar{t} - \bar{x} = \bar{v}$, $\bar{t} + \bar{x} = \bar{u}$, we find [Aichelburg and Sexl (1971)]

$$ds^2 = -\,d\bar{v}\,d\bar{u} + d\bar{y}^2 + d\bar{z}^2 + 4\mu\left[\frac{1}{|\,\bar{v}\,|} - \delta(\bar{v})\ln(\bar{y}^2 + \bar{z}^2)\right]d\bar{v}^2. \tag{7.6.11}$$

Using the coordinate transformation $du = d\bar{u} - 4\mu\,d\bar{v}/|\,\bar{v}\,|$, $v = \bar{v}$, the above expression is converted to the form

$$ds^2 = -\,du\,dv + d\bar{y}^2 + d\bar{z}^2 - 4\mu\,\delta(v)\ln(\bar{y}^2 + \bar{z}^2)\,dv^2. \tag{7.6.12}$$

We see from this form of the metric that the corresponding gravitational field is a special case of an axially symmetrical plane gravitational wave concentrated on the surface $v = 0$ separating two flat half-spaces, $v > 0$ and $v < 0$. The limiting procedure carried out above changes the algebraic type of the Weyl tensor; metric (7.6.12) is of type N, instead of the type D characterizing the original metric (7.6.7).[15] The corresponding curvature tensor vanishes everywhere except the surface $v = 0$ where its nonzero components have singularities of type $\delta(v)$.[16]

If there are two ultra-relativistic black holes moving toward each other on parallel courses, their gravitational field being concentrated in two plane waves described by metrics (7.6.12), the interaction between them disturbs these gravitational waves which are scattered in the transmission through each other. D'Eath (1978) was able to show that if the impact parameter is comparable with the quantity $M\gamma^2$, where M is the characteristic mass of black holes, and γ is the typical Lorentz factor in the center-of-mass reference frame, narrow beams of gravitational radiation of solid angles $\sim \gamma^2$ and characteristic power ~ 1 (in units of c^5/G) appear along the directions of motion of the black holes. The radiation is caused by the rapidly changing acceleration of a black hole at the moment of passage at the shortest distance from its counterpart.

If the impact parameter is comparable with the quantity $\mu = M\gamma$, gravitational radiation along the direction of motion remains weak up to angles $\theta \leq \gamma^{-1}$. At larger

[15] This property of the Schwarzschild metric was discovered by Pirani (1959). Later Penrose (1976) proved that a general-type spacetime has this property. Namely, the geometry of the surrounding spacetime tends to that of a flat gravitational wave for an observer whose velocity tends to the speed of light, and who uses the time parameter $\gamma\tau$, where $\gamma = (1 - v^2)^{-1/2}$, and τ is the proper time. In the limit $\gamma \to \infty$, the world line of this observer is a null geodesic, and $\gamma\tau$ is the affine parameter along this geodesic.

[16] A similar procedure can be used to obtain the ultra-relativistic limit for the Kerr metric [Lousto and Sánchez (1992), Hayashi and Samura (1994), Balasin and Nachbagauer (1995, 1996)]. It should be stressed that if one studies the limit $\alpha \to 0$ for a metric depending on a parameter α, one can always make a coordinate transformation depending on α before taking the limit. As a result, there may exist different limiting metrics. [For a general discussion of this point, see Geroch (1969).] Such an ambiguity is also present in (7.6.11). Since $|\,v\,|^{-1}$ is a singular function, the limit requires more accurate definition. The singular coordinate transformation used above serves to provide such a definition. A general discussion of this point as well as the discussion of the structure of stress-energy tensor distributions generating the ultrarelativistic Kerr and Schwarzschild geometry can be found in [Balasin and Nachbagauer (1995)].

angles (in the interval $\gamma^{-1} \ll \theta \ll 1$), the total energy of gravitational radiation per unit solid angle is about $dE/d\Omega \approx 0.248\, M\gamma/2\pi$. D'Eath (1978) assumed that the radiation of two black holes of equal mass in head-on collision is fairly isotropic and concluded that the efficiency of the transformation of the energy of black holes, 2μ, into the radiation energy, ΔE, is close to 25%: $\Delta E/2\mu \approx 0.248$ [Smarr *et al.* (1976); see also Smarr (1977, 1979), Bovyn (1983)].

A general restriction on the maximal efficiency of conversion of the energy of black holes in head-on collision into gravitational radiation energy ΔE can be found on the basis of Hawking's theorem [Penrose (1974), Smarr *et al.* (1976)]. If the two black holes have identical mass M and their center-of-mass velocity towards each other at infinity equals v, the maximal efficiency $\varepsilon = \Delta E/2\mu$ is

$$\varepsilon \leq 1 - \sqrt{\frac{1-v^2}{2}}\,. \tag{7.6.13}$$

The estimates given above show that the actual efficiency for ultra-relativistic black holes is about 25 % of the maximal value $\varepsilon(v=1) = 1$. If non-relativistic black holes collide, the efficiency is reduced by nearly two orders of magnitude (see Section 7.7).

7.6.3 Momentarily static black holes configurations

An interesting possibility for studying the interaction between black holes is found in analyzing so-called *momentarily static configurations* that describe a system of interacting black holes at the moment when all of them are at rest [Misner and Wheeler (1957), Misner (1960, 1963), Lindquist (1963), Brill and Lindquist (1963), Gibbons (1972), Bowen and York (1980), Kulkarni *et al.* (1983), Bowen *et al.* (1984), Kulkarni (1984)]. This possibility arises because the spacetime metric $g_{\mu\nu}$ at the moment of temporal symmetry can be so chosen that the conditions $g_{0i} = 0$, $\partial_0 g_{00} = \partial_0 g_{ij} = 0$ are satisfied. The extrinsic curvature on this slice $K_{ij} = 0$ and the three-dimensional geometry of space can be found by solving the constraint equations (A.78)-(A.79) which in the vacuum spacetime imply that

$$^{(3)}R = 0\,. \tag{7.6.14}$$

Here $^{(3)}R$ is the scalar of the three-dimensional curvature of metric h_{ij}. In the presence of an electromagnetic field, only the components $E_i = F_{i0}$ are nonzero at the moment of temporal symmetry. The initial conditions at this moment of time for the Einstein-Maxwell system of equations are

$$^{(3)}R = 2E_i\,E^i, \qquad E^i{}_{;i} = 0\,. \tag{7.6.15}$$

Equations (7.6.15) have the following solution describing a system of N interacting charged black holes [Lindquist (1963), Brill and Lindquist (1963)]:

$$dl^2 \equiv h_{ij}\,dx^i\,dx^j = (\,\chi\psi)^2(dx^2 + dy^2 + dz^2)\,,$$

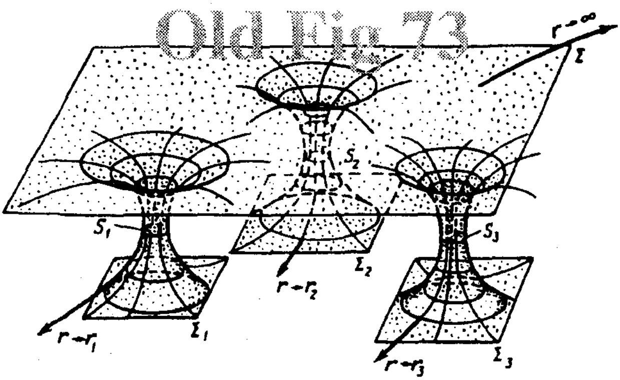


Figure 7.3: Momentarily static configuration of three interacting black holes (two-dimensional spatial section of spacetime).

$$E_i = [\ln(\chi/\psi)]_{,i}\,, \tag{7.6.16}$$

where

$$\chi = 1 + \sum_{i=1}^{N} \frac{\alpha_i}{|\mathbf{r} - \mathbf{r}_i|}, \quad \psi = 1 + \sum_{i=1}^{N} \frac{\beta_i}{|\mathbf{r} - \mathbf{r}_i|},$$

$$\alpha_i > 0\,, \quad \beta_i > 0\,, \quad \mathbf{r} = (x, y, z)\,. \tag{7.6.17}$$

(The two-dimensional section of the corresponding metric is schematically shown in Figure 7.3 for the case of $N = 3$.)

The mass M_i and charge Q_i of the ith black hole, determined by the asymptotic behavior of the solution at infinity on the sheet Σ_i (as $\mathbf{r} \to \mathbf{r}_i$), are

$$M_i = \alpha_i + \beta_i + \sum_{\substack{i,j=1\\ i\neq j}}^{N} \frac{\alpha_i \beta_j + \alpha_j \beta_i}{|\mathbf{r}_i - \mathbf{r}_j|}\,, \tag{7.6.18}$$

$$Q_i = \beta_i - \alpha_i + \sum_{\substack{i,j=1\\ i\neq j}}^{N} \frac{\beta_i \alpha_j - \alpha_i \beta_j}{|\mathbf{r}_i - \mathbf{r}_j|}\,. \tag{7.6.19}$$

The mass M and charge Q of a system of interacting black holes, determined by the asymptotic behavior of the solution at infinity on the sheet Σ (as $r \to \infty$), are

$$M = \sum_{i=1}^{N} (\alpha_i + \beta_j)\,, \quad Q = \sum_{i=1}^{N} (\beta_j - \alpha_i)\,. \tag{7.6.20}$$

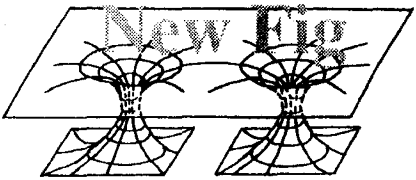


Figure 7.4: Topology of the Brill-Lindquist space.

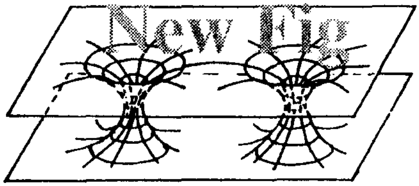


Figure 7.5: Topology of the Misner space.

It is easily shown that

$$Q = \sum_{i=1}^{N} Q_i \,, \quad M_{\text{int}} \equiv M - \sum_{i=1}^{N} M_i = -\sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{\alpha_i \beta_j + \alpha_j \beta_i}{|\mathbf{r}_i - \mathbf{r}_j|} < 0 \,. \tag{7.6.21}$$

These relations show that the charges of the black holes sum up additively while the gravitational mass defect due to the interaction between black holes reduces the total mass of the system below the sum of the masses of the component holes. For a discussion of the properties of the two-dimensional surfaces corresponding to the position of the apparent horizons at the moment of time symmetry for a system of interacting black holes, see Brill and Lindquist (1963), Gibbons (1972), Bishop (1984).

If $\beta_i = \alpha_i$, solution (7.6.16) describes a system of uncharged black holes [see Misner and Wheeler (1957), Misner (1960, 1963), Gibbons (1972)]. A special solution of the initial value problem describing two non-rotating uncharged black holes of equal masses momentarily at rest plays an important role in the numerical simulations of black hole collisions (see next section) [Brill and Lindquist (1963)]. The corresponding *Brill-Lindquist metric* written in the cylindrical coordinates (ρ, ϕ, z) $(\rho = \sqrt{x^2 + y^2}$, $\tan\phi = y/x)$ is

$$dl^2 = \psi^4 \left(d\rho^2 + \rho^2 d\phi^2 + dz^2 \right) , \tag{7.6.22}$$

$$\psi = \psi_{\text{BL}} \equiv 1 + \frac{M}{4} \left[\frac{1}{\sqrt{\rho^2 + (z - z_0)^2}} + \frac{1}{\sqrt{\rho^2 + (z + z_0)^2}} \right] . \tag{7.6.23}$$

The parameter M is the total mass of the system measured at spatial infinity (ADM mass), while z_0 determines the initial separation of black holes. The Brill-Lindquist geometry represents two throats. Each of the throats connects the upper sheet to a separate lower sheet (see Figure 7.4).

Another frequently used solution of the initial value problem was obtained by Misner (1960, 1963). The *Misner metric* is also of the form (7.6.22) with

$$\psi = \psi_{\text{M}} \equiv 1 + \frac{m}{4} \sum_{n=1}^{\infty} \frac{1}{\sinh(n\mu)} \left[\frac{1}{\sqrt{\rho^2 + (z - z_n)^2}} + \frac{1}{\sqrt{\rho^2 + (z + z_n)^2}} \right], \tag{7.6.24}$$

where $z_n = (m/4)\coth(n\mu)$. This solution is topologically two Euclidean sheets with isometric geometries joined by two throats representing non-rotating equal-mass black holes resting on the z-axis (see Figure 7.5).

The two throats are defined by $(z_{\rm throat} \pm z_0)^2 + \rho^2_{\rm throat} = (m/4)^2 \sinh^{-2}\mu$, where $z = \pm z_0 = \pm(m/4)\coth\mu$ are the black hole centers. The parameters m and μ are related to the ADM mass M

$$M = m \sum_{n=1}^{\infty} \frac{1}{\sinh(n\mu)} . \tag{7.6.25}$$

The proper distance L along the spacelike geodesic connecting the throats is

$$L = \frac{m}{2}\left[1 + 2\mu \sum_{n=1}^{\infty} \frac{n}{\sinh(n\mu)}\right] . \tag{7.6.26}$$

For fixed m, increasing μ decreases the total mass of the black holes and increases the separation between them. If the throats are close enough together (for $\mu < 1.36$) [Smarr *et al.* (1976)], a common apparent horizon surrounds them both. In this case, the system represents one highly excited black hole.

Misner (1960, 1963) and Lindquist (1963) generalized solution (7.6.16) to the case of a system of black holes replaced by a set of "wormholes" (a two-dimensional section of such a space is schematically shown in Figure 7.6 in the case of a single "wormhole"). The metric of the space at the time-symmetry moment for the uncharged wormhole is [Misner (1960)]

$$dl^2 = \Phi^4 \left[d\chi^2 + d\theta^2 + \sin^2\theta\, d\phi^2\right] , \quad -\pi < \chi < \pi , \tag{7.6.27}$$

$$\Phi = \left(\frac{m}{4}\right)^{1/2} \sum_{n=-\infty}^{\infty} \left[\cosh(\chi + 2n\mu) - \cos\theta\right]^{-1/2} . \tag{7.6.28}$$

The total ADM mass M for the system is

$$M = m \sum_{n=1}^{\infty} [\sinh(n\mu)]^{-1} , \tag{7.6.29}$$

and the length L of the shortest closed geodesic through the wormhole is

$$L = \frac{m}{\pi} K(k')\, E(k) , \quad k' = (1 - k^2)^{1/2} , \tag{7.6.30}$$

where K and E are elliptic integrals of the first and second kind, respectively, and the modulus k is fixed by the condition

$$\mu = \pi K(k)/K(k') . \tag{7.6.31}$$

The charged version of this solution was found by Lindquist (1963).

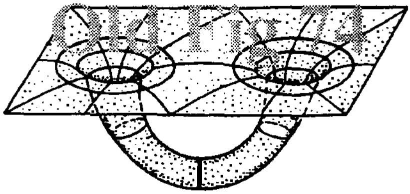


Figure 7.6: Time-symmetric initial data of "wormhole" type. As before, one of the coordinates is suppressed.

It should be emphasized again that, in general, a system of black holes described by solution (7.6.16) cannot be at rest at all times. The situation when $\alpha_i = 0$ and the condition $M_i = Q_i$ holds for all black holes is an exception [Hartle and Hawking (1972)].

The above-described method of constructing the spacetime geometry for a system of black holes at the moment of time symmetry can be extended to the case of rotating black holes [Bowen and York (1980), Bowen *et al.* (1984), Kulkarni (1984)]. On the possibility of equilibrium stationary axially symmetric configurations composed of rotating black holes, see Oohara and Sato (1981), Kihara and Tomimatsu (1982), Tomimatsu and Kihara (1982), Sato (1983), Tomimatsu (1983), Yamazaki (1983a,b), Bičák and Hoenselaers (1985), and the bibliography therein.

7.7 Black Hole Collisions

7.7.1 Head-on collision. Numerical Results

The collision of two black holes is a problem of great importance both for the theory and for astrophysical applications. Coalescence of black holes produces gravitational waves. These sources of gravitational radiation are considered among the most important ones likely to be observed by the Laser Interferometric Gravitational Wave Observatory (LIGO) [see e.g., Abramovici *et al.* (1992), Thorne (1997a,b)]. LIGO is a system of large Michelson interferometer detectors of gravitational radiation presently under construction in the United States. LIGO is a part of a global network that includes the French-Italian VIRGO detector (under construction near Pisa) and other detectors in various stages of planning. The information obtained by detecting gravitational waves from coalescing black hole binaries should allow one to reconstruct the astrophysical parameters of the system, and might provide the first direct information on the properties of spacetime near black holes. Moreover, details of the gravitational wave profiles can be used as a potential test of the Einstein equations in the domain of strong gravitational fields.

In order to extract the required information from the observations and to make comparison with the theory, it is necessary to perform accurate calculations of the gravitational waveform from coalescing black holes. The latter problem inevitably includes the study of the self-consistent dynamics of coalescing black holes.

The two-body problem in general relativity is unsolved. In order to illuminate its difficulties, it is instructive to compare it with the two-body problem in electrodynamics. The equations of motion of two massive point-like charges can be solved analytically only if we neglect their radiation and retardation effects in their interaction. Under these conditions dynamical properties of the electromagnetic field are not important, and one can determine the field at a given moment of time by solving an eliptic constraint equation. In the presence of the effects of retardation and radiation the dynamical nature of the electromagnetic field becomes important. This means that we can determine the motion of the particles only if we include the electromagnetic field as an important part of the system. For this reason (especially in the case of the relativistic motion of charges), even this electrodynamical problem becomes quite complicated and requires special approximations. The two-body problem for two colliding or coalescing black holes is essentially more complicated. The main reasons are the nonlinearity of the Einstein equations and their gauge invariance. An additional complication is that black holes are extended nonlocal objects. The approximation methods described in the Section 7.6 for interacting black holes, which were based on the existence of a small parameter, do not work when the colliding black holes come close to one another. Unfortunately, no general analytic techniques are available for obtaining strong-field, highly dynamical solutions in the absence of spatial symmetries. In such situations, the only recourse is numerical solution of the Einstein equations.

The numerical simulations performed until now are mainly restricted to the case of head-on collisions of two black holes. Because of the additional axial symmetry this problem is essentially two-dimensional. The head-on black hole collision was the earliest testbed for numerical relativity, the solution of the Einstein differential equations with numerical codes. The first numerical attack on this problem was made by Hahn and Lindquist (1964). Smarr and Eppley [Smarr (1979)] obtained the first numerical solution for the head-on collision of two black holes of equal mass. The starting point of their calculations was Misner's initial data (7.6.22) and (7.6.24) representing two non-rotating equal mass black holes colliding head on along the z axis. This choice of the initial data allows one to avoid problems connected with unimportant details of the dynamics of the collapsing matter forming the black holes. Smarr and Eppley showed that the black holes do coalesce and radiate gravitational waves with energy of order $10^{-3}M$, where M is the total mass of the system. Moreover, they demonstrated that the gravitational wave has a form similar to the ringing radiation (see Section 4.4). However, these numerical calculations proved to be very difficult due to inherent coordinate singularities and numerical instabilities.

Recently, these calculations were extended and refined by the NCSA group [Anninos, Hobill *et al.* (1993, 1995), Anninos, Bernstein *et al.* (1994, 1995), Anninos, Camarda *et al.* (1995), Anninos, Daues *et al.* (1995)]. In the works of these authors

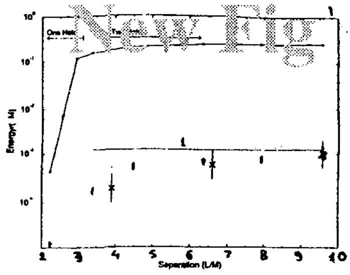


Figure 7.7: The total gravitation wave energy output for head-on collision of two non-rotating black holes. Each of black holes has mass M.

many of the numerical problems associated with colliding two black holes were overcome. In particular, they used a hybrid set of coordinates to resolve the axis and saddle point problems encounted by Smarr and Eppley. They developed a numerical code that can evolve black holes of equal mass with initial separation distances between $4M$ and $20M$, where M is a mass of the black hole. More modern analytic techniques were applied to the numerical data to extract waveforms and to compute the total energies emitted. The numerical simulations demonstrate that waveforms show similar behavior for all cases studied: The normal modes of the final black hole are excited and account for most of the emitted signal. The total energy radiated is of order $2 \times 10^{-3}M$. This radiated energy is far below the limit estimated by the area theorem. The results of the calculations for the radiated total energy and comparison with different approximations are presented in Figure 7.7 [Anninos, Hobill *et al.* (1995)]. In this figure solid bars show the result of the calculations with their errors. The crosses show earlier results by Smarr and Eppley with their approximate error bars.[17] The dashed line is the semianalytic estimate by Anninos, Hobill *et al.* (1995). The connected circles show an upper limit based on the area theorem.

Recall (see Section 4.6) that the amount of radiated energy in the case of radial fall of a *test* particle of mass m at the parabolic velocity into a black hole of mass M is given by the formula

$$\Delta E = 0.01\, m^2/M\,. \tag{7.7.1}$$

Note that the same formula reproduces the numerical result given above quite well if the reduced mass of two black holes, $M_H/2$, is substituted for m, and if we take into account that the total mass of the system is $M = 2M_H$. The solid line 1 is the result of this naive application of the point particle result after the reduced mass corrections are taken into account.

We already mentioned in the previous section that for $\mu < 1.36$ the apparent horizon surrounds both of the throats so that the initial data represent a single

[17] On another semi-analytic estimate of the total gravitational radiation in the head-on black hole collision see [Moreschi and Dain (1996)].

highly excited black hole. In fact, the event horizon is located outside the apparent horizon. It was shown that the initial data contain one black hole for $\mu < 1.8$.

The Misner initial data are special only in their mathematical convenience. For this reason, it is important to study how the obtained results depend on the particular choice of the initial conditions. This problem was analyzed by Abrahams and Price (1996b) who compared the results of the calculations for the Misner and Brill-Lindquist initial data. They concluded that the single difference between radiation from these two initial data lies in the relative strengths of multipoles. The head-on collision of non-rotating black holes of equal mass formed from the collapse of spherical balls of non-colliding particles was investigated by Shapiro and Teukolsky (1992).

In spite of the fact that supercomputers have advanced in speed by over 50,000 times since the first numerical calculations by Hahn and Lindquist (1964), the general problem of study of coalescence of two black holes in the absence of axial symmetry is still open. Full three-dimensional simulations of two coalescing black holes, which requires the power of the teraflop supercomputers, is the subject of a National Science Foundation High-Performance Computing and Communications Grand challenge project, termed the Binary Black Hole Alliance. Besides the study of head-on collisions of black holes which is a natural first step, a lot of work has already been done to overcome various difficulties of the numerical calculations in three dimensions.

One of these problems is connected with the singularities inside a black hole. The presence of extreme curvature near singularities does not allow one to use numerical schemes. To avoid this difficulty it was proposed to use so-called *singularity-avoiding slicing conditions* so that "equal time" slices wrap up around the singular regions [Smarr and York (1978); Eardley and Smarr (1979); Bardeen and Piran (1983)]. However, these conditions by themselves do not solve the problem, but simply serve to delay the breakdown of the numerical evolution. The slowdown of the evolution of the slices near the singularity results in their abrupt change near the horizon. Numerical simulations finally crash due to this pathological behavior of the slices.

In principle, the solution of this problem is evident. The black hole interior does not influence the exterior regions. That is why it is sufficient to use only those partial slices which lie outside the event horizon. The problem is that the event horizon itself can be determined only after the solution is obtained. For this reason, it was proposed to use an apparent horizon for this purpose. The main problem is to formulate some form of outgoing (i.e., into the black hole) boundary conditions for general relativity, where the waves are nonlinear, dispersive, and possess tails. It was demonstrated only recently that such horizon boundary conditions can be realized in numerical calculations [Seidel and Suen (1992); Anninos, Camarda *et al.* (1995)].

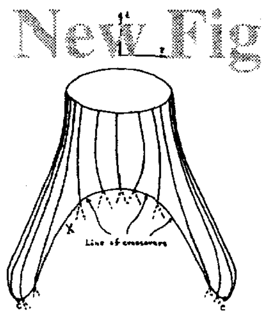


Figure 7.8: The event horizon formed by the merger of two non-rotating black holes.

7.7.2 Event horizon structure for colliding black holes

A quite unexpected application of numerical relativity is the study of the structure of the event horizon for colliding black holes [Matzner *et al.* (1995)]. This study demonstrated that earlier naive pictures of the horizon for colliding black holes, similar to Figure 5.9, give only a rough description and can be considerably refined.

The event horizon in a spacetime of a black hole can be found by numerically propagating null rays back in time [Hughes *et al.* (1994)]. A useful method of event horizon reconstruction in the numerical simulations was proposed by Anninos, Bernstein *et al.* (1995). If at late times the solution for a black hole is almost static, one can use the apparent horizon to approximate it and to define the approximate position of the event horizon in the earlier moments of time by integrating the null surface equation backward in time. A null surface defined by $f(t, x^i) = 0$ obeys the equation

$$g^{\mu\nu} f_{,\mu} f_{,\nu} = 0 . \tag{7.7.2}$$

This is a first-order partial differential equation. It was shown that numerically integrating an entire surface has several advantages with respect to the method of backward tracing of null geodesics.

The ability to determine accurately the apparent and event horizons for evolving black holes opens up the possibility of using the numerical methods as a tool to study the structure of the horizons. By combining numerical results with analytical methods, it is possible to arrive at the following qualitative picture (shown in Figure 7.8) of the event horizon structure for the head-on collision of two black holes [Matzner *et al.* (1995)].

This figure shows the formation of two black holes of equal mass. The horizon begins where its generating rays meet, either at caustics or at crossover points. In the presence of rotational symmetry the only stable (i.e., generic with respect to small perturbations preserving symmetry) caustics consist of cusps and folds.[18] The two points C are cusp-like vertices connected by the spacelike line X of crossover points which form a "seam" on the "trousers". The "seam" joins smoothly to the two null rays emanating from the cusps. A null ray which, on being propagated backward in time, crosses the crossover line X meets there another null ray. Earlier in time both of the intersecting rays are in the region visible from infinity. In the exactly axisymmetric case the crossover line is in fact a caustic because it is a focal line of a circle in the (x, y)-plane. However, under a perturbation away from axial symmetry this line is broadened into a two-dimensional spacelike crossover surface.

[18]The classification and study of caustics is one of the important applications of catastrophe theory [see e.g., Arnold (1990)]. Many details of the spacetime geometry of the elementary stable caustics are discussed in [Friedrich and Stewart (1983)].

Chapter 8

Black Hole Electrodynamics

8.1 Introduction

Black hole electrodynamics is defined as the theory of electrodynamic processes that can occur *outside* the event horizon, in the external space accessible to observation by distant observers. At first glance, black hole electrodynamics is quite trivial. Indeed, the electromagnetic field of a stationary black hole (of a given mass M) is determined unambiguously by its electric charge Q and rotation parameter a. If the charged black hole does not rotate, its electromagnetic field reduces to the radial electric field of the charge Q and is static. Any multipoles higher than the monopole are absent.

If the black hole rotates, the electromagnetic field has the form (3.6.3). The field is stationary, but now the rotation of the black hole, first, induces a magnetic field and, second, distorts the geometry of space and generates the higher-order electric (and magnetic) moments of the fields. However, these higher-order moments are determined unambiguously by the quantities M, a, and Q, not being in any way independent, as one would find in the case of ordinary bodies.

In astrophysics, the electric charge of a black hole (as of any other celestial body) cannot be high (see Section 3.6). The magnetic field must also be very weak: The dipole magnetic moment of a black hole is $\mu^* = Qa$. There can be no other stationary electromagnetic field inherent in a black hole. In this sense, the electrodynamics of, say, radio pulsars is definitely much richer than that of the intrinsic fields of black holes. A radio pulsar is a rapidly rotating *neutron star* of approximately solar mass possessing a gigantic "frozen-in" magnetic field of about 10^{12} Gauss. Rotation induces high electric fields that "rip" charges off the surface of the star and accelerate them to high energies. Energetic particles moving along curved magnetic field lines radiate hard curvature photons, which generate secondary electron-positron pairs in a strong magnetic field [Goldreich and Julian (1969), Sturrock (1971), Ruderman and Sutherland (1975), Arons (1983), Kardashev *et al.* (1984), Gurevich and Istomin (1985)]. As a result, a complicated magnetosphere of the pulsar is created, and a host of other related phenomena are generated [Michel (1991), Beskin, Gurevich, and

Istomin (1993)].

Black holes have neither strong magnetic nor electric fields (see Ginzburg (1964), Ginzburg and Ozernoi (1964), and Chapter 3), nor a surface from which charges could be ejected. Complex electrodynamic processes are thus impossible. However, if a black hole is placed in an *external* electromagnetic field, and if charges can be produced in its surroundings, the situation changes dramatically, and complex electrodynamics does appear. It is this aspect that we mean when black hole electrodynamics is discussed.

The case important for astrophysical applications is that of external magnetic (not electric) fields and rarefied plasma in which a black hole is embedded. In this system a regular magnetic field arises, for example, as it gets cleansed of magnetic loops which fall into a black hole [Thorne, Price, and Macdonald (1986)]. A regular magnetic field can also be generated in an accretion disk by the dynamo-action [Pudritz (1981a,b), Camenzind (1990)].

8.2 Maxwell's Equations

8.2.1 Black hole electrodynamics and membrane paradigm

We will consider electromagnetic fields against the background of a given metric; that is, we assume that these fields are not sufficiently strong to affect the metric. As a rule, this condition is met in astrophysics.[1]

Electrodynamic equations in four-dimensional form, using the tensor $F_{\alpha\beta}$, can hardly suggest anything to our intuition. They are difficult enough to apply to even modestly difficult concrete problems of physics. Thorne and Macdonald (1982) [see also Macdonald and Thorne (1982), Thorne, Price, and Macdonald (1986)] have rewritten these equations using the (3+1)-split for the exterior spacetime of a rotating black hole (see Section 3.2). Their formalism operates with familiar concepts: field strength, charge density, electric current density, and so forth, "absolute" space and unified "time". The equations of electrodynamics are written in a form similar to their form in a flat spacetime in a Lorentz reference frame. As a result, one can use not only well-developed methods of electrodynamics but also "rely" on conventional notions and the intuition stemming from experience of solving such problems. Furthermore, the work of Thorne and Macdonald cited above makes use of the so-called "membrane" interpretation of a black hole (see Section 3.7). The main point of this approach is that for a distant observer (who does not fall into the black hole), the boundary of a black hole can often be thought of as a thin physical membrane characterized by special electromagnetic, thermodynamic, and mechanical properties. This interpretation will be described in more detail in Section 8.3 [see also Thorne (1986),

[1]The slow change in the parameters of the black hole as a result of electromagnetic processes is analyzed later (Section 8.4).

Price and Thorne (1986), Thorne, Price, and Macdonald (1986)]. Of course, in reality there is no membrane. This concept needs very careful handling, and one must constantly bear in mind that it is only conditional and convenient for solving certain problems. The methods outlined above make it relatively simple to apply black hole electrodynamics to astrophysics even when the astrophysicist is not a specialist in relativity. For a review of this field, see Thorne *et al.* (1986).

This section presents the results of the papers [Thorne and Macdonald (1982), Macdonald and Thorne (1982)] for axially symmetric black holes cited above and subsequent development of these ideas. All physical quantities introduced below are three-dimensional vectors (or tensors) which are characterized by a position occupied in the "absolute" three-dimensional space (outside the black hole) and by the absolute unified "time" t (see Section 3.2). These are quantities that a locally no-rotating observer measures by conventional instruments.[2]

8.2.2 Maxwell's equations in (3+1)-form

We introduce the following notation for electrodynamic physical quantities measured by locally non-rotating observers: $\mathbf{E}$ is the electric field strength, $\mathbf{B}$ is the magnetic field strength, ρ_e is the electric charge density, and $\mathbf{j}$ is the electric current density. Denote by ϖ the norm of a Killing vector $\xi^{\mu}_{(\phi)}$ reflecting the axial symmetry of spacetime (3.2.1)

$$\varpi \equiv \sqrt{g_{\phi\phi}} = \sqrt{\frac{A}{\Sigma}} \sin\theta \,. \tag{8.2.1}$$

Here A and Σ are defined by (3.2.2). We denote by $\mathbf{e}_{\hat{\phi}}$ a three-dimensional unit vector in the direction of the Killing vector $\xi^{\mu}_{(\phi)}$. By using these notations, Maxwell's equations can be written in the following form:[3]

$$\nabla \mathbf{E} = 4\pi \rho_e \,, \tag{8.2.2}$$

$$\nabla \mathbf{B} = 0 \,, \tag{8.2.3}$$

$$\nabla \times (\alpha \mathbf{B}) = \frac{4\pi\alpha \mathbf{j}}{c} + \frac{1}{c}[\dot{\mathbf{E}} + \mathcal{L}_{\beta}\mathbf{E}] \,, \tag{8.2.4}$$

$$\nabla \times (\alpha \mathbf{E}) = -\frac{1}{c}[\dot{\mathbf{B}} + \mathcal{L}_{\beta}\mathbf{B}] \,. \tag{8.2.5}$$

---

[2]We recall that locally non-rotating observers are observers which possess zero angular momentum (see Sections 3.3 and 3.7). They are also known as ZAMO's [Thorne, Price, and Macdonald (1986)].

[3]In view of the applications where the electrodynamic formulas of this chapter are used in astrophysics, we write them in a conventional system of units, retaining c.

Here

$$\boldsymbol{\beta} = \omega \varpi \mathbf{e}_{\hat{\phi}}, \tag{8.2.6}$$

$\alpha = (\Sigma\Delta/A)^{1/2}$ is the lapse function given by (3.3.12), and ω is the angular velocity of rotation (in time t) of locally non-rotating observers [see (3.3.13)]. The notation $\mathcal{L}_{\beta}\mathbf{E}$ is used for the Lie derivative of a vector $\mathbf{E}$ along $\boldsymbol{\beta}$

$$\mathcal{L}_{\beta}\mathbf{E} \equiv (\boldsymbol{\beta}\,\boldsymbol{\nabla})\,\mathbf{E} - (\mathbf{E}\,\boldsymbol{\nabla})\,\boldsymbol{\beta}. \tag{8.2.7}$$

This Lie derivative describes how the vector $\mathbf{E}$ varies with respect to the field $\boldsymbol{\beta}$ (see Appendix A.7). $\mathcal{L}_{\beta}\mathbf{E}$ vanishes when the origin and end of the vector $\mathbf{E}$ are "glued" in a displacement by $\boldsymbol{\beta}\,d\phi$. A dot denotes differentiation with respect to t and $\boldsymbol{\nabla}$ is the three-dimensional (covariant) gradient operator in the curved "absolute" space.

Equations (8.2.2)–(8.2.3) have a familiar form, while that of equations (8.2.4)–(8.2.5) is slightly unusual. The following differences are evident. A function α has appeared because the physical time flows differently at different points of space while the equations are written in terms of the global "time" t (recall that the acceleration of free fall, $\mathbf{a}$, is related to α in the reference frame of locally non-rotating observers by the formula $\mathbf{a} = -c^2\,\boldsymbol{\nabla}\ln\alpha$). Furthermore, the expressions in brackets in (8.2.4) and (8.2.5) are "Lie-type" derivatives (with respect to time) for the set of locally non-rotating observers who move in the absolute space and for whom $d\mathbf{x}/dt = \boldsymbol{\beta}$. These expressions thus correspond to total derivatives with respect to the times of $\mathbf{E}$ and $\mathbf{B}$, respectively, with the motion of locally non-rotating observers taken into account.

Electrodynamic equations become especially lucid and convenient for the analysis of specific problems if written in integral form [see, e.g., Pikel'ner (1961)]. Here we give only one of such integral expression (we will need it later) for the external space of the black hole; namely, Faraday's law:

$$\int_{\partial A^*(t)} \alpha \left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right) d\mathbf{l} = -\frac{1}{c}\frac{d}{dt}\int_{A^*(t)} \mathbf{B}\,d\boldsymbol{\Sigma}. \tag{8.2.8}$$

Here $d\boldsymbol{\Sigma}$ is the vector of a surface element, with the vector length equal to the surface area of the element; $A^*(t)$ is a two-dimensional surface that does not intersect the horizon and is bounded by the curve $\partial A^*(t)$; $d\mathbf{l}$ is an element of $\partial A^*(t)$, and $\mathbf{v}$ is the physical velocity of $A^*(t)$ or $\partial A^*(t)$ relative to local non-rotating observers.

8.3 Stationary Axisymmetric Electrodynamics

8.3.1 Invariant variables

A rotating black hole and the space outside are stationary and axially symmetric. The motion of matter around a black hole can very often be regarded, with high accuracy, as stationary and axisymmetric as well. It is then natural to assume that the electromagnetic field too has these properties.

In this section we assume that these conditions are satisfied.[4] Then the derivatives with respect to time t vanish, and the Lie derivatives of vectors $\mathbf{E}$ and $\mathbf{B}$ take the form

$$\mathcal{L}_\beta \mathbf{E} = -(\mathbf{E}\,\nabla\omega)\,\varpi\,\mathbf{e}_{\hat\phi}\,, \qquad \mathcal{L}_\beta \mathbf{B} = -(\mathbf{B}\,\nabla\omega)\,\varpi\,\mathbf{e}_{\hat\phi}\,. \tag{8.3.1}$$

It is found that under stationarity and axial symmetry, the directly measured values of $\mathbf{E}$, $\mathbf{B}$, ρ_e, and $\mathbf{j}$ are expressible in terms of three scalar functions that can be chosen in the following manner.

Let ∂A^* be a closed coordinate line for constant r and θ in the absolute three-space, and A^* be a two-dimensional surface, bounded by ∂A^*, that does not intersect the black hole. This surface necessarily intersects the axis of symmetry. We denote by $d\Sigma$ the surface element on A^* and choose its orientation so that at the axis of symmetry it is directed along z-axis (i.e., from the black hole in the northern hemisphere, and to the black hole in the southern one). Then the three scalar functions mentioned above are:

1. Total magnetic flux Ψ across A^*

$$\Psi \equiv \int_{A^*} \mathbf{B}\,d\mathbf{\Sigma} \tag{8.3.2}$$

2. Total current I inside the loop ∂A^* (taken with reversed sign)

$$I \equiv -\int_{A^*} \alpha\,\mathbf{j}\,d\mathbf{\Sigma} \tag{8.3.3}$$

3. Electric potential

$$\mathcal{U}_0 \equiv -\alpha\,\Phi - \mathbf{A}\,\boldsymbol{\beta} \tag{8.3.4}$$

In the latter relation Φ is the scalar potential, and $\mathbf{A}$ is the vector potential. The quantities I and Ψ depend on the choice of the position of the loop ∂A^* but are independent of the shape of A^* (we assume the black hole to have zero magnetic charge). In other words, Ψ, I, and $\mathcal{U}_0$ are functions of the (r,θ)-coordinates. We demonstrate now that the electric and magnetic fields can be expressed in terms of these new invariant field variables.

Before expressing $\mathbf{E}$ and $\mathbf{B}$ in terms of Ψ, I, and $\mathcal{U}_0$, we decompose the fields into poloidal (superscript P) and toroidal (superscript T) components that are perpendicular and parallel to the vector $\mathbf{e}_{\hat\phi}$, respectively:

$$\mathbf{E}^T \equiv (\mathbf{E}\,\mathbf{e}_{\hat\phi})\,\mathbf{e}_{\hat\phi}\,, \qquad \mathbf{E}^P = \mathbf{E} - \mathbf{E}^T, \tag{8.3.5}$$

[4]The case of a non-stationary field is discussed, e.g., by Macdonald and Suen (1985), Thorne (1986), Park and Vishniak (1989, 1990).

$$\mathbf{B}^T \equiv (\mathbf{B}\,\mathbf{e}_{\hat\phi})\,\mathbf{e}_{\hat\phi}\,, \qquad \mathbf{B}^P = \mathbf{B} - \mathbf{B}^T. \tag{8.3.6}$$

The current density $\mathbf{j}$ is also decomposed into poloidal ($\mathbf{j}^{\mathbf{P}}$) and toroidal ($\mathbf{j}^{\mathbf{T}}$) components.

Faraday's law (8.2.8) and the stationarity condition imply $\mathbf{E}^T = 0$. Equation (8.2.3) and the condition of axial symmetry of $\mathbf{B}$ imply

$$\nabla\,\mathbf{B}^T = 0\,, \qquad \nabla\,\mathbf{B}^P = 0\,; \tag{8.3.7}$$

that is, the poloidal and toroidal magnetic lines of force can be treated independently (as not terminating anywhere).

Now we can give expressions for all electromagnetic quantities in terms of Ψ, I, and $\mathcal{U}_0$:

$$\mathbf{E}^P = \alpha^{-1}\left(\nabla\mathcal{U}_0 + \frac{\omega}{2\pi c}\,\nabla\Psi\right), \tag{8.3.8}$$

$$\mathbf{E}^T = 0\,, \tag{8.3.9}$$

$$\mathbf{B}^P = \frac{\nabla\Psi \times \mathbf{e}_{\hat\phi}}{2\pi\varpi}\,, \tag{8.3.10}$$

$$\mathbf{B}^T = -\,\frac{2I}{\alpha\, c\,\varpi}\,\mathbf{e}_{\hat\phi}\,, \tag{8.3.11}$$

$$\mathbf{j}^P = \frac{\mathbf{e}_{\hat\phi} \times \nabla I}{2\pi\alpha\,\varpi}\,, \tag{8.3.12}$$

$$j^T \equiv \mathbf{j}^T\mathbf{e}_{\hat\phi} = \frac{\varpi}{4\pi\alpha}\left\{-\,c\,\nabla\left[\frac{\alpha\,\nabla\Psi}{2\pi\,\varpi^2}\right] + \frac{1}{\alpha}\,\nabla\omega\left(\nabla\mathcal{U}_0 + \frac{\omega\,\nabla\Psi}{2\pi c}\right)\right\}, \tag{8.3.13}$$

$$\rho_e = \frac{1}{4\pi}\,\nabla\left[\frac{1}{\alpha}\left(\nabla\mathcal{U}_o + \frac{\omega\,\nabla\Psi}{2\pi c}\right)\right]. \tag{8.3.14}$$

Note that the last three equations can be treated as differential equations for determining Ψ, I, $\mathcal{U}_0$ (and hence $\mathbf{E}$ and $\mathbf{B}$ as well) provided the field sources $\mathbf{j}^P$, $\mathbf{j}^T$, and ρ_e are assumed to be given as stationary and axisymmetric but otherwise arbitrary functions. Note also that in the stationary and axisymmetric case the current $\mathbf{j}$ must be prescribed in such a way that the condition $\nabla(\alpha\,\mathbf{j}) = 0$ (the charge conservation law) is satisfied; that is, $\alpha\,\mathbf{j}^P$ must be divergence-free.

8.3.2 Fields in the plasma surrounding a black hole

Consider now the physical conditions in the plasma surrounding the black hole. In the case of most importance for astrophysics, the conductivity of the plasma is so high that the electric field in the reference frame comoving with the plasma vanishes, and the magnetic lines of force are "frozen" into the plasma. In this case, the electric and magnetic fields in an arbitrary reference frame are perpendicular to each other (degenerate fields):

$$\mathbf{E}\,\mathbf{B} = 0. \tag{8.3.15}$$

Note that this condition is only some approximation and generally a small longitudinal electric field is present. To solve problems concerning the configuration of fields, currents, and charge distributions, it is only necessary that the inequality

$$|\,\mathbf{E}\,\mathbf{B}\,| \ll |\,\mathbf{E}^2 - \mathbf{B}^2\,| \tag{8.3.16}$$

is satisfied instead of (8.3.15). Small deviations from the exact equation (8.3.15) in the neighborhood of a black hole may prove to be important for a number of astrophysical processes [e.g., see Kardashev *et al.* (1983), Beskin, Istomin, and Pariev (1993), Horiuchi, Mestel, and Okamoto (1995)].

The equation (8.3.9) shows that the field $\mathbf{E}$ is purely poloidal. The condition (8.3.15) implies that $\mathbf{E}$ can be expressed as the vector product of $\mathbf{B}^P$ by a vector $-\mathbf{v}^F/c$ which is a function of only r and θ and is parallel to $\mathbf{e}_{\hat{\phi}}$:

$$\mathbf{E} \equiv \mathbf{E}^P = -\frac{\mathbf{v}^F}{c} \times \mathbf{B}^P. \tag{8.3.17}$$

Recall that $\mathbf{E}$ and $\mathbf{B}$ are fields measured by locally non-rotating observers. It is clear that for $|\,\mathbf{v}^F| < c$ the equation (8.3.17) implies that an observer moving at a velocity $\mathbf{v}^F$ with respect to locally non-rotating observers measures only a magnetic field. The electric field is zero for this observer owing to the Lorentz transformation. Therefore, $\mathbf{v}^F$ can be interpreted as the linear velocity of the points of a magnetic line of force with respect to locally non-rotating observers. Field $\mathbf{E}$ is completely induced by this motion. If the vector $\mathbf{v}^F$ is written in the form

$$\mathbf{v}^F = \left(\frac{\Omega^F - \omega}{\alpha}\right) \varpi\, \mathbf{e}_{\hat{\phi}}\,, \tag{8.3.18}$$

then Ω^F is angular velocity of the points on the lines of force of the poloidal magnetic field in the "absolute" space. The relations (8.3.17) and (8.3.18) remain also formally valid for $|\,\mathbf{v}^F| > c$.

The surface obtained when a magnetic line of force rotates around the symmetry axis is called the *magnetic surface*. The quantity Ψ is obviously constant on this surface. By substituting (8.3.17)–(8.3.18) into the Maxwell's equations (8.2.5), we find that Ω^F depends only on Ψ

$$\Omega^F = \Omega^F(\Psi)\,. \tag{8.3.19}$$

This means that each line of force revolves around the black hole as a whole at the angular velocity Ω^F which is constant in t in "absolute" space. Finally, by comparing relations (8.3.17), (8.3.18), and (8.3.8), we obtain

$$\frac{d\mathcal{U}_0}{d\Psi} = -\frac{\Omega^F}{2\pi c}. \tag{8.3.20}$$

Thus, $\mathcal{U}_0$ is also a function of Ψ and the equation (8.3.20) allows one to define it if $\Omega^F(\Psi)$ is known. In what follows we use $\Omega^F(\Psi)$ as the basic quantity instead of $\mathcal{U}_0$. Equations (8.3.13) for j^T and (8.3.14) for ρ_e get somewhat simplified

$$j^T = -\frac{\varpi}{8\pi^2\alpha}\left[c\,\nabla\left(\frac{\alpha\,\nabla\Psi}{\varpi^2}\right) + \frac{1}{\alpha c}\left(\Omega^F - \omega\right)\nabla\Psi\,\nabla\omega\right], \tag{8.3.21}$$

$$\rho_e = -\frac{1}{8\pi^2 c}\,\nabla\left[\left(\frac{\Omega^F - \omega}{\alpha}\right)\nabla\Psi\right]. \tag{8.3.22}$$

We first consider the simplest case in which inertial (and gravitational) forces acting on the plasma are small in comparison with electromagnetic forces. The configuration of fields and currents is then such that currents in the reference frame comoving with the plasma are parallel to magnetic lines of force, and no Lorentz force acts on moving charges. Such fields are known as *force-free fields.* In an arbitrary reference frame the condition of existence of a force-free field is

$$\rho_e\,\mathbf{E} + \frac{1}{c}\mathbf{j}\times\mathbf{B} = 0. \tag{8.3.23}$$

This equation implies that $I = I(\Psi)$. To prevent possible misunderstanding, we stress that usually more than one magnetic surface correspond to a given value of the flux Ψ. That is why the current I as well as the other integrals of motion which will be introduced later and which are constant on the magnetic surfaces generally are not single-valued functions of Ψ. However, locally relations like $I = I(\Psi)$ have a well-defined meaning.

Note that the condition (8.3.23) is definitely violated somewhere in the exterior space of the black hole. Indeed, the outer magnetic field in normal conditions survives in the space around the black hole because the ends of magnetic lines of force are "frozen" into the sufficiently dense massive plasma which exists somewhat further away and which has "transported" the magnetic field to the black hole. Condition (8.3.15) is met in this plasma but condition (8.3.23) is not. Gravitation (and inertia) holds this plasma in the vicinity of the black hole, together with the magnetic field "frozen" into it. The lines of force of the field pass from the dense plasma into the region of much more rarefied plasma where condition (8.3.23) is satisfied. Some of these lines of force go around the black hole, and some go through it. For example, this is the case in the widely discussed model of disk accretion on a black hole (Figure 8.1).

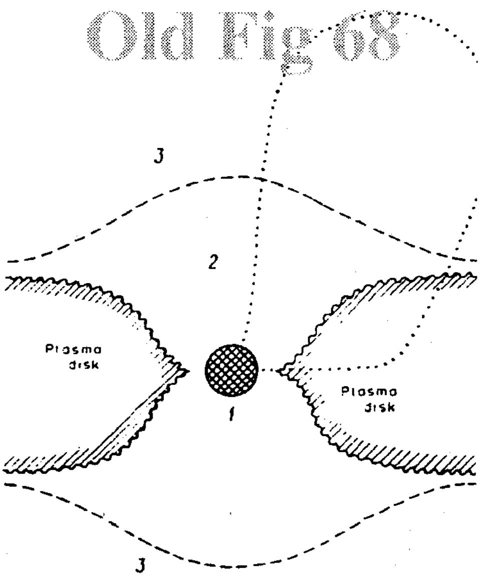


Figure 8.1: Schematic representation of disk accretion to a black hole: 1—rotating black hole, 2—region of force-free field (8.3.23), 3—"acceleration region" where conditions (8.3.15) and (8.3.23) are violated. The dashed curve is the boundary between the regions 2 and 3. The dotted curve is an example of electric current line.

If condition (8.3.23) were not violated somewhere and if the dense plasma did not counteract the outward pressure of the magnetic field, this pressure would drive an outward motion of the lines of force together with the rarefied plasma.

Condition (8.3.15) is likely to be violated far from the black hole (region 3 in Figure 8.1), where the magnetic field is sufficiently weak, while the inertial forces become relatively large (see the next section).

The most important and salient fact is that in the force-free approximation Ψ, I, and $\mathcal{U}_0$ are no longer arbitrary and independent; hence, ρ_e, j^T, and the divergenceless part of $\alpha\, j^P$ are likewise not arbitrary and independent as they would be if stationary and axial symmetry were the only constraints. All arbitrariness in choosing them has now been eliminated. The necessary (and sufficient) condition for a force-free field to exist is that Ψ satisfies an equation called the *stream equation*:

$$\nabla\left\{\frac{\alpha}{\varpi^2}\left[1-\frac{(\Omega^F-\omega)^2\varpi^2}{\alpha^2c^2}\right]\nabla\Psi\right\}$$

$$+\frac{(\Omega^F-\omega)}{\alpha c^2}\frac{d\Omega^F}{d\Psi}(\nabla\Psi)^2+\frac{16\,\pi^2}{\alpha c^2\varpi^2}\,I\,\frac{dI}{d\Psi}=0\,. \qquad (8.3.24)$$

The equation (8.3.24) follows directly from the force-free condition (8.3.23). Indeed, the force-free condition after substitution into it the expressions for the charge density ρ_e [(8.3.22)] and the electric current $\mathbf{j}$ [(8.3.12) and (8.3.21)] coincides (up to the common factor $\nabla\Psi$) with (8.3.24).

The stream equation (8.3.24) is a nonlinear elliptic equation for the potential Ψ in which the current $I(\Psi)$ and the angular velocity $\Omega^F(\Psi)$ are to be specified independently. If Ψ, $\Omega^F(\Psi)$, and $I(\Psi)$ are chosen so as to satisfy (8.3.24), then $\mathbf{E}$ for a region with a force-free field is found from (8.3.17) after substituting into it (8.3.18) and (8.3.10); $\mathbf{B}$ is found from (8.3.10) and (8.3.11); $\mathbf{j}^P$ from (8.3.12), and $\mathbf{j}^T$ and ρ_e from (8.3.21) and (8.3.22), respectively.

8.3.3 Magneto-hydrodynamic approximation

In the more general case when the mass of the particles cannot be neglected and the force-free condition (8.3.23) is not satisfied, a more general approach based on the one-fluid magneto-hydrodynamical (MHD) approximation can be applied. The applications of this approach to the strong gravitational field of black holes were considered by Phinney (1983a,b), Camenzind (1986a,b,c, 1987), Punsly and Coronity (1990a,b), Takahashi *et al.* (1990), Punsly (1991), Nitta, Takahashi, and Tomimatsu (1991), Hirotani *et al.* (1992), Beskin and Pariev[5] (1993). In this case, Maxwell's equations are to be supplemented by the continuity equation and the equation of state.

We assume that the number of particles is conserved and the continuity equation reads

$$\nabla(\alpha\, n\, \mathbf{u}) = 0\,, \tag{8.3.25}$$

where $\mathbf{u} = \gamma\,\mathbf{v}/c$; $\gamma = (1 - v^2/c^2)^{-1/2}$; $\mathbf{v}$ is the three-velocity of matter, and n is the matter density in the comoving reference frame. In the MHD-approximation the "frozen-in condition" (8.3.15) takes the form

$$\mathbf{E} + (\mathbf{v}/c) \times \mathbf{B} = 0\,. \tag{8.3.26}$$

This equation and (8.2.3) allow one to write the general solution of the continuity equation (8.3.25)

$$\mathbf{u} = \frac{\eta}{\alpha\, n}\,\mathbf{B} + \frac{\gamma}{\alpha}\,\frac{(\Omega^F - \omega)\,\varpi}{c}\,\mathbf{e}_{\hat\phi}\,. \tag{8.3.27}$$

The parameter η is the "integration constant" depending on Ψ, $\eta = \eta(\Psi)$. In other words, the value of η is constant on the magnetic surface. The relation (8.3.27) in particular shows that the poloidal component $\mathbf{u}^P$ of $\mathbf{u}$ is

$$\mathbf{u}^P = \frac{\eta}{\alpha\, n}\,\mathbf{B}^P. \tag{8.3.28}$$

[5]The consideration below in this section is based on the results presented in this last paper.

To specify the equation of state, it is convenient to use the pressure P and the entropy per particle s as thermodynamical variables. The corresponding thermodynamic potential is the specific entalpy μ

$$\mu = \mu(P, s) = (\rho_m + P)/n \,, \tag{8.3.29}$$

where ρ_m is the energy density of matter. The first law of thermodynamics implies [see e.g., Landau and Lifshitz (1959)]

$$d\mu = \frac{1}{n}\, dP + T\, ds \,. \tag{8.3.30}$$

The matter density n and the temperature T are defined as

$$n = n(P, s) = \left[\left(\frac{\partial \mu}{\partial P} \right)_s \right]^{-1} , \tag{8.3.31}$$

$$T = T(P, s) = + \left(\frac{\partial \mu}{\partial s} \right)_P . \tag{8.3.32}$$

The above relations allow one to express μ, T, and P as functions of n and s

$$\mu = \mu(n, s) \,, \qquad T = T(n, s) \,, \qquad P = P(n, s) \,. \tag{8.3.33}$$

We also make an additional assumption that the matter flow is isentropic

$$(\mathbf{u}\boldsymbol{\nabla})\, s = 0 \,. \tag{8.3.34}$$

In the axisymmetric case, equations (8.3.28) and (8.3.34) yield

$$s = s(\Psi) \,. \tag{8.3.35}$$

In the MHD-approximation, the current I fails to be constant on the magnetic surface. Nevertheless, the following two combinations $E(\Psi)$ and $L(\Psi)$ containing I are constant on the magnetic surface

$$E(\Psi) = \frac{\Omega^F I}{2\pi} + \mu\, c^2 \eta \left(\alpha\, \gamma + \frac{\omega\, \varpi}{c}\, u_{\hat{\phi}} \right) , \tag{8.3.36}$$

$$L(\Psi) = \frac{I}{2\pi} + \mu\, c\, \eta\, \varpi\, u_{\hat{\phi}} \,. \tag{8.3.37}$$

Here $u_{\hat{\phi}} = \mathbf{u}\, \mathbf{e}_{\hat{\phi}}$. These relations reflect the conservation of the total energy and ϕ-component of the momentum in the system. The first term on the right-hand side of (8.3.36) and (8.3.37) corresponds to the contribution of the electromagnetic field, while the others are connected with the contribution of particles. Thus, in the one-fluid MHD approximation there exist five quantities which are constant on the magnetic surface: $\Omega^F(\Psi)$, $E(\Psi)$, $L(\Psi)$, $\eta(\Psi)$, and $s(\Psi)$. We call these quantities integrals of motion.

The next step is to show that for known poloidal field $\mathbf{B}^P$ (i.e., for known potential Ψ) and for given integrals of motion one can reconstruct the toroidal magnetic field $\mathbf{B}^T$, the matter density n, and velocity $\mathbf{v}$. For this purpose, we use the conservation laws (8.3.36) and (8.3.37) which together with the ϕ-component of (8.3.27) allow one to express the electric current I, the Lorentz factor γ, and $u_{\hat{\phi}}$ as follows

$$\frac{I}{2\pi} = \frac{\alpha^2 L - (\Omega^F - \omega)(E - \omega L)\,(\varpi^2/c^2)}{\alpha^2 - (\Omega^F - \omega)^2\,(\varpi^2/c^2) - \mathcal{M}^2}\,, \tag{8.3.38}$$

$$\gamma = \frac{1}{\alpha\,\mu\,c^2\,\eta}\,\frac{\alpha^2(E - \Omega^F L) - \mathcal{M}^2(E - \omega L)}{\alpha^2 - (\Omega^F - \omega)^2\,(\varpi^2/c^2) - \mathcal{M}^2}\,, \tag{8.3.39}$$

$$u_{\hat{\phi}} = \frac{1}{\mu\,c\,\eta\,\varpi}\,\frac{(E - \Omega^F L)(\Omega^F - \omega)\,(\varpi^2/c^2) - L\mathcal{M}^2}{\alpha^2 - (\Omega^F - \omega)^2\,(\varpi^2/c^2) - \mathcal{M}^2}\,, \tag{8.3.40}$$

Here we denote

$$\mathcal{M}^2 \equiv \frac{4\,\pi\mu\,\eta^2}{n}\,. \tag{8.3.41}$$

We emphasize that $\mathcal{M}$ remains finite at the horizon provided the field and matter fluxes are regular there.[6] According to the equation (8.3.33), $\mu = \mu(n, s)$, and hence the relation (8.3.41) allows one to express n (and hence μ) as a function of η, s, and $\mathcal{M}$. This means that besides the integrals of motion only one additional quantity ($\mathcal{M}$) enters the expressions for I, γ, and $u_{\hat{\phi}}$.

We now show that if $\Psi = \Psi(r, \theta)$ is known, then one can define $\mathcal{M}$ in terms of integrals of motion. For this purpose, one can use the relation $\gamma^2 - \mathbf{u}^2 = 1$. After substitution of expressions (8.3.39) for γ, (8.3.40) for $u_{\hat{\phi}}$, and (8.3.28), (8.3.10) for u^P into this relation, one gets

$$\frac{K}{(\alpha\,\mu\,c\,\eta\,\varpi\mathcal{A})^2} - \frac{\mathcal{M}^4(\nabla\Psi)^2}{64\,\pi^4(\alpha\,\mu\,\eta\,\varpi)^2} = 1\,, \tag{8.3.42}$$

where

$$\mathcal{A} = \alpha^2 - (\Omega^F - \omega)^2\,(\varpi^2/c^2) - \mathcal{M}^2\,, \tag{8.3.43}$$

and

$$K = \alpha^2(\varpi^2/c^2)(E - \Omega^F L)^2(\mathcal{A} - \mathcal{M}^2) + \mathcal{M}^4[(\varpi^2/c^2)(E - \omega L)^2 - \alpha^2 L^2]. \tag{8.3.44}$$

[6] By introducing the Alfven velocity $u_A = B^P(4\,\pi\,n\,\mu)^{-1/2}$, we can also write $\mathcal{M}$ in the form $\mathcal{M} = \alpha\,u^P/u_A$. Thus, $\mathcal{M}$ is (up to the factor α) the Mach number calculated for the poloidal velocity u^P with respect to the Alfven velocity u_A.

The equation (8.3.42) defines $\mathcal{M}$ as the function of $(\nabla\Psi)^2$ and integrals of motion

$$\mathcal{M} = \mathcal{M}\left[(\nabla\Psi)^2, E, L, \Omega^F, \eta, s\,\right]. \tag{8.3.45}$$

We recall that

$$(\nabla\Psi)^2 = 4\pi^2\,\varpi^2\,(\mathbf{B}^P)^2. \tag{8.3.46}$$

Thus, for a known poloidal field $\mathbf{B}^P$ the relations (8.3.33), (8.3.38)–(8.3.41), and (8.3.45) define the toroidal magnetic field $\mathbf{B}^T$, the matter density n, and velocity $\mathbf{v}$.

The stream equation (8.3.24) defining the magnetic flux Ψ can be also generalized to the MHD-approximation. In this approximation, the force-free equation (8.3.23), defining the structure of the poloidal magnetic field, is to be replaced by the following equation, which is the poloidal component of the covariant conservation law

$$n\left[(\mathbf{u}\nabla)(\mu\,\mathbf{u})\right]^P + \frac{n}{\alpha}\,\mu\,\gamma\left[\frac{\varpi\,u_{\hat\phi}}{c}\,\nabla\omega + \gamma\,\nabla\alpha\right]$$
$$= -\,\nabla P + \rho_e\,\mathbf{E} + \left[(\mathbf{j}/c)\times\mathbf{B}\right]^P, \tag{8.3.47}$$

where P is the pressure; n is the density, and $\mathbf{u} = \gamma\,\mathbf{v}$. As a result, the stream equation takes the form

$$\begin{aligned}
&\mathcal{A}\left[\nabla\left(\frac{1}{\alpha\,\varpi^2}\,\nabla\Psi\right) + \frac{1}{D\alpha\,\varpi^2(\nabla\Psi)^2}\,\nabla^i\Psi\,\nabla^j\Psi\,\nabla_i\,\nabla_j\Psi\right] \\
&+\frac{1}{\alpha\,\varpi^2}\,\nabla'\mathcal{A}\,\nabla\Psi - \frac{\mathcal{A}}{2D\alpha\,\varpi^2(\nabla\Psi)^2}\,\nabla'F\,\nabla\Psi + \frac{(\Omega^F-\omega)}{\alpha\,c^2}\,\frac{d\Omega^F}{d\Psi}\,(\nabla\Psi)^2 \\
&+\frac{32\,\pi^4}{\alpha\,\varpi^2c^2\mathcal{M}^2}\,\frac{\partial}{\partial\Psi}\left(\frac{G}{\mathcal{A}}\right) - 16\,\pi^3\alpha\,\mu\,n\,\frac{1}{\eta}\,\frac{d\eta}{d\Psi} - 16\,\pi^3\alpha\,n\,T\,\frac{ds}{d\Psi} = 0\,.
\end{aligned} \tag{8.3.48}$$

Here

$$D = \frac{1}{\mathcal{M}^2}\left[\mathcal{A} + \frac{16\,\pi^2 I^2}{c^2(\nabla\Psi)^2} - \frac{\mathcal{A}}{(u^P)^2}\,\frac{a_s^2}{c^2-a_s^2}\right], \tag{8.3.49}$$

$$F = \frac{64\,\pi^4}{\mathcal{M}^4}\left(\frac{K}{c^2\mathcal{A}^2} - \alpha^2\,\varpi^2\,\mu^2\,\eta^2\right), \tag{8.3.50}$$

$$G = \alpha^2(E-\Omega^F L)^2\,(\varpi^2/c^2) + \alpha^2\mathcal{M}^2L^2 - \mathcal{M}^2(E-\omega L)^2\,(\varpi^2/c^2)\,, \tag{8.3.51}$$

$$(u^P)^2 = \frac{\mathcal{M}^4(\nabla\Psi)^2}{64\,\pi^4(\alpha\,\mu\,\eta\,\varpi)^2}\,, \tag{8.3.52}$$

and a_s is the sound velocity

$$a_s^2 = c^2\,\frac{1}{\mu}\left(\frac{\partial P}{\partial n}\right)_s\,. \tag{8.3.53}$$

In the equation (8.3.48) the quantities μ, n, T, and a_s are to be expressed in terms of $\mathcal{M}$ and integrals of motions. The gradient ∇' denotes the action of ∇ under the condition that $\mathcal{M}$ is fixed. The derivative $\partial/\partial\Psi$ in the expression $\partial/\partial\Psi(G/\mathcal{A})$ acts only on the integrals of motion $\Omega^F(\Psi)$, $E(\Psi)$, $L(\Psi)$, $\eta(\Psi)$, while other variables in a $G/\mathcal{A}$ are considered as constants. Finally, in the obtained relation the quantity $\mathcal{M}$ must be expressed by means of (8.3.45). (For more details, see [Nitta, Takahashi, and Tomimatsu (1991), Beskin and Pariev (1993)]).

The stream equation (8.3.48) is the desired equation for the poloidal field, which contains only the magnetic flux Ψ and the five integrals of motion depending on it. When the mass of particle can be neglected ($\mu = 0$), this equation reduces to the force-free stream equation (8.3.24).

8.3.4 Singular surfaces

MHD-flow, described by the equation (8.3.48), is characterized by the following singular surfaces, in which the matter velocity becomes equal to the velocity of the electromagnetic waves. These surfaces are:

1. *Alfven surfaces* defined by the condition $\mathcal{A} = 0$. On these surfaces the total velocity of matter u is equal to the Alfven velocity $u_A = B(4\pi\mu n)^{-1/2}$.
2. *Fast and slow magneto-sonic surfaces* defined by the condition $D = 0$. On these surfaces the matter velocity is equal to the velocity of the fast or slow MHD-wave.
3. *Slow magnetosonic cusp surfaces* defined by condition $D = -1$.

Unlike the force-free equation (8.3.24), the equation (8.3.48) is of mixed type: It is elliptic for $D > 0$, $D < -1$ and is hyperbolic for $-1 < D < 0$. In the case when the density of matter n remains finite at the horizon of a black hole, one has

$$D(r_+) = -1 \,. \tag{8.3.54}$$

If in the region where plasma is generated its velocity is less than the fast magneto-sonic velocity (and hence $D > 0$), then it is evident that between this region and the horizon there exists a fast magneto-sonic singular point. It should be stressed that the above-described singular surfaces are well known in solar physics [Weber and Davis (1967), Mestel (1968), Michel (1969), and Sakurai (1985)] and in pulsar physics [Okamoto (1978), Kennel, Fujimura, and Okamoto (1983), Bogovalov (1989), Takahashi (1991), Mestel and Shibata (1994)]. In the case of black holes, besides similar singular surfaces connected with ejection of matter, there is a new additional family of singular surfaces related to the accretion of matter in a strong gravitational field.

To conclude this section, we emphasize that the above consideration does not take into account a number of processes which might be important for real systems. Plasma interaction with the radiation field, the radiation of the plasma itself, and electron-positron pair creation in the magnetic field of a black hole are among them [Novikov and Thorne (1973), Shapiro and Teukolsky (1983), Begelman, Blandford, and Rees (1984), Blandford (1990)].

8.4 Membrane Interpretation and "Stretched" Horizon

8.4.1 Boundary conditions at the event horizon.

Black hole electrodynamics treats only processes outside the event horizon. Generally speaking, the solution of the electrodynamic equations requires that boundary conditions far from the black hole are to be supplemented with boundary conditions on its surface. Formally, this resembles the situation in pulsar electrodynamics where boundary conditions on the surface of the neutron star must also be specified.

Nevertheless, the two situations are in principal different. In contrast to a neutron star, a black hole has no material surface differing from the surrounding space. For black holes the role of the boundary conditions is played by the obvious physical requirement that the region of spacetime lying inside the black hole cannot affect the processes outside it, and all physical observables measured at the horizon in a freely falling reference frame are to be finite.

The event horizon is generated by null geodesics which are bicharacteristics of Maxwell's equations. For this reason, the regularity conditions allow an attractive mathematical formulation [Znajek (1978) and Damour (1978)]. Namely, it has been found that the corresponding boundary conditions can be written in a very clear form. This happens if one assumes that the surface of the black hole has a *fictitious* surface electric charge density σ^H which compensates for the flux of electric field across the surface and a *fictitious* surface electric current $\mathbf{i}^H$ which closes tangent components of the magnetic fields at the horizon. This interpretation is used in the *membrane formalism* [Thorne (1986), Thorne *et al.* (1986)]. A pedagogical introduction to this formalism was given in Section 3.7. Now we discuss its application to black hole electrodynamics.

We recall that in the (3+1)-formalism the horizon of a stationary black hole is a two-dimensional surface of infinite gravitational redshift, $\alpha = 0$ (see Chapter 3). The redshifted gravitational acceleration $\alpha\,\mathbf{a} \equiv -c^2\,\alpha\,\boldsymbol{\nabla}\ln\alpha$ remains finite at the horizon

$$(\alpha\,\mathbf{a})_H = -\kappa\,\mathbf{n}\,, \tag{8.4.1}$$

where $\mathbf{n}$ is a unit vector pointing orthogonally out of the horizon, and κ is the surface gravity. In calculations near the horizon it is convenient to introduce a coordinate system (α, λ, ϕ), where λ is a proper distance along the horizon from the north pole toward the equator. In these coordinates the metric of the absolute three-space near the horizon takes the form

$$ds^2 = (c^2/\kappa)^2\, d\alpha^2 + d\lambda^2 + \varpi^2\, d\phi^2, \tag{8.4.2}$$

and the unit vectors along the "toroidal" (ϕ), "poloidal" (λ), and "normal" (α) directions are

$$\mathbf{e}^{\mu}_{\hat{\phi}} \frac{\partial}{\partial x^{\mu}} = \varpi^{-1} \frac{\partial}{\partial \phi}, \quad \mathbf{e}^{\mu}_{\hat{\lambda}} \frac{\partial}{\partial x^{\mu}} = \frac{\partial}{\partial \lambda}, \quad \mathbf{n}^{\mu} \frac{\partial}{\partial x^{\mu}} = \frac{\kappa}{c^2} \frac{\partial}{\partial \alpha}. \tag{8.4.3}$$

Macdonald and Thorne (1982) formulated the conditions at the horizon as follows:

1. Gauss' law $\quad \mathbf{E}\,\mathbf{n} \equiv E_{\perp} \to 4\pi\,\sigma^{H}$ (8.4.4)
2. Charge conservation law $\quad \alpha\,\mathbf{j}\,\mathbf{n} \to -\dfrac{\partial \sigma^{H}}{\partial t} - {}^{(2)}\nabla \mathbf{i}^{H}$ (8.4.5)
3. Ampere's law $\quad \alpha\,\mathbf{B}_{\parallel} \to \mathbf{B}^{H} \equiv \left(\dfrac{4\pi}{c}\right) \mathbf{i}^{H} \times \mathbf{n}$ (8.4.6)
4. Ohm's law $\quad \alpha\,\mathbf{E}_{\parallel} \to \mathbf{E}^{H} \equiv R^{H}\,\mathbf{i}^{H}$ (8.4.7)

In these relations the symbol $\to$ indicates approach to the black hole horizon along the trajectory of a freely falling observer; ${}^{(2)}\nabla$ is the two-dimensional divergence at the horizon, and $\mathbf{B}_{\parallel}$ and $\mathbf{E}_{\parallel}$ are the magnetic and electric field components tangent to the horizon. $R^{H} \equiv 4\pi/c$ is the effective *surface resistance of the event horizon* ($R^{H} =$ 377 Ohm). The lapse function α in conditions (8.4.5)–(8.4.7) reflects the slowdown in the flow of physical time for locally non-rotating observers in the neighborhood of the black hole.

The values of $\mathbf{E}^{H}$ and $\mathbf{B}^{H}$ at the horizon are finite, and $\alpha \to 0$. Hence, taking into account the conditions given above, we arrive at the following properties of fields at the horizon:

1. $\mathbf{E}_{\perp}$ and $\mathbf{B}_{\perp}$ are finite at the horizon. (8.4.8)
2. $|\,\mathbf{E}_{\parallel}\,|$ and $|\,\mathbf{B}_{\parallel}\,|$ generally diverge at the horizon as α^{-1}. (8.4.9)
3. $|\,\mathbf{E}_{\parallel} - \mathbf{n} \times \mathbf{B}_{\parallel}\,| \propto \alpha \to 0$ at the horizon. (8.4.10)

Condition (8.4.10) signifies that for locally non-rotating observers, the electromagnetic field at the horizon acquires (in the general case) the characteristics of an electromagnetic wave sinking into the black hole at infinite blue shift.

In the presence of matter, one needs also to require that the matter density n in a freely-falling reference frame remains finite at the horizon. It is interesting to note that in the MHD-approximation the latter condition automatically implies the condition (8.4.10).

The conditions listed above make it possible to imagine quite clearly how electromagnetic processes will affect the properties of the black hole, slowly varying its parameters (the change is slow because we have assumed from the beginning that the electromagnetic field is relatively weak and the plasma is rare enough; see Section 8.1).

8.4.2 Slow change of black hole parameters

The change in the energy Mc^2 and angular momentum J of a black hole are equal to the total flux of energy and angular momentum across the horizon, respectively. All quantities are considered in the global time t.

The electromagnetic field contribution to the change of the black hole parameters can be written as

$$\frac{d(Mc^2)}{dt} = \int_H \left\{ \Omega^H \left[\sigma^H \mathbf{E}^H + (\mathbf{i}^H/c) \times \mathbf{B}_\perp \right] \varpi \, \mathbf{e}_{\hat{\phi}} + \mathbf{E}^H \, \mathbf{i}^H \right\} d\Sigma^H, \tag{8.4.11}$$

$$\frac{dJ}{dt} = \int_H \left(\sigma^H \mathbf{E}^H + (\mathbf{i}^H/c) \times \mathbf{B}_\perp \right) \varpi \, \mathbf{e}_{\hat{\phi}} \, d\Sigma^H. \tag{8.4.12}$$

Here $d\Sigma^H$ is an element of the horizon area. The first term in the braces in (8.4.11) describes the change in the rotational energy of the black hole, and the second term gives the change due to the "heating" of the black hole by the surface current.

We thus find that the boundary conditions at the event horizon, described in this section, make it possible to model a black hole for electrodynamic problems in the exterior space as an imaginary surface with specific electromagnetic properties. The surface can carry surface charge and electric currents. This visually clear picture helps greatly in solving specific problems. As we already mentioned, this approach is called the membrane formalism (see Section 3.7).

We wish to emphasize again that there is no real (material) surface, no charges, and no current at the black hole boundary. Note also that the fields $\mathbf{E}$ and $\mathbf{B}$ that a locally non-rotating observer measures at the event horizon, differ drastically from $\mathbf{E}^H$ and $\mathbf{B}^H$ that appear in the boundary conditions: The reason is the factor α in the definition of $\mathbf{E}^H$ and $\mathbf{B}^H$ [see (8.4.6) and (8.4.7)]. This factor appears (recall the formulation of Maxwell's equations) as a result of using the "global" time t.

In the membrane paradigm the boundary conditions are fixed on the stretched horizon that lies in the immediate vicinity and outside the event horizon (see Figure 3.8). The stretched horizon is defined by the equation $\alpha = \text{const}$, where $\alpha \ll 1$. The exact value of α, and hence the position of the stretched horizon, is not specified. It is important to emphasize that the quantities $\mathbf{E}^H$ and $\mathbf{B}^H$ are defined in such a way that in the limit $\alpha \to 0$ they do not depend on the real position of the stretched horizon.

Let us return to the true event horizon. As in the preceding section, we introduce "special" physical conditions. First, we assume that the problem is stationary and axisymmetric. The fictitious surface current $\mathbf{i}^H$ and electric field $\mathbf{E}^H$ at the horizon are then completely poloidal, while the magnetic field $\mathbf{B}^H$ is toroidal. The values of these quantities at the horizon are

$$\mathbf{i}^H = \frac{I^H}{2\pi \, \varpi_H} \, \mathbf{e}_{\hat{\lambda}}, \tag{8.4.13}$$

$$\mathbf{E}^H = \frac{2I^H}{c\,\varpi_H}\,\mathbf{e}_{\hat\lambda}, \tag{8.4.14}$$

$$\mathbf{B}^H = -\frac{I^H}{2\pi\,\varpi_H}\,\mathbf{e}_{\hat\phi}, \tag{8.4.15}$$

where I^H and ϖ_H are the values of I and ϖ at the horizon. Furthermore, the poloidal magnetic field measured by non-rotating observers intersects the event horizon at right angles. Recall that the toroidal component diverges at the horizon.

In the particular case of the MHD-approximation, the conditions (8.4.14) and (8.4.15) imply

$$\frac{(E - \Omega^H L)}{\mathcal{M}_H^2 + (\Omega^F - \Omega^H)^2\,(\varpi_H^2/c^2)} = -\frac{c^2}{8\pi^2\,\Sigma_H^{1/2}\,\varpi_H}\left(\frac{\partial\Psi}{\partial\theta}\right)_H . \tag{8.4.16}$$

For regular flow ($\mathcal{M}(r_+) \neq 0$), this condition follows directly from the $\alpha \to 0$ limit of the algebraic constraint (8.3.42). Moreover, the solutions of the equation (8.3.48) without any unphysical singularity automatically satisfy the boundary condition (8.4.16) at the event horizon. According to the definitions (8.3.8) and (8.3.10), the electric field $\mathbf{E}^H$ at the horizon is directly expressible in terms of $\mathbf{B}_\perp$:

$$\mathbf{E}^H = -\frac{1}{c}\,(\Omega^F - \Omega^H)\,\varpi\,\mathbf{e}_{\hat\phi} \times \mathbf{B}_\perp . \tag{8.4.17}$$

In the MHD-approximation, the change of the energy and angular momentum of a black hole can be simply expressed in terms of the integrals of motion E and L taken at the surface of a black hole:

$$\frac{d(Mc^2)}{dt} = -\frac{1}{c}\int_0^\pi E|_H \left(\frac{d\Psi}{d\theta}\right)_H d\theta, \tag{8.4.18}$$

$$\frac{dJ}{dt} = -\frac{1}{c}\int_0^\pi L|_H \left(\frac{d\Psi}{d\theta}\right)_H d\theta . \tag{8.4.19}$$

On the black hole surface the quantities $E|_H$ and $L|_H$ are well-defined (single-valued) functions of θ, and hence the integrals have a well-defined meaning.

In the force-free approximation, the following "principle of least action" is found to hold [Macdonald and Thorne (1982)]. The lines of the poloidal magnetic field that intersect the horizon have a distribution that ensures extremal total surface energy $\mathcal{E}$ of the tangential electromagnetic field at the horizon. The expression for $\mathcal{E}$ has the form

$$\mathcal{E} = \frac{1}{8\pi}\int_H \left[(\mathbf{B}^H)^2 + (\mathbf{E}^H)^2\right] d\Sigma^H, \tag{8.4.20}$$

where the integration is carried out over the horizon.

Now the boundary condition (8.4.16) can be rewritten in the form

$$4\pi\, I(\Psi) = \left[\Omega^H - \Omega^F(\Psi)\right] \frac{\varpi_H}{\Sigma_H^{1/2}} \left(\frac{\partial\Psi}{\partial\theta}\right)_H . \tag{8.4.21}$$

Then equations (8.4.18) and (8.4.19) take the form

$$\frac{d(Mc^2)}{dt} = \int_H \frac{\Omega^F(\Omega^F - \Omega^H)}{4\pi c} \frac{A^H \sin^2\theta}{\Sigma_H} (B_\perp)^2\, d\Sigma^H , \tag{8.4.22}$$

$$\frac{dJ}{dt} = \int_H \frac{(\Omega^F - \Omega^H)}{4\pi c} \frac{A^H \sin^2\theta}{\Sigma_H} (B_\perp)^2\, d\Sigma^H . \tag{8.4.23}$$

The angular momentum and energy lost by the black hole are transferred along the lines of force of the poloidal field in the force-free region to those "regions" where condition (8.3.23) is violated.

Note that if $\Omega^F = 0$ (i.e., if the magnetic lines of force are, say, frozen into plasma far from the black hole and this plasma *does not participate* in the rotation around the hole), then $dM = 0$. This means that the total mass of the black hole is conserved and $dJ < 0$. In other words, the rotation of the black hole slows down. The entire energy of rotation transforms into the mass of the black hole (the so-called irreducible mass; see Section 7.1) so that nothing escapes.

If the parameters of the black hole are fixed, its angular velocity Ω^F is determined by the boundary condition far from the black hole in the external plasma. The situation in realistic astrophysical conditions will be discussed in Section 8.6.

8.5 Electromagnetic Fields in Vacuum Near a Black Hole

Before beginning the description of the magnetosphere of a rotating black hole (it is formed via accretion of magnetized gas; see the next section), we illustrate the above analysis by the solutions of the following problems on electromagnetic fields in vacuum:

1. Electric charge in the vacuum in the Schwarzschild metric [Copson (1928), Linet (1976), Hanni and Ruffini (1973)]
2. Magnetic field in the vacuum in the Kerr metric, uniform at infinity [Wald (1974b), Bičák and Dvořák (1976), King and Lasota (1977), Damour (1978), Thorne and Macdonald (1982)]

8.5.1 Electric field of a point-like charge

We begin with problem 1. Let a point-like charge q be at rest in the Schwarzschild coordinates at $r = b$, $\theta = 0$. We already discussed this problem in Section 7.2, where the explicit expressions for vector potential A_μ and field strength $F_{\mu\nu}$ were given. Now we demonstrate how the same problem can be solved in the (3+1)-approach.

The problem reduces to solving system (8.3.12)–(8.3.14) with a δ-function for ρ_e and $j^P = j^T = 0$. Conditions (8.3.12) and (8.3.13) are satisfied when $\Psi = 0$ and $I = 0$. Expressions (8.3.10) and (8.3.11) then imply that the external magnetic field is absent. External currents also vanish. Expressions (8.4.5) and (8.4.6) show that the surface current at the horizon vanishes as well, $\mathbf{i}^H = 0$. As follows from condition (8.4.10), at the horizon $\mathbf{E}_{\parallel} \to 0$ so that electric lines of force intersect the horizon at right angles. The total flux of $\mathbf{E}$ across the horizon is zero (the black hole is uncharged). With these boundary conditions, $\mathcal{U}_0$ is found from the solution (8.3.14) which for $\omega = 0$ to be of the form

$$\nabla(\alpha^{-1}\,\nabla\mathcal{U}_0) = 4\pi\,\rho_e\,. \tag{8.5.1}$$

By solving this equation for a point-like charge q located on the symmetry axis at a point $r = b$ and using the equation (8.3.8), one finds the following expression for $\mathbf{E}^P$:

$$\begin{aligned}\mathbf{E}^P = \frac{q}{b\,r^2}\Bigg\{&M\left(1 - \frac{b - M + M\cos\theta}{R}\right)\\ &+ \frac{r[(r-M)(b-M) - M^2\cos\theta]}{R^3}[r - M - (b-M)\cos\theta]\Bigg\}\,\mathbf{e}_{\hat r}\\ &+ \frac{q(b-2M)(1-2M/r)^{1/2}\sin\theta}{R^3}\,\mathbf{e}_{\hat\theta}\,,\end{aligned} \tag{8.5.2}$$

where $\mathbf{e}_{\hat r}$ and $\mathbf{e}_{\hat\theta}$ are unit vectors along the directions of r and θ, respectively, and

$$R \equiv \left[(r-M)^2 + (b-M)^2 - M^2\sin^2\theta - 2(r-M)(b-M)\cos\theta\right]^{1/2}. \tag{8.5.3}$$

(Throughout this section, with the exception of the final formulas, we set $G = 1$, $c = 1$)

The pattern of electric lines of force is shown in Figure 8.2. The charge surface density at the boundary of the black hole follows from (8.4.4):

$$\sigma^H = \frac{q\,[M(1+\cos^2\theta) - 2(b-M)\cos\theta]}{8\pi b\,[\,b - M(1+\cos\theta)\,]^2}\,. \tag{8.5.4}$$

Let us bring the charge closer to the horizon ($b \to 2M$). At a distance $r \gg b - 2M$ from the horizon, the lines of force become practically radial, and the field strength tends to q/r^2. With the exception of a narrow region close to the horizon, the general picture is almost the same as for a charge placed at the center of the black hole.

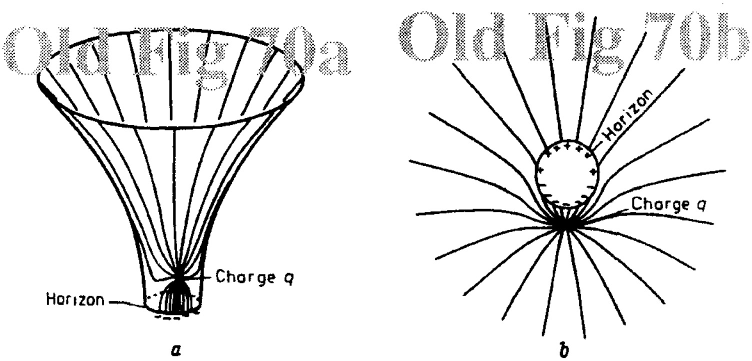


Figure 8.2: The electric-field lines of force of a test charge q at rest in the Schwarzschild metric, in a ϕ = const section: (a) lines of force on a curved surface whose geometry coincides with the section ϕ = const of the Schwarzschild metric; (b) the same lines projected on a plane ("bird's-eye view"). The distribution of the fictitious surface charge σ^H is shown on the horizon. The charge q is assumed to be positive.

8.5.2 Black hole in a homogeneous magnetic field

Now we will give, without detailed justification, the solution of problem 2. A rotating black hole is placed in a magnetic field B_0, uniform at infinity. In the Kerr metric, the magnetic field is given by the expression

$$\mathbf{B} = \frac{B_0}{2A^{1/2}\sin\theta}\left[\frac{\partial X}{\partial\theta}\,\mathbf{e}_{\hat{r}} - \Delta^{1/2}\,\frac{\partial X}{\partial r}\,\mathbf{e}_{\hat{\theta}}\right], \tag{8.5.5}$$

where $X \equiv (A - 4a^2Mr)(\sin^2\theta)/\Sigma$, and A is given by (3.2.2).

The electric field induced by the rotation of the black hole is proportional to a:

$$\mathbf{E} = -\frac{B_0\,a\,A^{1/2}}{\Sigma}\left\{\left[\frac{\partial(\alpha^2)}{\partial r} + \frac{M\sin^2\theta}{\Sigma}(A - 4a^2Mr)\,\frac{\partial}{\partial r}\left(\frac{r}{A}\right)\right]\mathbf{e}_{\hat{r}}\right.$$
$$\left. + \Delta^{-1/2}\left[\frac{\partial(\alpha^2)}{\partial\theta} + \frac{Mr\sin^2\theta}{\Sigma}(A - 4a^2Mr)\,\frac{\partial}{\partial\theta}\left(\frac{1}{A}\right)\right]\mathbf{e}_{\hat{\theta}}\right\}. \tag{8.5.6}$$

As in problem 1, $\mathbf{E}^H$, $\mathbf{B}^H$, and $\mathbf{i}^H$ are absent.[7] Formulas (8.4.12), (8.4.11) imply that the angular momentum $\mathbf{J}$ of a black hole and its mass M remain invariant. But

[7] If we calculate the components of the field $E_\perp$ and $B_\perp$ at the event horizon (see Section 8.4), we find that both are proportional to $r_+ - M$, where $r_+ = M + (M^2 - a^2)^{1/2}$. Hence, we obtain for a black hole rotating at maximum angular velocity, with $a_{\max} = M$, that $E_\perp = 0$ and $B_\perp = 0$ at the horizon; that is, the lines of force of the axisymmetric field do not pass through the black hole (the case of an asymmetric field is treated later). Problems in black hole electrodynamics were also treated by Lèaute and Linet (1976, 1982), Misra (1977), Gal'tsov and Petukhov (1978), Linet (1976, 1977a,b, 1979), Demianski and Novikov (1982), Bičák and Dvořák (1980)

the angular momentum changes if the magnetic field is tilted at an angle to the black hole axis. The result is as follows.

Let a magnetic field $\mathbf{B}$, uniform at infinity, be tilted at an angle to the direction of the angular momentum $\mathbf{J}$. Decompose $\mathbf{J}$ into a component $\mathbf{J}_{\parallel}$ parallel to the field and a component $\mathbf{J}_{\perp}$ perpendicular to it. The following formulas give $\mathbf{J}$ as a function of time:

$$\mathbf{J}_{\parallel} = \text{const}\,, \tag{8.5.7}$$

$$\mathbf{J}_{\perp} = \mathbf{J}_{\perp}(t=0)\,\exp(-\,t/\tau)\,, \tag{8.5.8}$$

where

$$\tau = \frac{3\,c^5}{2GMB^2} = 10^{36}\,\text{yr}\left(\frac{M}{M_{\odot}}\right)^{-1}\left(\frac{B}{10^{-5}\ \text{Gauss}}\right)^{-2}. \tag{8.5.9}$$

The component $\mathbf{J}_{\perp}$ of the black hole angular momentum is thus completely lost with time. (But the time of disintegration is fantastically long!) The rotational energy connected with $\mathbf{J}_{\perp}$ thus is transformed in a static magnetic field into the "irreducible" mass of the black hole, while the component $\mathbf{J}_{\parallel}$ remains unaltered.

The final state of the black hole is subject to Hawking's theorem which states that the stationary state must be axially symmetric. Press (1972) points out that if the external magnetic field (or any other field) is not axisymmetric, the black hole finally loses its angular momentum (by Hawking's theorem). If the field $\mathbf{B}$ varies smoothly on a scale much greater than the black hole size, $\mathbf{J}$ can again be decomposed into $\mathbf{J}_{\parallel}$ and $\mathbf{J}_{\perp}$ with respect to the field direction in the black hole neighborhood. In order of magnitude, the decrease of $\mathbf{J}_{\perp}$ is again determined by formula (8.5.8), while $\mathbf{J}_{\parallel}$ decreases according to the formula

$$\frac{d\mathbf{J}_{\parallel}}{dt} \approx -\frac{\mathbf{J}_{\parallel}}{\tau}\,O\!\left(\frac{r_g}{R}\right), \tag{8.5.10}$$

where R is the scale of non-uniformity of the field.

We have mentioned in the note 7 to page 310 that if $a = a_{\text{max}}$, an axisymmetric magnetic field does not intersect the black hole horizon. It is possible to show that if a magnetic field B_0, uniform at infinity, is tilted with respect to the rotation axis, then the flux through one half of the horizon[8] for the field component $B_{0\perp}$ is maximal when $a = a_{\text{max}}$ [Bičák (1983) and Bičák and Janiš (1985)]. This flux equals

$$\Psi^{H}_{\text{max}} \approx 2.25\,B_{0\perp}\,\pi M^2. \tag{8.5.11}$$

Finally, consider a *non-rotating* black hole placed in a strong magnetic field B_0, uniform at infinity [Bičák (1983)]. Let the field be so strong that its self-gravitation

[8]Of course, the total flux of magnetic field across any closed surface, including the horizon, equals zero. In speaking about the flux through a black hole, one considers the incoming (or the outgoing) lines of force.

has to be taken into account. It is then found that for a black hole of fixed mass M, there is a critical field $B_{0,\mathrm{cr}}$ at which the flux Ψ^H through one half of the event horizon is maximal:

$$B_{0,\mathrm{cr}} = c^4\, G^{-3/2}\, M^{-1} = 2.5 \times 10^{13}\,\mathrm{Gauss} \times \left(\frac{10^6 M_\odot}{M}\right), \tag{8.5.12}$$

$$\Psi^H_{\mathrm{max},1} = 2\pi\, G^{1/2}\, M. \tag{8.5.13}$$

The flux across the horizon cannot be greater than $\Psi^H_{\mathrm{max},1}$.

Concerning a rotating electrically charged black hole in an external magnetic field, see Dokuchayev (1987).

8.6 Magnetosphere of a Black Hole

8.6.1 Magnetospheric models

The problems discussed in the preceding section illustrate some important properties of electric and magnetic fields in the neighborhood of a black hole. However, these problems can hardly be used for the description of the actual electrodynamic processes expected to take place in astrophysical conditions. We have already mentioned that the reason for this is the no-vacuum nature of the fields in the vicinity of black holes. The space is always filled with rarefied plasma (so that the fields become force-free) or even more complicated situations could arise.

Recall that the fields in the neighborhood of rotating magnetized neutron stars (pulsars) cannot be considered as vacuum fields either. A complicated pulsar magnetosphere is formed. For more details, see e.g., the books [Michel (1991), Beskin, Gurevich, and Istomin (1993)].

By analogy to pulsars, the region of magnetized plasma around a black hole is called the *black hole magnetosphere.* The complexity and diversity of processes in this region stand in the way of developing a complete magnetospheric theory. In fact, we still lack an acceptably complete theory of a pulsar's magnetosphere, despite the great expenditure of effort and time that has been put into it.

We shall not discuss all aspects of the theory of the black hole magnetosphere here. We shall only look at the important aspects of the electrodynamic processes that occur in the black hole neighborhood and are black-hole-specific. Correspondingly, we restrict the presentation to the simplest model. Only the processes caused by the black hole itself are covered. (For discussion of other problems, such as the motion of plasma in the accreted gas disk formed around the black hole and the problem of jet formation, which are not discussed in the book, see e.g., [Shapiro and Teukolsky (1983), Blandford and Payne (1982), Heyvaerts and Norman (1989), Li, Chiuen, and Begelman (1992), Pelletier and Pudritz (1992)].)

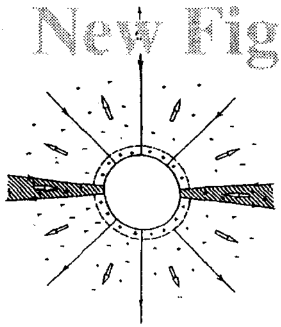


Figure 8.3: The model of a black hole magnetosphere considered by Blandford and Znajek (1977). Arrows indicate the longitudinal electric current $\mathbf{j}^P$. Dashed lines separate regions with positive and negative charge density ρ_e.

We choose the stationary axisymmetric model of the force-free magnetosphere of a black hole inducing disc accretion of magnetized gas (the magnetosphere schematically shown in Figure 8.1). This model was analyzed by Blandford (1976), Blandford and Znajek (1977), Macdonald and Thorne (1982), and Phinney (1983a,b); see also Lovelace (1976), Lovelace *et al.* (1979), Thorne and Blandford (1982), Rees *et al.* (1982), Macdonald (1984), Okamoto (1992), Beskin, Istomin, and Pariev (1983, 1992).

For the condition of force-free field existence, (8.3.23), to be satisfied in the neighborhood of a black hole, it is necessary to have rarefied plasma with electric currents flowing along magnetic lines of force. Charges for these currents crossing the black hole must be constantly replenished because they sink into the hole ,and obviously, charges cannot flow back out.[9] Mechanisms of free charge generation thus have to exist in the neighborhood of the black hole. Such mechanisms were analyzed by Blandford and Znajek (1977), Kardashev *et al.* (1983), and Beskin, Istomin, and Pariev (1993). We remark, without going into the details, that these mechanisms require a small component of the electric field, parallel to the magnetic field. This component is so small that inequality (8.3.16) is not violated.

Two examples of a magnetosphere of a black hole are shown in Figures 8.3 and 8.4. They correspond to the solutions of the force-free equation (8.3.24) for a slowly rotating black hole ($\Omega^H r_+/c \ll 1$). Arrows indicate the longitudinal electric current $\mathbf{j}^P$. Dashed lines separate regions with positive and negative charge density ρ_e.

The solution presented in Figure 8.3 corresponds to the following choice of inte-

[9] The total charge of the black hole in the stationary solution cannot be changed by this process because the total numbers of charges of opposite signs, sinking into the black hole, are equal to each other.

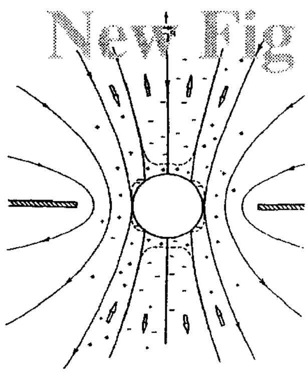


Figure 8.4: The model of a black hole magnetosphere in which the magnetic field near the black hole is close to homogeneous [Beskin, Istomin, and Pariev (1992)]. In this case the closure of electric current takes place where the magnetic field lines cross the horizon at $\theta \sim \pi/2$.

grals of motion [Blandford and Znajek (1977)]

$$\Omega^F(\Psi) = \frac{1}{2}\,\Omega^H\,, \tag{8.6.1}$$

$$I(\Psi) = \pm\,\frac{\Omega^F}{4\pi}\left(2\Psi - \Psi^2/\Psi_0\right), \tag{8.6.2}$$

where the sign + (−) is to be taken for the upper (lower) hemisphere. It can be shown that for a slowly rotating black hole, up to terms which remain small everywhere outside the horizon, the solution to the equation (8.3.24) is of the form

$$\Psi = \Psi_0\,(1 - |\cos\theta\,|)\,. \tag{8.6.3}$$

Thus, $\Psi_0 = \Psi^H_{\rm max}$ is the maximal magnetic flux through the horizon. Such a split-monopole magnetic field can be generated by toroidal currents flowing in the accretion disk provided that the latter (formally) extends up to the horizon. For this solution the influence of the rotating charged plasma is totally compensated by the longitudinal current I. The charge density ρ_e (8.3.22) changes sign at $r = 2^{1/3}\,r_+$. The magnetic field is always (even at the distances lying outside the *light cylinder* $R_L = c/\Omega^F$) larger than the electric field. The validity of the solution (8.6.3) (under the condition (8.6.2)) at large distances ($r \gg R_L$) was shown by Michel (1973) for the case of a pulsar magnetosphere. It is interesting to emphasize that the expression (8.6.3) is the exact vacuum solution of Maxwell's equations (8.2.4) and (8.3.10) for a non-rotating black hole.

The solution presented in Figure 8.4 corresponds to the case when the magnetic field near the black hole is close to homogeneous, while it is monopole-like at large distances [Beskin, Istomin, and Pariev (1992)]. Such a magnetic field can be generated in the disc with finite inner radius. The longitudinal current closure takes place not inside the disk (as in the Blandford-Znajek solution) but near the points where the magnetic field lines cross the horizon at $\theta \sim \pi/2$. It can be shown that the electric current obeys the following condition at the horizon

$$\int_H \mathbf{j}\, d\Sigma^H = 0\,, \tag{8.6.4}$$

and hence the electric charge of the black hole is not changed (see also Okamoto (1992)). The surface separating the regions with positive and negative charge densities (shown by dashed lines) has a more complicated form. One can argue that the domain where plasma is generated inside the magnetosphere lies near this surface [Beskin, Istomin, and Pariev (1992)].

8.6.2 Efficiency of the power-generation process near a rotating, magnetized black hole

We discuss now the efficiency of the power-generation process near a rotating, magnetized black hole. Consider a thin tube of lines of force that pass through the black hole. This tube rotates around the black hole at a constant angular velocity Ω^F (see Section 8.3). In the force-free approximation, formula (8.4.22) shows that the rotational energy of the black hole is extracted at the rate

$$\mathcal{P} = -\frac{d(Mc^2)}{dt} = \int_H \frac{\Omega^F(\Omega^H - \Omega^F)}{4\pi c}\, \frac{A^H \sin^2\theta}{\Sigma_H}\, B_\perp^2\, d\Sigma^H. \tag{8.6.5}$$

This energy is transferred along magnetic lines of force into region 3 (see Figure 8.1) where the force-free condition is violated; energy is pumped into accelerated particles, and so forth.

In this region, the particles exert a backreaction on the line of force, owing to their inertia, and thus determine Ω^F. If the inertia is large, the angular velocity Ω^F is small ($\Omega^F \ll \Omega^H$); in the limit, $\Omega^F \to 0$. The power $\mathcal{P}$ of the above "engine" is quite low, as follows from (8.6.5). Otherwise (i.e., when the inertia of the particles in the region 3 is low), $\Omega^F \to \Omega^H$, and (8.6.5) again gives low power. The power is maximal when $\Omega^F = \Omega^H/2$.

Macdonald and Thorne (1982) demonstrated that this condition is very likely to be implemented in the described model. Their arguments run as follows. The angular velocity of non-rotating observers far from the black hole can be assumed zero. That is why the velocity $\mathbf{v}^F$ of the points of the lines of force far from the symmetry axis (with respect to non-rotating observers) is much greater than the speed of light:

$$|\,\mathbf{v}^F| = \Omega^F \varpi \gg c\,. \tag{8.6.6}$$

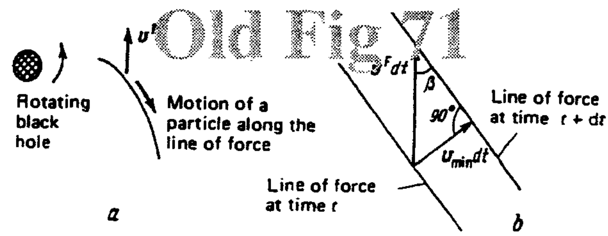


Figure 8.5: (a) Schematics of motion of a charged particle along a magnetic field line of force rotating around a black hole. (b) The position of a segment of a magnetic line of force in the plane of the vectors $\mathbf{B}^P$ and $\mathbf{B}^T$ at the moments t and $t + dt$. If the velocity vector of a particle that slides outward in the absolute space along a line of force is perpendicular to this line, the velocity of the particle is minimal.

Charged particles cannot move at a velocity greater than c. However, staying on a line of force and sliding along it outward (see Figure 8.5a), they can have velocities less than $\mathbf{v}^F$.

Figure 8.5b shows a segment of a line of force in the plane defined by the vectors $\mathbf{B}^P$ and $\mathbf{B}^T$. There is an optimal velocity of sliding along the line of force (which, in turn, moves at a velocity $\mathbf{v}^F$) such that the total velocity of the particle in the absolute space is minimal (this is clear from Figure 8.5b). The angle β is the angle between the direction of the line of force and the direction of the ϕ coordinate; hence, β is found from the relation

$$\sin\beta = \frac{|\mathbf{B}^P|}{\sqrt{(\mathbf{B}^P)^2 + (\mathbf{B}^T)^2}}\,. \tag{8.6.7}$$

Using (8.6.7), we find $|\mathbf{v}_{\min}|$:

$$|\mathbf{v}_{\min}| = \frac{|\mathbf{v}^F|}{\sqrt{1 + (\mathbf{B}^T)^2/(\mathbf{B}^P)^2}}\,. \tag{8.6.8}$$

It is found that the condition $|\mathbf{v}_{\min}| \approx c$ is equivalent to the maximal energy condition $\Omega^F \approx \Omega^H/2$. To show this, we need expressions for $|\mathbf{v}^F|$, $|\mathbf{B}^T|$, and $|\mathbf{B}^P|$ far from both the black hole and the symmetry axis. For $|\mathbf{v}^F|$ we find

$$|\mathbf{v}^F| = \Omega^F\,\varpi\,. \tag{8.6.9}$$

Making use of formulas (8.3.11), (8.4.14), (8.4.17), and definition (8.3.2), we obtain

$$|\mathbf{B}^T| \approx \frac{(\Omega^H - \Omega^F)\Psi}{\pi c\,\varpi}\,. \tag{8.6.10}$$

Finally, formula (8.3.10) and relation $|\boldsymbol{\nabla}\Psi| \approx 2\Psi/\varpi$ yield

$$|\mathbf{B}^P| \approx \frac{\Psi}{\pi c\,\varpi^2}\,. \tag{8.6.11}$$

Substituting (8.6.9)-(8.6.11) into (8.6.8), we find

$$|\mathbf{v}_{\min}| \approx \frac{c\,\Omega^F}{\Omega^H - \Omega^F}\,. \tag{8.6.12}$$

The last formula implies that $|\mathbf{v}_{\min}| \approx c$ when $\Omega^F = \Omega^H/2$. If $\Omega^F \ll \Omega^H/2$ and $|\mathbf{v}_{\min}| \ll c$, then the inertia of particles in the region 3 is small, and Ω^F increases until the velocity $|\mathbf{v}_{\min}|$ grows close to c. If $|\mathbf{v}_{\min}| \gg c$, then particles cannot stay long on the lines of force, and their backreaction on the field reduces Ω^F until we obtain $|\mathbf{v}_{\min}| \approx c$.[10]

These are the reasons why it is likely that $\Omega^F \approx \Omega^H/2$ and the rate of extraction of rotational energy from the black hole (8.6.5) is almost optimal.

In order of magnitude, the power of the "electric engine" outlined above is

$$\mathcal{P} \approx \left(10^{39}\,\frac{\text{erg}}{\text{s}}\right)\left(\frac{M}{10^6\,M_\odot}\right)^2\left(\frac{a}{a_{\max}}\right)^2\left(\frac{B}{10^4\,\text{Gauss}}\right)^2. \tag{8.6.13}$$

Here B is the magnetic field strength in the neighborhood of the black hole. Sometimes this electric engine is described in terms of electrical engineering [Blandford (1979), Znajek (1978), Damour (1978), Macdonald and Thorne (1972), Thorne and Blandford (1982), Phinney (1983a)]. We will give the expressions for quantities at the black hole horizon using this notation.

8.6.3 Black hole as a unipolar inductor

Equipotential curves at the horizon are the lines of constant θ since the field $\mathbf{E}^H$ is meridional [see (8.4.14)]. Hence, the potential difference between two equipotential lines (marked by 1 and 2) are [see also (8.4.17)]

$$\begin{aligned}\Delta U^H &= \int_1^2 \mathbf{E}^H d\mathbf{l} = \left(\Omega^H - \Omega^F\right)\frac{\Delta\Psi}{2\pi c}\\ &\approx (10^{17}\,\text{V})\left(\frac{M}{10^6\,M_\odot}\right)\left(\frac{B}{10^4\,\text{Gauss}}\right)\left(\frac{a}{a_{\max}}\right),\end{aligned} \tag{8.6.14}$$

where $d\mathbf{l}$ is an element of distance along a meridian on the black hole surface, and $\Delta\Psi$ is the difference between the values of Ψ on the equipotential 1 and 2. The approximate equality in (8.6.14) is written for the condition $\Omega^F \approx \Omega^H/2$, maximal Ω^H, and the equipotentials 2 and 1 corresponding to the equatorial and polar regions, respectively.

[10] A similar analysis of the slipping of particles along the magnetic lines of force at the black hole horizon shows that the condition $|v_{\min}| \approx c$ corresponds to the boundary conditions (8.4.14) and (8.4.17).

On the other hand, ΔU^H can be written in terms of the surface current $\mathbf{i}^H$ and resistance:

$$\Delta U^H = R^H\, \mathbf{i}^H\, \Delta \mathrm{l}, \tag{8.6.15}$$

where $\Delta\mathrm{l}$ is the distance along the meridian between the equipotentials 2 and 1. Substituting expression (8.4.13) for $\mathbf{i}^H$,we obtain

$$\Delta U^H = \frac{I\, R^H\, |\Delta \mathbf{l}|}{2\pi\, \varpi_H} = I\, \Delta Z^H, \tag{8.6.16}$$

where

$$\Delta Z^H \equiv \frac{R^H\, |\Delta \mathbf{l}|}{2\pi\, \varpi_H} \tag{8.6.17}$$

is the total resistance between the equipotential lines 2 and 1. (If the equipotentials 2 and 1 correspond to the equator and to $\theta \approx \pi/4$, the integration of (8.6.17) yields $\Delta Z^H \approx 30\,\mathrm{Ohm}$.)

Formulas (8.6.14) and (8.6.17) permit the conclusion that in this model the rotating black hole acts as a battery with e.m.f. of order

$$(10^{17}\,\mathrm{V}) \left(\frac{M}{10^6\, M_\odot}\right) \left(\frac{B}{10^4\,\mathrm{Gauss}}\right) \tag{8.6.18}$$

and internal resistance of about 30 Ohm.

This mechanism (and a number of its variants) has been employed in numerous papers for the explanation of the activity of the nuclei of galaxies and quasars [for example, see Ruffini and Wilson (1975), Blandford(1976), Blandford and Znajek (1977), Blandford and Rees(1978), Blandford (1979), Lovelace *et al.* (1979), Kardashev *et al.* (1983), Rees (1982), Phinney (1983a,b), Begelman, Blandford, and Rees (1984), Novikov and Stern (1986), Camenzind (1986a,b, 1987), Punsly and Coronity (1990a, b), Takahashi *et al.* (1990), Punsly (1991), Nitta, Takahashi, and Tomimatsu (1991), Hirotani *et al.* (1992), Okamoto (1992), Beskin, Istomin, and Pariev (1993), Beskin and Pariev (1993), Horiuchi, Mestel, and Okamoto (1995), Beskin (1997)].

Chapter 9

Astrophysics of Black Holes

9.1 Introduction

The black hole phenomenon is undoubtedly one of the most striking and intriguing of those predicted by theorists. However, a paramount question is: "Do black holes exist in the Universe, or are they only an abstract concept of the human mind?" In principle, a black hole could be built artificially. However, this meets such grandiose technical difficulties that it looks impossible, at least in the immediate future. In fact, the artificial construction of a black hole looks even more problematic than an artificial creation of a star. Thus, we have to conclude that the physics of black holes, as well as the physics of stars, is the physics of celestial bodies. Stars definitely exist, but what may one say about the existence of astrophysical black holes?

Modern astrophysics considers three possible types of black holes in the Universe:

- *Stellar black holes*; that is, black holes of stellar masses which were born when massive stars died
- *Supermassive black holes* with masses up to $10^9 M_\odot$ and greater at centers of galaxies ($M_\odot = 2 \times 10^{33}$g is the solar mass)
- *Primordial black holes* that might appear from inhomogeneities at the very beginning of the expansion of the Universe. Their masses can be arbitrary, but primordial black holes with $M \leq 5 \times 10^{14}$g would have radiated away their mass by the Hawking quantum process in a time $t \leq 10^{10}$ years (the age of the Universe)[1]. Only primordial black holes with mass $M > 5 \times 10^{14}$g could exist in the contemporary Universe.

[1] We shall discuss the Hawking process in Chapter 10. Here we just mention that because of the quantum instability of vacuum in a strong gravitational field, a black hole is a source of quantum radiation. It radiates all species of physical particles in the same manner as a heated body of size $r_g = 2GM_H/c^2$ with temperature $T_H \approx 10^{-26} K(M_\odot/M_H)$. As a result of this process, the mass M_H of a black hole decreases. For rough estimation of the time of the black-hole complete evaporation one can use the relation $t \approx (M_H/5 \times 10^{14}\text{g})^3 \times 10^{10}$yr.

There is enough material for a separate book on each of the types of astrophysical black holes. In this chapter we are only able to summarize briefly the main ideas of these branches of astrophysics and give some references to material where further details can be found. The history of the idea of black holes and the related astrophysics has been described by Israel (1987). General problems of the astrophysics of black holes are discussed by: Zel'dovich and Novikov (1971b), Novikov and Thorne (1973), Shapiro and Teukolsky (1983), Blandford (1987), Lamb (1991), Begelman and Rees (1996).

We would like to note also that the style of this chapter differs from the style of other parts of the book. It is closer to the style adopted in astrophysics. We give only order of magnitude estimates (as is typical for astrophysics), and discuss available observational data. In this chapter we do not assume that $G = c = 1$ and write explicitly G and c in formulas. Our main attention is focused on the possible observational manifestations of black holes.

9.2 The Origin of Stellar Black Holes

"When all the thermo-nuclear sources of energy are exhausted a sufficiently heavy star will collapse" - this is the first sentence of the abstract of a remarkable paper by Oppenheimer and Snyder (1939). Every statement of this paper accords with ideas that remain valid today. The authors conclude the abstract by the following sentence: "... an external observer sees the star shrinking to its gravitational radius." This is the modern prediction of the formation of black holes when massive stars die.

How heavy should a star be to turn into a black hole? The answer is not simple. A star which is not massive enough ends up either as a white dwarf or a neutron star. There are upper limits on the masses of both these types of celestial bodies. For white dwarfs it is the *Chandrasekhar limit*, which is about $(1.2 - 1.4) \times M_{\odot}$ [see Shapiro and Teukolsky (1983), Kippenhahn and Weigert (1990), Kawaler (1997)]. For neutron stars it is the *Oppenheimer-Volkoff limit* [Oppenheimer and Volkoff (1939)]. The exact value of this limit depends on the equation of state at matter densities higher than the density of nuclear matter $\rho_0 = 2.8 \times 10^{14}\text{g cm}^{-3}$. The modern theory gives for the maximal mass of a non-rotating neutron star the estimate $(2-3) \times M_{\odot}$ [see Baym and Pethick (1979), Lamb (1991), Srinivasan (1997)].[2]

Rotation can increase the maximal mass of a non-rotating neutron star only slightly [up to 25%; Friedman and Ipser (1987), Haswell *et al.* (1993)]. Thus, one can believe that the upper mass limit for neutron stars should not be greater than $M_0 \approx 3M_{\odot}$ [Lamb (1991), Cowley (1992), McClintock (1992), Glendenning (1997)]. If a star at the very end of its evolution has a mass greater than M_0, it must turn into

[2]Some authors have discussed the possible existence of so-called "quark stars", "hadronic stars", and "strange stars" [see Alcock, Farhi, and Olinto (1986), Bahcall, Lynn, and Selipsky (1990), Madsen (1994), Glendenning (1997)]. At present, there is no evidence for such stars.

a black hole. However, this does not mean that all normal stars (on the "main sequence" of the Hertzsprung-Russell diagram, [Bisnovatyi-Kogan (1989), Kippenhahn and Weigert (1990)]) with masses $M > M_0$ are black hole progenitors. The point is that the final stages of evolution of massive stars are poorly understood. Steady mass loss, catastrophic mass ejection, and even disruption in supernovae explosions are possible [see Kippenhahn, Weigert (1990)]. These processes can considerably reduce the mass of a star at the end of its evolution. Thus, the initial mass of black hole progenitors could be substantially greater than M_0.

There are different estimates for the minimal mass M_* of a progenitor star that still forms a black hole. For example, $M_* \approx 10 M_\odot$ [Shapiro and Teukolsky (1983)], $M_* \approx 30 M_\odot$ [Lipunov (1987)], or even $M_* \geq 40 M_\odot$ [van den Heuvel and Habets (1984), Schild and Maeder (1985)]. Note that the evolution of stars in close binary systems differs from the evolution of single stars because of mass transfer from one star to another [see Novikov (1974), Masevich and Tutukov (1988), Lipunov, Postnov, and Prokhorov (1996)]. The conclusions about masses of black hole progenitors in this case could be essentially different (see discussion at the end of Section 9.5). In the remaining part of this section we shall focus our attention on the fate of single stars.

One can try to estimate how many black holes have been created by stellar collapse in our Galaxy during its existence. For stars on the "main sequence" in the solar neighborhood, the spectrum of stellar masses at present is roughly known from observational data together with the theory of stellar structure. The lifetime of massive stars is less than 10^{10} years (the age of the Galaxy). We assume a constant birthrate and the same constant deathrate for massive stars during the lifetime of the Galaxy. Now if we suppose that all stars with $M > M_*$ on the "main sequence" (progenitors) must turn into black holes, we can calculate the birthrate of black holes in the solar neighborhood. If this rate is the same everywhere we can estimate the total number of black holes in the Galaxy and the total mass of all stellar black holes in it.

Much work has been devoted to such estimates. However, because of great uncertainties progress since the pioneering works [Zwicky (1958), Schwarzschild (1958), Hoyle and Fowler (1963), Novikov and Ozernoy (1964), and Hoyle *et al.* (1964)] has been very slow. A review of more recent estimates is given by Shapiro and Teukolsky (1983). For rough estimation of the rate of black hole formation in the Galaxy one can use the following relation [see, for example, Novikov (1974)]

$$\frac{dN}{dt} \approx 0.1 \left(\frac{M_*}{3M_\odot}\right)^{-1.4} \mathrm{yr}^{-1}. \tag{9.2.1}$$

At present, we can probably repeat the conclusion of Novikov and Thorne (1973): "For stars with masses greater than ~ 12 to $30 M_\odot$ the (supernova) explosion may produce a black hole. If this tentative conclusion is correct, then no more than ~ 1 per cent of the (visible) mass of the Galaxy should be in the form of black holes today;

and new black holes should be created at a rate not greater than ~ 0.01 per year." Thus, the total number of stellar black holes in the Galaxy may be $N \approx 10^8$ or less. The estimates of Brown and Bethe (1994) give $N \approx 10^9$. It is thus plausible that a large number of black holes exist. If so, what is their observational appearance?

9.3 Stellar Black Holes in the Interstellar Medium

The most important physical process which leads to observable manifestations of a black hole presence is gas accretion [Zel'dovich (1964), Salpeter (1964)]. If a black hole is in a gaseous nebula, the gas will be falling in the gravitational field of the black hole.

Assume that a Schwarzschild black hole is at rest with respect to the gas. We discuss adiabatic accretion first. In the stationary accretion pattern the amount of matter falling upon the black hole per unit time, $\dot{M}$, and other basic parameters are determined by the gas properties and the gravitational field at large distances (greatly exceeding the gravitational radius r_g). At such distances the gravitational field may be treated as a Newtonian one. Properties of the flux close to the black hole will be considered later.

In Newtonian theory the problem was solved by Bondi (1952). Pioneering studies of radial accretion by astrophysical black holes were performed by Shvartsman (1971) and Shapiro (1973a,b). The general scenario is described by Novikov and Thorne (1973) and Shapiro and Teukolsky (1983).

The gas flow is governed by two equations:

- *Conservation of mass*

$$4\pi r^2 \rho u = \dot{M} = \text{const}. \tag{9.3.1}$$

This equation implies that the left-hand side does not depend on r.

- *The Euler equation*

$$u\,\frac{du}{dr} = -\frac{1}{\rho}\frac{dP}{dr} - \frac{GM_{\rm H}}{r^2}. \tag{9.3.2}$$

In the above relations $M_{\rm H}$ is the mass of the black hole; P and ρ are the pressure and mass density of the gas; u is its radial velocity, and $\dot{M}$ is the rate at which the black hole accretes mass. We assume that ρ and P are related by the adiabatic law:

$$P = K\rho^{\Gamma}, \tag{9.3.3}$$

K and Γ are constants. The adiabatic sound velocity is given by $a = (\Gamma P/\rho)^{1/2}$. For the conditions in the accretion flux $\Gamma \approx 1.4$. A solution of the system (9.3.1)–(9.3.3) determines the parameters of the flow. The general picture of accretion is as follows. A crucial role is played by the *sonic radius*:

$$r_s \approx \frac{GM_H}{a_\infty^2}\,, \tag{9.3.4}$$

where a_∞ is the sound velocity at infinity. At $r > r_s$ the gravitation of the black hole practically does not change the parameters of the gas and the velocity of the gas is much smaller than a_∞. At a distance of the order of r_s the velocity u of the gas flow increases and becomes equal to the local sound velocity a. The local value of a at this r is greater than a_∞, but still of the same order of magnitude as a_∞. At distances $r \ll r_s$, the velocity $u \gg a$, and the gas is nearly in free fall. In this region

$$\rho \approx 0.66\,\rho_\infty\,(r_s/r)^{3/2}\,, \tag{9.3.5}$$

$$T \approx 0.76\,T_\infty\,(r_s/r)^{(3/2)(\Gamma-1)}\,. \tag{9.3.6}$$

The numerical coefficients in these relations correspond to $\Gamma \approx 1.4$ near the sonic radius. Here ρ_∞ and T_∞ are the mass density and the temperature at infinity. The rate of accretion is

$$\dot{M} \approx 4\pi\, r_s^2\, \rho_\infty\, a_\infty \approx \left(10^{11}\mathrm{g/s}\right)\left(\frac{M_H}{M_\odot}\right)^2\left(\frac{\rho_\infty}{10^{-24}\mathrm{g/cm^3}}\right)\left(\frac{T_\infty}{10^4 K}\right)^{-3/2}. \tag{9.3.7}$$

So far we have discussed simplified adiabatic hydrodynamic accretion. Now we consider various processes that can complicate the picture. For accretion on a stellar black hole, thermal bremsstrahlung is many orders of magnitude smaller than the luminosity of ordinary stars, and hence is negligible. However, the influence of interstellar magnetic fields on accretion and the radiation from infalling gas can be significant. Estimates [see Bisnovatyi-Kogan (1989)] show that the total luminosity of the accreting black hole due to synchrotron radiation of electrons in magnetic field which is frozen into the falling gas is

$$L_{\mathrm{synch}} \approx \left(3\times 10^{30}\frac{\mathrm{erg}}{\mathrm{s}}\right)\times \begin{cases} \left(\frac{M_H}{M_\odot}\right)^3\left(\frac{\rho_\infty}{10^{-24}\mathrm{g/cm^3}}\right)^2\left(\frac{T_\infty}{10^4 K}\right)^{-3}, & \xi < 1\,, \\ 10\left(\frac{M_H}{M_\odot}\right)^2\left(\frac{\rho_\infty}{10^{-24}\mathrm{g/cm^3}}\right)\left(\frac{T_\infty}{10^4 K}\right)^{-3/2}, & \xi \gg 1\,, \end{cases}$$

where

$$\xi \approx \left(\frac{M_H}{M_\odot}\right)^{1/2}\left(\frac{T_\infty}{10^4 K}\right)^{-3/4}\left(\frac{\rho_\infty}{10^{-24}\mathrm{g/cm^3}}\right). \tag{9.3.8}$$

Most of the radiation comes from the high-temperature, strong-field region near the black hole. For typical conditions the radiation spectrum has a broad maximum at the frequency $\nu \sim 7 \times 10^{14}$ Hz. The full spectrum of the radiation is depicted in the book by Bisnovatyi-Kogan (1989).

An important measure of the accretion rate into a black hole is provided by *Eddington critical luminosity*

$$L_{\rm E} = 4\pi \frac{G M_{\rm H}\, \mu\, m_{\rm p}\, c}{\sigma_{\rm T}} = (1.3 \times 10^{38} {\rm erg/s})\, \mu \left(\frac{M_{\rm H}}{M_\odot}\right). \tag{9.3.9}$$

Here μ is the molecular weight per electron, $m_{\rm p}$ is the rest mass of the proton, and $\sigma_{\rm T}$ is the *Thomson cross section.* This is the luminosity at which the radiation pressure exactly balances the gravitational force of mass $M_{\rm H}$ (for a fully ionized plasma). Eddington (1926) derived this value in a discussion of stellar equilibrium. Zel'dovich and Novikov (1964b) introduced it in the context of black holes in quasars and accretion.

If $L \ll L_E$, gravity dominates, and photon pressure is not important. Comparing expressions $L_{\rm synch}$ and (9.3.9) demonstrates that $L \ll L_{\rm E}$ for the case under consideration.

Some aspects of spherically symmetric accretion were recently reanalyzed by Park (1990a,b). The influence of radiation from the inner regions on the flux far away is discussed by Ostriker *et al.* (1976), Cowie, Ostriker and Stark (1978), Wandel, Yahil and Milgrom (1984), Park (1990a,b).

Near the boundary of the black hole, the gas is practically in a state of free fall. The moment of crossing the horizon is in no way peculiar, as seen by an observer falling together with the gas. For typical conditions, $\rho \simeq (6 \times 10^{-12} {\rm g/cm^3})\,(r_{\rm g}/r)^{3/2}$ and $T \simeq (10^{12} K)\,(r_{\rm g}/r)$. The radiation which reaches infinity from an infalling gas element will, of course, be sharply cut off as the gas element crosses the horizon.

A review of works on possible pair creation and solutions with shocks close to the event horizon was presented by Lamb (1991). Some important details of radial accretion are discussed in the papers by Krolik and London (1983), Colpi, Marachi and Treves (1984), Schultz and Price (1985). In the case of a rotating black hole, the dragging of inertial frames in its vicinity is essential. This dragging swings the accreting gas into orbital rotation around the black hole. This effect is known as *Bardeen-Petterson effect* [Bardeen and Petersen (1975)]. For its discussion, see also Thorne *et al.* (1986). Stationary accretion of gas with zero angular momentum onto a Kerr black hole was considered by Beskin and Pidoprigora (1995), and Beskin and Malishkin (1996).

Spherically symmetric accretion onto a Schwarzschild black hole is probably only of academic interest as a testbed for theoretical ideas. It is of little relevance for interpretations of the observational data. More realistic is the situation where a black hole moves with respect to the interstellar gas. Once again, the main processes that are responsible for the rate of accretion take place far away from the boundary of the black hole, in the regions where one can use the Newtonian approximation. The corresponding problem was analyzed by Hoyle and Littleton (1939), Bondi and

Hoyle (1944), Bondi (1952), Salpeter (1964). For a review see Novikov and Thorne (1973) and Bisnovatyi-Kogan (1989).

When a black hole moves through the interstellar medium with a speed v much greater than the sound speed a_∞, a shock front must develop around the black hole. If the remaining kinetic energy of a gas element behind the shock front exceeds the potential energy in the gravitational field of the black hole, this element will escape the pull of the hole. Otherwise, it will be captured by the black hole. The corresponding accretion rate is

$$\dot{M} \approx \frac{4\pi G^2 M_{\rm H}^2 \rho_\infty}{v^3} \approx (10^{11}\mathrm{g/s}) \left(\frac{M_{\rm H}}{M_\odot}\right)^2 \left(\frac{\rho_\infty}{10^{-24}\mathrm{g/cm}^3}\right) \left(\frac{v}{10\frac{\mathrm{km}}{\mathrm{s}}}\right)^{-3} . \quad (9.3.10)$$

Note that this expression for the accretion rate for supersonic motion, up to a constant of order of 1, coincides with the expression (9.3.7) for the accretion rate on a black hole at rest provided one replaces the speed of the hole v in the former formula by the sound speed a_∞ at infinity and takes into account that $T_\infty \sim a_\infty^2$.

To conclude the discussion of accretion onto isolated stellar black holes, we note that the luminosity of these objects is low, and it is rather difficult to distinguish their observational appearance from some other astrophysical objects at great distances. For this reason, no such black holes have yet been identified. On the possible strategy for searching for isolated black holes of stellar mass in the solar neighborhood see Heckler and Kolb (1996).

9.4 Disk Accretion onto Black Holes

For the purpose of finding and investigating black holes, two specific cases of accretion are of particular importance: accretion in binary systems and accretion onto supermassive black holes that probably reside at the centers of galaxies. In both cases, the accreting gas has specific angular momentum $\tilde{L} \gg r_{\rm g}c$. As a result, the gas elements circle around the black hole in Keplerian orbits, forming a disk or a torus around it. The inner edge of the disk is in the region of the last stable circular orbit. Close to a rotating black hole the Lense-Thirring precession drags the gas around in the equatorial plane [see Bardeen and Petterson (1975)]. Viscosity plays a crucial role in the accretion. It removes angular momentum from each gas element, permitting it to gradually spiral inward toward the black hole. When the gas reaches the inner edge of the disk, it spirals down into the black hole, practically, in a state of free fall. At the same time, the viscosity heats the gas, causing it to radiate. Probable sources of viscosity are turbulence in the gas disk and random magnetic fields. Unfortunately, we are nowhere near a good physical understanding of the effective viscosity (for the discussion of the numerical approach to the problem, see Brandenburg *et al.* (1995, 1996)). Large-scale magnetic fields can also play an important role in the physics of accretion.

The properties of the accreting disk are determined by the rate of gas accretion. It is worth noting the following. In the problem of spherical accretion, which was discussed in Section 9.3, the rate of accretion is determined by solving the equations governing the gas flow. In the case of disk accretion, the accretion rate is an independent (external) parameter. It is determined by the evolution of a binary system or by the conditions in a galactic nucleus. As a result, the rate of accretion can be much higher than in the spherical case, and observational manifestations of such black holes are much more prominent.

In what follows, we use the dimensionless parameter $\dot{m} \equiv \dot{M}/\dot{M}_{\mathrm{E}}$, where $\dot{M}_E$ is the so-called *critical accretion rate*:

$$\dot{M}_{\mathrm{E}} = L_{\mathrm{E}}\, c^{-2}\,, \qquad (9.4.1)$$

determined by the Eddington critical luminosity L_{E} given by (9.3.9). As long as we do not consider the innermost parts of the disk, relativistic effects are not important for the physical processes involved in accretion.

Lynden-Bell (1969) was the first to propose the model of gaseous disk accretion onto a black hole. Shakura (1972), Pringle and Rees (1972), Shakura and Sunyaev (1973) built Newtonian models of accretion disks. Novikov and Thorne (1973) and Page and Thorne (1974) formulated the theory of disk accretion in the framework of general relativity. The first models were rather simple. They focused on the case of moderate rate of accretion $\dot{m} < 1$. Subsequently, theories for $\dot{m} \sim 1$ and $\dot{m} \gg 1$ were developed. They take into account complex processes in radiative plasmas and various types of instabilities. A review is given by Lamb (1991), Abramowicz *et al* (1996), and Narayan (1997). However, these processes have no direct relation (or at least almost no relation) to the specific properties of spacetime in the vicinity of a black hole.[3] Thus, we will not describe them in detail here. We only give some theoretical estimates of observational characteristics of accreting disks around black holes in binary systems and galactic nuclei.

The main source of luminosity for disk accretion is the gravitational energy that is released when gas elements in the disk spiral down. In the simplest models, most of the gravitational energy is released, generating most of the luminosity, from the inner parts of the disk. The total energy radiated by the gas element must be equal to the gravitational binding energy of the element when it is at the last stable circular orbit. For a mass m_* this energy is (see Chapters 2 and 3)

$$E_{\mathrm{bind}} \approx \begin{cases} 0.057\, m_*\, c^2 & \text{for a non-rotating hole}, \\ 0.42\, m_*\, c^2 & \text{for a maximally rotating hole}. \end{cases} \qquad (9.4.2)$$

The total luminosity L of the disk can be obtained if we substitute the accretion rate

[3]See a note at the end of this section.

$\dot{M}$ instead of m_* into this formula. As a result, we have

$$L \approx \begin{cases} \left(3 \times 10^{36}\ \frac{\text{erg}}{\text{s}}\right)\left(\dfrac{\dot{M}}{10^{-9}M_\odot/\text{yr}}\right) & \text{for a non-rotating hole}, \\ \left(3 \times 10^{37}\ \frac{\text{erg}}{\text{s}}\right)\left(\dfrac{\dot{M}}{10^{-9}M_\odot/\text{yr}}\right) & \text{for a maximally rotating hole}. \end{cases} \tag{9.4.3}$$

The accretion rate $\dot{M}$ is an arbitrary external parameter, which is determined by the source of gas (for example, by the flux of gas from the upper atmosphere of the companion star in a binary system). We normalized $\dot{M}$ by the value $\dot{M}_0 = 10^{-9}M_\odot/\text{yr}$ because this is probably the typical rate at which a normal star is dumping gas onto a companion black hole. In this model, the accreting gas is assumed to be relatively cool, with its temperature much less than the virial temperature corresponding to the potential energy in the gravitational field. As estimates show [see Lamb (1991)], a *geometrically thin disk* (with heights $h \ll r$) might be formed under these conditions. This is the so-called *standard disk model* [Shakura and Sunyaev (1973), Novikov and Thorne (1973)]. In this model, the electron and ion temperatures are equal, and the disk is effectively *optically* thick. The temperature of the gas in the inner parts of the disk reaches $T \approx 10^7 - 10^8$K. In this region, electron scattering opacity modifies the emitted spectrum so that it is no longer the black body spectrum. Instead, the total spectrum of the disk radiation is a power law $F \sim \omega^{1/3}$ with an exponential cut off at high frequencies [Shakura and Sunyaev (1973)]. The innermost regions of such "standard" disks are probably unstable. The thin accretion disk model is unable to explain the hard spectra observed in accretion flows around black holes in many observable cases [see Narayan (1997)].

A few types of hot accretion flow models have been proposed. One of them is a model with a hot corona above a standard thin accreting disk (see Haardt and Maraschi (1993), Fabian (1994), and references therein). In a model proposed by Thorne and Price (1975), the ions in the inner region are hot $T_i \approx 10^{11}$K, but the electrons are considerably cooler $T_e \approx 10^9$K. This inner disk is thicker than in the "standard" model and produces most of the X-ray emission. Shapiro, Lightman and Eardley (1976) showed that the models with hot ions and cooler electrons are optically thin.

Further development of the theory of disk accretion led to more sophisticated models. Abramowicz *et al.* (1988, 1989) have demonstrated that when the luminosity reaches the critical one (corresponding to $\dot{m} \equiv \dot{M}/\dot{M}_E$ of the order of unity), radiation pressure in the inner parts of the disk dominates the gas pressure, and the disk is thermally and viscously unstable. For especially big $\dot{m}$, the essential part of the energy of the plasma is lost by advection into the black hole horizon because the radiation is trapped in the accreting gas and is unable to escape [Begelman and Maier (1982), Abramowicz *et al.* (1988)]. This process stabilizes the gas flow against perturbations. Advection can also be important for smaller $\dot{m}$. For high mass-accretion rates, the height of the accretion disk becomes comparable to its radius. In modern models, the radial pressure gradients and the motion of gas elements along radius are taken into account. In the innermost parts of the disk and down to the black hole, the flow of gas is supersonic.

Recently, a new class of optically thin hot disk solutions has been discovered [Narayan and Yi (1994)]. In this model, most of the viscously dissipated energy is advected with the accreting gas, with only a small fraction of the energy being radiated. This occurs because the gas density is so low that the radiative efficiency is very poor. These models were named advection-dominated. They have been applied successfully to a few concrete celestial objects [see Narayan (1997)]. Probably, the future development of the models allows one to use them as "proof" that the accreting objects are indeed black holes [see Narayan, McClintock and Yi (1996)].

In conclusion, we note that in some models of disk accretion electron-positron pair production can be important. A review of these aspects is given by Lamb (1991).

For new developments in the theory of disk accretion, see also Björnsson and Svensson (1991, 1992), Narayan and Yi (1994), Abramowicz *et al.* (1995), and Artemova *et al.* (1996). Many aspects of the physics of accretion disks, including the development of instabilities, were discussed in the works by Mineshige and Wheeler (1989), Wheeler, Soon-Wook Kim and Moscoso (1993), and Moscoso and Wheeler (1994). We believe that new models involving recent developments in plasma physics will play a key role in the modern astrophysics of black holes.

9.5 Evidence for Black Holes in Stellar Binary Systems

Probably the best evidence that black holes exist comes from studies of X-ray binaries. Galactic X-ray sources were discovered by Giacconi *et al.* (1962). Hayakawa and Matsuoko (1964) pointed out that X-rays might be produced by accretion of gas in close binary systems. However, they discussed accretion into the atmosphere of a normal companion star, rather than onto a compact companion. Novikov and Zel'dovich (1966) were the first to point out that accretion onto compact relativistic objects (neutron stars and black holes) in binary systems should produce X-rays. They also inferred that Sco X-1, which had been just discovered, might be a collapsed star in a state of accretion. After that the observational data were analyzed by Shklovsky (1967a,b). Models for X-ray sources have been discussed in some detail by Prendergast and Burbidge (1968). They argue that the gas flow forms a disk around the compact object, with an approximately Keplerian velocity distribution.

A new era started in December 1970 when the X-ray satellite "UHURU" was launched. This satellite has provided much new data about the sources [see Giacconi *et al.* (1972)]. For a description of the early observational evidence for black holes in stellar binary systems, see Novikov (1974). In this section we summarize the best documented observational evidence for black holes in stellar binary systems [for an overview see McClintock (1992), Cowley (1992), Bender *et al.* (1995), Cherepashchuk (1996), and Wheeler(1997)].

The arguments that are used to prove that an X-ray binary contains a black hole are as follows:

1. The X-ray emitting object in a binary system is very compact, and therefore cannot be an ordinary star. Thus, it is either a neutron star or a black hole. This argument comes mainly from analysis of the features of emitted X-rays.
2. Analysis of the observational data allows one to determine the orbital motion in the binary system and makes it possible to obtain the mass of the compact object. The data on the observed velocity of the optical companion star is of the most importance. Note that Newtonian theory is always sufficient for the analysis. The technique of weighing stars in binaries is well known in astronomy. If the mass of the compact component is greater than the maximal possible mass of neutron stars $M_0 \approx 3M_\odot$ (see Section 9.2), then it is a black hole.

It is worth noting that this evidence is somewhat indirect because it does not confront us with the specific relativistic effects that occur near black holes and which are peculiar to black holes alone. However, it is the best that modern astronomy has proposed so far. In spite of these circumstances, we believe that the logic of the arguments is reliable enough.

According to the generally accepted interpretation, we have the necessary observational confirmation only for a few systems at the present time. For these systems we have strong reasons to believe that the compact X-ray emitting companions are black holes. Historically, for a long time three objects (Cyg X-1, LMC X-3, and A0620-00) have been considered as the most probable candidates for black holes. Now (in 1997) there are probably ten such objects. Some characteristics of these leading black hole candidates are summarized in Table 9.1 [according to Cherepashchuk (1996)].

The most plausible masses of the compact objects in these systems are considerably larger than $M_0 \approx 3M_\odot$. Table 9.2 provides estimates of the minimum masses of three historically first candidates. These estimates were obtained by varying methods and have various degrees of reliability. In most cases, even the minimum mass is greater than M_0. Thus, these three objects are strong black hole candidates.

The strongest candidates are those which have a *dynamical lower limit* of the mass of the compact object (or so-called *mass function* $f(M)$) greater than $3M_\odot$. From this point of view, the strongest candidates are GS 2023+338 ($f(M) = 6.5M_\odot$), GS 2000+25 ($f(M) = 5M_\odot$), and XN Oph 1977 ($f(M) = 4M_\odot$).

For the discussion of black hole candidates in binary systems, see McClintock (1992), Cowley (1992), Tutukov and Cherepashchuk (1993), Bender *et al.* (1995), Cherepashchuk (1996), and references therein.

System	Spectral type of the optical companion	Orbital period (days)	Mass of the compact companion (in $M_\odot$)	Mass of the optical companion (in $M_\odot$)	X - ray luminosity (erg/sec)
Cyg X-1 (V 1357 Cyg)	O9.7Iab	5.6	7–18	20–30	$\sim 8 \times 10^{37}$
LMC X-3	B(3-6)II-III	1.7	7–11	3–6	$\sim 4 \times 10^{38}$
LMC X-1	O(7-9)III	4.2	4–10	18–25	$\sim 2 \times 10^{38}$
A0620-00 (V616 Mon)	K(5-7)V	0.3	5–17	~ 0.7	$\leq 10^{38}$
GS 2023+338 (V 404 Cyg)	K0IV	6.5	10–15	0.5–1.0	$\leq 6 \times 10^{38}$
GRS 1121-68 (XN Mus 1991)	K(3-5)V	0.4	9–16	0.7–0.8	$\leq 10^{38}$
GS 2000+25 (QZ Vul)	K(3-7)V	0.3	5.3–8.2	~ 0.7	$\leq 10^{38}$
GRO J0422+32 (XN Per 1992= =V518 Per)	M(0-4)V	0.2	2.5–5.0	~ 0.4	$\leq 10^{38}$
GRO J1655-40 (XN Sco 1994)	F5IV	2.6	4–6	~ 2.3	$\leq 10^{38}$
XN Oph 1977	K3	0.7	5–7	~ 0.8	$\leq 10^{38}$

Table 9.1: Black hole candidates in binary systems [Cherepashchuk (1996)].

Cyg X-1	Ref	LMC X-3	Ref.	A0620-00	Ref.
3	(1,2)	6	(5)	3.2	(11)
7	(3)	2.5	(6)	2.90 ± 0.8	(12)
$3.4\left(\frac{d}{2Kpc}\right)^2$	(4)	2.3	(7)	4.5	(12)
		6	(7)	3.30 ± 0.95	(13)
		3	(8,9)	6.6	(13)
		4	(10)	3.1 ± 0.2	(14)
				3.82 ± 0.24	(15)
				6	(16)
				4.16 ± 0.1	(17)

Table 9.2: Estimates of the minimum mass (in $M_\odot$) for the compact objects. (In estimate (4) d is the distance from the Solar system). Numbers in Ref. refer to the following publications:

1. Webster and Murdin (1972)
2. Bolton (1972a,b; 1975)
3. Gies and Bolton (1986)
4. Paczynski (1974)
5. Paczynski (1983)
6. Mazeh *et al.* (1986)
7. Cowley *et al.* (1983)
8. Kuiper *et al.* (1988)
9. Bochkarev *et al.* (1988)
10. Treves *et al.* (1990)
11. McClintock (1988)
12. McClintock and Remillard (1990)
13. Johnston *et al.* (1989)
14. Johnston and Kulkarni (1990)
15. Haswell and Shafter (1990)
16. Cowley *et al.* (1990)
17. Haswell *et al.* (1993)

The total number of systems that are frequently mentioned as possible candidates for black holes of stellar mass is about 20. All seriously discussed candidates are X-ray sources in binary systems. Some of them are persistent, other are transient. For example, Cyg X-1 and LMC X-1 are persistent, and A0620-00 is transient. Begelman and Rees (1996) summarized the present status as follows: "There is also overwhelming evidence for black holes in our own galaxy, formed when ordinary massive stars die, each weighting a few times as much as the Sun". Most of experts now agree with this unambiguous conclusion.

During the more than 25 years since the discovery of the first black hole candidate Cyg X-1 only a few new candidates have been added. This is in contrast to the rapid increase of the number of identified neutron stars. At present, many hundreds of neutron stars have been identified in the Galaxy. About 100 of them are in binary systems [Lamb (1991)]. One might conclude that black holes in binary systems are exceedingly rare objects. This is not necessarily true, however. The small number of identified black hole candidates may as well be related to the specific conditions which are necessary for their observable manifestation [for a summary, see Cowley (1992)]. According to estimations of van den Heuvel (1983), the evolutionary stage when a black hole binary continuously radiates X-rays may last only 10^4 years. We can thus detect it only during this short period. In effect, the population of black hole binaries may be much larger than what we can presently observe. Such systems may be as common as neutron star binaries. An estimate of the formation rate for black hole binaries has been given by van den Heuvel and Habets (1984), and Lipunov *et al.* (1994).

9.6 Supermassive Black Holes in Galactic Centers

Since the middle of this century astronomers have come across many violent or even catastrophic processes associated with galaxies. These processes are accompanied by powerful releases of energy and are fast not only by astronomical but also by earthly standards. They may last only a few days or even minutes. Most of such processes occur in the central parts of galaxies, the galactic nuclei.

About one per cent of all galactic nuclei eject radio-emitting plasma and gas clouds, and are themselves powerful sources of radiation in the radio, infrared, and especially, the "hard" (short wavelength) ultraviolet, X-ray and gamma regions of the spectrum. In some cases, the full luminosity of the nucleus is $L \approx 10^{47}$erg/s, and it is millions of times the luminosity of the nuclei of quieter galaxies, such as ours. These objects were called *active galactic nuclei* (AGN). Practically all the energy of their activity and of the giant jets released by galaxies originates from the centers of their nuclei.

Quasars form a special subclass of AGN. Their characteristic property is that their total energy release is hundreds of times greater than the combined radiation of all the stars in a large galaxy. At the same time, the average linear dimensions of the radiating regions are small: a mere one-hundred-millionth of the linear size of a galaxy. Quasars are the most powerful energy sources registered in the Universe to date [Blandford (1987), Wallinder (1993), Begelman and Rees (1996), and Peterson (1997)]. What processes are responsible for the extraordinary outbursts of energy from AGN and quasars?

Learning about the nature of these objects involves measuring their sizes and masses. This is not at all easy. The central emitting regions of AGN and quasars are so small that a telescope views them just as point sources of light. Fortunately, quite soon after the discovery of the quasar 3C 273 it was shown that its brightness changes. Sometimes it changes very rapidly, in less than a week. After this discovery, even faster variability (on timescales of a few hours or less) was detected in other galactic nuclei. From these variations one could estimate the dimensions of the central parts of the nuclei that are responsible for the radiation. The conclusion was that these regións are not more than a few light-hours in diameter. That is, they are comparable to the solar system in size.

In spite of the rather small linear dimensions of quasars and many galactic nuclei, their masses turned out to be enormous. They were first estimated by Zel'dovich and Novikov (1964b), who used formula (9.3.9). For quasi-static objects, the luminosity cannot be essentially greater than L_{E}. A comparison of the observed luminosity with the expression (9.3.9) gives an estimate of the lower limit of the central mass. In some quasars this limit is $M \approx (1 - 100) \times 10^7 M_{\odot}$.[4] These estimates are supported by data on the velocities within the galactic nuclei of stars, and gas clouds accelerated in the gravitational fields of the center of the nuclei. We will discuss this in the next section.

Great mass but small linear dimensions prompt the guess that there could be a black hole. Zel'dovich and Novikov (1964b) and Salpeter (1964) suggested that the centers of quasars and AGN could harbor supermassive black holes. This would account for all the extraordinary properties of these objects. A further crucial step was made by Lynden-Bell (1969). He developed and applied a theory of thin accretion disks orbiting around massive black holes. Now it is generally accepted that in AGN there are supermassive black holes with accretion gas (and maybe also dust) disks. One of the most important facts implied by the observations, especially by means of radio telescopes, is the existence of directed jets from the nuclei of some active galaxies. For some of the objects there is evidence that the radio components move away from the nucleus at ultrarelativistic velocities. The existence of an axis of ejection strongly suggests the presence of some stable compact gyroscope, probably

[4]It is interesting to note that the estimates of masses of black holes in AGN: $M \approx 10^7 M_{\odot}$ and more, are close to the estimates of masses of supermassive invisible "stars" in the Universe (black holes in our terminology) which were made at the end of the 18th century by Michell and Laplace. They speculated on supermassive stars, which would generate a gravitational field strong enough to trap light rays, and thus make them invisible.

a rotating black hole. In some cases, one can observe evidence that there is also a precession of this gyroscope. An essential role in the physics of processes in the centers of AGN is probably played by black hole electrodynamics. We discussed this in Chapter 8. Reviews of the physics of AGN are given by Lamb (1991), Svensson (1994), Rees (1996), and Peterson (1997).

In the model of a supermassive black hole with an accretion disk for AGN, one requires sources of fuel - gas or dust. The following sources have been discussed: gas from a nearby galactic companion (the result of interaction between the host galaxy and the companion), interstellar gas of the host galaxy, disruption of stars by high velocity collisions in the vicinity of a black hole, disruption of stars by the tidal field of the black hole and some others [see Shlosman, Begelman, and Franc (1990)]. Excellent reviews of all problems of supermassive black holes in galactic centers, including the problem of their origin and evolution, are given by Rees (1990a,b) and Haehnelt and Rees (1993). See also discussions in Fukugita and Turner (1996).

Clearly, the processes taking place in quasars and other galactic nuclei are still a mystery in many respects. But the suggestion that we are witnessing the activity of a supermassive black hole with an accretion disk seems rather plausible.

9.7 Dynamical Evidence for Black Holes in Galactic Nuclei

So far we have considered supermassive black holes as the most probable explanation for the phenomena associated with the activity of nuclei of some galaxies. Are there more conclusive lines of evidence for the presence of such black holes?

First of all, one can expect that massive black holes should exist, not only in active galactic nuclei but, also in the centers of "normal" galaxies (including nearby galaxies and our own Milky Way) [Rees (1990a)]. They are quiescent because they are now starved of fuel (gas). Observations show that galactic nuclei were more active in the past. Thus, "dead quasars" (massive black holes without fuel) should be common at the present epoch.

How can these black holes be detected? It has been pointed out that black holes produce cusp-like gravitational potentials, and hence they should produce cusp-like density distributions of the stars in the central regions of galaxies. Some authors have argued that the brightness profiles of the central regions of particular galaxies imply that they contain black holes. Kormendy (1993) emphasized that arguments based only on surface brightness profiles are inconclusive. The point is that a high central number density of stars in a core with small radius can be a consequence of dissipation, and a cusp-like profile can be the result of anisotropy of the velocity dispersion of stars. Thus, these properties taken alone are not sufficient evidence for the presence of a black hole.

The reliable way to detect black holes in the galactic nuclei is analogous to the case of black holes in binaries. Namely, one must prove that there is a large dark mass in a small volume, and that it can be nothing other but a black hole. In order to obtain such a proof, we can use arguments based on both stellar kinematics and surface photometry of the galactic nuclei. The mass M inside radius r can be inferred from observed data through the formula [Kormendy (1993)]:

$$M(r) = \frac{v^2 r}{G} + \frac{\sigma_r^2 r}{G}\left[-\frac{d\ln I}{d\ln r} - \frac{d\ln\sigma_r^2}{d\ln r} - \left(1 - \frac{\sigma_\theta^2}{\sigma_r^2}\right) - \left(1 - \frac{\sigma_\phi^2}{\sigma_r^2}\right)\right], \qquad (9.7.1)$$

where I is the brightness, v is the rotation velocity, σ_r, σ_θ, σ_ϕ are the radial and the two tangential components of the velocity dispersion. These values must all be obtained from observations. More complex formulas are used for more sophisticated models.

Now we can consider the mass-to-light ratio M/L (in solar units) as a function of radius. This ratio is well known for different types of stellar populations. As a rule, this ratio is between 1 and 10 for elliptical galaxies and globular clusters (old stellar population dominates there). If for some galaxy the ratio M/L is almost constant at rather large radii (and has a"normal" value between 1 and 10) but rises rapidly (toward values much larger than 10) as one approaches the galactic centre, then there is evidence for a central dark object (probably a black hole).

As an example, consider galaxy NGC 3115 which is at a distance of 9.2Mpc from us [Kormendy and Richstone (1992)]. For this galaxy $M/L \approx 4$ and almost constant over a large range of radii $r > 4''$ (in angular units). This value is normal for a bulge of this type of galaxy. At radii $r < 2''$ the ratio M/L rises rapidly up to $M/L \approx 40$. If this is due to a central dark mass added to a stellar distribution with constant M/L, then $M_{\rm H} = 10^{9.2\pm0.5} M_\odot$.

Is it possible to give another explanation of the large mass-to-light ratio in the central region of a galaxy? We cannot exclude the possibility that a galaxy contains a central compact cluster of dim stars. But it is unlikely. The central density of stars in the galaxy NGC 3115 is not peculiar. It is the same as in the centers of globular clusters. The direct observational data (spectra and colors) of this galaxy do not give any evidence of a dramatic population gradient near the center. Thus, the most plausible conclusion is that there is a central massive black hole.

Unfortunately, it is difficult to detect massive black holes in giant elliptical galaxies with active nuclei, where we are almost sure black holes must exist because we observe their active manifestation [Kormendy (1993)]. The reason for this is a fundamental difference between giant elliptical galaxies (the nuclei of some of them are among the most extreme examples of AGN), dwarf elliptical galaxies, and spiral galaxies. Dwarf elliptical galaxies rotate rapidly, and star velocity dispersions are nearly isotropic. Giant elliptical galaxies do not rotate significantly and they have anisotropic velocities. It is not so easy to model these dispersions. This leads to uncertainties in equation (9.7.1). Furthermore, giant elliptical galaxies have large cores and shallow brightness profiles. Consequently, the projected spectra are dominated by light from large radii, where a black hole has no effect.

The technique described above has been used to search for black holes in galactic nuclei. Another possibility is to observe rotational velocities of gas in the vicinity of the galactic center. So far (the middle of 1997) black hole detections have been reported for the following galaxies: M32, M31, NGC 3377, NGC 4594, The Milky Way, NGC3115, M87, and NGC 4258 [for review, see Kormendy (1993), Miyoshi *et al.* (1995), Kormendy and Richstone (1995)]. Some evidence for a supermassive black hole in NGC 4486B was reported by Kormendy *et al.* (1997). This galaxy is a low-luminosity companion of M87. NGC 4486B as well as M31 has a double nucleus. A possible explanation of this based on the hypothesis about the black holes in the centers of these galaxies is proposed in [Tremaine (1995) and Kormendy *et al.* (1997)].

Special investigations were performed in the case of the galaxy M87 [see Dresler (1989) for a review of earlier works and Lauer *et al.* (1992). This is a giant elliptical galaxy with an active nucleus and a jet from the center. At present, there is secure stellar-dynamical evidence for a black hole with mass $M \approx 3 \times 10^9 M_\odot$ in this galaxy. The Hubble Space Telescope has revealed a rotating disk of gas orbiting the central object in the galaxy [Ford *et al.* (1994), Harms *et al.* (1994)]. The estimated mass of the central object is $M = 2.4 \times 10^9 M_\odot$. The presence of a black hole in M87 is especially important for our understanding of the nature of the central regions of galaxies because in this case we observe also the activity of the "central engine".

Radioastronomical observations of the nucleus of the galaxy NGC 4258 are of special interest [Miyoshi *et al.* (1995)]. Using radio interferometric techniques of observing of maser lines of molecules of water in gas clouds orbiting in the close vicinity of the nucleus, the observers obtained an angular resolution 100 times better that in the case of observations by the Hubble Space Telescope. The spectral resolution is 100 times better as well. According to the interpretation of the observations, the center of NGC 4258 harbors a thin disk which was measured on scales of less than one light-year. The mass of the central object is $3.6 \times 10^7 M_\odot$. In the opinion of Begelman and Rees (1996): "It represents truly overwhelming evidence for a black hole... NGC 4258 is the system for which it is hardest to envisage that the mass comprises anything but a single black hole".

In conclusion, we list in Table 9.3 estimates of masses of black holes in the nuclei of some galaxies [Kormendy (1993), van der Marel (1995), Miyoshi *et al.* (1995)].

Progress in this field is very rapid and in the near future our knowledge about evidence for supermassive black holes in galactic nuclei will become more profound. Finally, we want to mention the possibility of the formation of binary supermassive black hole systems through merging of galaxies [for example, see Rees (1990b), Polnarev and Rees (1994), and Valtonen (1996)]. The radiation of gravitational waves from such a system leads to decay of the orbits of the black holes. Eventually they coalesce. The final asymmetric blast of gravitational radiation may eject the merged

Galaxy	Mass of black hole (in $M_\odot$)
M31	2×10^7
M32	$(2-5) \times 10^6$
Milky Way	3×10^6
NGC 4594	10^9
NGC 3115	10^9
NGC 3377	10^8
M87	2.4×10^9
NGC 4258(M106)	3.6×10^7

Table 9.3: Estimated masses of black holes in some galactic nuclei.

hole from the galaxy. (We will discuss gravitational waves from binary systems in more detail in Section 9.9.)

9.8 Primordial Black Holes

Let us now consider the possible existence of *primordial black holes*. The smaller the mass, the greater is the density to which matter must be compressed in order to create a black hole. Powerful pressure develops at high densities, counteracting the compression. As a result, black holes of mass $M \ll M_\odot$ cannot form in the contemporary Universe. However, the density of matter at the beginning of expansion in the Universe was enormously high. Zel'dovich and Novikov (1967a,b), and then Hawking (1971) hypothesized that black holes could have been produced at the early stages of the cosmological expansion of the Universe. Such black holes are known as *primordial*. Very special conditions are needed for primordial black holes to be formed. Lifshitz (1946) proved that small perturbations in a homogeneous isotropic hot Universe (with an equation of state $p = \varepsilon/3$) cannot produce appreciable inhomogeneities. A hot Universe is stable under small perturbations [see Bisnovatyi-Kogan *et al.* (1980)]. The large deviations from homogeneity that are required for the formation of primordial black holes must exist from the very beginning in the metric describing the Universe (i.e., the gravitational field had to be strongly inhomogeneous) even though the spatial distribution of matter density close to the beginning of the cosmological expansion was very uniform. When the quantity $l = ct$, where t is the time elapsed since the Big Bang, grows in the course of expansion to a value of the order of the linear size of an inhomogeneity of the metric, the formation of a black hole (of the mass contained in the volume l^3 at time t) becomes possible. Formation of black holes with masses substantially smaller than stellar masses is thus possible provided that such holes were created at a sufficiently early stage (see below).

Primordial black holes are of special interest because Hawking quantum evaporation (see Chapter 10) is important for small-mass black holes, and only primordial black holes can have such masses. (Note that quantum evaporation of massive and even supermassive black holes may be of essential importance only for the very distant future of the Universe.)

First of all, the following two questions arise:

1. How large must the deviations from the metric of a homogeneous isotropic model of the Universe be for black holes to be born?
2. What is the behavior of the accretion of the surrounding hot matter onto the created hole and how does the accretion change the mass of the hole?

The second question arises because of a remark made in the pioneering paper of Zel'dovich and Novikov (1967a,b): If a stationary flux of gas onto the black hole builds up, the black hole mass grows at a catastrophically fast rate. But if such stationary accretion does not build up immediately after the black hole is formed, accretion is quite negligible at later stages because the density of the surrounding gas in the expanding Universe falls off very rapidly.

Both questions can be answered via numerical modeling. For the case of spherical symmetry, the required computation was carried out by Nadejin *et al.* (1977, 1978) and Novikov and Polnarev (1979, 1980). The main results of these computations are as follows. The dimensionless amplitude of metric perturbations, $\delta g^{\alpha}_{\beta}$, necessary for the formation of a black hole, is about 0.75–0.9. The uncertainty of the result reflects the dependence on the perturbation profile. Recall that the amplitude of a metric perturbation is independent of time as long as $l = ct$ remains much smaller than the linear size of the perturbed region. If $\delta g^{\alpha}_{\beta}$ is less than 0.75–0.9, the created density perturbations transform into acoustic waves when $l = ct$ has increased to about the size of the perturbation. This answers the first of the questions formulated above.

As for the second question, the computations show that the black hole incorporates 10 % to 15 % of the mass within the scale $l = ct$ at the moment of formation. This means that the accretion of gas onto the newborn black hole cannot become catastrophic. Computations confirm that the gas falling into the black hole from the surrounding space only slightly increases the mass.

If in the early history of the Universe there were periods when pressure was reduced for a while, then pressure effects were not important. Formation of primordial black holes (PBHs) under these conditions was discussed by Khłopov and Polnarev (1980) and Polnarev and Khlopov (1981). Other mechanisms of primordial black hole formation include cosmic phase transitions [Sato *et al.* (1981), Maeda *et al.* (1982), Kodama *et al.* (1982), Hawking *et al.* (1982), Kardashev and Novikov (1983), Naselsky and Polnarev (1985), Hsu (1990)] and collapse of loops of cosmic strings [Hawking (1989), Polnarev and Zembowicz (1991), Polnarev (1994)]. A review of the problem is given by Carr (1992).

Hawking's discovery (1974) that black holes can evaporate by thermal emission (see Chapter 10) made the study of PBH formation and evaporation of considerable astrophysical interest. The evaporation effects of small black holes are potentially observable. Their absence from observational searches therefore enables powerful limits to be placed upon the structure of the very early Universe [Zel'dovich and Novikov (1967a,b), Carr and Hawking (1974), Novikov *et al.* (1979), Carr (1983), Carr *et al.* (1994)].

For stellar and supermassive black holes, Hawking radiation is negligible, but for sufficiently small PBHs, it becomes a dominant factor in the black hole evolution and very important for their possible observational manifestations. Using the formulas of Chapter 10,[5] it is easy to conclude that PBHs of $M \approx 5 \times 10^{14}$g or less would have evaporated entirely in the 10^{10} years or so of the Universe's history. PBHs initially a little more massive than this will still be evaporating in the present-day Universe. The rate of their evaporation is large enough so that the stream of energetic particles and radiation they would emit can be turned into a strong observational limit on their presence.

Searches for PBHs attempt to detect either a diffuse photon (or some other particle) background from a distribution of PBHs, or the final emission stage of individual black holes. Using the theoretical spectra of particles and radiation emitted by evaporating black holes of different masses (see Chapter 10), one can calculate the theoretical backgrounds of photons and other particles produced by a distribution of PBHs emitting over the lifetime of the Universe. The level of this background depends on the integrated density of PBHs with initial masses in the considered range.

A comparison of the theoretical estimates with the observational cosmic ray and γ-ray backgrounds places an upper limit on the integrated density of PBHs with initial masses in this range. According to estimates of MacGibbon and Carr (1991), this limit corresponds to $\approx 10^{-6}$ of the integrated mass density of visible matter in the Universe (matter in the visible galaxies). The comparison of the theory with other observational data gives weaker limits [for reviews see Halzen *et al.* (1991), Carr (1992), Coyne (1993)].

The search for high energy gamma-ray bursts as direct manifestations of the final emission of evaporating (exploding) individual PBHs has continued for more than 20 years. No positive evidence for the existence of PBHs has been reported [see Cline and Hong (1992, 1994)]. For future searches see Heckler (1997).

[5]See a footnote on page 319.

A population of PBHs whose influence is small today may have been more important in the earlier epochs of the evolution of the Universe. Radiation from PBHs could perturb the usual picture of cosmological nucleosynthesis, distort the microwave background, and produce too much entropy in relation to the matter density of the Universe. As we mentioned above, limits on the density of PBHs, now or at earlier times, can be used to provide information on the homogeneity and isotropy of the very early Universe, when the black holes were formed. For reviews see Novikov *et al.* (1979) and Carr *et al.* (1994).

The final state of the black hole evaporation is still unclear (see Chapter 15). There is a possibility that the endpoint of the black hole evaporation is a stable relic [Markov (1965, 1966)]. The possible role of such relics in cosmology was first discussed by MacGibbon (1987). For a recent review see Barrow *et al.* (1992).

In this chapter we have reviewed the current search for evidence for black holes in the Universe. The conclusion is the following. Now, in 1997, we are almost 100 percent certain that black holes of stellar masses exist in binary systems. We can probably say the same about supermassive black holes in the centers of many galaxies. So far there is no evidence for the existence of PBHs in the Universe, however.

9.9 Black Holes and Gravitational Wave Astronomy[6]

As should be clear from the previous sections of this chapter, we have good reasons to believe that black holes exist in the Universe. But some caution is perhaps still in order since all the evidence discussed above is indirect. The main reason for this is simple: In most situations, a black hole will be obscured by gas clouds, and the region close to the event horizon will not be transparent to electromagnetic waves. Any photon that reaches us will have been scattered many times as it escapes from the strong-field regime. Only when the black hole is surrounded by a thin accretion disk, can we hope to get a clear view of the horizon region. Does this mean that we typically face a hopeless situation, and that we will rarely be able to observe black holes directly? The answer to this question is no, but the main hope for such observations rests on the future identification of gravitational waves.

That gravitational waves exist is an unambiguous prediction of general relativity. In contrast to electromagnetic waves that are mainly generated in processes involving individual atoms, gravitational waves are created by bulk motion. The strongest gravitational waves will come from processes involving massive and compact objects. Hence, black holes will be the strongest sources of gravitational waves in the Universe. Most dynamical processes involving a black hole should give rise to very characteristic gravitational waves (see the discussion in Chapter 4). Furthermore, the gravitational

[6]This section is written jointly with N.Andersson.

waves interact weakly with matter so they can escape the region close to a black hole almost unaffected by a possible accretion disk. Thus, the gravitational waves provide an ideal probe of the strong field regime of general relativity.

Decades of efforts to detect the gravitational waves that should (theoretically) bathe the Earth have been crowned with very little success, however. Most likely, this is because the gravitational waves that reach us are incredibly weak. Typical waves are expected to induce a relative displacement $h \sim \Delta L/L \sim 10^{-21}$ in matter. But there is great hope that observations will soon become possible, and that new large-scale interferometric detectors will make gravitational-wave astronomy a reality. Such large-scale detectors are presently under construction in both the USA (the LIGO project) and Europe (the French-Italian VIRGO collaboration and the German-British GEO effort) [for a review, see Thorne (1994a, 1995, 1997c)]. At the same time, there are advanced plans for a similar, but much larger, detector in space. This project – which has been given the acronym LISA – could lead to observations of the supermassive black holes that seem to reside in galactic nuclei.

Binary systems consisting of compact objects (black holes or neutron stars) provide the most promising sources for detectable gravitational waves. In fact, we already know that the behaviour of one such system – the binary pulsar PSR 1913+16 – is in impressive agreement with the predictions of Einstein's theory [Taylor (1994)]. This convinces us that gravitational waves exist. As a binary system loses energy due to the emission of gravitational waves, the orbital separation shrinks. Finally, when the separation is of the order of a few times the total mass of the system (in relativistic units), the two bodies will merge. In the case of PSR 1913+16, this will happen some 10^8 years into the future. At some stage of the evolution the frequency of the gravitational waves produced by the system will become larger than 10 Hz. Then the waves enter the frequency bandwidth where the large-scale interferometers will be sensitive and should, in principle, become detectable. For a binary system which consists of two black holes with masses of ten solar mass each, this comprises the last few seconds of the binary evolution.

There are several reasons why inspiraling binaries are considered as promising sources for gravitational-wave detection. It is estimated that the new detectors will be able to see black hole binaries out to the edge of the Universe, to distances with the redshift $Z \sim 1$. In that volume of space there should be several events per year. Furthermore, compact binaries are expected to be very clean systems, and spend many revolutions at a given orbital separation. This makes them "easily" modeled. For example, it has been estimated that tidal effects are negligible up to the point where the bodies are about to merge. It is also believed that radiation reaction will circularize the system before it enters the detector bandwidth. This means that, for the purpose of wave-detection, the system can be accurately modeled as two point masses with a number of internal properties (mass, spin, quadrupole moment etcetera) in adiabatically shrinking circular orbits.

For the main part of the binary inspiral, one need not specify the constituents of the binary. The signal from a binary containing at least one black hole will initially be indistinguishable from the radiation from a pair of neutron stars. The difference between neutron stars and black holes becomes important only in the final stages. For a direct identification of a black hole one must consequently be able to model the last few cycles (when the distance between the bodies becomes comparable to the size of a black hole) of the inspiral very accurately. But from the observational data for the earlier part of the inspiral one should be able to infer the individual masses (as one has been able to do in the case of PSR 1913+16). If one of the masses (or both) is larger than the maximum allowed mass for a neutron star, it seems likely that it is a black hole. The gravitational waves from the late stages of evolution of a black hole binary will probe the strong gravitational field close to the event horizon. Such observations will put our understanding of black hole physics to the test.

A simple estimate for the amplitude of gravitational waves from a binary system can be written as

$$h \approx \frac{r_{g1} r_{g2}}{dD} , \tag{9.9.1}$$

where r_{gi} are the Schwarzschild radii of the two bodies; d is the distance between them, and D represents the distance to the observer. For a system of two black holes of ten solar mass each in the Virgo cluster, typical values would be $r_{g1} = r_{g2} \approx 30$ km, $d \approx 200$ km and $D \approx 3 \times 10^{20}$ km, leading to a gravitational-wave strain $h \approx 1.5 \times 10^{-20}$ on Earth.

It is clear that the problem of observing gravitational waves provides a tremendous challenge. In part this is a challenge for theoretical astrophysics. To detect a gravitational-wave signal, and extract astrophysical data from it, one must integrate the detector output against a theoretical template. Although one will not require very accurate theoretical predictions to *detect* the signal in the first place, more detailed information can be *extracted* through this process by using better theoretical waveforms. It is thus imperative that we have a detailed understanding of inspiraling binary systems.

For a system of two neutron stars one can estimate that the gravitational-wave signal undergoes some 10^4 cycles as it sweeps through the frequency bandwidth of an interferometric detector. At the same time, the orbital velocity v/c ranges from approximately 0.1 to 0.4 of the velocity of light. For a binary of black holes of ten solar mass, the number of cycles is a few hundred, and the final velocities of black holes are relativistic. Thus, it is clear that satisfactory accuracy cannot be achieved unless relativistic effects are incorporated in the calculations. But how accurate must a theoretical waveform be if the overlap integral is not to be seriously depleted? Unfortunately, there is a significant reduction as soon as the theoretical template goes a single cycle out of phase. Thus, we need an accuracy in the phase of order 10^{-4}. This translates into a need for relativistic corrections to at least of order $(v/c)^4$.

How can such corrections be calculated? Much effort has gone into slow-motion approximations of the equations of general relativity. The basic idea is that if v/c

is a typical velocity inside the matter source, then $v/c \ll 1$. The approximation that follows is usually referred to as "post-Newtonian" theory [Will (1994)]. To date, post-Newtonian waveform calculations have been carried out through second post-Newtonian order $O((v/c)^4)$ beyond the standard quadrupole formula [Blanchet *et al.* (1995)]. Since a discussion of these calculations is beyond the scope of this book, we refer the interested reader to Will (1994) for further details.

An alternative to full post-Newtonian studies is provided by black hole perturbation theory (see Chapter 4 for further discussion). At present, such calculations have been carried out to order $(v/c)^8$ [Tagoshi and Nakamura (1994), Tagoshi and Sasaki (1994)]. The main conclusions are as follows: The convergence of the post-Newtonian expansion is very poor [Cutler, Finn, Poisson and Sussman (1993), Poisson (1995)]. The existence of the black hole event horizon will not affect the outgoing radiation until very high order; $O((v/c)^{18})$ [Sasaki (1994)], and black hole absorption will be an effect of order $(v/c)^8$ [Poisson and Sasaki (1995)]. Consequently, effects associated with the horizon are expected to be irrelevant for gravitational-wave observations of inspiraling binaries.

On the other hand, the presence of the horizon will be important at the late stages, especially when the two bodies merge. The burst of gravitational waves emitted when the two bodies collide should be detectable, but these waves are much harder to model theoretically than those from the inspiral phase. For studies of the merger one must resort to numerical simulations. So far, only the head-on collision of two, initially static black holes has been considered in detail, but considerable effort is directed towards the general problem of inspiraling binaries (see Section 7.7). In a general scheme a "post-Newtonian" approximation will provide initial data for a fully relativistic simulation of the final stage of a binary systems evolution.

Part II

Further Developments

Chapter 10

Quantum Particle Creation by Black Holes

10.1 Quantum Effects in Black Holes

10.1.1 Introduction

So far, our description has completely ignored the quantum aspects of the interaction of matter and physical fields with black holes. Quantum effects are, indeed, insignificant for black holes of about a solar mass (and greater). However, for a small-mass (primordial) black hole these effects cannot be ignored; in fact, they qualitatively change the black hole's evolution. They are also likely to become important for those regions inside a black hole where the classical theory predicts a spacetime singularity.

From a general point of view the application of quantum ideas to gravity is the subject of *quantum gravity*, a branch of physics whose development is far from complete. Strictly speaking, we have at present no self-consistent theory that allows us to describe all quantum aspects of black holes. Nevertheless, fortunately as often happens in physics, there exist special regimes and approximations where the theory provides us with reliable predictions.

When discussing quantum effects in black holes, it is convenient to split these into the following three categories:

- Particle creation
- Vacuum polarization
- Quantum fluctuations of the metric

Considering *particle creation*, one can assume that the gravitational field of a black hole is fixed and is classical so that this type of effects is described by the quantum theory of fields propagating in a given external background. The *vacuum polarization* effects imply that the Einstein equations are modified due to quantum corrections, which can be taken into account by a perturbation expansion in a series of powers of the Planck constant $\hbar$. For a black hole of mass much greater than the *Planck mass*

$$m_{\rm Pl} = \sqrt{\frac{\hbar c}{G}} \approx 2.177 \times 10^{-5} {\rm g}, \qquad (10.1.1)$$

the curvature outside the black hole is much smaller than the *Planck curvature*

$$\mathcal{R}_{\rm Pl} = \frac{c^3}{\hbar G} \approx 3.829 \times 10^{65} {\rm cm}^{-2}. \qquad (10.1.2)$$

Under these conditions, it is often sufficient to consider only the lowest order quantum corrections and to use the so-called *semi-classical approximation.* The effects of *quantum fluctuations of the metric* are important when the curvature of spacetime becomes comparable with $\mathcal{R}_{\rm Pl}$, or when one considers spatial regions of size comparable with the *Planck length*

$$l_{\rm Pl} = \sqrt{\frac{\hbar G}{c^3}} \approx 1.616 \times 10^{-33} {\rm cm} \qquad (10.1.3)$$

and time intervals comparable with the *Planck time*

$$t_{\rm Pl} = \sqrt{\frac{\hbar G}{c^5}} \approx 5.591 \times 10^{-44} {\rm s}\,. \qquad (10.1.4)$$

These effects will be important for studying the structure of singularities inside the black hole or final states of quantum black hole evaporation.

For quantum particle creation by an external classical field the required theory is sufficiently well developed. The semi-classical approximation that takes into account the backreaction of quantum effects on the gravitational field of a black hole is also a basically well posed problem. But technically it is quite complicated so until now we have only partial results concerning the self-consistent description of the quantum evolution of a black hole. The physics of very strong gravitational fields at Planck curvatures and quantum fluctuations of the metric at the Planck scales are still open problems. Both of these problems require knowledge of physics at Planck scales,which is now one of the greatest challenges in theoretical physics. For this reason, our knowledge of the final states of an evaporating black hole and the structure of spacetime in the black hole interior is only fragmentary, and is based on a number of additional assumptions and hypotheses.

In this and subsequent chapters, we shall discuss the role of quantum effects in black hole physics. Particle creation by black holes is the subject of the present chapter, while vacuum polarization and quantum fluctuations of the metric, as well as the application of quantum mechanics to black holes as total objects, are considered in the next chapter. But before developing the formalism, we shall make a few general remarks and order-of-magnitude estimates. One can regard this section as an introduction for pedestrians into the "quantum physics of black holes".[1] For convenience, we preserve and explicitly display the fundamental constants c, $\hbar$, and G in all expressions throughout this section.

10.1.2 Particle creation

Current quantum concepts lead to the view that the physical vacuum (i.e., the state without real particles) is a quite complex entity [see, e.g., Aitchison (1985)]. Virtual ("short-lived") particles are constantly created, interact with one another, and are then annihilated in this vacuum. Usually the vacuum is stable in the absence of external fields; that is, the processes considered here never produce real ("long-lived") particles. However, in an external field, some virtual particles may acquire sufficient energy to become real. The result is in effect the quantum creation of particles from the vacuum by an external field.

Let us estimate the probability of particle creation in a static external (not necessarily gravitational) field. Let Γ be the field strength and g be the charge of the particles that are created. For particles of a virtual pair the probability to find one particle at the distance l from another is proportional to $\exp(-l/\lambda_m)$, where $\lambda_m = \hbar/mc$ is the Compton length for a particle of mass m. In other words, the probability is exponentially small at distances greater than λ_m.[2] The same exponential factor enters the expression for the probability of creating a pair of real particles. The probability amplitude of a pair creation by a static field contains a δ-function reflecting the conservation of energy. The amplitude is non-vanishing only if the separation l is such that the work $g\Gamma l$ done over this distance by the field equals $2mc^2$. This implies for the probability of pair creation the following estimate: $w \sim \exp(-2m^2c^3/\hbar g\Gamma)$. This estimate can be made more accurate if we realize that the process we are considering is a tunneling process, and the quantity that stands in the exponent is nothing but the Euclidean action calculated for this tunneling "motion" of a particle. For electrically charged particles this results in an additional factor $\pi/2$ in the exponent. Thus, we have for the probability w of particle creation in a field of strength Γ the expression

$$w = A\exp(-\beta\pi m^2c^3/\hbar g\Gamma), \tag{10.1.5}$$

where the constant β (a dimensionless constant of order unity) and the pre-exponential factor A depend on subtler characteristics of the field.

A well-known example of the creation of particles in an external field is the production of electron-positron pairs by a strong electric field. The following expression

[1] See also Frolov(1976b, 1978b, 1983b).

[2] This conclusion follows directly from the behavior of the propagators for massive particles for spacelike point separation.

derived by Schwinger (1951) holds for the rate of particle production by a uniform electric field in unit volume per unit time:

$$\frac{dN}{dt\,dV} = \frac{e^2E^2}{\pi^2\hbar^2 c}\sum_{n=1}^{\infty}\frac{1}{n^2}\exp(-\pi m^2c^3\,n/eE\,\hbar)\,, \tag{10.1.6}$$

where E is the electric field strength. It is readily seen that if the field strength is much less than the *critical electric field* $E_{\rm cr} = \pi m^2c^3/e\,\hbar$, this formula agrees with (10.1.5), the numerical coefficient β being equal to 1.

If we naively apply the relation (10.1.5) to particle creation by a static gravitational field, we can make an important observation. Namely, in the gravitational interaction the role of the charge is played by the mass of the particle. This means that the probability of creation of a particle by a static gravitational field can be represented in the form

$$w = A\exp(-m\,c^2/\theta), \quad \text{with} \;\; \theta = \frac{\hbar\Gamma}{\beta\pi c}. \tag{10.1.7}$$

In other words, the probability of creating particles of mass m has a Boltzmann-type behavior with an effective temperature (written in energy units) equal to θ.

Consider now the possibility of particle creation by a static gravitational field. We know that the energy of a particle in such a field is $E = -p_\mu\,\xi^\mu$, where p^μ is the four-momentum of the particle, and ξ^μ is the Killing vector. For massive (massless) particles the momentum p^μ is a future-directed timelike (null) vector. For this reason, the energy E of a particle is always positive in the regions where the Killing vector is future directed and timelike. If both of the created particles were created in such a region, their total energy would not vanish so that this process is forbidden by energy conservation. This means that a static gravitational field can create particles only if there exists a region where the Killing vector (which is future-directed and timelike at asymptotic infinity) becomes spacelike. This region lies inside the surface where $\xi^2 = 0$, i.e., inside the Killing horizon. But we know that a Killing horizon in a static spacetime is necessarily the event horizon (see Section 6.3). These simple arguments show that one can expect particle creation in a static spacetime only if it contains a black hole. Similar arguments apply with the same conclusion in the more general case when the spacetime is stationary.

Hawking (1974, 1975) proved that a vacuum is unstable in the presence of a black hole. Moreover, he showed that a black hole creates and emits particles as if it were a black body heated to the temperature

$$T_H = \hbar\,\kappa/2\pi c\,k\,, \tag{10.1.8}$$

where κ is the surface gravity that characterizes the "strength" of the gravitational field in the vicinity of the black hole surface.[3] The Boltzmann constant k is introduced

[3]Strictly speaking, the radiation spectrum of a black hole deviates from the black body spectrum owing to the effects of scattering on the gravitational field. We discuss this in detail later (see Section 10.4).

in order to have the usual units for temperature.

For a Schwarzschild black hole, $\kappa = c^4/4GM$, and the expression for the *Hawking temperature* is

$$T_H = \frac{\hbar c^3}{8\pi G M k}. \tag{10.1.9}$$

This is one of the most beautiful relations in theoretical physics since it contains all the fundamental constants. The Hawking result implies that if one neglects the scattering of created particles by the gravitational field of the black hole, then the probability of creation of a particle with energy E measured at infinity is

$$w \sim \exp(-E/k T_H). \tag{10.1.10}$$

The previous estimate (10.1.7) is consistent with this result provided one takes the strength of the field Γ to coincide with the surface gravity of the black hole, and choose $\beta = 2$.

The Hawking effect, although similar to effects of particle creation in an electric fields, nevertheless has a very important difference. Namely, since the states of negative energy are located inside the black hole, only one of the created particles can appear outside the hole and reach a distant observer. This means that as a result of such a processes, an observer outside the black hole at late times has access to only a part of the total quantum system. As a consequence, the state of particles outside the black hole is always described by a *density matrix* even if initially (before the formation of the black hole) the state of the quantum system was pure. This conclusion has far-reaching consequences, which we shall discuss in Chapter 15.

The particles of Hawking radiation are created uniformly in time, so that some time after the black hole formation the flux of these particles is constant. At this late time the gravitational field of the black hole is static or stationary.[4] That is why estimates based on the stationary field approach gave correct results.

There exists another approach to the problem of particle creation by a black hole. Of course, it gives the same result, but in this approach the role of *zero-point fluctuations* becomes more transparent [see, e.g., Gerlach (1976)].

Let us consider the process of formation of a spherically-symmetric black hole which is schematically shown in Figure 10.1. Before the formation of the black hole the spacetime is practically flat. Let us assume that the initial state of the quantum field is the *vacuum* in the standard Minkowski spacetime. In order to quantize the field (say an electromagnetic one), it is convenient to decompose it into modes, i.e., freely propagating (normalized) waves. One can characterize a mode i by its frequency

[4] In our consideration we neglect the change of the parameters of the black hole due to its quantum radiation. We discuss these effects later. Now note only that the corresponding time-scale for an appreciable change of the black hole parameters is large when the mass of the black hole is much greater than the Planck mass. For such black holes the parameters change so slowly that we can neglect these effects in our considerations.

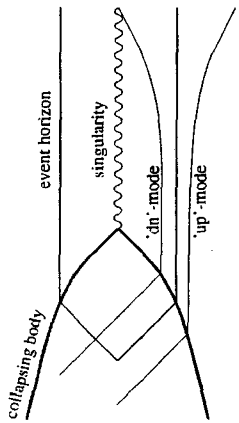


Figure 10.1: Propagation of zero-point fluctuations in a black hole spacetime.

ω_i so that the time-dependent amplitudes $a_i(t)$ of the mode decomposition obey the oscillator-like equation

$$\ddot{a}_i + \omega_i^2\, a_i = 0\,. \tag{10.1.11}$$

In the process of quantization the mode amplitudes a_i become operators. One may say that each of the field modes behaves like a *quantum oscillator.* In the vacuum state the average expectation value of the quantized field vanishes, but because of the uncertainty relations, the averages of quantities quadratic in amplitudes do not vanish. In other words, there exist fluctuations of the quantized field near its zero averaged value. These fluctuations are known as *zero-point fluctuations.* As a result of these fluctuations, the vacuum energy is

$$E = \sum_i \frac{1}{2}\,\hbar\,\omega_i\,, \tag{10.1.12}$$

where the summation is taken over all modes enumerated by index i, and ω_i is the frequency of the mode i. In the presence of real quanta this expression is modified to

$$E = \sum_i \left(n_i + \frac{1}{2}\right)\hbar\omega_i\,, \tag{10.1.13}$$

where n_i is the number of quanta in the state i. Since the number of different modes is infinite, both of the expressions (10.1.12) and (10.1.13) are divergent. In order to obtain a finite (observable) quantity, the corresponding relations are renormalized by subtracting the vacuum energy (10.1.12), i.e., imposing the condition that the energy of the ground (vacuum) state in Minkowski spacetime vanishes.

In the presence of a time-dependent external fields the frequencies ω_i become time-dependent, and an analogue of the effect of *parametric excitation* is possible. This effect occurs if the Fourier transform of the time-dependent part of $\omega_i(t)$ contains frequencies close to the frequency ω_i itself. Under this condition, the initial amplitude of the oscillator can be substantially magnified. In quantum theory, this is interpreted as particle creation. The particle creation is exponentially suppressed if the condition of resonance is not satisfied.

Now we return to Figure 10.1 and apply these arguments to the problem of particle creation by a black hole. Consider a wavepacket which originally has frequency ω and which propagates in the gravitational field of a collapsing body. We assume that it passes through the collapsing body and crosses its boundary at the moment when the boundary is close to the horizon (we denote this packet as an "up"-mode). Because of the interaction with the gravitational field, the frequency of this mode after it passes through the collapsing body is redshifted and becomes $\omega' \sim \omega \exp(-\kappa u/c)$, where κ is the surface gravity of the black hole, and u is the instant of retarded time when the packet reaches a distant observer located in the black hole exterior. (For simplicity, we assume that the quantum field is massless and its quanta propagate with the velocity of light).

The gravitational field of the collapsing matter which forms the black hole is time-dependent. The characteristic time scale for the change of the field is c/κ. It can be shown that the modes with $\omega' \sim \kappa/c$ obey the condition of parametric resonance, and hence are amplified. Even if there were no real particles in this mode initially, real particles would be present in the final state. For the particular mode chosen the created particles appear at the instant of retarded time $u = (c/\kappa)\ln(\omega c/\kappa)$. But since in the initial Minkowski space there exist zero-point fluctuations of arbitrarily high frequencies, the created particles are emitted by the black hole at arbitrarily late moments of the retarded time u, so that a black hole is a permanent source of quantum radiation. The created particles have characteristic frequency κ/c. For modes with outgoing frequencies $\omega' \gg \kappa/c$ the resonance condition is not satisfied, and the parametric excitation of these modes is exponentially suppressed. In particular, this holds for all the modes of massive particles with $m \gg \hbar\kappa/c^3$.

It is also possible to show that the flux of created particles is constant in time. In fact, the number dN of zero-point modes with the initial frequency ω falling in during the interval dv of advanced time is of order $dN \sim \omega\, dv$. The frequency and the value of the retarded time u for the outgoing mode can be found from relations

$$\omega' \sim \omega \exp(-\kappa u/c)\,, \qquad v \sim (c/\kappa)\exp(-\kappa u/c)\,. \tag{10.1.14}$$

If we take into account that creation takes place when $\omega' \sim \kappa/c$, we can conclude that the number of created particles during interval du of retarded time u is $dN \sim (\kappa/c)\, du$. This means that the rate of creation of particles $dN/du \sim \kappa/c$ is constant in time. The above consideration was for massless particles, but a similar conclusion remains valid for massive particles also.

The above arguments might be helpful but certainly they are not rigorous. In the next section we describe rigorous calculations of particle production by black holes and make precise the conclusions drawn above. We shall see that the general picture presented here is qualitatively true, but it requires additional specifications. In particular, we shall show that the single-oscillator picture considered here is oversimplified. In reality mode decomposition of the field equations in the spacetime of a black hole allows one to reduce the problem to a system of three coupled oscillators. If one neglects the scattering of modes by the gravitational field, this system reduces to a system for two coupled oscillators. Basically, the main conclusion concerning particle creation remains the same, with the only difference being that particles are created in pairs so that the "up"-mode particle created in the black hole exterior is always accompanied by the "dn"-mode particle, propagating inside the black hole (see Figure 10.1). This is exactly the same picture as considered before, when we used the static field approach. The remarkable fact is that both (static and time-dependent) approaches for the black hole problem result in the same conclusions, while for other physical systems these regimes possess quite different properties. This happens because particle creation in black holes takes place in a region that is located close to the event horizon, which is a null surface.

Although the discovery by Hawking (1974, 1975) of the effect of quantum evaporation of black holes was quite unexpected for specialists working in black hole physics, at the time of the discovery there were already publications where the potential importance of quantum effects had been discussed.

In 1970 Markov and Frolov considered charged-particle creation by electrically charged black holes. It was the first paper where the importance of quantum effects in black hole physics was emphasized. It was demonstrated that quantum creation of pairs of charged particles in the field of a charged black hole reduces its electric charge practically to zero. If the potential on the black hole surface, Q/r_+, is sufficiently high for pair production ($eQ/r_+ > mc^2$, where m is the electron mass, and e is its charge), and the charge of the black hole exceeds $\hbar c/e$, the conditions under which the uniform field approximation is valid are met; hence, Schwinger's relation (10.1.6) can be used to estimate the rate of production of charged particles by the field of a charged black hole.

A similar phenomenon of quantum production of particles [see Zel'dovich (1971, 1972, 1973), Starobinsky (1973), Misner (1972), Unruh (1974)] takes place in the gravitational field of a rotating black hole. Recall that superradiance, discussed in Section 7.1, is a purely classical phenomenon. This is evident from the fact that the enhancement coefficient does not depend on Planck constant. Like any other classical process, superradiance can be described in quantum terms. In these terms, the phenomenon consists of an increase in the number of quanta in the reflected wave as compared with that in the incident wave. Indeed, in classical terms, the energy of a wave of a fixed frequency is proportional to the square of its amplitude; in quantum terms, it is proportional to the number of quanta. Correspondingly, the enhanced amplitude of the wave of constant frequency signifies an increased number of quanta of the field.

A quantum analogue can be found for the classical phenomenon of superradiance: spontaneous creation of particles from the vacuum in the gravitational field of a rotating black hole. The amplitude of those modes of zero-point fluctuations that meet the enhancement (superradiance) condition grows continuously; this growth manifests itself in the creation of real quanta of the field.

The effect of the production of quanta in the field of a rotating black hole can be described in a somewhat different manner which brings out more clearly the role of the ergosphere. For the creation of a real particle which is ejected from a black hole without violating the law of energy conservation, it is necessary for the second particle of the virtual pair to gain negative energy. This becomes possible if this particle is in the ergosphere and possesses a certain value of angular momentum.

The work necessary for the transformation of a virtual particle into a real one is done by the gravitational field of the black hole. The real particles in superradiant modes leaving the black hole necessarily possess an angular momentum whose direction coincides with that of the black hole. The result is a flux of particles outside the rotating black hole which carries away its energy and angular momentum. The characteristic frequency of this radiation is of order of the angular velocity Ω^H of the black hole rotation, and the total energy and momentum fluxes are

$$-\frac{dE}{dt} \sim \hbar\left(\Omega^H\right)^2, \qquad -\frac{dJ}{dt} \sim \hbar\Omega^H, \tag{10.1.15}$$

The maximum angular velocity for a given mass M, $\Omega^H = c^3/2GM$, is achieved for $J = GM^2/c$, when the black hole is extremal. The rates of energy and angular momentum dissipation by a rotating black hole of mass M due to spontaneous production of particles in its field do not exceed the following values:[5]

$$-\frac{dE}{dt} \sim \hbar\frac{c^6}{G^2M^2} \sim 10^{-17}\,\frac{\text{erg}}{\text{s}}\left(\frac{M_\odot}{M}\right)^2, \tag{10.1.16}$$

$$-\frac{dJ}{dt} \sim \hbar\frac{c^3}{GM} \sim 10^{-22}\,\text{erg}\left(\frac{M_\odot}{M}\right). \tag{10.1.17}$$

These estimates show that the effect is appreciable only for black holes of small mass (much smaller than the solar mass). Note that the above formulas are valid for the production of massless particles (photons, neutrinos, gravitons); massive particles are born at much lower rates.

---

[5]Note that the relations (10.1.15)–(10.1.17), based on dimensional analysis, give a rough estimate only. The properties of the radiation from rotating black holes are discussed in Sections 10.4 and 10.5. Suffice it to remark here that the intensity of this radiation essentially depends on the particle spin. The values of dE/dt and dJ/dt for gravitons, for $J = GM^2/c$, are less than those given by (10.1.16)–(10.1.17) by an order of magnitude, and those for neutrinos, by three orders of magnitude.

If a black hole simultaneously possesses an electric charge Q and angular momentum J, particle production results in a reduction of both angular momentum and charge. If the energy of the new-born particle is ε, angular momentum j, and charge e, these parameters satisfy the inequality implied by condition (7.1.24) for superradiant modes:

$$\varepsilon \leq \Omega^H j + \Phi^H e \,, \tag{10.1.18}$$

where Ω^H is the angular velocity and Φ^H is the electric potential of the black hole.

Because the energy, angular momentum, and charge carried away by the created particles satisfy the same restriction (10.1.18) valid for the radiation parameters in superradiance, it is readily verified that the black hole area does not decrease in this process (see Section 7.1). This result shows that the irreducible mass of the black hole does not decrease in such processes either. Particles are created at the expense of the electrostatic energy or rotational energy stored in the black hole. This distinguishes these processes from the Hawking effect in which the energy required for the creation of particles reduces the (classical) "irreducible" mass. In the process of Hawking radiation the surface area decreases. In reality these processes take place simultaneously, and for a black hole having charge and angular momentum, the processes that remove angular momentum and electric charge are accompanied by the process that reduces the surface area of the black hole.

10.1.3 Black hole evaporation

In the Hawking effect one of the created particles (namely, the particle that has negative energy) is created beneath the event horizon, while the other, with positive energy is created outside the horizon. Hawking radiation carries away the energy, and as a result the mass of the black hole decreases. This observation, based on energy conservation, implies that there must be a flux of negative energy through the horizon into the black hole. This can happen only if the quantum average of the stress-energy tensor $\langle T_{\mu\nu}\rangle$ violates the *weak energy condition.* In our discussion of the main results of classical black hole physics in the first part of the book, we often assumed that the weak energy condition is satisfied. There were reasons to believe in this while we were dealing with classical systems and classical processes. Now, when we begin the consideration of the quantum aspects of black hole physics, we might expect that some of the results proved earlier are not directly applicable. The most important consequence is the violation of the Hawking area theorem in the quantum domain [Markov (1974)]. In the process of quantum particle creation the mass (and hence the surface area) of a black hole decreases. This process is known as the *evaporation of a black hole.*

A black hole radiates as a black body with temperature T_H and surface area $\mathcal{A}_H = 4\pi r_g^2$. Hence, the rate of mass loss can be estimated as $\dot{M} \sim \sigma T_H^4 \mathcal{A}_H$. Here

$\sigma = \pi^2 k^4/(60\,\hbar^3 c^2)$ is the Stefan-Boltzmann constant. This estimate gives

$$-\frac{dM}{dt} \sim b \left(\frac{m_{\rm Pl}}{M}\right)^2 \left(\frac{m_{\rm Pl}}{t_{\rm Pl}}\right) N, \tag{10.1.19}$$

where $b = (30 \times 8^3 \times \pi)^{-1} \approx 2.59 \times 10^{-6}$, and N is the number of states and species of particles that are radiated. This expression reproduces qualitatively the result of the exact calculations (10.5.32). By using this estimate, one can conclude that the "lifetime" of a black hole with respect to the process of quantum evaporation is[6]

$$t_H \sim \frac{1}{3b} t_{\rm Pl} \left(\frac{M}{m_{\rm Pl}}\right)^3 \sim 2.37 \times 10^{-24}\,{\rm s} \left(\frac{M}{1{\rm g}}\right)^3 N^{-1}. \tag{10.1.20}$$

As the mass of the black hole decreases in the process of evaporation, its temperature as well as the number of species of particles that can be emitted grow. As a result of both effects, the last stage of the evaporation looks like an explosion. At this stage the mutual interaction of the created particles may become important. The spectrum of the radiated particles and other details at this stage depend on the concrete model used for the high energy physics. Whether a black hole evaporates completely, or there remains a remnant of Planck mass is an open question which requires knowledge of physics at Planck scales for its solution (for a more detailed discussion of these questions, see Chapter 15).

10.1.4 Vacuum polarization

The result of the action of an external field on the vacuum goes beyond the effect of particle production. In fact, even those virtual particles that do not gain sufficient energy to become real, and hence again disappear, are nevertheless subject to the action of the external field during their short lifetime, and thus move differently than if there were no external field. As a result, the contribution of such virtual particles to various local physical observables (e.g., to the mean value of the energy-momentum tensor $\langle T_{\mu\nu}\rangle$) depends on the strength and other characteristics of the external field. In other words, the change in the vacuum expectation values of local observables in an external field in comparison with the original values in the absence of the field (this difference is what can be measured) is a function of the external field. This effect of the dependence of quantum averages of local observables on the external field is called *vacuum polarization*[7]. This effect may also occur even in the case where for some reason the external field does not create particles.

[6]Using (10.5.32), it is easy to show that it takes a time comparable with the time since the Big Bang for a black hole of mass 5×10^{14}g to evaporate completely. We used this estimate in Section 9.8.

[7]The expression *vacuum polarization* was first proposed in the quantum electrodynamics. An electric field has different action on electrons and positrons. For this reason, the vacuum where virtual electron-positron pairs are present behaves up to some extend as a usual polarizable medium. It should be emphasized that, according to the equivalence principle, the gravitational field acts

The distinction between real and virtual particles has a precise meaning in the absence of an external field. However, it ceases to be unambiguous in a region of spacetime where the external field is strong. This ambiguity leads to a difficulty in defining a particle in strong gravitational fields [for a discussion of this problem, see, e.g., Birrell and Davies (1982)]. As a result, it is not always possible to separate the contributions of real and virtual particles to the mean values of local observables, or to give a precise answer to the question about the exact place or exact moment of time of creation of a particle. Uncertainties arising here are manifestations of the general uncertainty relations inherent in quantum mechanics.

One of the manifestations of the vacuum polarization effect is a change in the equations describing the mean expectation value $\langle\Phi\rangle$ of a physical (not necessarily gravitational) field Φ produced by an external source J. The field Φ of the source J changes the state of virtual vacuum particles interacting with this field. The resulting additional quantum polarization corrections in the equation for $\langle\Phi\rangle$ take into account the backreaction of the state of the virtual particles on the original field Φ. The quantum process of creation and annihilation of virtual pairs being random, the "instantaneous" value of the field Φ does not coincide with the mean expectation value $\langle\Phi\rangle$; the field undergoes *quantum fluctuations*. Hence, the description of the field in terms of its mean expectation values has a limited range of applicability. This description is acceptable in situations in which quantum fluctuations are small in comparison with the mean expectation value of the field.

These general remarks on the possible manifestations of the quantum nature of physical fields and particles are entirely relevant when quantum effects in black holes are considered. In this case, the role of the external source producing the field is played by the massive body whose collapse results in the formation of the black hole.

Let us make some order of magnitude estimates which should clarify the possible role of vacuum polarization effects in black holes. It can be shown [DeWitt (1965)] that after quantum effects have been taken into account, the average gravitational field $\langle g\rangle$ is described by an equation arising from variation of the quantity

$$W[\langle g\rangle] = \frac{1}{16\pi}\int \mathcal{L}_{\text{eff}}(\langle g\rangle)\, d^4v\,, \tag{10.1.21}$$

called the *effective action*. In the absence of quantum effects (for $\hbar = 0$), the effective action coincides with Einstein's action. An expansion of the type

$$\mathcal{L}_{\text{eff}} = R + \hbar\,\mathcal{L}^{(1)} + \cdots \tag{10.1.22}$$

identically on both particles of a virtual pair. For this reason, literally speaking, the effect of the polarization is absent. Nevertheless, we follow the tradition and use the term "vacuum polarization" to describe the change of the physical characteristics of the vacuum under the action of any (including gravitational) field. It should be emphasized that the difference in the motion of particles of vacuum pairs induced by the gravitational fields is connected with tidal forces which are determined by the curvature of spacetime.

can be used in the general case to analyze $\mathcal{L}_{\text{eff}}$. Using arguments based on dimensional analysis, one can expect that to the lowest approximation in $\hbar$ (in the so-called one-loop approximation) the quantum corrections to $\mathcal{L}_{\text{eff}}$ are of order l_{Pl}^2/L^4, where L is the characteristic radius of curvature of spacetime. The first term in the expansion being of order $\mathcal{R} \sim 1/L^2$, the quantum polarization effects may substantially change Einstein's equations when the curvature is comparable with the Planck curvature $\mathcal{R}_{\text{Pl}}$.

In the case of the Schwarzschild metric, this condition is met for $r \sim r_1 = l_{\text{Pl}}(r_g/l_{\text{Pl}})^{1/3}$, which is deep inside the event horizon provided the black hole mass M is greater than the Planck mass. Hence, for $M \gg m_{\text{Pl}}$ quantum effects appreciably change the mean field $\langle g \rangle$ in comparison with the classical solution only when $r < r_1$. These effects are negligibly small outside the black hole.

If all terms in the equations for the mean field $\langle g \rangle$, except the Einstein term (corresponding to $\hbar = 0$), are carried to the right-hand side, the resulting expression on the right-hand side (which is nonzero only if $\hbar \neq 0$) can be interpreted as the vacuum expectation value $\langle T_\mu^\nu \rangle$ of the energy-momentum tensor of those physical fields whose contributions are included in the effective action. This set of fields must include the gravitational field itself. The characteristic values of the components $\langle T_\mu^\nu \rangle$ at the event horizon of a Schwarzschild black hole are of order $\hbar c/r_g^4$. Note that if $M \gg m_{\text{Pl}}$, $\langle T_\mu^\nu \rangle$ affects the external geometry of the black hole only slightly. Nevertheless, this small change can result in substantial qualitative changes in the global properties of the solutions describing the black hole over sufficiently long times. Thus, the negative energy density flux across the event horizon of an evaporating black hole, which accompany the Hawking radiation, ultimately results in reducing the event horizon to a Planck size (perhaps even to its complete disappearance). It is readily verified that the expectation value $\langle T_\mu^\nu \rangle$ for the energy flux across the event horizon, accompanying the Hawking effect, is again of order $\hbar c/r_g^4$.

These arguments show that as long as the mass of a black hole is much greater than the Planck mass, and we are dealing with time intervals much shorter than the characteristic time of black hole evaporation, the backreaction of the created matter and vacuum polarization are negligible. The effects of quantum fluctuations are also negligible under these conditions. In order to describe the black hole geometry, one can use a solution of classical Einstein's equations as a zeroth-order approximation in $\hbar$. The quantum corrections of first order in $\hbar$ can be obtained by making one-loop calculations on a given classical background and solving the Einstein equation with $\langle T^{\mu\nu} \rangle$ as its right-hand side. One can use iterations to get a self-consistent solution with quantum corrections.

10.1.5 Quantum fluctuations of the metric

The following simple arguments give a qualitative estimate of the contribution of quantum fluctuation effects in black holes. Assume that a fluctuation of the metric occurs in a spacetime region of characteristic size l so that the value of g deviates from the expectation value $\langle g \rangle$ by δg. The curvature in this region changes by a quantity of order $\delta g/(l^2\langle g \rangle)$, and the value of the action S for the gravitational field correspondingly changes by

$$\delta S \sim \frac{\delta g}{\langle g \rangle} l^2 \frac{c^2}{G} . \tag{10.1.23}$$

The probability of such a quantum fluctuation is appreciable only if $\delta S \sim \hbar$. Therefore, the following estimate is obtained for the value of the metric fluctuation $\delta g/\langle g \rangle$ in a spacetime region of size l:

$$\frac{\delta g}{\langle g \rangle} \sim \frac{l_{\mathrm{Pl}}^2}{l^2} , \tag{10.1.24}$$

where l_{Pl} is the Planck length [Wheeler (1962), Harrison (1970)]. The metric fluctuations are thus generally insignificant on a large scale although their amplitude reaches unity on the Planck scale. The mean-field approximation is definitely justified for the description of black holes with a mass much greater than the Planck mass m_{Pl}. It can be expected [York (1983)] that the quantum gravitational fluctuations outlined above lead to an effect resembling quantum "trembling" of the event horizon. The amplitude of "trembling" δr_{g} of the gravitation radius has the following form for a spherically symmetric black hole, as we find from (10.1.24):

$$\delta r_{\mathrm{g}} \sim l_{\mathrm{Pl}}^2/r_{\mathrm{g}} . \tag{10.1.25}$$

It is interesting to note that although this quantity is small for the black holes we discuss here (those with $M \gg m_{\mathrm{Pl}}$), the very existence of this effect changes qualitatively (as viewed by a distant observer) the idealized classical description of the collapse and fall of a body into a black hole. The formally infinite expression

$$\Delta l \sim \int_{r_{\mathrm{g}}}^{R} \frac{dr}{1 - r_{\mathrm{g}}/r} \tag{10.1.26}$$

for the duration of this process by the clock of a distant observer must now be replaced with a finite quantity $\Delta t \sim r_{\mathrm{g}} \ln(r_{\mathrm{g}}/l_{\mathrm{Pl}})$, as a result of the replacement $r_{\mathrm{g}} \to r_{\mathrm{g}} + \delta r_{\mathrm{g}}$ in the lower limit of integration.

A similar estimate can be obtained if we take into account the quantum nature of the motion of a particle falling into the black hole or if we try to evaluate the accuracy with which the position of a falling body can be localized in the vicinity of the event horizon using the method of scattering of the waves of a physical field on this body. In the latter case, the restriction (10.1.25) arises because the size δr of the wavepacket, whose energy is less than the black hole mass M, must be greater than the quantity $\delta r \sim \hbar/Mc$.

10.2 Particle Creation by Black Holes

10.2.1 General theory

To prove the results mentioned in the preceding section in connection with the quantum production of particles in black holes, and to extract more detailed information about these quantum phenomena, we will need a sufficiently developed mathematical apparatus of quantum field theory in curved spacetime.

Formally, the problem of particle creation in black holes is a particular case of the more general problem of particle creation in an external field. The standard scheme of constructing the theory is as follows. One chooses an external field in such a way that the field in the distant past and in the distant future vanishes. The concepts of particle and vacuum admit unambiguous definitions in these in- and out-regions. Thus, the lowest-energy state of the system is usually chosen as the vacuum. The system undergoes evolution in the external field. If under the action of the external field the state obtained as a result of the time evolution of the in-vacuum is not identical to the out-vacuum state, then particles are created. The operator relating the in- and out-states is known as the *S-matrix*. This operator contains complete information on the processes of particle creation, their scattering, and annihilation in the external field.

The problem of particle creation in black holes has two very important aspects which necessitate a modification of the standard scheme. First, it is impossible to "switch off" in a natural manner the gravitational field of the new-born black hole in the future, although the gravitational field in the past (before the collapse began) can be considered weak in a physically realistic formulation of the problem of collapse and all the states in the in-region can be determined. A decrease in the black hole mass enhances the surface gravity, instead of reducing it, and hence enhances the intensity of radiation. This is why the "switching off" of the gravitational field of a black hole by formally reducing its mass does not yield the desired result.

The second point is more important. Namely, a distant observer can record the state of only those particles that escape from the hole. The created particles that sink into the black hole are "invisible" to this observer. When the results of observations outside a black hole are described, one has to average over "invisible" particle states. In other words, the observer outside the black hole invariably deals with only a part of the total quantum system. In accordance with the general principles of quantum mechanics, black hole radiation is described by a *density matrix* even if the initial state (before the formation of the black hole) was quantum-mechanically pure. Note that the necessary averaging involves just those states that correspond to the "particles" not leaving the region of the strong field. It is for these particles that the very concept of "particle" is poorly defined since the field of the black hole cannot be "switched off". Fortunately, the result of averaging, describing the state of radiation of the black hole, is independent of the arbitrariness in the definition of these "invisible" states.

To summarize, we remark that the problem of interest in this section, the calculation of the characteristics of quantum radiation of a black hole, is naturally broken into two stages: (i) calculation of the S-matrix operator, and (ii) averaging it over the part of the out-states that correspond to "invisible" particles. In this section we calculate the S-matrix. The problem of density-matrix calculation for the black hole radiation is left for the next section.

The general formalism of constructing the S-matrix for problems in an arbitrary (not necessarily gravitational) external field is given in Appendices H.1–H.3. The main steps are:

- Define in- and out-bases. To construct these bases one considers complex solutions of the field equations and splits them into parts that are of positive and negative frequencies with respect to the chosen time parameter either in the past (for the in-basis) or in the future (for the out-basis).
- Find the Bogoliubov coefficients, relating the in- and out-basis solutions. These coefficients are defined by equation (H.26).
- Calculate the S-matrix. The explicit expression for the S-matrix in terms of the Bogoliubov coefficients is given by equations (H.34)–(H.35).

Now that the general formal scheme has been outlined, we turn to its adaptation to the description of quantum effects in black holes. Special attention is focused on massless fields. The case of massless fields is the most important because, on the one hand, massless fields give the main contribution to the quantum radiation of black holes, and on the other hand, this analysis serves as a good approximation in describing the creation of massive particles in the case when the Hawking temperature of the black hole is much higher than the mass of these particles so that the ultra-relativistic approximation is valid for their description. The effect of the mass of the field will be discussed later.

10.2.2 Modes and bases

Model

For the sake of simplicity, we begin the analysis with the theory of a massless neutral scalar field φ in the spacetime of an uncharged non-rotating black hole. We will later return to discussing the effects of black hole charge and rotation as well as the effects of spin, mass, and charge of the particles on the process of their creation.

To make our considerations more concrete and simple, we first consider a model in which the collapsing body is represented by a massive thin shell collapsing with the velocity of light. The metric of such a spacetime can be written in the form

$$ds^2 = -\left(1 - \frac{2M(v)}{r}\right) dv^2 + 2\, dv\, dr + r^2\, d\omega^2 \,, \qquad (10.2.1)$$

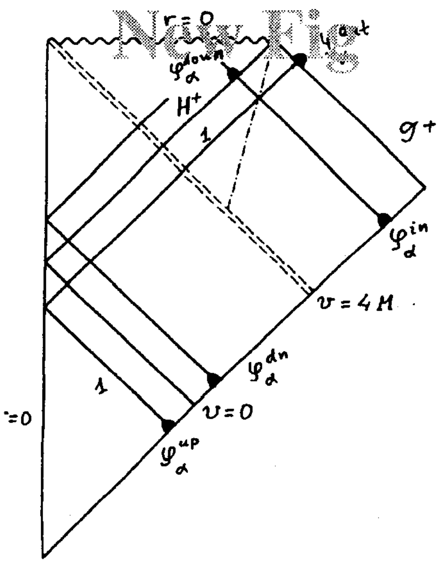


Figure 10.2: Penrose-Carter conformal diagram for the spacetime of an uncharged non-rotating black hole created by the spherically-symmetric collapse of a massive thin null shell.

where $d\omega^2$ is the metric on the unit sphere, and $M(v) = M\vartheta(v - v_0)$. Inside the collapsing shell (for $v < v_0$) the spacetime is flat, while outside (for $v > v_0$) the spacetime geometry is described by the Schwarzschild metric, M being the mass of the resulting black hole (see Figure 10.2).

Consider now a radial null ray that reaches $\mathcal{J}^+$ at a moment u of retarded time (ray 1 in Figure 10.2). Let us trace this ray back in time and denote by $v = U(u)$ the moment of advanced time when the ray began its propagation from $\mathcal{J}^-$. We choose the origin of the v coordinate so that the radial null ray sent from $\mathcal{J}^-$ at $v = 0$ reaches $r = 0$ at the moment of event horizon formation. A simple consideration shows that for this choice the parameter $v_0 = 4M$.

The equation of the outgoing null ray in the region outside the collapsing shell (in the Schwarzschild region) is

$$u = v - 2r_* = \text{const}\,, \tag{10.2.2}$$

where $r_* = r - 2M + 2M\ln[(r - 2M)/2M]$ is the tortoise coordinate. The ray crosses the collapsing shell $v = 4M$ at the radius $r = R$, where R is defined by the equation

$$u = 8M - 2R - 4M\ln[(R - 2M)/2M]. \tag{10.2.3}$$

Null rays that reach $\mathcal{J}^+$ at late time ($u \to \infty$) cross the collapsing shell at a radius close to $2M$, so that for these late-time rays we have

$$u \approx -4M \ln[(R-2M)/2M]. \tag{10.2.4}$$

Let us consider now an incoming radial null ray that was emitted from $\mathcal{J}^-$ at a moment $v < 0$. It propagates in a flat spacetime until it crosses the collapsing shell from inside. For this ray $v = t + r = \text{const}$ so that it reaches $r = 0$ at the moment of time $t = v$. After passing through the origin $r = 0$, it continues its propagation as an outgoing ray, and its equation (until it reaches the shell) is $u = t - r = v = \text{const}$. The ray crosses the shell at a radius R given by the relation

$$v = -2(R - 2M). \tag{10.2.5}$$

Comparing (10.2.4) and (10.2.5), we get the following expression

$$u \approx -\kappa^{-1} \ln(-\kappa\, v) \tag{10.2.6}$$

relating the moment v when a late-time ray left $\mathcal{J}^-$ and the moment u when it reaches $\mathcal{J}^+$. Here $\kappa = (4M)^{-1}$ is the surface gravity of the black hole.

Modes

A massless scalar field φ obeys the wave equation

$$\Box\varphi = 0. \tag{10.2.7}$$

In a spherically symmetric spacetime we can write its solutions, using a decomposition into modes

$$\varphi_{\ell m} = \frac{u_\ell(t,r)}{r}\, Y_{\ell m}(\theta,\phi), \tag{10.2.8}$$

where $Y_{\ell m}(\theta,\phi)$ are standard spherical harmonics obeying the normalization conditions

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta\, \bar{Y}_{\ell' m'}(\theta,\phi)\, Y_{\ell m}(\theta,\phi) = \delta_{\ell\ell'}\, \delta_{mm'}. \tag{10.2.9}$$

If we write the metric of a spherically symmetric spacetime in the form ($A, B = 0, 1$)

$$ds^2 = \gamma_{AB}\, dx^A\, dx^B + r^2\, d\Omega^2, \tag{10.2.10}$$

then $u_{\ell m}$ obeys the equation

$$\left({}^2\Box - W_\ell\right) u_\ell = 0, \tag{10.2.11}$$

where ${}^2\Box = (-\gamma)^{-1/2}\,\partial_A[(-\gamma)^{1/2}\,\gamma^{AB}\,\partial_B]$ is the two-dimensional "$\Box$" operator for the metric γ_{AB}, and

$$W_\ell = \frac{\ell(\ell+1)}{r^2} + \frac{{}^2\Box r}{r}. \tag{10.2.12}$$

For the metric (10.2.1)

$$\frac{{}^2\Box r}{r} = \frac{2M}{r^3}\,\vartheta(v - v_0). \tag{10.2.13}$$

In the Schwarzschild region (i.e., for $v > v_0$) equation (10.2.11) is identical to equation (4.2.4). In this region we can write u_ℓ in the form $u_{\ell\omega} = \exp(-i\omega t)\,u_\ell(r,\omega)$, where $u_\ell(r,\omega)$ is a solution to equation (4.2.7)

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - V_\ell(r)\right] u_\ell(r,\omega) = 0, \tag{10.2.14}$$

$$V_\ell(r) = \left(1 - \frac{2M}{r}\right) W_\ell. \tag{10.2.15}$$

Our aim now is to define useful sets of solutions (modes) that will allow us to introduce convenient in- and out-bases for a quantum field $\hat{\varphi}$. Let us consider, at first, special solutions, which we denote $\varphi^{\mathrm{out}}_{\ell m\omega}(x) = u^{\mathrm{out}}_{\ell\omega}(t,r)\,Y_{\ell m}(\theta,\phi)$. These solutions are singled out by the property that they vanish on H^+ and have the following asymptotic behavior near $\mathcal{J}^+$

$$\varphi^{\mathrm{out}}_{\ell m\omega} \sim \frac{1}{r}\,\Phi^{\mathrm{out}+}_{\ell m\omega} = \frac{1}{\sqrt{4\pi\omega}}\,\frac{e^{-i\omega u}}{r}\,Y_{\ell m}(\theta,\phi). \tag{10.2.16}$$

We call $\Phi^{\mathrm{out}+}_{\ell m\omega}$ the *image* on $\mathcal{J}^+$ of the field $\varphi^{\mathrm{out}}_{\ell m\omega}$.

Denote by

$$<\varphi^1,\varphi^2> = i\int_\Sigma \mathrm{d}\sigma^\mu\,(\bar{\varphi}^1\,\varphi^2_{,\mu} - \varphi^2\,\bar{\varphi}^1_{,\mu}) \tag{10.2.17}$$

the inner product for a pair φ^1, φ^2 of solutions to equation (10.2.7). Here $\mathrm{d}\sigma^\mu$ is a future-directed vector of the surface element of Σ. The inner product (10.2.17) does not depend on the particular choice of the Cauchy surface Σ. For massless fields in a spacetime of a black hole, this surface can be moved to the future so that it coincides with $\mathcal{J}^+ \cup H^+$

$$<\,,\,> = <\,,\,>_{H^+} + <\,,\,>_{\mathcal{J}^+}. \tag{10.2.18}$$

If Φ^{1+} and Φ^{2+} are images of φ^1 and φ^2 on $\mathcal{J}^+$, then

$$<\varphi^1,\varphi^2>_{\mathcal{J}^+} = i\int_{\mathcal{J}^+} (\bar{\Phi}^{1+}\,\partial_u\Phi^{2+} - \Phi^{2+}\,\partial_u\bar{\Phi}^{1+})\,du\,d\Omega, \tag{10.2.19}$$

where $d\Omega = \sin\theta\, d\theta\, d\phi$.

Using relations (10.2.16), (10.2.18), and (10.2.19), one can show that the following normalization conditions are valid for φ^{out}

$$< \varphi_J^{\text{out}}, \varphi_{J'}^{\text{out}} > = - < \bar\varphi_J^{\text{out}}, \bar\varphi_{J'}^{\text{out}} > = \delta_{JJ'}\,, \qquad < \varphi_J^{\text{out}}, \bar\varphi_{J'}^{\text{out}} > = 0\,. \tag{10.2.20}$$

Here, and hereafter, we use the notation $J = \{\ell, m, \omega\}$, $J' = \{\ell', m', \omega'\}$, and

$$\delta_{JJ'} \equiv \delta(\omega - \omega')\, \delta_{\ell\ell'}\, \delta_{mm'}\,. \tag{10.2.21}$$

In what follows, it is convenient to consider as the basis functions not $\varphi_{\omega\ell m}$ but wavepacket-type solutions constructed of $\varphi_{\omega\ell m}$ [Hawking (1975)]. To achieve this, we fix a real number δ $(0 < \delta \ll 1)$, define $\epsilon = \delta\kappa$ (κ is the surface gravity of the black hole), and introduce the notation

$$\varphi_{jn\ell m} = \epsilon^{-1/2} \int_{j\epsilon}^{(j+1)\epsilon} e^{2\pi i n\omega/\epsilon}\, \varphi_{\omega\ell m}\, d\omega\,. \tag{10.2.22}$$

In this relation j is a non-negative integer and n is an integer. Hereafter we denote the collective index $\{jn\ell m\}$ by a single letter α.[8] The wavepackets $\varphi_\alpha^{\text{out}}$ have frequencies in the interval from $j\epsilon$ to $(j+1)\epsilon$. Their images on $\mathcal{J}^+$ reach a maximum close to the moment $u = 2\pi n/\epsilon$ of retarded time; the packet width is $\Delta u \sim 2\pi/\epsilon$. The wavepackets constructed from φ_J^{out} solutions obey the following normalization conditions

$$< \varphi_\alpha^{\text{out}}, \varphi_{\alpha'}^{\text{out}} > = - < \bar\varphi_\alpha^{\text{out}}, \bar\varphi_{\alpha'}^{\text{out}} > = \delta_{\alpha\alpha'}, \qquad < \varphi_\alpha^{\text{out}}, \bar\varphi_{\alpha'}^{\text{out}} > = 0\,, \tag{10.2.23}$$

where

$$\delta_{\alpha\alpha'} \equiv \delta_{jj'}\, \delta_{nn'}\, \delta_{\ell\ell'}\, \delta_{mm'}\,. \tag{10.2.24}$$

We introduce another set of solutions φ_J^{in} that are specified by their initial values (images) on $\mathcal{J}^-$

$$\Phi_J^{\text{in}-}(v,\theta,\phi) = \lim_{\substack{r\to\infty \\ v,\theta,\phi=\text{const}}} \left[r\, \varphi_J^{\text{in}}(r,v,\theta,\phi)\right]. \tag{10.2.25}$$

(The image Φ_J^- of the solution on $\mathcal{J}^-$ is defined in a similar way to the image on $\mathcal{J}^+$ with the obvious change of u to v.) Namely, we put

$$\Phi_J^{\text{in}-}(v,\theta,\phi) = \frac{e^{-i\omega v}}{\sqrt{4\pi\omega}}\, Y_{\ell m}(\theta,\phi)\,. \tag{10.2.26}$$

We denote by $\varphi_\alpha^{\text{in}}$ the wavepackets constructed from φ_J^{in} according to (10.2.22).

[8] When there is a summation over the collective indices such as α or J, we always indicate it explicitly. Hence, the Einstein rule of an implicit summation over the repeated indices is not extended to this case.

Let us focus our attention on late-time wavepackets $\varphi_\alpha^{\rm in}$; that is, wavepackets with $n \geq N \gg 1$. Such a wavepacket originated from $\mathcal{J}^-$ at $v > 0$ and at all times is propagating in a static Schwarzschild geometry. It reaches the peak of the potential barrier, located at $r \approx 3M$. Part R_α of this wavepacket is scattered to $\mathcal{J}^+$, and the other part T_α penetrates through the barrier and crosses the horizon. The part that is scattered has the image $R_\alpha \Phi_\alpha^{\rm out+}$ on $\mathcal{J}^+$. When the other part that passes through the potential barrier reaches the horizon, it is of the form $T_\alpha\, \varphi_\alpha^{\rm down}|_{H^+}$. Here $\varphi_\alpha^{\rm down}$ are wavepackets constructed from so-called *DOWN*-modes $\varphi_J^{\rm down}$, which are specified by the conditions that they vanish on $\mathcal{J}^+$ and have the following form at the horizon H^+:

$$\varphi_J^{\rm down}(v,\theta,\phi) = \frac{1}{\sqrt{4\pi\omega}}\, \frac{e^{-i\omega v}}{r_+}\, Y_{\ell m}(\theta,\phi)\,. \tag{10.2.27}$$

The inner product for two solutions φ^1 and φ^2 that vanish on $\mathcal{J}^+$ can be expressed in terms of integrals over the event horizon H^+. The three-surface element $d\Sigma^\mu$ of the future event horizon H^+, with normal vector in the inward radial direction written in Eddington-Finkelstein coordinates (v,r,θ,ϕ) reads

$$d\sigma^\mu = l_H^\mu\, r_+^2\, dv\, d\Omega, \tag{10.2.28}$$

where v is an advanced time; $l_H^\mu = (1,0,0,0)$, and $d\Omega = \sin\theta\, d\theta\, d\phi$. Thus, we have

$$< \varphi^1, \varphi^2 >_{H^+} = i r_+^2 \int_{H^+} \left(\bar{\varphi}^1_{H^+}\, \partial_v \varphi^2_{H^+} - \varphi^2_{H^+}\, \partial_v \bar{\varphi}^1_{H^+}\right) dv\, d\Omega\,. \tag{10.2.29}$$

It can be verified that the functions $\{\varphi_\alpha^{\rm down}, \bar{\varphi}_\alpha^{\rm down}\}$ satisfy the following normalization conditions:

$$< \varphi_\alpha^{\rm down}, \varphi_{\alpha'}^{\rm down} > = \delta_{\alpha\alpha'}\,, \qquad < \varphi_\alpha^{\rm down}, \bar{\varphi}_{\alpha'}^{\rm down} > = 0\,. \tag{10.2.30}$$

The thus defined late-time *IN*-, *OUT*-, and *DOWN*-modes are related by[9]

$$\varphi_\alpha^{\rm in} = R_\alpha\, \varphi_\alpha^{\rm out} + T_\alpha\, \varphi_\alpha^{\rm down}\,. \tag{10.2.31}$$

This is a direct analogue of the relations

$$\varphi_J^{\rm in} = R_J\, \varphi_J^{\rm out} + T_J\, \varphi_J^{\rm down} \tag{10.2.32}$$

which are valid in the spacetime of an eternal black hole (cf. Section 4.2.2). Equation (10.2.32) yields (10.2.31) if the parameter ϵ in the wavepacket definition (10.2.22) is chosen sufficiently small, so that the *reflection amplitude* R_J and *absorption amplitude* T_J remain practically constant in the frequency intervals of width ϵ. Under this condition, we also have [cf. (4.2.13)]

$$|\, R_\alpha\,|^2 + |\, T_\alpha\,|^2 = 1\,. \tag{10.2.33}$$

[9]Note that there is no summation over repeated index α (see a footnote at page 366.)

We would like to stress that globally the geometry of an eternal black hole differs from the geometry of a black hole that is formed as the result of a collapse. That is why the global behavior of modes φ_J in these spaces is also different. Nevertheless, for late-time wavepackets the relation (10.2.31) is valid for both spaces.

We define now a new set of modes $\varphi_\alpha^{\mathrm{up}}$ by requiring that they are orthogonal to $\varphi_\alpha^{\mathrm{in}}$,

$$< \varphi_\alpha^{\mathrm{up}}, \varphi_{\alpha'}^{\mathrm{in}} > = 0\,, \tag{10.2.34}$$

admit the decomposition

$$\varphi_\alpha^{\mathrm{up}} = t_\alpha\, \varphi_\alpha^{\mathrm{out}} + r_\alpha\, \varphi_\alpha^{\mathrm{down}}\,, \tag{10.2.35}$$

and are normalized by the condition

$$< \varphi_\alpha^{\mathrm{up}}, \varphi_{\alpha'}^{\mathrm{up}} > = \delta_{\alpha\alpha'}\,. \tag{10.2.36}$$

These conditions imply that the following relations hold:

$$|\, r_\alpha\,|^2 + |\, t_\alpha\,|^2 = 1\,, \qquad t_\alpha\, \bar{R}_\alpha + r_\alpha\, \bar{T}_\alpha = 0\,. \tag{10.2.37}$$

As the consequence of the relations (10.2.33) and (10.2.37), one has

$$|\, R_\alpha\,| = |\, r_\alpha\,|, \qquad |\, T_\alpha\,| = |\, t_\alpha\,|\,. \tag{10.2.38}$$

Let us trace back in time a wavepacket $\varphi_\alpha^{\mathrm{up}}$ until it crosses the massive null shell and reaches $\mathcal{J}^-$. The relation (10.2.6) shows that the late-time wavepacket that began from $\mathcal{J}^-$ is blueshifted by the factor $\sim \exp(\kappa u)$ with respect to the wavepacket which arrives at $\mathcal{J}^+$ at the moment u. Under these conditions, one can use the *geometrical optics approximation* [Hawking (1975)] to describe the evolution into the past of $\varphi_\alpha^{\mathrm{up}}$ in the region inside the null shell. This allows one to show that the image $\Phi_\alpha^{\mathrm{up}-}$ at $\mathcal{J}^-$ is a wavepacket (10.2.22) constructed from the following expression

$$\Phi_{\omega\ell m}^{\mathrm{up}-} = \frac{1}{\sqrt{4\pi\omega}}\, \mathrm{e}^{i\omega \ln(-\kappa v)/\kappa}\, Y_{\ell m}(\theta,\phi)\, \vartheta(-v)\,. \tag{10.2.39}$$

We include the Heaviside step function $\vartheta(-v)$ to stress that $\Phi^{\mathrm{up}-}$ vanishes for $v > 0$. The wavepacket $\varphi_\alpha^{\mathrm{up}}$ emitted from $\mathcal{J}^-$ crosses the shell and passes through the potential barrier from inside. The relation (10.2.35) shows that it is partly scattered by the barrier, so that a portion $r_\alpha\, \varphi_\alpha^{\mathrm{down}}$ of it enters the horizon. The other part $t_\alpha\, \varphi_\alpha^{\mathrm{out}}$ penetrates the barrier and reaches $\mathcal{J}^+$ (see Figure 10.2).

The final set of modes that we need is denoted by $\varphi_\alpha^{\mathrm{dn}}$. Their images on $\mathcal{J}^-$, $\Phi_\alpha^{\mathrm{dn}-}$, are constructed by means of the transformation (10.2.22) from the following functions:

$$\Phi_{\omega\ell m}^{\mathrm{dn}-}(v,\theta,\phi) = \bar{\Phi}_{\omega\ell m}^{\mathrm{up}-}(-v,\theta,\phi) = \frac{1}{\sqrt{4\pi\omega}}\, \mathrm{e}^{-i\omega \ln(\kappa v)/\kappa}\, \bar{Y}_{\ell m}(\theta,\phi)\, \vartheta(v)\,. \tag{10.2.40}$$

We shall use the introduced set of modes $\varphi_\alpha^{\mathrm{in}}$, $\varphi_\alpha^{\mathrm{out}}$, $\varphi_\alpha^{\mathrm{down}}$, $\varphi_\alpha^{\mathrm{up}}$, and $\varphi_\alpha^{\mathrm{dn}}$ for the construction of simple and convenient in- and out-bases.

Late-time in- and out-bases

Let us denote by $\varphi_\alpha^{\mathrm{d}}$ and $\varphi_\alpha^{\mathrm{p}}$ the linear combinations of $\varphi_\alpha^{\mathrm{up}}$ and $\varphi_\alpha^{\mathrm{dn}}$ modes,

$$\varphi_\alpha^{\mathrm{d}} = c_\alpha \varphi_\alpha^{\mathrm{dn}} + s_\alpha \bar{\varphi}_\alpha^{\mathrm{up}}, \qquad \varphi_\alpha^{\mathrm{p}} = c_\alpha \varphi_\alpha^{\mathrm{up}} + s_\alpha \bar{\varphi}_\alpha^{\mathrm{dn}}, \tag{10.2.41}$$

$$s_\alpha = w_\alpha (1 - w_\alpha^2)^{-1/2}, \qquad c_\alpha = (1 - w_\alpha^2)^{-1/2}, \tag{10.2.42}$$

where $w_\alpha = \exp[-\pi\omega_j/\kappa]$. The modes $\varphi_\alpha^{\mathrm{d}}$ and $\varphi_\alpha^{\mathrm{p}}$ have images on $\mathcal{J}^-$ that are of positive frequency with respect to the advanced time v. To prove this statement, it is sufficient to note that the functions $\varphi_\alpha^{\mathrm{d}}$ and $\varphi_\alpha^{\mathrm{p}}$ are obtained by applying transformation (10.2.22) to solutions that manifest the following dependence on v on $\mathcal{J}^-$:

$$F_{\omega\ell m}(v) = \theta(v)\, \mathrm{e}^{-iq\ln\kappa v} + \theta(-v)\, \mathrm{e}^{-\pi q}\, \mathrm{e}^{-iq\ln(-\kappa v)}. \tag{10.2.43}$$

In this formula $q = \omega/\kappa$ for D-modes and $q = -\omega/\kappa$ for P-modes. The fact that $F_{\omega\ell m}(v)$ for $-\infty < \omega < \infty$ contains only positive frequencies with respect to the advanced time v follows from the Lemma proved in Appendix H.4.2 .

We choose as in-basis the following positive-frequency solutions

$$\boldsymbol{\Phi}_\alpha^{\mathrm{in}} = \begin{pmatrix} \varphi_\alpha^{\mathrm{in}} \\ \varphi_\alpha^{\mathrm{d}} \\ \varphi_\alpha^{\mathrm{p}} \end{pmatrix} \tag{10.2.44}$$

with $\alpha = \{j, n, \ell, m\}$ and $n \geq N$, and augment them to a complete orthonormalized basis with an arbitrary set of positive-frequency functions defined on $\mathcal{J}^-$. Likewise, we form the out-basis by augmenting the set of functions

$$\boldsymbol{\Phi}_\alpha^{\mathrm{out}} = \begin{pmatrix} \varphi_\alpha^{\mathrm{out}} \\ \varphi_\alpha^{\mathrm{down}} \\ \varphi_\alpha^{\mathrm{dn}} \end{pmatrix} \tag{10.2.45}$$

to a complete orthonormalized system. Similar bases were introduced by Wald (1975). In what follows we refer to (10.2.44) and (10.2.45) as *Wald's bases.*

We would like to emphasize that the modes $\boldsymbol{\Phi}_\alpha^{\mathrm{in}}$ are of positive frequency with respect to advanced time v on $\mathcal{J}^-$. As we shall see, in the corresponding vacuum state there are no particles incoming onto a black hole from infinity (from $\mathcal{J}^-$). As for the definition of $\boldsymbol{\Phi}_\alpha^{\mathrm{out}}$, the modes $\varphi_\alpha^{\mathrm{out}}$ are of positive frequency with respect to the retarded time u on $\mathcal{J}^+$. In the corresponding out-vacuum there are no particles outgoing to infinity. Both types of particles (incoming from $\mathcal{J}^-$ and outgoing to $\mathcal{J}^+$) are well-defined since the gravitational field in the asymptotic regions is weak. A particle interpretation for the modes $\varphi_\alpha^{\mathrm{down}}$ and $\varphi_\alpha^{\mathrm{dn}}$ (which vanish on $\mathcal{J}^+$) is not so straightforward.

In the general case, the matrices of Bogoliubov coefficients relating in- and out-bases are infinite-dimensional. For Wald's bases, the matrices of Bogoliubov coefficients factorize, so that the problem basically becomes three-dimensional. The price

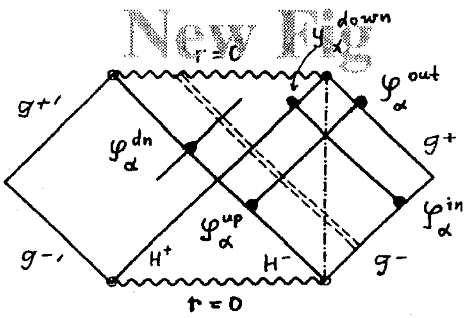


Figure 10.3: Penrose-Carter conformal diagram for the spacetime of the eternal version of a black hole shown in Figure 10.2.

for this is that there is no clear physical interpretation of the vacuum and particle states on H^+ which correspond to the modes $\varphi_\alpha^{\text{down}}$ and $\varphi_\alpha^{\text{dn}}$. But, fortunately, for the problem under consideration, where we are interested in the description of observables only in the black hole exterior, the specific choice of these modes is irrelevant. We shall demonstrate this later (see Section 10.3).

Modes and bases in the eternal version of a black hole

We are now ready to discuss the Bogoliubov transformations, but before this we shall make some remarks concerning the possibility of further simplification of the problem. In the problem we are considering, where a black hole is formed as a result of gravitational collapse, the gravitational field is time-dependent. Moreover, there exist two physically different regions: one inside the collapsing matter, and the other outside it. At late times the geometry is static and is described by the Schwarzschild metric. We may consider a new spacetime; that is, the eternal Schwarzschild black hole, which has the same mass M as the original black hole. We call this geometry the *eternal version of a black hole.* Figure 10.3 shows the Penrose-Carter conformal diagram of the eternal version of the black hole presented in Figure 10.2. Since the late-time regions of both spaces are isometric, one can identify *IN*-, *UP*-, *OUT*-, and *DOWN*-modes, propagating in these regions. These modes can be continued in the spacetime of the eternal version of the black hole beyond the late-time region, as solutions of the field equations.

The so-defined solutions $\varphi_\alpha^{\text{in}}$, $\varphi_\alpha^{\text{up}}$, $\varphi_\alpha^{\text{out}}$, and $\varphi_\alpha^{\text{down}}$ in the spacetime of the eternal version of a black hole are precisely the wavepackets constructed for *IN*-, *UP*-, *OUT*-, and *DOWN*-modes, that we discussed in Section 4.2. We know that any two of these set of modes are sufficient to construct a complete basis in the black hole exterior. For the construction of Wald's bases for the eternal black hole we need an additional set of modes, which we denoted by $\varphi_\alpha^{\text{dn}}$. For this purpose, consider a discrete isometry that maps the black hole exterior I onto the region I'. In Kruskal coordinates (U, V, θ, ϕ)

this transformation is given by relations $U \to -U$, $V \to -V$. The solutions $\varphi_J^{\rm dn}$ are defined as[10]

$$\varphi_{\ell m\omega}^{\rm dn}(U,V,\theta,\phi) = \bar{\varphi}_{\ell m\omega}^{\rm up}(-U,-V,\theta,\phi)\,. \tag{10.2.46}$$

The coordinates V and U are the affine parameter coordinates on the future H^+ and past H^- horizons, respectively. The Kruskal coordinates V and U are related to the ingoing (v) and outgoing (u) Eddington-Finkelstein coordinates by

$$v = -\kappa^{-1}\ln \kappa V, \quad u = -\kappa^{-1}\ln(-\kappa U)\,. \tag{10.2.47}$$

The relation between u and U is identical to (10.2.6). This is why the modes $\varphi^{\rm d}$ and $\varphi^{\rm p}$ constructed from $\varphi^{\rm up}$ and $\varphi^{\rm dn}$ in the spacetime of the eternal version of the black hole by rules (10.2.41)–(10.2.42) are of positive frequency with respect to the affine parameter U on H^-.

One can introduce in- and out-bases in the eternal version of the black hole by defining them in accordance with Eqs.(10.2.44) and (10.2.45). It is evident that the relations between in- and out-bases at late times in both of the spaces (the original black hole and its eternal version) are identical. In other words, all the results concerning quantum particle creation by a black hole at late times can be obtained by making calculations for the eternal version of the black hole. In the latter case, the spacetime metric is much simpler analytically, so that the calculations are simplified. It should also be noted that the standard positive frequency definition for incoming waves in the original spacetime with black hole formation corresponds to a special prescription in the spacetime of the eternal version. Namely, the solutions corresponding to the in-basis are of positive frequency with respect to the advanced time v on $\mathcal{J}^-$ and of positive frequency with respect to the affine parameter U on H^-. The vacuum state in the spacetime of the eternal version of the black hole defined by using such a definition of positive frequencies is known as the *Unruh vacuum*. The above consideration implies that for calculations in the late-time regime the Unruh vacuum in the eternal version of the black hole is equivalent to the choice of the standard Minkowski vacuum as the initial state in the past in the spacetime with black hole formation.

To define in- and out-bases we used a simple model of black hole formation. But more detailed analysis shows that the results obtained are model independent provided we are interested only in the late-time observables. The basic formula (10.2.6)

[10] In order to construct a complete basis in the spacetime of the eternal black hole besides the modes defined in the black hole exterior (in the right wedge of the diagram 10.3), one must also introduce modes defined in the left wedge. It is covenient to do it by doubling the external modes by using relations similar to (10.2.46), that is, by the mirror-like reflection of external modes with respect to the origin ($U = V = 0$). The notation "dn" which is the mirror-like image of "up" serves to remind of this procedure. We use the same convention for other set of modes. For example, modes which are dual to "down"-, "out"-, "in"-, "n"-, and "p"-modes are denoted as "umop", "fno", "ui", "u", and "d", respectively.

for late-time radial null rays evidently remains valid if one considers instead of the shell an arbitrary spherically symmetric distribution of collapsing matter. Moreover, one can generalize the above consideration to the case of a rotating charged black hole [Hawking (1975)]. In particular, one has again identical expressions for quantum observables at late-time both in the physical spacetime and in the spacetime of the eternal version of a black hole.

10.2.3 Bogoliubov transformations and S-matrix

Bogoliubov transformations

One can use the modes (10.2.44) and (10.2.45) to decompose the quantum field operator $\hat{\varphi}$ (see Appendix H.3)

$$\begin{aligned} \hat{\varphi} &= \sum_{\alpha}{}' \left[\hat{\mathbf{a}}_{\text{in},\alpha}\, \mathbf{\Phi}^{\text{in}}_{\alpha} + \hat{\mathbf{a}}^{\dagger}_{\text{in},\alpha}\, \bar{\mathbf{\Phi}}^{\text{in}}_{\alpha}\right] + \dots \\ &= \sum_{\alpha}{}' \left[\hat{\mathbf{a}}_{\text{out},\alpha}\, \mathbf{\Phi}^{\text{out}}_{\alpha} + \hat{\mathbf{a}}^{\dagger}_{\text{out},\alpha}\, \bar{\mathbf{\Phi}}^{\text{out}}_{\alpha}\right] + \dots . \end{aligned} \tag{10.2.48}$$

We denote by $\hat{\mathbf{a}}_{\text{in},\alpha} = (\hat{a}_{\text{in},\alpha}, \hat{a}_{\text{d},\alpha}, \hat{a}_{\text{p},\alpha})$ and $\hat{\mathbf{a}}^{\dagger}_{\text{in},\alpha} = (\hat{a}^{\dagger}_{\text{in},\alpha}, \hat{a}^{\dagger}_{\text{d},\alpha}, \hat{a}^{\dagger}_{\text{p},\alpha})$ the sets of operators corresponding to annihilation and creation of in-particles in modes *IN*, *D*, and *P*, respectively. We use matrix notation so that, e.g.,

$$\hat{\mathbf{a}}_{\text{in},\alpha}\, \mathbf{\Phi}^{\text{in}}_{\alpha} = \hat{a}_{\text{in},\alpha}\, \varphi^{\text{in}}_{\alpha} + \hat{a}_{\text{d},\alpha}\, \varphi^{\text{d}}_{\alpha} + \hat{a}_{\text{p},\alpha}\, \varphi^{\text{p}}_{\alpha} .$$

Because of the imposed normalization conditions the operators of annihilation and creation $\{\hat{a}_{\Lambda,\alpha}, \hat{a}^{\dagger}_{\Lambda,\alpha}\}$ obey the standard commutation relations (H.15)

$$\left[\hat{a}_{\Lambda,\alpha}, \hat{a}^{\dagger}_{\Lambda',\alpha'}\right] = \delta_{\Lambda,\Lambda'}\, \delta_{\alpha,\alpha'} . \tag{10.2.49}$$

Similar notations are used for out-particles. The symbol $\sum'$ in (10.2.48) indicates that the summation is performed over the late-time modes, while the remaining terms, which depend on the details of the black hole formation, are denoted by dots. The states of in- and out-vacua are defined as

$$\mathbf{a}_{\text{in},\alpha}|0;\text{in}\rangle = 0, \quad \mathbf{a}_{\text{out},\alpha}|0;\text{out}\rangle = 0 . \tag{10.2.50}$$

We do not specify early-time modes. Strictly speaking, in- and out-vacua depend on their choice, but in what follows we consider only those physical quantities (such as the flux of created particles at late times) which do not depend on the specification of early-time modes.

The important property of Wald's bases is that the Bogoliubov transformation matrices for late times factorize. The transformation (3×3)-matrices $\mathbf{A}_{\alpha}$ and $\mathbf{B}_{\alpha}$ which relate the sets of in-basis modes $\mathbf{\Phi}^{\text{in}}_{\alpha}$ and out-basis modes $\mathbf{\Phi}^{\text{out}}_{\alpha}$

$$\mathbf{\Phi}^{\text{in}}_{\alpha} = \mathbf{A}^{+}_{\alpha}\, \mathbf{\Phi}^{\text{out}}_{\alpha} - \mathbf{B}'_{\alpha}\, \bar{\mathbf{\Phi}}^{\text{out}}_{\alpha} , \tag{10.2.51}$$

are easily determined from relations (10.2.32), (10.2.35), and (10.2.41); they are

$$\mathbf{A} = \begin{pmatrix} \bar{R} & 0 & c\,\bar{t} \\ \bar{T} & 0 & c\,\bar{r} \\ 0 & c & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & -s\,\bar{t} & 0 \\ 0 & -s\,\bar{r} & 0 \\ 0 & 0 & -s \end{pmatrix}. \tag{10.2.52}$$

(We use notations $(\ldots)^+$ and $(\ldots)'$ for hermitian conjugation and transposition of matrices, respectively. In relations (10.2.52) we omit the indices α for simplicity.)

Using Wald's bases thus yields an explicit expression for the Bogoliubov transformation coefficients which determine the relation between in- and out-basis functions at large values of $n \geq N$.

S-matrix

The following expression for the S-matrix operator can be obtained using the general formulas (H.34) and (H.35)

$$\hat{S} = \left(\prod_{\alpha}{}' \hat{S}_\alpha\right) \times \hat{S}_\bullet, \quad \hat{S}_\alpha = e^{iW_\alpha^0} N_{\text{out}} \left[\exp(\hat{F}_\alpha)\right], \tag{10.2.53}$$

where

$$\hat{F}_\alpha = \frac{1}{2}\,\hat{\mathbf{a}}_{\text{out},\alpha}\,\mathbf{\Lambda}_\alpha\,\hat{\mathbf{a}}'_{\text{out},\alpha} + \hat{\mathbf{a}}^\dagger_{\text{out},\alpha}\,(\boldsymbol{M}_\alpha - \boldsymbol{I})\,\hat{\mathbf{a}}'_{\text{out},\alpha} + \frac{1}{2}\,\hat{\mathbf{a}}^\dagger_{\text{out},\alpha}\,\boldsymbol{V}_\alpha\,\hat{\mathbf{a}}^+_{\text{out},\alpha}\,, \tag{10.2.54}$$

$$\mathbf{\Lambda}_\alpha = \boldsymbol{A}_\alpha^{-1}\,\boldsymbol{B}_\alpha\,, \qquad \boldsymbol{V}_\alpha = -\,\bar{\boldsymbol{B}}_\alpha\,\boldsymbol{A}_\alpha^{-1}\,, \qquad \boldsymbol{M}_\alpha = \boldsymbol{A}_\alpha^{-1\,\prime}\,,$$

$$e^{iW_\alpha^0} = \theta\,[\det(\boldsymbol{A}_\alpha^+\,\boldsymbol{A}_\alpha)]^{-1/4}, \qquad |\theta| = 1\,. \tag{10.2.55}$$

In these relations $\dagger$ is a hermitian conjugation of operators, $\bar{()}$ is complex conjugation, and $()'$ is transposition of matrices (see Appendix H). We use compressed matrix notations. For example, in these notations

$$\hat{\mathbf{a}}^\dagger_{\text{out},\alpha}\,\boldsymbol{V}_\alpha\,\hat{\mathbf{a}}^+_{\text{out},\alpha}$$

$$= (\hat{a}^\dagger_{\text{out},\alpha}, \hat{a}^\dagger_{\text{down},\alpha}, \hat{a}^\dagger_{\text{dn},\alpha}) \begin{pmatrix} V_\alpha^{\text{out,out}} & V_\alpha^{\text{out,down}} & V_\alpha^{\text{out,dn}} \\ V_\alpha^{\text{down,out}} & V_\alpha^{\text{down,down}} & V_\alpha^{\text{down,dn}} \\ V_\alpha^{\text{dn,out}} & V_\alpha^{\text{dn,down}} & V_\alpha^{\text{dn,dn}} \end{pmatrix} \begin{pmatrix} \hat{a}^\dagger_{\text{out},\alpha} \\ \hat{a}^\dagger_{\text{down},\alpha} \\ \hat{a}^\dagger_{\text{dn},\alpha} \end{pmatrix}$$

The expression (10.2.53) implies that the S-matrix for late-time states is also factorized. We denoted by $\hat{S}_\bullet$ the part of the S-matrix which depends on details of the black hole formation, and is unimportant for us. We recall the reader that the S-matrix in (10.2.53) is expressed in terms of out-operators, and $N_{\text{out}}[\ldots]$ stands for the operation of normal ordering with respect to out-operators. This means that after expansion of the exponent into a series one must move all the creation operators to the left of the annihilation operators. In the process of reordering, the operators $\hat{\mathbf{a}}^{\dagger}_{\text{out},\alpha}$ and $\hat{\mathbf{a}}_{\text{out},\alpha}$ inside $\mathcal{N}_{\text{out}}[\ldots]$ are to be considered as commuting with one another.[11]

For completeness we present here also the explicit expressions for the matrices $\mathbf{V}$, $\mathbf{M}$, and $\mathbf{\Lambda}$:

$$\mathbf{V} = \begin{pmatrix} 0 & 0 & st/c \\ 0 & 0 & sr/c \\ st/c & sr/c & 0 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} R & 0 & t/c \\ T & 0 & r/c \\ 0 & 1/c & 0 \end{pmatrix},$$

$$\mathbf{\Lambda} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -s/c \\ 0 & -s/c & 0 \end{pmatrix}. \tag{10.2.56}$$

In these relations s_α and c_α are defined by (10.2.42), and t_α (T_α) and r_α (R_α) are absorption and reflection coefficients, respectively.

The S-matrix contains complete information concerning the quantum creation, annihilation, and scattering of particles by the black hole. The factorization property allows one to deal with each of the set α of late-time modes independently. From a mathematical point of view, this corresponds to the situation where the total Hilbert space of states of the system is the direct product of Fock spaces enumerated by α. In each of these Fock spaces we can define the α-vacuum $|0_\alpha\rangle$ as the state without α-particles; that is, the state which is annihilated by $\hat{\mathbf{a}}_\alpha$. In order to distinguish the quantities and objects corresponding to in- and out-bases, we include (as before) the corresponding indices in the notations.

The factorization property of the S-matrix implies

$$|0_\alpha; \text{in}\rangle = \hat{S}_\alpha |0_\alpha; \text{out}\rangle = e^{iW^0_\alpha} \exp[\frac{1}{2}\, \hat{\mathbf{a}}^{\dagger}_{\text{out},\alpha}\, \boldsymbol{V}_\alpha\, \hat{\mathbf{a}}^{+}_{\text{out},\alpha}] |0_\alpha; \text{out}\rangle\,. \tag{10.2.57}$$

This result directly follows from Eqs.(10.2.53)–(10.2.54). The terms containing $\mathbf{\Lambda}$ and $\mathbf{M}$ do not enter the expression in the right-hand side of (10.2.57). This happens because such terms necessarily contain at least one annihilation operator which (due to normal ordering) acts directly on the vacuum $|0_\alpha; \text{out}\rangle$. Relation (10.2.57) explicitly demonstrates that out-particles are created from the initial vacuum state. Moreover, the particles are always created in pairs, the creation of a particle in an OUT-mode being always accompanied by the creation of a particle in a DN-mode. The latter always propagates inside the black hole. For the chosen Wald's bases, the effect of particle creation can be described as the quantum analogue of parametric excitation for a system of three coupled oscillators. A similar formula which expresses $|0_\alpha; \text{out}\rangle$ in terms of a superposition of in-particles can also be easily obtained.

[11] Quite often the normal ordering operation is denoted as : : . We use the notation $\mathcal{N}_{\text{out}}[\ldots]$ to make explicit that the normal ordering is performed with respect to out-operators of creation and annihilation. Later we shall use also normal ordering with respect to in-operators, which we denote $\mathcal{N}_{\text{in}}[\ldots]$.

In quantum optics the state of a system of coupled oscillators which is connected to the vacuum by a relation similar to (10.2.57) is known as a *squeezed state*, $\hat{S}_\alpha$ being the *squeezing operator* [see, e.g., Peřina, Hradil, and Jurčo (1994)]. By using this terminology, one can say that as a result of particle production the black hole squeezes the initial vacuum state [Grishchuk and Sidorov (1990)].

Rate of particle creation by black holes

In the following sections we shall describe powerful tools which allow one to obtain explicit expressions for different characteristics of black hole radiation. Now, in order to give a simple example, we derive the rate of particle creation.

Substituting (10.2.51) into the first line of (10.2.48) and comparing the obtained result with the second line of (10.2.48), one gets

$$\begin{aligned} \hat{\mathbf{a}}_{\mathrm{out},\alpha} &= \hat{\mathbf{a}}_{\mathrm{in},\alpha}\,\mathbf{A}^{+}_{\alpha} - \hat{\mathbf{a}}^{\dagger}_{\mathrm{in},\alpha}\,\mathbf{B}^{+}, && (10.2.58)\\ \hat{\mathbf{a}}^{\dagger}_{\mathrm{out},\alpha} &= -\hat{\mathbf{a}}_{\mathrm{in},\alpha}\,\mathbf{B}^{'}_{\alpha} + \hat{\mathbf{a}}^{\dagger}_{\mathrm{in},\alpha}\,\mathbf{A}^{'}. && (10.2.59)\end{aligned}$$

The average number n^{out}_α of particles created in the mode $\varphi^{\mathrm{out}}_\alpha$ is

$$n^{\mathrm{out}}_\alpha = \langle \mathrm{in};0|\hat{a}^{\dagger}_{\mathrm{out},\alpha}\hat{a}_{\mathrm{out},\alpha}|0;\mathrm{in}\rangle = (\mathbf{B}_\alpha\,\mathbf{B}^{+}_{\alpha})_{\mathrm{out,\,out}}\,. \qquad (10.2.60)$$

In the latter expression, it is assumed that one first multiplies the matrices $\mathbf{B}$ and $\mathbf{B}^{+}$ and then takes the (out, out)–component of the product. Simple calculations show that n^{out}_α can be written in the form [Hawking (1975)]

$$n^{\mathrm{out}}_\alpha = \frac{|\,t_\alpha\,|^2}{\exp(2\pi\,\omega_\alpha/\kappa) - 1}\,. \qquad (10.2.61)$$

Here κ is the surface gravity of the black hole. It is easy to see that n^{out}_α has a form similar to the expression for the thermal emission of photons by a black body of temperature $\theta = \kappa/(2\pi)$. The additional factor in the numerator reflects the fact that only a fraction of the emitted particles equal to $|\,t_\alpha\,|^2$ reaches infinity, while the other particles are reflected back by the gravitational and centrifugal potential.

10.2.4 Rotating black holes and higher spins

So far we have considered a quantum scalar massless field in the gravitational field of a non-rotating black hole. The quantization scheme can be easily generalized to the case of a rotating black hole and massless fields of higher spins. The most important cases are electromagnetic waves ($s = 1$) and gravitational perturbations ($s = 2$). The mode expansion of the solutions for waves of these fields propagating in the Kerr metric can be found in Appendix G, while the formal scheme of quantization is given in Appendix H.4. We shall not reproduce here the details which can be found in these appendices, but only describe the main steps.

The quantization scheme for higher spins in many aspects repeats the scheme for the scalar field. The easiest way to construct the bases for the electromagnetic field and for gravitational perturbations is to begin directly with the spacetime of the eternal version of the Kerr black hole. The complete set of modes in the black hole exterior is $\{a_\mu^\Lambda, \bar{a}_\mu^\Lambda\}_{\ell m\omega P}$ for the electromagnetic field and $\{h_{\mu\nu}^\Lambda, \bar{h}_{\mu\nu}^\Lambda\}_{\ell m\omega P}$ for gravitational perturbations. Here Λ is either $\{\text{in}, \text{up}\}$ or $\{\text{out}, \text{down}\}$, and $P = \pm 1$ is an additional parameter corresponding to the two linearly independent physical polarization states of electromagnetic and gravitational waves.

When the concrete choice of the field is not important, we use the notation $\varphi_A = \varphi_A(x)$ for the field, and $\varphi_J^\Lambda = \varphi_A^\Lambda(J;x)$ for its modes, assuming that the collective index J is $\{\ell, m, \omega, P\}$. The modes φ_J^Λ contain the factor $\exp(-i\omega t)$. The explicit form of these solutions is given in Appendix G.

Modes φ_J^{in} as before are solutions that vanish on H^- and have image[12] $\sim e^{-i\omega v + im\tilde{\phi}}$ on $\mathcal{J}^-$, where $(v, r, \theta, \tilde{\phi})$ are the *Kerr ingoing coordinates* (see Appendix D.7). The *IN*-modes are normalized as $<\varphi_J^{\text{in}}, \varphi_{J'}^{\text{in}}> = \delta_{JJ'}$ provided $\omega > 0$. The *OUT*-modes are defined similarly with the obvious change of past null infinity and horizon to the future ones.

One defines φ_J^{up} as solutions that vanish on $\mathcal{J}^-$ and have the following behavior on H^-: $\varphi_J^{\text{up}}|_{H^-} \sim e^{-i\varpi u + im\tilde{\phi}_H}$. Here $(u, \theta, \tilde{\phi}_H)$ are coordinates on the past horizon H^- chosen so that $\theta, \tilde{\phi}_H = \text{const}$ on the generator, and u is the Killing time parameter along the generator. We also define $\varpi = \omega - m\,\Omega^H$, where Ω^H is the angular velocity of the black hole. The *UP*-modes are normalized as $<\varphi_J^{\text{up}}, \varphi_{J'}^{\text{up}}> = \delta_{JJ'}$ provided $\varpi > 0$. The *DOWN*-modes are defined similarly with the obvious change of past null infinity and horizon to the future ones. The so-introduced *IN*-, *UP*-, *OUT*-, and *DOWN*-modes are identical to the modes we discussed in Section 4.2.

The main problem in the case of a rotating black hole is connected with superradiant modes for which $\omega\varpi < 0$. For example, consider equation (10.2.32) with $0 < \omega < m\,\Omega_H$. Both of the modes *IN* and *OUT* with $\omega > 0$ considered on H^+ have *negative* frequency ϖ, while *DOWN*-modes are defined only for positive ϖ. One cannot simply change φ_J^{down} to $\bar{\varphi}_J^{\text{down}}$ since the latter function has a wrong dependence on v and $\tilde{\phi}_H$. The trick that does work is to use for the superradiant modes in (10.2.32) not φ_J^{down} but $\bar{\varphi}_{\tilde{J}}^{\text{down}}$. Here and later $\tilde{J}$ denotes the collective index $\{\ell, -m, -\omega, P\}$. To make our formulas valid for both superradiant and non-superradiant cases simultaneously, we introduce the notation

$$\tilde{\varphi}_J^\bullet = \begin{cases} \varphi_J^\bullet & \text{if} \quad \omega\varpi > 0 \,, \\ \bar{\varphi}_{\tilde{J}}^\bullet & \text{if} \quad \omega\varpi < 0 \,. \end{cases} \tag{10.2.62}$$

Here the symbol $\bullet$ stands for any index specifying the type of modes. Using this

[12] We focus our attention on the dependence of solutions on time t and angular ϕ coordinates since only this dependence is important for us at the moment.

notation, the relations (10.2.32) and (10.2.35) for the Kerr geometry take the form

$$\varphi_J^{\text{in}} = R_J\, \varphi_J^{\text{out}} + T_J\, \tilde{\varphi}_J^{\text{down}}\,, \tag{10.2.63}$$

$$\tilde{\varphi}_J^{\text{up}} = t_J\, \varphi_J^{\text{out}} + r_J\, \tilde{\varphi}_J^{\text{down}}\,. \tag{10.2.64}$$

To define *DN*-modes, one writes the Kerr metric in the Kruskal-like coordinates (U, V, θ, ϕ_+) [see (D.73)]. The Kerr metric in these coordinates, (D.74), is invariant under the isometry transformation: $U \to -U$, $V \to -V$. This transformation maps the region I onto I' (see Figure 6.4). We define φ_J^{dn} by means of relation (H.54). Finally, we define

$$\varphi_J^{\text{d}} = c_J\, \varphi_J^{\text{dn}} + s_J\, \bar{\tilde{\varphi}}_J^{\text{up}}\,, \tag{10.2.65}$$

$$\varphi_J^{\text{p}} = c_J\, \varphi_J^{\text{up}} + s_J\, \bar{\tilde{\varphi}}_J^{\text{dn}}\,, \tag{10.2.66}$$

where

$$c_J = [1 - \exp(-2\pi\varpi\sigma_J/\kappa)]^{-1/2}\,, \qquad s_J = [\exp(2\pi\varpi\sigma_J/\kappa) - 1]^{-1/2}\,. \tag{10.2.67}$$

Each of the modes *UP*, *DN*, *D*, and *P* that enter relations (10.2.65) and (10.2.66) are defined for $\varpi > 0$. By using the Lemma proved in Appendix H.4.2, one can show that the *D*- and *P*-modes are of positive frequency with respect to the affine parameter U on H^-.

Equations (10.2.64)–(10.2.66) allow one to write

$$\tilde{\varphi}_J^{\text{d}} = s_J\, \bar{t}_J\, \bar{\tilde{\varphi}}_J^{\text{out}} + s_J\, \bar{r}_J\, \bar{\tilde{\varphi}}_J^{\text{down}} + c_J\, \tilde{\varphi}_J^{\text{dn}}\,, \tag{10.2.68}$$

$$\tilde{\varphi}_J^{\text{p}} = c_J\, t_J\, \varphi_J^{\text{out}} + c_J\, r_J\, \tilde{\varphi}_J^{\text{down}} + s_J\, \bar{\tilde{\varphi}}_J^{\text{dn}}\,. \tag{10.2.69}$$

All the above relations can be rewritten in the "wavepacket-like" form. One can simply substitute $J \to \alpha$ provided the parameter ϵ that enters definition (10.2.22) of these wavepackets is chosen small enough so that the coefficients R_J, T_J, c_J, and s_J are practically constant in the frequency interval $(\omega, \omega + \epsilon)$.

For non-superradiant modes we use the sets (10.2.44) and (10.2.45) as in- and out-bases. For superradiant modes the definition is slightly different. Namely, for *D*-, *P*-, *DOWN*-, and *DN*-modes instead of φ_α one uses $\varphi_{\tilde{\alpha}}$; that is, wavepackets constructed from $\varphi_{\tilde{J}}$. For this definition of in- and out-bases we again have factorization. The field operator $\hat{\varphi}$ can be written as

$$\hat{\varphi} = \sum_J \left[\hat{\mathrm{a}}_{\text{in},J}\, \Phi_J^{\text{in}} + \hat{\mathrm{a}}_{\text{in},J}^{\dagger}\, \bar{\Phi}_J^{\ \text{in}}\right] = \sum_J \left[\hat{\mathrm{a}}_{\text{out},J}\, \Phi_J^{\text{out}} + \hat{\mathrm{a}}_{\text{out},J}^{\dagger}\, \bar{\Phi}_J^{\ \text{out}}\right]. \tag{10.2.70}$$

where $\{\hat{\mathrm{a}}_{\text{in},J}, \hat{\mathrm{a}}_{\text{in},J}^{\dagger}\}$ $(\{\hat{\mathrm{a}}_{\text{out},J}, \hat{\mathrm{a}}_{\text{out},J}^{\dagger}\})$ are operators of annihilation and creation of in- (out-) particles. The in-vacuum defined by relations

$$\hat{\mathrm{a}}_{\text{in},J}|0;\text{in}\rangle = 0 \tag{10.2.71}$$

is called the *Unruh vacuum* state.

The Bogoliubov matrices $\mathbf{A}$ and $\mathbf{B}$ for non-superradiant modes coincide with (10.2.52), while for superradiant modes ($\sigma \equiv \mathrm{sign}(\omega\varpi) < 0$) they have a slightly different form (see Frolov (1986) and Appendix H.5)

$$\mathbf{A}^{\sigma<0} = \begin{pmatrix} \bar{R} & s\bar{t} & 0 \\ 0 & 0 & cr \\ 0 & c & 0 \end{pmatrix}, \quad \mathbf{B}^{\sigma<0} = \begin{pmatrix} 0 & 0 & -c\bar{t} \\ -T & -sr & 0 \\ 0 & 0 & -s \end{pmatrix}. \tag{10.2.72}$$

The S-matrix is given by expressions (10.2.53)–(10.2.55), where for non-superradiant modes matrices $\mathbf{V}$, $\mathbf{M}$, and $\mathbf{\Lambda}$ are given by (10.2.56), while for superradiant modes they are [cf. (H.91), Frolov (1986)]

$$\mathbf{\Lambda}^{\sigma<0} = \begin{pmatrix} 0 & 0 & t/(cr) \\ 0 & 0 & -s/c \\ t/(cr) & -s/c & 0 \end{pmatrix},$$

$$\mathbf{V}^{\sigma<0} = \begin{pmatrix} 0 & t/r & 0 \\ t/r & 0 & s/(cr) \\ 0 & s/(cr) & 0 \end{pmatrix}, \tag{10.2.73}$$

$$\mathbf{M}^{\sigma<0} = \begin{pmatrix} 1/\bar{R} & 0 & 0 \\ 0 & 0 & 1/(cr) \\ -s\bar{t}/(c\bar{R}) & 1/c & 0 \end{pmatrix}.$$

To summarize, we have an expression for the S-matrix that is valid for the Kerr geometry and massless fields of spin s=0,1,2. The generalization to the charged case is straightforward. Similar expressions can be obtained for massless fermions [neutrino ($s = 1/2$), gravitino ($s = 3/2$)] [Bolashenko and Frolov (1989)].

10.3 Density Matrix and Generating Functional

10.3.1 Density matrix

The next step is the calculation of the density matrix that describes black hole radiation. Suppose we are interested in expectation values of the type $\langle\hat{F}\rangle = \langle 0;\mathrm{in}|\hat{F}|0;\mathrm{in}\rangle$, where $\hat{F} = F\left(\hat{a}^{\dagger}_{\mathrm{out},\alpha}, \hat{a}_{\mathrm{out},\alpha}\right)$ is an operator depending only on the operators of creation and annihilation of the *OUT*-particles in late-time modes α. These quantities describe observables registered by a distant observer some time after black hole formation. In order to calculate $\langle\hat{F}\rangle$, it is convenient to express $|0;\mathrm{in}\rangle$ in terms of out-operators of creation and annihilation acting on the out-vacuum

$$|0;\mathrm{in}\rangle = \hat{S}|0;\mathrm{out}\rangle, \quad \langle 0;\mathrm{in}| = \langle 0;\mathrm{out}|\hat{S}^{+}, \tag{10.3.1}$$

where $\hat{S}$ is the S-matrix operator (10.2.53). Expressed in this form, $|0;\text{in}\rangle$ is evidently a superposition of many-particle states of out-particles. In our problem these particles are naturally divided into two categories. Particles of the first category (called "visible") are those that are created by $\hat{a}^\dagger_{\text{out},\alpha}$ operators. Particles of the other category (called "invisible") are those that are created by $\hat{a}^\dagger_{\text{down},\alpha}$ and $\hat{a}^\dagger_{\text{dn},\alpha}$ operators. Only "visible" particles reach $\mathcal{J}^+$ and contribute to $\langle \hat{F} \rangle$. This means that we are effectively dealing with a subsystem of the total system, and in accordance with the general rules of quantum mechanics, the state of the subsystem is described by a *density matrix*. In the application to our particular problem this means that (see Appendix H.6)

$$\langle \hat{F} \rangle \equiv \langle 0;\text{in}|\hat{F}\,|0;\text{in}\rangle = \text{Tr}_{\text{out}}(\hat{\rho}\hat{F})\,. \tag{10.3.2}$$

Here $\hat{\rho}$ is the density matrix that depends only on the operators of creation and annihilation of OUT-particles. The symbol Tr_{out} indicates that tracing is over only OUT-particle states. It is possible to show that the density matrix $\hat{\rho}$ does not depend on the choice of the basis for "invisible" states. Namely, it is invariant under the Bogoliubov transformations in the subspace of "invisible" particles (see Appendix H.6).

In the general case, the density matrix $\hat{\rho}$ that describes observables on $\mathcal{J}^+$ depends on the finer details of black hole formation. However, if we are interested only in the values of observables on $\mathcal{J}^+$ at sufficiently late retarded time, these details are found to be unimportant, so that the values assumed by these observables depend only on the parameters of the resulting stationary black hole. The density matrix $\hat{\rho}$ can be written in the form

$$\hat{\rho} = \left(\prod_\alpha{}'\hat{\rho}_\alpha\right) \times \hat{\rho}_\bullet\,. \tag{10.3.3}$$

As before, the prime indicates that the product must be taken only over late-time modes α with $n \geq N \gg 1$, while $\hat{\rho}_\bullet$ stands for the part of the density-matrix depending on the details of the black hole formation. The factorizability of the late-time part of the density matrix is a consequence of the factorization property of the S-matrix. The calculations show [Frolov (1983a, 1986)] that the operators $\hat{\rho}_\alpha$ have the same form for both superradiant and non-superradiant modes:

$$\hat{\rho}_\alpha = Q_\alpha\, N_{\text{out}}\left[\exp[-\,Q_\alpha\,\hat{n}_\alpha\,]\right] = Q_\alpha \exp[\ln(1-Q_\alpha)\,\hat{n}_\alpha]\,, \tag{10.3.4}$$

where $\hat{n}_\alpha = \hat{a}^\dagger_{\text{out},\alpha}\,\hat{a}_{\text{out},\alpha}$, and

$$Q_\alpha = \frac{1-w_\alpha^2}{1-w_\alpha^2\,|\,R_\alpha\,|^2}\,, \qquad w_\alpha = \exp(-\pi\varpi_\alpha/\kappa)\,. \tag{10.3.5}$$

In the expression (10.3.4) $N_{\text{out}}[\ldots]$ denotes the operation of normal ordering with respect to the operators $\hat{a}_{\text{out},\alpha}$ and $\hat{a}^\dagger_{\text{out},\alpha}$. The last equality in (10.3.4) takes into account the well-known relation

$$N\left[\exp\,\tau\hat{a}^\dagger\,\hat{a}\,\right] = \exp\left[\,\ln(1+\tau)\,\hat{a}^\dagger\,\hat{a}\right] \tag{10.3.6}$$

[see, e.g., Klauder and Sudarshan (1968)]. Note that for both superradiant and non-superradiant modes $Q_\alpha > 0$.

Another useful representation of the density matrix $\hat{\rho}$, which directly follows from (10.3.4), is

$$\hat{\rho}_\alpha = \sum_{n_\alpha} \rho_{n_\alpha} |n_\alpha\rangle\langle n_\alpha| , \tag{10.3.7}$$

where $|n_\alpha\rangle = (n!)^{-1/2} (\hat{a}^\dagger_{\text{out},\alpha})^{n_\alpha} |0_\alpha\rangle$ is a state with n_α particles in the OUT-mode α, and

$$\rho_{n_\alpha} = \frac{(w_\alpha^{-2} - 1)(\Gamma_\alpha)^{n_\alpha}}{(w_\alpha^{-2} - 1 + \Gamma_\alpha)^{n_\alpha+1}} . \tag{10.3.8}$$

Here $\Gamma_\alpha = \sigma_\alpha |T_\alpha|^2 \equiv 1 - |R_\alpha|^2$, and $w_\alpha = \exp(-\pi\varpi_\alpha/\kappa)$. The expression (10.3.8) for non-rotating black holes was derived by Hawking (1976b). A similar expression (10.3.7) with

$$\rho_{n_\alpha} = (\Gamma_\alpha)^{n_\alpha} (\mathrm{e}^{2\pi\omega/\kappa} - \varepsilon)^\varepsilon (\mathrm{e}^{2\pi\omega/\kappa} - \varepsilon + \varepsilon\,\Gamma_\alpha)^{-(n_\alpha+\varepsilon)} \tag{10.3.9}$$

is valid for $\hat{\rho}_\alpha$ in the case of fermions as well, but ε (which for bosons is 1) then equals -1, and n_α can assume only the values 0 and 1 [Wald (1975)].

If one neglects the scattering of modes by the gravitational field and puts $R_\alpha = 0$, expression (10.3.4) becomes

$$\hat{\rho}_\alpha = \mathrm{e}^{-\hat{\mathcal{H}}_\alpha/\theta} , \tag{10.3.10}$$

where $\hat{\mathcal{H}}_\alpha$ is the free Hamiltonian describing OUT-particles escaping to $\mathcal{J}^+$:

$$\hat{\mathcal{H}}_\alpha = \omega_\alpha\, \hat{a}^\dagger_{\text{out},\alpha}\, \hat{a}_{\text{out},\alpha} , \tag{10.3.11}$$

and

$$\theta = \kappa/2\pi \tag{10.3.12}$$

is the *Hawking temperature* of the black hole.

10.3.2 Generating functional

Method of generating functional

Expression (10.3.4) for the density matrix $\hat{\rho}$ makes it possible to calculate the expectation values of the observables on $\mathcal{J}^+$. The calculations are usually straightforward but may be quite long. Moreover, for each new observable the main part of these calculations must be redone. There exists a method, known in quantum field theory and quantum statistics as the method of the *generating functional*, which permits substantial simplification of the calculations for an important class of observables.

We first illustrate this method by simple examples and afterwards describe its applications to studying quantum effects in black holes. Suppose we need to calculate the average number of particles in a state described by some density matrix $\hat{\rho}$

$$\langle \hat{n} \rangle = \mathrm{Tr}\left(\hat{a}^\dagger \hat{a}\ \hat{\rho}\right). \tag{10.3.13}$$

Instead of this let us calculate the following quantity (a *generating function*)

$$Z(\psi, \bar{\psi}) = \mathrm{Tr}\left[\hat{K}(\psi, \bar{\psi})\, \hat{\rho}\right], \tag{10.3.14}$$

where $\hat{K}(\psi, \bar{\psi}) = \exp(\psi \hat{a}^\dagger) \exp(\bar{\psi} \hat{a})$. Note now that a differentiation with respect to the parameter ψ in (10.3.13) moves the operators $\hat{a}^\dagger$ down from the exponent, while a differentiation with respect to the parameter $\bar{\psi}$ moves down the operators $\hat{a}$. Then we have

$$\langle \hat{n} \rangle = \left[\frac{\partial^2 Z}{\partial \psi\, \partial \bar{\psi}}\right]_{\psi = \bar{\psi} = 0}. \tag{10.3.15}$$

Similarly, if the operator $\hat{F}$ is represented in the normal form

$$\hat{F} = \sum_{n,m} F_{nm}\, (\hat{a}^\dagger)^n\, (\hat{a})^m, \tag{10.3.16}$$

then

$$\langle \hat{F} \rangle = \left[F(\partial_\psi, \partial_{\bar{\psi}})\, Z(\psi, \bar{\psi})\right]_{\psi = \bar{\psi} = 0}, \tag{10.3.17}$$

where

$$F(\partial_\psi, \partial_{\bar{\psi}}) = \sum_{n,m} F_{nm}\, \frac{\partial^{n+m}}{\partial \psi^n\, \partial \bar{\psi}^m}. \tag{10.3.18}$$

This means that it is sufficient to calculate just once the quantum average of a special operator $\hat{K}$ depending on parameters. All other averages for various observables can be obtained by partial differentiation of the generating function with respect to its parameters.

This method can be generalized. For example, one can consider the function

$$Z(\mu) = \mathrm{Tr}(e^{-\mu \hat{n}} \hat{\rho}), \tag{10.3.19}$$

with $\hat{n} = \hat{a}^\dagger\, \hat{a}$. Then a subsequent differentiation of the generating function $Z(\mu)$ with respect to μ gives the averages $\langle \hat{n} \rangle$, $\langle\, \hat{n}^2 \rangle$, and so on. A similar procedure can be used for the calculation of averages

$$\langle m; \mathrm{in} | \hat{F}\, | m; \mathrm{in} \rangle. \tag{10.3.20}$$

One begins with the generating function

$$Z(\gamma, \bar{\gamma}) = \langle 0; \text{in}| e^{\bar{\gamma}\hat{a}_{\text{in}}} \hat{F} e^{\gamma \hat{a}^{\dagger}_{\text{in}}} |0; \text{in}\rangle . \qquad (10.3.21)$$

The average (10.3.20) can be easily obtained by partial differentiation of $Z(\gamma, \bar{\gamma})$ with respect to γ and $\bar{\gamma}$.

To summarize, if one knows the generating function, then the calculation of a wide class of matrix elements is reduced to simple differentiation with respect to parameters which are the arguments of the generating function. In the application to a system with an infinite number of degrees of freedom the generating function depends on an infinite number of parameters, and it is usually called a *generating functional*.

Generating functional for quantum effects in black holes

We return now to our problem: the construction of the generating functional for quantum effects in black holes. Let us consider the following function

$$\begin{aligned} Z_\alpha &\equiv Z[\psi_\alpha, \mu_\alpha; \gamma_\alpha] \\ &= \langle 0_\alpha; \text{in}| \exp(\bar{\gamma}_\alpha \hat{a}_{\text{in},\alpha}) N_{\text{out}} \Big[\exp\Big(-\mu_\alpha \hat{a}^{\dagger}_{\text{out},\alpha} \hat{a}_{\text{out},\alpha} \\ &+ \psi_\alpha \hat{a}^{\dagger}_{\text{out},\alpha} + \bar{\psi}_\alpha \hat{a}_{\text{out},\alpha}\Big)\Big] \exp(\gamma \hat{a}^{\dagger}_{\text{in},\alpha})|0_\alpha; \text{in}\rangle . \end{aligned} \qquad (10.3.22)$$

We define the *generating functional for quantum effects in a black hole*

$$Z[\psi, \mu; \gamma] = \prod_\alpha{}' Z_\alpha . \qquad (10.3.23)$$

Here, as before, the prime indicates that the product is taken over the late-time modes. The functional Z depends on the three infinite sets of parameters: μ_α, $\{\psi_\alpha, \bar{\psi}_\alpha\}$, and $\{\gamma_\alpha, \bar{\gamma}_\alpha\}$, enumerated by the collective index α. The explicit expression for Z_α can be calculated and has the following form [Frolov (1983a, 1986)]:

$$Z_\alpha = \frac{Q_\alpha}{C_\alpha} \exp\left(\frac{P_\alpha}{C_\alpha}\right) . \qquad (10.3.24)$$

where,

$$Q_\alpha = (1 - w_\alpha^2)\left(1 - w_\alpha^2 \,|\, R_\alpha \,|^2\right)^{-1} , \qquad (10.3.25)$$

$$C_\alpha = Q_\alpha + \mu_\alpha - Q_\alpha \mu_\alpha , \qquad (10.3.26)$$

$$\begin{aligned} P_\alpha = &\left(C_\alpha - \mu_\alpha Q_\alpha \,|\, R_\alpha \,|^2\right) \bar{\gamma}_\alpha \gamma_\alpha + Q_\alpha R_\alpha \gamma_\alpha \bar{\psi}_\alpha \\ &+ Q_\alpha \bar{R}_\alpha \bar{\gamma}_\alpha \psi_\alpha + (1 - Q_\alpha)\bar{\psi}_\alpha \psi_\alpha . \end{aligned} \qquad (10.3.27)$$

Here R_α is the reflection amplitude, and w_α is defined in (10.3.5).

How to use the generating functional

We describe now some general rules which establish the relation of the generating functional $Z[\psi, \mu; \gamma]$ to the main quantities of physical interest, which characterize quantum effects in black holes.

We use the following notations:

$$D_\beta^{l_\beta} = \frac{1}{(l_\beta)!} \frac{\partial^{l_\beta}}{\partial \gamma_\beta^{l_\beta}} \frac{\partial^{l_\beta}}{\partial \bar{\gamma}_\beta^{l_\beta}}, \qquad \Delta_\beta^{k_\beta} = \frac{1}{(k_\beta)!} \frac{\partial^{k_\beta}}{\partial \psi_\beta^{k_\beta}} \frac{\partial^{k_\beta}}{\partial \bar{\psi}_\beta^{k_\beta}}, \tag{10.3.28}$$

and

$$|\beta_1, \dots, \beta_m; \text{in}\rangle = \hat{a}^\dagger_{\text{in},\beta_1} \dots \hat{a}^\dagger_{\text{in},\beta_m} |0; \text{in}\rangle . \tag{10.3.29}$$

Let $\hat{F} = F[\hat{a}^\dagger_{\text{out}}, \hat{a}_{\text{out}}]$ be an operator given in the normal form

$$F[\hat{a}^\dagger_{\text{out}}, \hat{a}_{\text{out}}] = \sum_{m,m'} \sum_{\substack{\alpha_1,\dots,\alpha_m \\ \alpha'_1,\dots,\alpha'_{m'}}} F^{mm'}_{\alpha_1,\dots,\alpha_m;\alpha'_1,\dots,\alpha'_{m'}} \hat{a}^\dagger_{\text{out},\alpha_1} \dots \hat{a}^\dagger_{\text{out},\alpha_m} \hat{a}_{\text{out},\alpha'_1} \dots \hat{a}_{\text{out},\alpha'_{m'}}, \tag{10.3.30}$$

then we define

$$F\left[\partial_\psi, \partial_{\bar{\psi}}\right] = \sum_{m,m'} \sum_{\substack{\alpha_1,\dots,\alpha_m \\ \alpha'_1,\dots,\alpha'_{m'}}} F^{mm'}_{\alpha_1,\dots,\alpha_m;\alpha'_1,\dots,\alpha'_{m'}} \frac{\partial^{m'}}{\partial \psi_{\alpha_1} \dots \partial \psi_{\alpha_m}} \frac{\partial^m}{\partial \bar{\psi}_{\alpha'_1} \dots \partial \bar{\psi}_{\alpha'_{m'}}}. \tag{10.3.31}$$

Theorem 1: *Let $\hat{A} = A(\hat{n}_{\alpha_1}, \dots, \hat{n}_{\alpha_k})$ be a function of the operators $\hat{n}_\alpha \equiv \hat{a}^\dagger_{\text{out},\alpha} \hat{a}_{\text{out},\alpha}$ Then*

$$\langle \beta_1, \dots, \beta_m; \text{in}| A(\hat{n}_{\alpha_1}, \dots, \hat{n}_{\alpha_k}) |\beta_1, \dots, \beta_m; \text{in}\rangle = \left\{ D^1_{\beta_1} \dots D^1_{\beta_m} A\left(\frac{\partial}{\partial \lambda_{\alpha_1}}, \dots, \frac{\partial}{\partial \lambda_{\alpha_k}}\right) Z[\psi = 0, \mu = 1 - e^\lambda; \gamma] \right\}_{\lambda=\gamma=0} . \tag{10.3.32}$$

Theorem 2: *Let $F[\hat{a}^\dagger_{\text{out}}, \hat{a}_{\text{out}}]$ be an operator given in the normal form (10.3.30). Then one has*

$$\langle \beta_1, \dots, \beta_m; \text{in}| F(\hat{a}^\dagger_{\text{out},\alpha}, \hat{a}_{\text{out},\alpha}) |\beta_1, \dots, \beta_m; \text{in}\rangle = \left\{ D^1_{\beta_1} \dots D^1_{\beta_m} F\left[\partial_\psi, \partial_{\bar{\psi}}\right] Z[\psi, \mu = 0; \gamma] \right\}_{\psi=\gamma=0}, \tag{10.3.33}$$

where $F\left[\partial_{\psi}, \partial_{\bar{\psi}}\right]$ *is given by (10.3.31).*

Theorem 3: *Let* $P(k_{\alpha_1}, \ldots, k_{\alpha_n} | l_{\alpha_1}, \ldots, l_{\alpha_n})$ *be the probability that* k_{α_1} *quanta are emitted by the black hole to the infinity* $\mathcal{J}^+$ *in the mode* $\varphi^{\text{out}}_{\alpha_1}, \ldots,$ *and* k_{α_n} *quanta in the mode* $\varphi^{\text{out}}_{\alpha_n}$, *assuming there were initially* l_{α_1} *quanta incoming from* $\mathcal{J}^-$ *in the mode* $\varphi^{\text{in}}_{\alpha_1}, \ldots,$ *and* l_{α_n} *quanta in the mode* $\varphi^{\text{in}}_{\alpha_n}$ $(k_{\alpha_i} \geq 0,\ l_{\alpha_i} \geq 0,\ k_{\alpha_i} + l_{\alpha_i} > 0)$. *The following equalities then hold:*

$$P(k_{\alpha_1}, \ldots, k_{\alpha_n} | l_{\alpha_1}, \ldots, l_{\alpha_n}) = P(k_{\alpha_1} \,|\, l_{\alpha_1}) \ \ldots \ P(k_{\alpha_n} \,|\, l_{\alpha_n}), \tag{10.3.34}$$

$$P(k_\alpha \,|\, l_\alpha) = (D^{l_\alpha}_\alpha \, \Delta^{k_\alpha}_\alpha \, Z_\alpha[\psi_\alpha, 1; \gamma_\alpha])_{\gamma_\alpha = \psi_\alpha = 0} \,. \tag{10.3.35}$$

The proof of these theorems can be found in [Frolov (1986)]. The theorems allow one to calculate the expectation values of observables, correlation functions, and probability distributions for processes in the field of a black hole by means of differentiation of the generating functional Z defined by (10.3.23)-(10.3.27) with respect to its parameters. Note that the generating functional is completely determined if in addition to the surface gravity κ and angular velocity Ω^H of the black hole, one also knows the coefficients of reflection, R_α, of wavepackets by the black hole. A similar generating functional for fermions can be found in [Bolashenko and Frolov (1988, 1989)].

10.4 Particular Cases

10.4.1 Hawking effect

In this and the following sections we consider different characteristics of the quantum radiation of black holes and demonstrate how these characteristics can be obtained by using the theorems of the previous section.

Consider first the rate of particle production. Let a quantum field occupy the vacuum state before the black hole was formed. After the formation, the black hole becomes a source of radiation. The mean number of particles radiated in the *OUT*-mode α and recorded by a distant observer is given by the following expression:

$$\langle \hat{n}_\alpha \rangle_0 \equiv \langle 0; \text{in} | \hat{n}_\alpha | 0; \text{in} \rangle$$

$$= \left\{ \frac{\partial}{\partial \lambda_\alpha} Z_\alpha \left[0, 1 - e^{\lambda_\alpha}; 0\right] \right\}_{\lambda_\alpha = 0} = \frac{\Gamma_\alpha}{\exp(\varpi_\alpha / \theta) - 1}, \tag{10.4.1}$$

where $\hat{n}_\alpha = \hat{a}^\dagger_{\text{out},\alpha} \hat{a}_{\text{out},\alpha}$, $\Gamma_\alpha = \sigma_\alpha \,|\, T_\alpha \,|^2$, $\varpi_\alpha = \omega_\alpha - m_\alpha \Omega^H$, and $\theta = k/2\pi$ is the black hole temperature. For a non-rotating black hole (10.4.1) reproduces the earlier obtained expression (10.2.61). Note that if the black hole is rotating, then even though the denominator of (10.4.1) vanishes at $\varpi_\alpha = 0$, the mean number of particles created in such modes remains finite because of the simultaneously vanishing of $|T_\alpha|^2$. It is also clear that reversal of the sign of the denominator at $\varpi_\alpha < 0$ is accompanied by reversal of sign of σ_α so that, on the whole, expression (10.4.1) is always positive.

As the absorption coefficient of the packet $\varphi_{\text{in},\alpha}$ of a stationary black hole is independent of the moment of time $v = 2\pi n/\epsilon$ at which this packet was emitted, the number of particles reaching $\mathcal{J}^+$ is independent of retarded time u. In other words, the new-born black hole becomes a source of a stationary flux of radiation.[13] We have already mentioned above that the appearance of a stationary flux can be interpreted as a corollary of a spontaneous process of particle pair creation in a stationary gravitational field close to the event horizon. One of the particles of each created pair sinks into the black hole, while the other escapes to infinity.

One can obtain more detailed information about this process by calculating the matrix elements $\langle 1_{\Lambda_1\alpha} 1_{\Lambda_2\alpha}; \text{out}|0; \text{in}\rangle$. Here $\Lambda \in \{\text{out}, \text{down}, \text{dn}\}$ is the type of the out-particle. These matrix elements are amplitudes of probability of creation of a pair of out-particles from the in-vacuum state. Because of the factorization property of the S-matrix the both of the created particles must be in the same mode α. Simple calculations using (10.2.57) give

$$\langle 1_{\Lambda_1} 1_{\Lambda_2}; \text{out}|0; \text{in}\rangle = e^{iW^0}\, V^{\Lambda_1\Lambda_2}\,.$$

Here the phase factor $\exp(iW^0)$ is given by (10.2.55), and $V^{\Lambda_1\Lambda_2}$ are the elements of the matrix $\mathbf{V}$. (For simplicity, we omit the mode index α). Using expression (10.2.56) for $\mathbf{V}$, we obtain that for non-superradiant modes ($\sigma < 0$) the only non-vanishing matrix elements are

$$V^{\text{out, dn}} = \frac{s}{c}\,t \quad \text{and} \quad V^{\text{down, dn}} = \frac{s}{c}\,r\,.$$

In the absence of scattering ($r = 0$) particles are created as the correlated pairs. The creation of an *OUT*-particle is accompanied by the creation of a *DN*-particle. The latter has negative energy and always propagates inside the black hole. In the presence of scattering ($r \neq 0$) besides this process there is also the creation of (*DOWN-DN*)-pairs.

For superradiant modes the process of particle creation is slightly different. Namely, the only non-vanishing matrix elements of $\mathbf{V}$ are

$$V^{\text{out, down}} = \frac{t}{r} \quad \text{and} \quad V^{\text{down, dn}} = \frac{s}{c}\,\frac{1}{r}\,;$$

[13] Unruh (1981) pointed out that the phenomenon of quantum creation of particles in black holes permits a hydrodynamic analogy. If the flow of the fluid is such that a closed two-dimensional surface separates the region of subsonic and supersonic flow (Laval nozzle), the system (a *dumb hole*) must radiate thermal-spectrum phonons.

see (10.2.73). Thus, the creation of an *OUT*-particle is always accompanied by the creation of a *DOWN*-particle. Both of these particles can propagate in the black hole exterior. This process is possible because *DOWN*-particles are created inside the ergosphere and have negative energy. The creation of a *DOWN*-particle is accompanied by the creation of a *DN*-particle inside the black hole.

10.4.2 Stimulated radiation

Let m particles in the mode α fall on a black hole from $\mathcal{J}^-$. Then we make use of (10.3.32) and obtain the following expression for the mean number of particles in the state α at $\mathcal{J}^+$:

$$\begin{aligned} \langle \hat{n}_\alpha \rangle_m &= \langle m_\alpha ; \text{in} | \hat{n}_\alpha | m_\alpha ; \text{in} \rangle \\ &= \{ D_\alpha^{m_\alpha} [(-1 + Q_\alpha^{-1} + | R_\alpha |^2 \bar{\gamma}_\alpha \gamma_\alpha) e^{\bar{\gamma}_\alpha \gamma_\alpha}] \}_{\gamma=0} \\ &= \langle \hat{n}_\alpha \rangle_0 + m_\alpha | R_\alpha |^2 . \end{aligned} \tag{10.4.2}$$

The term $\langle \hat{n}_\alpha \rangle_0$ describes the spontaneous creation of particles from the vacuum; it is given by expression (10.4.1). Bekenstein and Meisels (1977) noticed that the resulting expression can be rewritten in the equivalent form

$$\langle \hat{n}_\alpha \rangle_m = A_\alpha^{\downarrow} + B_\alpha^{\downarrow} m_\alpha + (1 - B_\alpha^{\uparrow}) m_\alpha , \tag{10.4.3}$$

and the quantities

$$A_\alpha^{\downarrow} = B_\alpha^{\downarrow} = \langle \hat{n}_\alpha \rangle_0 = \frac{\Gamma_\alpha}{\exp(\varpi_\alpha/\theta) - 1} , \tag{10.4.4}$$

$$B_\alpha^{\uparrow} = \frac{\Gamma_\alpha}{1 - \exp(-\varpi_\alpha/\theta)} \tag{10.4.5}$$

can be interpreted as analogues of Einstein's coefficients for processes in black holes. The term $A_\alpha^{\downarrow}$ describes the spontaneous creation of particles from the vacuum, $B_\alpha^{\uparrow}$ acts as the absorption coefficient, and $B_\alpha^{\downarrow} m_\alpha$ describes the *stimulated emission* of particles in the mode α.[14]

[14]Expression (10.4.2) for the number of particles emitted by a black hole can be easily generalized to the case where besides incoming particles in an *IN*-mode there exist also initial particles in the *D*- and *P*-modes. For a non-rotating black hole the corresponding expression is

$$\langle \hat{n}_\alpha \rangle = \langle \hat{n}_\alpha \rangle_0 + m_\alpha | R_\alpha |^2 + \frac{| t_\alpha |^2}{\exp(2\pi\omega_\alpha/\kappa) - 1} \left[\exp(2\pi\omega_\alpha/\kappa)\, n_\alpha^{\mathrm{p}} + n_\alpha^{\mathrm{d}} \right] ,$$

where n_α^{d} and n_α^{p} are the number of incoming particles in *D*- and *P*-modes, respectively. This expression implies that the number of outgoing particles depends also on n_α^{d} and n_α^{p}. This effect is sometimes also called stimulated emission [see e.g., Müller and Lousto (1994)]. Note however, that in order to change the number of outgoing particles at frequency ω, one must send in a wavepacket in either a *D*- or a *P*-mode with the frequency $\sim \omega \exp \kappa u$, where u is the moment when the *OUT*-particle reaches $\mathcal{J}^+$. This exponentially large blueshift makes such stimulated emission processes virtually impossible.

The quantity $1 - B_\alpha^\uparrow$ can be interpreted as the probability of scattering of the mode α by the black hole. Using (10.4.2) and (10.4.3), we rewrite the expression for $|R_\alpha|^2$ in the form

$$|R_\alpha|^2 = (1 - B_\alpha^\uparrow) + B_\alpha^\downarrow . \tag{10.4.6}$$

This formula shows that the quantity $|R_\alpha|^2$, characterizing the scattering of a particle in mode α by the black hole, is the sum of the probability of scattering of the incoming particle, $1 - B_\alpha^\uparrow$, and the probability of stimulated emission of radiation in this mode, $B_\alpha^\downarrow$. The coefficients $B_\alpha^\uparrow$ and $B_\alpha^\downarrow$ are related as

$$B_\alpha^\downarrow = B_\alpha^\uparrow \exp(-\varpi_\alpha/\theta) . \tag{10.4.7}$$

The coefficient of stimulated emission for superradiant modes, $B_\alpha^\downarrow$, *exceeds* the absorption coefficient $B_\alpha^\uparrow$, so that $|R_\alpha|^2$ takes on values greater than unity.

Formulas (10.4.3)–(10.4.5) show that black hole radiation obeys the same laws as the radiation emitted by a hot body. However, an essential difference exists. The temperature of a black hole is determined by the same parameters (mass and angular momentum) that determine its geometrical size, while the temperature of an ordinary body and its size are independent parameters.

10.4.3 Scattering of coherent waves

We have already mentioned that there is a close relationship between the processes of propagation of a classical wave and individual quanta in an external field. Let us consider this relationship in the case of scattering by a black hole in more detail. Consider a classical wave incident on a black hole. From the quantum point of view, this wave can be described as a coherent ensemble of quanta, and characterized by the following coherent state:

$$|\gamma_\beta; \mathrm{in}\rangle = \exp\left(-\frac{1}{2}|\gamma_\beta|^2\right) \exp(\gamma_\beta\, \hat{a}^\dagger_{\mathrm{in},\beta})|0; \mathrm{in}\rangle . \tag{10.4.8}$$

The normalization condition for this state,

$$\langle\gamma_\beta; \mathrm{in}|\gamma_\beta; \mathrm{in}\rangle = 1 , \tag{10.4.9}$$

is implied by the following relation [see, e.g., Klauder and Sudarshan (1968)]:

$$\mathrm{e}^{\hat{X}}\, \mathrm{e}^{\hat{Y}} = \mathrm{e}^{[\hat{X},\hat{Y}]}\, \mathrm{e}^{\hat{Y}}\, \mathrm{e}^{\hat{X}} , \tag{10.4.10}$$

which holds for arbitrary operators $\hat{X}$ and $\hat{Y}$, such that $[\hat{X}, \hat{Y}]$ commutes with $\hat{X}$ and $\hat{Y}$. Simple checking confirms that

$$\hat{a}_{\mathrm{in},\beta} \exp\left(\gamma_\beta\, \hat{a}^\dagger_{\mathrm{in},\beta}\right) = \exp\left(\gamma_\beta\, \hat{a}^\dagger_{\mathrm{in},\beta}\right)(\hat{a}_{\mathrm{in},\beta} + \gamma_\beta) , \tag{10.4.11}$$

$$\exp(\bar{\gamma}_\beta\, \hat{a}_{\mathrm{in},\beta})\, \hat{a}^\dagger_{\mathrm{in},\beta} = \left(\hat{a}^\dagger_{\mathrm{in},\beta} + \bar{\gamma}_\beta\right) \exp(\bar{\gamma}_\beta\, \hat{a}_{\mathrm{in},\beta}) . \tag{10.4.12}$$

Using these relations and (10.2.48), it can be shown that the mean value of the field $\hat{\varphi}$ in the coherent state $|\gamma_\beta;\text{in}\rangle$ is

$$\varphi_{\gamma_\beta} \equiv \langle\gamma_\beta;\text{in}|\hat{\varphi}|\gamma_\beta;\text{in}\rangle = \gamma_\beta\,\varphi^{\text{in}}_\beta + \bar{\gamma}_\beta\,\bar{\varphi}^{\text{in}}_\beta\,. \tag{10.4.13}$$

The image $\Phi^-_{\gamma_\beta}$ of φ_{γ_β} on $\mathcal{J}^-$ is

$$\Phi^-_{\gamma_\beta} = \gamma_\beta\,\Phi^{\text{in}\,-}_\beta + \bar{\gamma}_\beta\,\bar{\Phi}^{\text{in}\,-}_\beta\,. \tag{10.4.14}$$

Relation (10.2.63) shows that the image $\Phi^+_{\gamma_\beta}$ of φ_{γ_β} on $\mathcal{J}^+$ is

$$\Phi^+_{\gamma_\beta} = \gamma_\beta\,R_\beta\,\Phi^{\text{out}\,+}_\beta + \bar{\gamma}_\beta\,\bar{R}_\beta\,\bar{\Phi}^{\text{out}\,+}_\beta\,. \tag{10.4.15}$$

In other words, the ratio of the squared amplitude of the scattered wave to the squared amplitude of the incident wave is $|\,R_\beta\,|^2$. This quantity is greater than unity (that is, amplification occurs) for those modes that satisfy the superradiance condition $\omega - m\,\Omega^H < 0$.

10.4.4 Probability distribution

The probability $P(k_\alpha\,|\,l_\alpha)$ of finding k_α scalar massless particles in the state $\varphi_{\text{out},\alpha}$ in the black hole radiation, assuming the flux incident on the hole contains l_α particles in the state $\varphi_{\text{in},\alpha}$, is given by the general expression (10.3.35). It can be shown that this expression can be transformed as follows [Bekenstein and Meisels (1977), Panangaden and Wald (1977)]:

$$P(k_\alpha\,|\,l_\alpha) = (l_\alpha)!\,(k_\alpha)!\,\frac{w_\alpha^{2k_\alpha}(1-w_\alpha^2)(1-|\,R_\alpha\,|^2)^{k_\alpha+l_\alpha}}{(1-w_\alpha^2\,|\,R_\alpha\,|^2)^{k_\alpha+l_\alpha+1}}$$

$$\times\sum_{n=0}^{\min(k_\alpha,l_\alpha)}\left\{\frac{1}{(l_\alpha-n)!\,(k_\alpha-n)!\,(n!)^2}\left[\frac{(1-w_\alpha^2)^2|\,R_\alpha\,|^2}{(1-|\,R_\alpha\,|^2)^2\,w_\alpha^2}\right]^n\right\}. \tag{10.4.16}$$

Analogous expressions for fermions can be found in [Bolashenko and Frolov (1989)].

It is easy to show that

$$e^{-(\varpi_\alpha/\theta)l_\alpha}\,P(k_\alpha\,|\,l_\alpha) = e^{-(\varpi_\alpha/\theta)k_\alpha}\,P(l_\alpha\,|\,k_\alpha)\,. \tag{10.4.17}$$

We will prove below that this condition ensures detailed equilibrium of the black hole with a cavity rotating at an angular velocity Ω^H and filled with black body radiation at temperature θ.

10.4.5 Black hole in a "thermal bath"

If there is no matter outside the black hole, Hawking radiation is the only process that changes the state of a stationary black hole. If there is matter or radiation outside the black hole, Hawking evaporation is accompanied by the process of accretion of this matter and radiation onto the black hole. It emerges that a particular matching of the parameters of the matter distribution to the black hole parameters produces an equilibrium situation in which the loss of particles through accretion in each mode is exactly compensated by the black hole radiation in this mode. In the simplest case of negligible interaction between different species of particles, these equilibrium conditions must obviously hold for each species of particles individually. We will now discuss the conditions of equilibrium of a black hole with a gas of massless quanta.

Assume that the density matrix $\hat{\rho}_{\text{in}}$ that describes the state of such particles outside a stationary black hole (on $\mathcal{J}^-$) is of the form

$$\hat{\rho}_{\text{in}} = \prod_{\alpha}{}' \hat{\rho}_{\text{in},\alpha} \,,$$

$$\hat{\rho}_{\text{in},\alpha} = \rho_{0,\alpha} N_{\text{in}} \left[\exp[-(1+\mu_\alpha)\, \hat{a}^\dagger_{\text{in},\alpha}\, \hat{a}_{\text{in},\alpha}] \right] . \tag{10.4.18}$$

Here $N_{\text{in}}\,[\ldots]$ is the normal ordering operator with respect to the operators $\hat{a}^\dagger_{\text{in},\alpha}$ and $\hat{a}_{\text{in},\alpha}$. It can be verified [Frolov (1986)] that under the given choice of initial condition (10.4.18) the density matrix $\hat{\rho}_{\text{out}}$ that describes the black hole radiation on $\mathcal{J}^+$ is given by the following expression:

$$\hat{\rho}_{\text{out}} = \prod_{\alpha}{}' \hat{\rho}_{\text{out},\alpha} \,,$$

$$\hat{\rho}_{\text{out},\alpha} = \rho_{0,\alpha} B_\alpha N_{\text{out}} \left[\exp[-(1+\mu_\alpha) B_\alpha\, \hat{a}^\dagger_{\text{out},\alpha}\, \hat{a}_{\text{out},\alpha}] \right] , \tag{10.4.19}$$

$$B_\alpha = \frac{1 - w_\alpha^2}{1 + \mu_\alpha - (\mu_\alpha + w_\alpha^2) |\, R_\alpha \,|^2} . \tag{10.4.20}$$

The equilibrium condition, signifying that the distribution of outgoing particles is identical to that of incoming ones, is equivalent to the condition $B_\alpha = 1$ which is satisfied if and only if

$$\mu_\alpha = -w_\alpha^2 = -\exp(-\varpi_\alpha/\theta) . \tag{10.4.21}$$

Assume that the black hole we consider is enclosed in a stationary axially symmetric, perfectly reflecting ("mirror") surface. Particles reflected by this surface conserve their angular momentum and energy. The action of this surface on wavepackets transforms the mode $\varphi_{\text{out},\alpha}$ into the mode $-\varphi_{\text{in},\alpha}$. If condition (10.4.21) is satisfied, the radiation in the cavity is in equilibrium with the black hole, and the density matrix corresponding to this equilibrium state is

$$\hat{\rho}_\theta = \prod_\alpha \hat{\rho}_\alpha \,,$$

$$\hat{\rho}_\alpha = \rho^0_\alpha \exp\left(-\frac{\varpi_\alpha}{\theta}\,\hat{a}^\dagger_{\text{out},\alpha}\,\hat{a}_{\text{out},\alpha}\right), \tag{10.4.22}$$

where ρ^0_α is a normalization constant, and $\hat{a}^\dagger_{\text{out},\alpha}$ and $\hat{a}_{\text{out},\alpha}$ are the operators of creation and annihilation for the OUT-mode α.

This result can be described somewhat differently. Since $\operatorname{Tr}\hat{\rho}_\alpha = 1$, the probability of finding k_α quanta in the mode α for the distribution (10.4.21) is

$$P(k_\alpha) = \frac{\mathrm{e}^{-(\varpi_\alpha/\theta)k_\alpha}}{1-\mathrm{e}^{-\varpi_\alpha/\theta}} \,. \tag{10.4.23}$$

If we take into account the equality (10.4.17) for the conditional probability $P(l_\alpha\,|\,k_\alpha)$, then the probability $P(l_\alpha\,|\,k_\alpha)\,P(k_\alpha)$ of finding l_α outgoing quanta and k_α incoming quanta in the mode α is equal to the probability $P(k_\alpha\,|\,l_\alpha)\,P(l_\alpha)$ of finding k_α outgoing and l_α incoming quanta. This means that the detailed balance condition is satisfied, ensuring thermodynamic equilibrium between the cavity and the black hole in the given mode.

We must emphasize that the density matrix (10.4.19) is normalizable and actually describes the real physical state only if $\mu_\alpha > -1$. The equilibrium condition (10.4.21) for superradiant modes $\varpi_\alpha \le 0$ contradicts the condition of normalizability of the density matrix $\hat{\rho}_{\text{in}}$. This result admits the following interpretation. Consider a rotating black hole enclosed in a mirror-walled cavity. Let an arbitrary (normalizable) distribution of particles exist in the chosen superradiant mode α at some moment of time. Then the scattering of this mode by the black hole increases the number of quanta in it. After reflection by the enclosing mirror, these quanta are again scattered on the black hole and their number is again increased. In other words, the system consisting of a black hole and a mirror-like boundary surrounding it acts as a generator for superradiant modes so that no stationary equilibrium distribution is possible for such modes.

The conclusion obtained above does not imply that in the general case a rotating black hole cannot be in equilibrium with the radiation gas inside the cavity. We need only to require that the size of this cavity is not too large $[r \le (\Omega^H)^{-1}]$ so that the system has no superradiant modes.

The following arguments support this conclusion. Note that the wave modes $\varphi_J \sim \exp(-i\omega t + im\phi)$ are eigenmodes for the operator $\eta^\mu\,\partial_\mu$, where

$$\eta = \xi_t + \Omega^H\,\xi_\phi \,, \tag{10.4.24}$$

and they satisfy the relation

$$\eta^\mu\,\partial_\mu\varphi_J = -i\,\varpi\,\varphi_J \,. \tag{10.4.25}$$

Assume that a surface enclosing the black hole coincides with the surface of $\boldsymbol{\eta}\cdot\boldsymbol{\eta} =$ const and that the angular velocity of this surface is Ω^H. An observer at rest on this surface has the velocity

$$u^\mu = \eta^\mu/U\,, \qquad U^2 \equiv -\boldsymbol{\eta}\cdot\boldsymbol{\eta}. \tag{10.4.26}$$

In this reference frame, the mode $\varphi_{\omega\ell m}$ has the frequency $\varpi' = \varpi/U$. Note that now the equilibrium density matrix (10.4.22) can be written in the form

$$\hat{\rho}_\alpha = \rho_{0,\alpha}\exp\left(-\frac{\varpi'_\alpha}{\theta_{\rm loc}}\,\hat{a}^\dagger_{{\rm out},\alpha}\,\hat{a}_{{\rm out},\alpha}\right), \tag{10.4.27}$$

where

$$\theta_{\rm loc} = \frac{\kappa}{2\pi U}\,. \tag{10.4.28}$$

This means that a rotating *absorbing* shell enclosing the black hole does not violate the equilibrium condition if and only if its temperature is $\theta_{\rm loc}$. If the shell surface does not coincide with the $U =$ const surface, the temperature $\theta_{\rm loc}$ on the shell at equilibrium is not constant. It is given by (10.4.28) in which

$$U = \Sigma^{-1/2}\left\{\Delta\left(1 - a\,\Omega^H\sin^2\theta\right)^2 - \sin^2\theta\left[a - \Omega^H(r^2+a^2)\right]^2\right\}^{1/2}. \tag{10.4.29}$$

Let us refer to the surface Σ_n outside the hole on which the condition

$$\boldsymbol{\eta}\cdot\boldsymbol{\eta} = 0 \tag{10.4.30}$$

is satisfied as the *null cylinder*. The equation of Σ_n for the Kerr metric in the Boyer-Lindquist coordinates is of the form ($r > r_+$)

$$|\sin\theta| = \frac{\sqrt{X^2 + 4a\,\Omega^H} - X}{2a\,\Omega^H}\,, \qquad X^2 = \frac{[\Omega^H(r^2+a^2) - a]^2}{\Delta}\,. \tag{10.4.31}$$

The temperature $\theta_{\rm loc}$, given by relations (10.4.28)-(10.4.29) is defined, real, and bounded in the domain between the event horizon H^+ and the "null cylinder" Σ_n. Correspondingly, the equilibrium state described by density matrix (10.4.22) is possible only if the shell enclosing the black hole also lies in this domain.

For a non-rotating black hole, equilibrium with the gas of radiation is possible regardless of the size of cavity.[15] The equilibrium requires that the temperature of radiation far from the black hole is equal to its Hawking temperature.

This conclusion on the possibility of thermal equilibrium between a black hole and a radiation gas, provided they have equal temperatures and angular velocities,[16] is of a general nature. Gibbons and Perry (1978) demonstrated that this conclusion can be extended to the case of interacting particles.

[15] Note that if the size of the radiation-filled cavity is sufficiently large, this equilibrium is generally unstable (for details, see Chapter 12).

[16] Provided the conditions that guarantee the absence of superradiant modes are met.

10.5 Energy, Angular Momentum, and Entropy Emission

10.5.1 Loss of energy and angular momentum

The energy and angular momentum flux of the field φ through a two-dimensional surface S in a stationary axisymmetric spacetime can be written in the form

$$\dot{E} \equiv \frac{dE}{dt} = \int_S dS\, T_{\mu\nu}[\varphi,\varphi]\, \xi^{\mu}_{(t)}\, n^{\nu}\,, \tag{10.5.1}$$

$$\dot{J} \equiv \frac{dJ}{dt} = \int_S dS\, T_{\mu\nu}[\varphi,\varphi]\, \xi^{\mu}_{(\phi)}\, n^{\nu}\,. \tag{10.5.2}$$

Here $T_{\mu\nu}[\varphi,\varphi]$ is the stress-energy tensor which is bilinear in the field φ; $\xi^{\mu}_{(t)}\partial_\mu = \partial_t$ and $\xi^{\mu}_{(\phi)}\partial_\mu = \partial_\phi$ are Killing vectors; dS is the surface element on S, and n^μ is a unit vector lying on the surface $t=$ const and normal to S.

Let us first calculate the average values of $\langle \dot{E}\,\rangle$ and $\langle \dot{J}\,\rangle$ for the quantum field $\hat{\varphi}$ in the state $|1_{\text{out},\alpha};\text{out}\rangle$ in which there exists only one out-particle in the mode $\varphi^{\text{out}}_{\alpha}$. As usual for this type of calculation we are interested in the change of the quantities with respect to their vacuum values. Only this difference is observable. Define

$$t^{(\alpha)}_{\mu\nu} = \langle 1_{\text{out},\alpha};\text{out}|\, T_{\mu\nu}[\hat{\varphi},\hat{\varphi}]|1_{\text{out},\alpha};\text{out}\rangle - \langle 0;\text{out}|\, T_{\mu\nu}[\hat{\varphi},\hat{\varphi}]|0;\text{out}\rangle\,. \tag{10.5.3}$$

Using the expansion (10.2.48) of the field operator $\hat{\phi}$ in terms of out-modes, and the commutation relations (10.2.49), one gets

$$t^{(\alpha)}_{\mu\nu} = T_{\mu\nu}[\varphi^{\text{out}}_{\alpha},\bar{\varphi}^{\text{out}}_{\alpha}] + T_{\mu\nu}[\bar{\varphi}^{\text{out}}_{\alpha},\varphi^{\text{out}}_{\alpha}]\,. \tag{10.5.4}$$

Since we are interested in the calculation of the outgoing fluxes of energy and angular momentum, we can move the surface S close to $\mathcal{J}^+$. Near $\mathcal{J}^+$ the late-time wavepacket $\varphi^{\text{out}}_{\alpha}$ is nonvanishing only in the retarded time interval $u \in (2\pi(n-1/2)/\epsilon, 2\pi(n+1/2)/\epsilon)$, and hence $t^{(\alpha)}_{\mu\nu}$ for this wavepacket also possesses this property. The integral of $t^{(\alpha)}_{\mu\nu}\,\xi^\mu\, n^\nu$ over the surface S and the retarded time u gives the total energy or angular momentum flux on $\mathcal{J}^+$. For the chosen normalization of the modes the total energy and angular momentum fluxes are ω_α and m_α, respectively. This result is not surprising since at infinity the particles propagate in a practically flat spacetime, where the energy ω and angular momentum m are the standard quantum numbers "enumerating" the states of a quantum.

After these remarks let us calculate the energy and angular momentum fluxes from a black hole. We assume that initially (before the formation of the black hole) there were no particles. Let us consider the quantity

$$\langle\, \hat{T}_{\mu\nu}\rangle \equiv \langle 0;\text{in}|\, T_{\mu\nu}[\hat{\varphi},\hat{\varphi}]|0;\text{in}\rangle - \langle 0;\text{out}|\, T_{\mu\nu}[\hat{\varphi},\hat{\varphi}]|0;\text{out}\rangle\,. \tag{10.5.5}$$

Using the second line of (10.2.48), we can write

$$\begin{aligned}\langle \hat{T}_{\mu\nu} \rangle &= \sum_{\alpha}{}' \Big\{ t^{(\alpha)}_{\mu\nu} \, \langle 0; \mathrm{in} | \hat{a}^{\dagger}_{\mathrm{out},\alpha} \, \hat{a}_{\mathrm{out},\alpha} | 0; \mathrm{in} \rangle \\ &+ T_{\mu\nu}[\varphi^{\mathrm{out}}_{\alpha}, \varphi^{\mathrm{out}}_{\alpha}] \, \langle 0; \mathrm{in} | \hat{a}_{\mathrm{out},\alpha} \, \hat{a}_{\mathrm{out},\alpha} | 0; \mathrm{in} \rangle \\ &+ T_{\mu\nu}[\bar{\varphi}^{\mathrm{out}}_{\alpha}, \bar{\varphi}^{\mathrm{out}}_{\alpha}] \, \langle 0; \mathrm{in} | \hat{a}^{\dagger}_{\mathrm{out},\alpha} \, \hat{a}^{\dagger}_{\mathrm{out},\alpha} | 0; \mathrm{in} \rangle \Big\} .\end{aligned} \tag{10.5.6}$$

Here we used relation (10.5.4) and the fact the averages of quantities that contain only one operator of creation or annihilation in a given mode must vanish.

We use now a special case of the Theorem 2 of Section 10.3; namely, the relation

$$\langle 0; \mathrm{in} | F(\hat{a}^{\dagger}_{\mathrm{out},\alpha}, \hat{a}_{\mathrm{out},\alpha}) | 0; \mathrm{in} \rangle = \left\{ F[\partial_{\psi}, \partial_{\bar{\psi}}] \; Z_{\alpha}[\psi_{\alpha}, \mu = 0; \gamma = 0] \right\}_{\psi_{\alpha}=0} . \tag{10.5.7}$$

According to (10.3.24)

$$Z_{\alpha}[\psi_{\alpha}, \mu = 0; \gamma = 0] = \exp\left[\frac{\Gamma_{\alpha}}{w_{\alpha}^{-2} - 1} \; \bar{\psi}_{\alpha} \psi_{\alpha} \right] . \tag{10.5.8}$$

Here $\Gamma_{\alpha} = 1 - | R_{\alpha} |^2 = \sigma_{\alpha} | T_{\alpha} |^2$. Relations (10.5.7) and (10.5.8) imply

$$\langle 0; \mathrm{in} | \hat{a}_{\mathrm{out},\alpha} \, \hat{a}_{\mathrm{out},\alpha} | 0; \mathrm{in} \rangle = \langle 0; \mathrm{in} | \hat{a}^{\dagger}_{\mathrm{out},\alpha} \, \hat{a}^{\dagger}_{\mathrm{out},\alpha} | 0; \mathrm{in} \rangle = 0 ,$$

$$\langle 0; \mathrm{in} | \hat{a}^{\dagger}_{\mathrm{out},\alpha} \, \hat{a}_{\mathrm{out},\alpha} | 0; \mathrm{in} \rangle = \frac{\Gamma_{\alpha}}{w_{\alpha}^{-2} - 1} . \tag{10.5.9}$$

By using these results, we obtain for $\langle \hat{T}_{\mu\nu} \rangle$ the following expression

$$\langle \hat{T}_{\mu\nu} \rangle = \sum_{\alpha}{}' t^{(\alpha)}_{\mu\nu} \, \frac{\Gamma_{\alpha}}{w_{\alpha}^{-2} - 1} . \tag{10.5.10}$$

We can now substitute this expression for $\langle \hat{T}_{\mu\nu} \rangle$ into (10.5.1) and (10.5.2) and find $\langle \dot{E} \rangle$ and $\langle \dot{J} \rangle$ for the quantum radiation from the black hole. In order to obtain the answer in a simpler form, let us note that a wavepacket $\varphi_{jn\ell m}$ with a given value of n reaches $\mathcal{J}^+$ in the interval of retarded time from $2\pi(n - \frac{1}{2})/\epsilon$ to $2\pi(n + \frac{1}{2})/\epsilon$. Hence,

$$\int_{u_1}^{u_2} \mathrm{d}u \langle \dot{E} \rangle = \sum_{n=n_1}^{n_2} \; \sum_{j,\ell,m,P} \frac{\omega_{\alpha} \Gamma_{\alpha}}{w_{\alpha}^{-2} - 1} , \tag{10.5.11}$$

where n_1 and n_2 the values of the index n corresponding to u_1 and u_2, respectively. In the late-time regime Γ_{α} does not depend on the moment of time u, and hence on n. Thus, we have ($\alpha = \{j, n, \ell, m, P\}$)

$$\frac{dE}{du} = \langle \dot{E} \rangle = \frac{\epsilon}{2\pi} \sum_{j,\ell,m,P} \frac{\omega_{\alpha} \Gamma_{\alpha}}{w_{\alpha}^{-2} - 1} . \tag{10.5.12}$$

Now we note that the quantities in this expression are smooth functions of frequency ω and vary little when ω changes in the interval from $j\epsilon$ to $(j+1)\epsilon$. Hence, the summation over j can be replaced by integration with respect to frequency ω:

$$\epsilon \sum_j (\ldots) = \int_0^\infty d\omega\, (\ldots)\,. \tag{10.5.13}$$

As a result, we get

$$\frac{dE}{du} = \frac{1}{2\pi} \int_0^\infty d\omega \sum_{l,m,P} \frac{\sigma_{\omega m}\, \omega\, |\, T_{\omega\ell m P}\,|^2}{\exp(\varpi/\theta) - \epsilon}\,. \tag{10.5.14}$$

Similarly, we get for the flux of angular momentum

$$\frac{dJ}{du} = \frac{1}{2\pi} \int_0^\infty d\omega \sum_{l,m,P} \frac{\sigma_{\omega m}\, m\, |\, T_{\omega\ell m P}\,|^2}{\exp(\varpi/\theta) - \epsilon}\,. \tag{10.5.15}$$

We include in the above relations the parameter ϵ, assuming that for bosons it takes the value $\epsilon = 1$. Relations (10.5.14) and (10.5.15) are also valid for fermions with $\epsilon = -1$. A similar expression holds for massive fields. The summation then extends to all quantum numbers that enumerate the state, and the lower limit of integration is the mass of the field μ.

For a non-rotating black hole the transmission coefficient T does not depend on m, and hence the average flux of angular momentum vanishes. For estimation of the energy flux from a non-rotating black hole one can use expression (2.9.2) for the capture cross-section for ultra-relativistic particles. This expression implies that particles with impact parameter b less than the critical impact parameter $b_{\rm cr} = 3\sqrt{3}\,M$ are captured by the black hole. Using the relation $\ell = \omega b$, one can rewrite the capture condition as

$$\ell < 3\sqrt{3}\,M\omega\,. \tag{10.5.16}$$

Assuming that all the modes obeying (10.5.16) are absorbed by the black hole, we can approximate $|\,T_{\omega\ell}\,|^2$ as [DeWitt (1975)]

$$|\,T_{\omega\ell}\,|^2 \approx \vartheta\left(3\sqrt{3}\,M\omega - \ell\right). \tag{10.5.17}$$

In this *DeWitt approximation* one gets

$$\sum_{\ell,m} |\,T_{\omega\ell}\,|^2 \approx \sum_{\ell=0}^{\infty} (2\ell+1)\, \vartheta\left(3\sqrt{3}\,M\omega - \ell\right) \approx 27M^2\omega^2\,. \tag{10.5.18}$$

In order to obtain this relation, one changes the summation over ℓ by integration. Strictly speaking, this procedure is not justified when $M\omega$ is small, so that only a few first terms in the sum are important. Nevertheless, the direct calculations show that the approximation (10.5.18) is quite good.[17]

Substituting (10.5.18) into (10.5.14) gives[18]

$$\frac{dE}{du} = \frac{27M^2}{2\pi} h(s) \int_0^\infty \frac{d\omega\, \omega^3}{\exp(8\pi M\omega) - \epsilon} = \frac{27}{2 \times 8^4 \times \pi^5 \times M^2} h(s)\, I_\epsilon . \quad (10.5.19)$$

We include in this relation a factor $h(s)$ which is equal to the number of different states (polarizations) of the spin-s field. We also define

$$I_\epsilon = \int_0^\infty \frac{dx\, x^3}{e^x - \epsilon} .$$

Calculation gives $I_1 = \pi^4/15$ and $I_{-1} = 7\pi^4/120$. Finally, we get the following expression for the rate of energy emission in the DeWitt approximation

$$\frac{dE}{du} \approx \frac{9}{10 \times 8^4 \times \pi} \frac{1}{M^2} [h_b + 7/8\, h_f] \approx \frac{7 \times 10^{-5}}{M^2} [h_b + 7/8\, h_f] . \quad (10.5.20)$$

Here h_b is the total number of polarization states of massless bosons, and h_f is the total number of polarization states of massless fermions.

A more accurate approximation for the rate of the energy loss can be obtained by using numerical calculations for $|T_{\omega\ell}|^2$. These calculations, performed by Page (1976a,b), give

$$\frac{dE}{du} \approx 4 \times 10^{-5} f \left(\frac{m_{Pl}}{M}\right)^2 \frac{m_{Pl}}{t_{Pl}} = 7.4 \times 10^{24} \left(\frac{M}{1\,\mathrm{g}}\right)^{-2} f\, \frac{\mathrm{g}}{\mathrm{s}} . \quad (10.5.21)$$

Here

$$f = 1.023\, h(1/2) + 0.420\, h(1) + 0.048\, h(2) , \quad (10.5.22)$$

where $h(s)$ is the number of distinct polarizations of spin-s particles.

---

[17] Sánchez (1978a, 1997) gives the following approximate expression for the total absorption cross-section $\sigma(\omega)$ as the function of frequency ω:

$$\sigma(\omega) = \pi\omega^{-2} \sum_{\ell,m} |T_{\omega\ell}|^2 \approx 27\pi M^2 - 2\sqrt{2}\, M \frac{\sin(2\sqrt{27}\pi\omega M)}{\omega} , \quad \omega M \geq 0.07 ,$$

$$\sigma(0) = 16\pi M^2 .$$

[18] In a simplified model when the spacetime has only two dimensions, the Hawking radiation of massless particles can be calculated exactly [Christensen and Fulling (1977)]. The derivation based on the Polyakov effective action was first given by Frolov and Vilkovisky (1983). Mukhanov, Wipf, and Zelnikov (1994) used the effective action approach for the calculation of the Hawking radiation in the four-dimensional spacetime.

Consider now a rotating black hole. If the black hole mass is large so that the temperature is very low, then for bosons

$$\frac{\sigma_J}{\exp(\varpi/\theta)-1} \approx \frac{1}{2}(1-\sigma_J)\,. \tag{10.5.23}$$

In this limit, only the modes satisfying the superradiance condition contribute to the radiation, and we have

$$\frac{dE}{du} = \frac{1}{2\pi}\sum_{l,m,P}\int_0^{m\,\Omega^H} d\omega\,\omega\,|\,T_{\omega\ell mP}\,|^2\,, \tag{10.5.24}$$

$$\frac{dJ}{du} = \frac{1}{2\pi}\sum_{l,m,P}\int_0^{m\,\Omega^H} d\omega\, m\,|\,T_{\omega\ell mP}\,|^2\,. \tag{10.5.25}$$

The emission of energy and angular momentum by a black hole changes its parameters. In accordance with the conservation laws, one has

$$\dot{M}^H = -\frac{dE}{du}\,, \qquad \dot{J}^H = -\frac{dJ}{du}\,. \tag{10.5.26}$$

These relations can be proved by comparison of the mode expansions for fluxes of energy and angular momentum on $\mathcal{J}^+$ with similar mode expansions on H^+ [see Frolov and Thorne (1989)].

10.5.2 Entropy of black hole radiation

The entropy S of a system described by a density matrix $\hat{\rho}$ is defined by the relation

$$S = -\operatorname{Tr}(\hat{\rho}\,\ln\hat{\rho})\,. \tag{10.5.27}$$

It is readily shown that $\hat{\rho}$ is expressible in the form

$$\hat{\rho} = \left(\prod_\alpha{}'\hat{\rho}_\alpha\right)\times\hat{\rho}_\bullet\,, \tag{10.5.28}$$

where $\hat{\rho}_\alpha$ is an operator depending only on the creation operators $\hat{a}^\dagger_{\mathrm{out},\alpha}$ and annihilation operators $\hat{a}_{\mathrm{out},\alpha}$ in a late-time OUT-mode α. Using (10.5.28), entropy S can be represented as a sum,

$$S = \sum_\alpha{}' S_\alpha + S_\bullet\,, \tag{10.5.29}$$

where

$$S_\alpha = -\operatorname{Tr}_\alpha(\hat{\rho}_\alpha\,\ln\hat{\rho}_\alpha)\,. \tag{10.5.30}$$

Here Tr_α denotes the operation of calculating the trace in the space generated by the action of $\hat{a}^\dagger_{\mathrm{out},\alpha}$ on the vacuum; S_α is the contribution of the late-time α-modes to the emission of entropy, and $S_\bullet$ is the entropy emitted before the late-time regime.

Using expression (10.3.4) for $\hat{\rho}_\alpha$ after transforming the summation over the wave-packets into integration over the frequency, we obtain

$$\frac{dS}{du} = \frac{1}{2\pi}\int_0^\infty d\omega \sum_{l,m,P}\left[\frac{\Gamma_\alpha}{z_\alpha - \varepsilon}\ln\left(\frac{z_\alpha-\varepsilon}{\Gamma_\alpha}+\varepsilon\right)+\varepsilon\ln\left(1+\frac{\varepsilon\Gamma_\alpha}{z_\alpha-\varepsilon}\right)\right], \quad (10.5.31)$$

where

$$\Gamma_\alpha = \sigma_\alpha\,|\,T_\alpha\,|^2, \qquad z_\alpha = \exp(\varpi_\alpha/\theta)\,, \qquad \varepsilon = 1\,. \quad (10.5.32)$$

This expression is also valid for Fermi fields if we set $\varepsilon = -1$. A similar expression holds for massive fields. The summation then extends to all quantum numbers that enumerate the state, and the integration begins from the mass of the field μ. The contribution of the neutrino ($s = \frac{1}{2}$), photon ($s = 1$), and gravitational ($s = 2$) fields to the entropy of the radiation of non-rotating black hole can be written in the form [Page (1983)]

$$\frac{dS}{du} = 10^{-3}M^{-1}\,[1.685\,h(1/2) + 0.634\,h(1) + 0.065\,h(2)]\,, \quad (10.5.33)$$

where $h(s)$ is the number of polarizations of the spin-s field.

10.5.3 Radiation of a charged rotating black hole

Evolution of black hole parameters in the process of its quantum evaporation

In the general case of a black hole also having, in addition to mass M and angular momentum J, an electric charge Q, the expression for the mean number of particles of mass μ, electric charge $q|\,e\,|$ ($q = \pm 1$), and spin s that the black hole creates, can be written as follows [Hawking (1975)]:

$$\langle n_J\rangle = \frac{\Gamma_J}{\exp[(2\pi\,\varpi_J)/\kappa] - (-1)^{2s}}\,. \quad (10.5.34)$$

Here the collective subscript J denotes the complete set of quantum numbers that must be specified to describe a mode. The set includes the subscript enumerating the particle species and also carrying information on particle spin s, frequency or energy ω, spheroidal quantum number ℓ, azimuthal quantum number m, polarization or helicity P, and the sign of the particle charge q. The quantity ϖ_J in (10.5.34) is

$$\varpi_J = \omega_J - m_J\,\Omega^H - q_J\,|\,e\,|\,\Phi^H\,, \quad (10.5.35)$$

where Ω^H and Φ^H are the angular velocity and electric potential of the black hole, and $\Gamma_J = 1 - |R_J|^2$, where R_J is the reflection amplitude of the incident wave J. The coefficient Γ_J becomes negative for superradiant modes of boson fields. Nevertheless, the expression on the right-hand side of (10.5.34) is always positive. For fermion fields, the Pauli principle implies that the mean number of particles scattered in a given mode cannot exceed unity, so that $|R_J|^2 \leq 1$ always. Superradiance is not possible for fermions [Martellini and Treves (1977), Iyer and Kumar (1979), Chandrasekhar (1979a,b)].

We denote by Σ_J the summation over all discrete quantum numbers and the integration over continuous ones in J,

$$\sum_J = \sum_{\ell, m, P, q} \frac{1}{2\pi} \int_\mu^\infty d\omega \,, \tag{10.5.36}$$

and then obtain the following formulas for the rate of change, due to quantum radiation, of mass, angular momentum, and charge of the black hole:

$$-\frac{dM}{dt} = \sum_J \omega_J \langle \hat{n}_J \rangle, \; -\frac{dJ}{dt} = \sum_J m_J \langle \hat{n}_J \rangle, \; -\frac{dQ}{dt} = |e| \sum_J q_J \langle \hat{n}_J \rangle. \tag{10.5.37}$$

The calculation of the contributions of individual particle species to the quantum radiation from a black hole reduces to the determination of the respective reflection coefficients for wave functions describing these particles. Relations (10.5.37) allow one to study how the parameters of the black hole change in the process of its quantum evaporation. We mention only some results that are of special physical interest.

Loss of electric charge

The loss of electric charge by a black hole was analyzed by Markov and Frolov (1970), Zaumen (1974), Carter (1974), Gibbons (1975), Nakamura and Sato (1976), Damour and Ruffini (1975), Page (1977), Ruffini (1979), Novikov and Starobinsky (1980c). The main result of this analysis is as follows. Isolated black holes of mass $M \leq Ge^2/m_e \approx 10^{15}\,$g ($m_e$ is the electron mass) shed their electric charge almost completely and very rapidly. The time necessary for radiating away the electric charge of black holes with mass $M < \sqrt{G}\, e\, m_{\rm Pl}^2/m_e^2 \approx 10^5\, M_\odot$ is much shorter than the characteristic time of evaporation of a black hole. Therefore, a black hole can be treated as electrically neutral during nearly the entire period of its evaporation.

Loss of angular momentum

Dimensional arguments show that, in the general case, a black hole can dissipate angular momentum over a time comparable with the black hole evaporation time. Carter (1974) hypothesized that in the course of evaporation, the ratio of angular momentum to squared mass of the black hole tends to a nonzero limiting value.

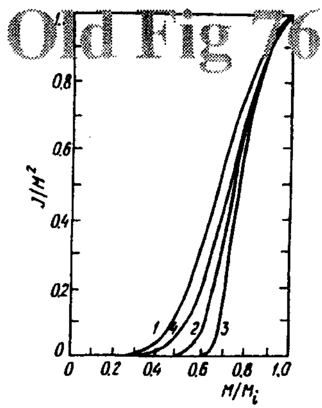


Figure 10.4: The evolution of the rotation parameter J/M^2 of a black hole in the course of evaporation (M_i is the initial and M is the current mass of the black hole). The figure shows the behavior of J/M^2 as a function of M/M_i in the case of only a single neutrino field (1), only the photon field (2), or only the graviton field (3), and in the real situation (4) (two species of neutrinos and anti-neutrinos, one photon, and one graviton species) [data from Page (1976b)].

However, Page's numerical results (1976b) refuted this conjecture. If the contribution of known massless particles (neutrinos, photons, and gravitons) is taken into account, the angular momentum is found to be dissipated several times faster than the black hole mass (Figure 10.4).

It is also found that the smaller the spin of massless particles, the greater their contribution to the radiation of mass and angular momentum from slowly rotating black holes ($J/M^2 \leq 0.6$, and the number of polarization states is assumed to equal 2 for all particles). The situation is the opposite for rapidly rotating black holes ($J/M^2 > 0.6$), the contribution being the greater, the higher the particle spin (Figure 10.5). This effect is in agreement with that of Starobinsky and Churilov (1973) on the spin dependence of superradiance.

Note also that a rotating black hole emits neutrinos asymmetrically. Neutrinos are more abundant in the radiation along the direction of rotation, while antineutrinos are more abundant in the opposite direction [Unruh (1973), Vilenkin (1979b), Leahy and Unruh (1979)]. This effect has an analogue in the radiation of photons and gravitons by a rotating black hole. Namely, there is an asymmetry in the emission of left- and right-hand-polarized quanta in a given direction. As a result, the electromagnetic and gravitational radiation of a rotating black hole are polarized [Dolgov *et al.* (1988b), Bolashenko and Frolov (1987, 1989)]. Dolgov *et al.* (1988a) pointed out that these effects are directly related to chiral anomalies in the gravitational field. On other effects in black holes due to chiral anomalies, see Gal'tsov (1986), Gal'tsov *et al.* (1988).

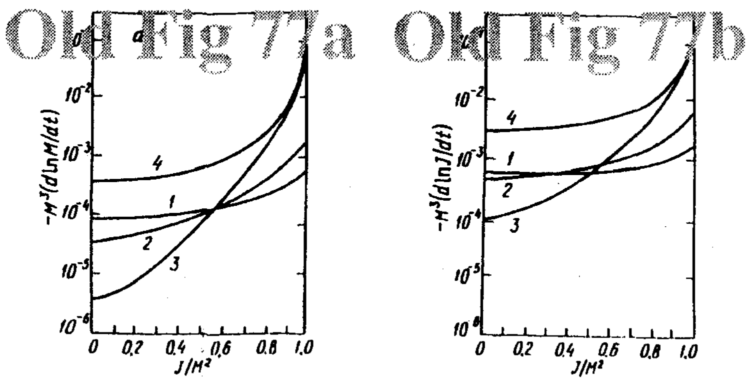


Figure 10.5: Energy (a) and angular momentum (b) radiative power of a black hole as a function of rotation parameter J/M^2. The contributions of a single neutrino species (1), photons (2), and gravitons (3) are plotted separately, as well as the total radiative power for massless particles in a real situation (4) (two species of neutrinos and anti-neutrinos, one photon, and one graviton) [data from Page (1976b)].

Black hole evaporation

A black hole emits radiation like a body heated to the temperature

$$\theta = \frac{\hbar\,\kappa}{2\pi\, c}\,.$$

This expression is written in the usual units, with c, G, and $\hbar$ restored. In this relation the temperature has the same dimension as energy, $\theta = k\,T_H$, where T_H is the black hole temperature in the standard units. For a non-rotating black hole one has

$$T_H = \frac{\hbar\, c^3}{8\pi G k M} \approx 10^{26}\, K\, (M/1\,\mathrm{g})^{-1}\,. \tag{10.5.38}$$

The deviation from a thermal spectrum occurs because the coefficient Γ_J is frequency-dependent. At high frequencies, the effective cross-section of the black hole is $27\pi\, G^2 M^2/c^4$ for all particles. The cross-section decreases at low frequencies, and is found to be strongly spin-dependent (cf. (4.2.29)). As spin increases, the contribution of particles to the total radiation of a non-rotating black hole diminishes [Page (1976a)] (Figure 10.6).

Black holes of mass $M > 10^{17}$ g can emit only massless particles: neutrinos (ν), photons (γ), and gravitons (g). Black holes of mass $5\times 10^{14}\,\mathrm{g} \le M \le 10^{17}\,\mathrm{g}$ can also emit electrons and positrons. Black holes of smaller mass can emit heavier elementary particles as well. The distribution of black hole products in different mass intervals is shown in Figure 10.7.

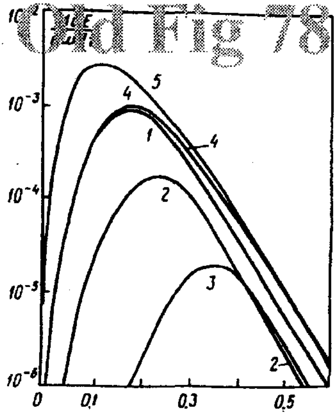


Figure 10.6: Power spectra from a non-rotating black hole. The curves plot the contributions of four neutrino species (1), photons (2), and gravitons (3), and the aggregate spectrum (4). For the sake of comparison, the black body radiation spectrum of these particles with cross-section $27\pi M^2$ is given (curve 5) [data from Page (1976a)].

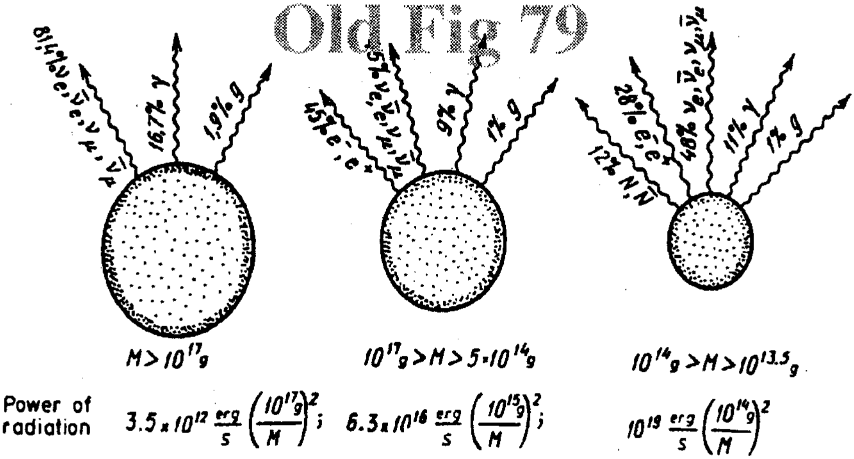


Figure 10.7: Quantum decay of a non-rotating black hole. The fractions of gravitons (g), photons (γ), neutrinos (ν) and other elementary particles are given in percent of the total number of particles emitted by black holes of different masses.

As a result of the quantum evaporation, the mass of the black hole decreases. Simple model calculations shows that for slow quasistatic change of the mass the temperature of the black hole also changes as M^{-1} [Volovich *et al.* (1976)]. Integrating equation (10.5.21) for the mass loss, one can get the lifetime of a black hole with respect to its quantum evaporation

$$t_{\mathrm{BH}} \approx 4.5 \times 10^{-26} f^{-1} \left(\frac{M}{1\,\mathrm{g}}\right)^3 . \tag{10.5.39}$$

Since the rate of emission grows rapidly as the mass of the black hole decreases, the most of the time t is occupied by the stage where the mass is close to the initial mass M. The factor f which enters (10.5.39) is the effective number of species of particles (10.5.22) that can be emitted by a black hole of mass M.

The lifetime of a black hole will not exceed the age of the Universe $t_{\mathrm{U}} \sim 5\times 10^{17}$s if its mass $M \leq M_{\mathrm{U}}$, where $M_{\mathrm{U}} \approx f^{1/3}\times 2.23\times 10^{14}$ g. Using (10.5.22), one gets $f \approx 9.12$ (for 2 types of neutrino, electron, muon, their antiparticles, photon, and graviton). Hence, $M_{\mathrm{U}} \approx 5 \times 10^{14}$ g. Such black holes could be born only as primordial holes at early stages in the evolution of the Universe. As for the possibility of observing quantum explosions of such small black holes and the products of their decay, see the discussion in section 9.8 and the references therein.

"External control" over the Hawking radiation

The radiative power and other characteristics of black hole radiation change when an external field is switched on. Thus, external factors could be used to affect the radiation of a black hole and, to some extent, control it. For example, when a black hole is "inserted" into an external gravitational potential, the intensity of its radiation at infinity diminishes in complete accordance with the reduction in temperature as measured by a distant observer [Zhuk and Frolov (1981), Geroch and Hartle (1982)]. On the effect of magnetic fields on particle creation in black holes, see Gibbons (1976), Gal'tsov (1980). An interesting example of the effect of external factors is the process of mining a black hole, which was analyzed by Unruh and Wald (1982) (see Section 12.3.2).

Chapter 11

Quantum Physics of Black Holes

11.1 Vacuum Polarization near Black Holes

11.1.1 Semi-classical approximation

Quantum radiation from an isolated black hole reduces its mass, and hence its surface area. This "violation" of Hawking's area theorem is explained by drawing the inevitable conclusion that the particle flux to infinity which carries away positive energy is accompanied by a flux of negative energy across the horizon into the black hole. This would be impossible in classical theory, with ordinary physical assumptions satisfied (energy dominance conditions). In quantum theory, negative energy density may arise in a region of space because an external field applied to the vacuum may either increase or reduce the local energy density of zero point fluctuations. It is this phenomenon, known as *vacuum polarization*, that is to be expected to play an important role in the neighborhood of black holes.

The process of evaporation of a black hole of a mass much greater than the Planck mass can be described using the *semi-classical approximation*. Assuming the fluctuations of the gravitational field to be small, we describe it in terms of the classical metric

$$g_{\mu\nu} = \langle \hat{g}_{\mu\nu} \rangle \tag{11.1.1}$$

satisfying the modified Einstein equations

$$G_{\mu\nu} = 8\pi \langle \hat{T}_{\mu\nu} \rangle , \tag{11.1.2}$$

where the right-hand side contains the expectation value of the energy-momentum tensor of the relevant quantized fields in the chosen state. If the characteristic radius of curvature in a region of spacetime, L, is much greater than the Planck length $l_{\mathrm{Pl}} = \sqrt{\hbar G/c^3}$, the calculation of $\langle \hat{T}_{\mu\nu} \rangle$ can make use of an expansion in a small parameter $\varepsilon = (l_{\mathrm{Pl}}/L)^2$ retaining only terms up to first order in ε. The first term of order ε^0 is just the expression for the energy-momentum tensor of the classical field, while the term of order ε^1, containing a factor $\hbar$, represents the main contribution of quantum effects (in the chosen approximation of $\varepsilon \ll 1$). This contribution describes the change in energy-momentum density due to the effect of the gravitational field on the state of virtual vacuum pairs. Other terms of higher order in ε describe additional contributions due to the interaction between the particles of a virtual pair as a result of emission and subsequent absorption of field quanta by these particles.[1] In the approximation linear in ε ("one-loop" approximation), virtual pairs of various fields can be treated as non-interacting. Therefore, the contributions of all fields to $\langle \hat{T}_{\mu\nu} \rangle$ are additive in this approximation, and thus can be studied independently.

The main problem in analyzing $\langle \hat{T}_{\mu\nu} \rangle$ stems from its divergence. Specifically, any calculation of the expectation value of a quantity containing a product of two or more field operators at the same point ($\hat{T}_{\mu\nu}$ has precisely this form) results in infinities. Such divergences, arising even in flat spacetime, are generated by vacuum zero-point fluctuations. The methods of extracting a finite, physically meaningful part of $\langle \hat{T}_{\mu\nu} \rangle$, known as renormalization procedures, were widely discussed in the literature in connection with the development of general field theory in curved spacetime and with its specific applications to cosmology and black hole physics. A detailed discussion of these problems can be found in DeWitt (1965, 1975), Grib *et al.* (1980), Christensen (1976, 1978), Birrell and Davies (1982), and Wald (1994). For this reason, we only briefly outline the renormalization procedure for $\langle \hat{T}_{\mu\nu} \rangle$ and go into some detail about those aspects of the vacuum polarization effects that are related to black hole specifics (e.g., the choice of the vacuum state). We also give the main results of calculating $\langle \hat{T}_{\mu\nu} \rangle^{\text{ren}}$ in the spacetime of a black hole.

11.1.2 Wald's axioms

In order to extract a physically meaningful part of $\langle \hat{T}_{\mu\nu} \rangle$, one needs at first to regularize this expression, that is, to make it finite by introducing some regularization parameter. There exist a considerable number of regularization methods, such as dimensional regularization, ζ-function method, Pauli-Villars regularization, n-wave regularization, adiabatic regularization, and the point-splitting approach. In each of the chosen regularization schemes there exists a procedure of extraction of those parts of $\langle \hat{T}_{\mu\nu} \rangle$ that are responsible for infinities in the process of removing the regularization. By subtracting these terms, one obtains an expression which remains finite. This quantity is called the *renormalized stress-energy tensor* and is denoted as $\langle \hat{T}_{\mu\nu} \rangle^{\text{ren}}$. The concrete procedure of renormalization differs from one method to another, but it is important, however, that the final results are essentially independent of the choice of renormalization method.

[1]When the Feynman diagram technique in quantum theory is used to calculate the indicated expectation values, the expansion in $\hbar$ coincides with an expansion in closed loops found in the corresponding diagrams. Terms of order $\hbar^0$ are described by diagrams without loops (the "tree" approximation), and those of order $\hbar^1$, by diagrams with a single loop ("one-loop" approximation).

In fact, Wald (1977, 1978a,b) proved the following result:

All renormalization methods for $\langle \hat{T}_{\mu\nu} \rangle$ *that:*

1. *Preserve general covariance* ($\nabla_\mu \langle \hat{T}^{\mu\nu} \rangle^{ren} = 0$)

2. *Satisfy natural causality requirements*

3. *Preserve the value of* $\langle \Psi | \hat{T}_{\mu\nu} | \Phi \rangle$ *for those states* $|\Psi\rangle$ *and* $|\Phi\rangle$ ($\langle \Psi | \Phi \rangle = 0$) *for which this quantity is finite*

4. *Agree with the standard procedure of normal ordering in flat spacetime*

result in expressions for $\langle \hat{T}_{\mu\nu} \rangle^{ren}$ *that can differ from one another only by a local conserved tensor constructed from the curvature tensor at a given point and its covariant derivatives.*

We shall use later the following important consequence of this result. Consider a massless field contribution to $\langle \hat{T}_{\mu\nu} \rangle^{\text{ren}}$. The theory does not contain the mass parameters. For this reason, the expression which has the correct dimensionality and describes the possible ambiguity in $\langle \hat{T}_{\mu\nu} \rangle^{\text{ren}}$ in the one-loop approximation is a sum of terms quadratic in curvature and terms linear in the second derivatives of curvature. The analysis shows that such a symmetric conserved second-rank tensor necessarily contains the Ricci tensor and vanishes when the latter is equal to zero. Hence, the above-mentioned ambiguity in the determination of $\langle \hat{T}_{\mu\nu} \rangle^{\text{ren}}$ is absent in the one-loop approximation for massless fields for a metric satisfying Einstein's vacuum equations ($R_{\mu\nu} = 0$).

11.1.3 Point-splitting method

To illustrate a procedure of the renormalization of $\langle \hat{T}_{\mu\nu} \rangle$ we consider the *point-splitting method* which is quite often used for practical calculations of $\langle \hat{T}_{\mu\nu} \rangle^{\text{ren}}$ in the gravitational field of black holes. The method is as follows. The stress-energy tensor $T_{\mu\nu}$ is a bilinear function of the fields so that it is formally possible to generalize $T_{\mu\nu}$ to the case where the arguments are different for each of the fields (to "split the points"). In the classical theory, $T_{\mu\nu}(x)$ appears as a limit of the corresponding expression $T_{\mu\nu'}(x, x')$ as $x' \to x$. When the expectation value of the operator $\hat{T}_{\mu\nu'}(x, x')$ in a given state, $T_{\mu\nu'}(x, x') = \langle \hat{T}_{\mu\nu'}(x, x') \rangle$, is calculated in the quantum theory, a simple transition to the limit $x' \to x$ is impossible because divergences appear. For this reason, the quantity $T_{\mu\nu'}(x, x')$ is "corrected" (renormalized), before taking the limit, by subtracting from it the standard expression $T^{\text{div}}_{\mu\nu'}(x, x')$. It is sufficient to calculate $T^{\text{div}}_{\mu\nu'}$ once for each of the fields. Christensen (1978) derived the expressions for $T^{\text{div}}_{\mu\nu'}$ in the cases of massless scalar and spinor fields and for the electromagnetic field.

Consider the field $\hat{\varphi}_A$. The index A enumerates the components of the field. When matrix elements $\langle\Psi|\hat{T}_{\mu\nu}(x)|\Phi\rangle^{\text{ren}}$ of the energy-momentum tensor of the field $\hat{\varphi}_A$ are calculated for the specified states $|\Phi\rangle$ and $|\Psi\rangle$, it is convenient to use Green's function

$$G_{AB'}(x,x') = i\,\frac{\langle\Psi|T(\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x'))|\Phi\rangle}{\langle\Psi|\Phi\rangle}\,. \tag{11.1.3}$$

The symbol T in (11.1.3) denotes the operation of T-product,

$$T(\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x')) = \vartheta(x,x')\,\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x') + \vartheta(x',x)\,\hat{\varphi}_{B'}(x')\,\hat{\varphi}_A(x)\,, \tag{11.1.4}$$

where $\vartheta(x,x')$ is the step function equal to 1 if x lies in the future relative to x', and equal to 0 otherwise. Using the commutation relations (H.9), it is easy to show that Green's function (11.1.3) for the field $\hat{\varphi}_A$, described by equation (H.2), satisfies the following equation:

$$D^{AC}G_{CB'}(x,x') = -\,\delta^A_{B'}\,\delta^4(x-x')\,, \tag{11.1.5}$$

where $\int\delta^4(x-x')\,d^4x = 1$. This Green's function is related to the so-called *Wightman function* $G^+_{AB'}(x,x')$ and *Hadamard function* $G^{(1)}_{AB'}(x,x')$:

$$G^+_{AB'}(x,x') = \frac{\langle\Phi|\,\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x')|\Psi\rangle}{\langle\Phi|\Psi\rangle}\,, \tag{11.1.6}$$

$$G^{(1)}_{AB'}(x,x') = G^+_{AB'}(x,x') + G^+_{B'A}(x',x)\,. \tag{11.1.7}$$

If the points x and x' are separated by a spacelike distance, then $\hat{\varphi}_A(x)$ and $\hat{\varphi}_{B'}(x')$ commute and $G^{(1)}_{AB'}(x,x') = -\,2i\,G_{AB'}(x,x')$. The values assumed by the energy-momentum tensor at the split points are

$$T_{\mu\nu'}(x,x') = \frac{1}{2}\,T^{AB'}_{\mu\nu'}(x,x')\,G^{(1)}_{AB'}(x,x')\,, \tag{11.1.8}$$

where $T^{AB'}_{\mu\nu'}(x,x')$ is the differential operator with respect to variables x and x', such that the quantity $T^{AB'}_{\mu\nu'}(x,x')\,\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x')$ coincides with $T_{\mu\nu'}(x,x')$. The explicit form of the operators $T^{AB'}_{\mu\nu'}$ for fields of different spin is given in Christensen's paper (1978).

In addition to $\langle\hat{T}_{\mu\nu}\rangle^{\text{ren}}$, one often considers quantities of type $\langle\hat{\varphi}^2_A\rangle^{\text{ren}}$ that describe the fluctuations of the field $\hat{\varphi}_A$. In the case of a scalar field $\hat{\varphi}$, the quantity $\langle\hat{\varphi}^2\rangle$ was analyzed in the field of a black hole in connection with the possibility of phase transition in its neighborhood. These transitions consist of the appearance of $\langle\hat{\varphi}(x)\rangle \neq 0$ [Hawking (1981), Fawcett and Whiting (1982), Moss (1985)]. The following simple expression for the renormalized quantity $\langle\hat{\varphi}^2\rangle^{\text{ren}}$ holds in a gravitational field described by Einstein's vacuum equations:

$$\langle\hat{\varphi}^2(x)\rangle^{\text{ren}} = \lim_{x'\to x}\left[-\,i\,G(x,x') - \frac{1}{8\pi^2\,\sigma(x,x')}\right], \tag{11.1.9}$$

where $\sigma(x,x') = \frac{1}{2}s^2(x,x')$, and $s(x,x')$ is the interval of geodesic distance between x and x'. The limit is taken so that x' tends to x along a spacelike direction.

The problem of calculating $\langle \hat{T}_{\mu\nu}\rangle^{\text{ren}}$ and $\langle \hat{\varphi}^2\rangle^{\text{ren}}$, which carry the information on energy-momentum density related to vacuum polarization and on vacuum fluctuations, thus reduces to performing a set of standard operations on the Green's function (11.1.3). As a result, the solution of the problem of vacuum polarization in black holes reduces to solving equation (11.1.5) in a given metric describing the spacetime of a black hole. The ambiguity in the choice of state used to carry out the averaging in (11.1.3) corresponds to the arbitrariness in the choice of the boundary conditions that unambiguously specify a particular Green's function.

11.1.4 Conformal trace anomaly

The renormalization of $\langle \hat{T}_{\mu\nu}\rangle$ preserves the covariance of the theory but can violate other symmetries. This phenomenon is known as the anomaly. An important example is the *conformal trace anomaly*.

Consider a classical theory with the action W. The theory is conformal invariant if the action is invariant under the conformal rescaling of the metric: $g_{\mu\nu}(x) \to \Omega(x)\, g_{\mu\nu}(x)$. For the infinitesimal conformal transformation $\Omega = 1+\omega(x)$, the change of the action is proportional to $\omega\, g_{\mu\nu}\,\delta W/\delta g_{\mu\nu}$. Because $\delta W/\delta g_{\mu\nu} \sim T^{\mu\nu}$, the conformal invariance of the theory implies that the trace T^μ_μ of the classical stress-energy tensor $T^{\mu\nu}$ vanishes.

An important distinctive feature of the quantum theory is that the trace of $\langle \hat{T}_{\mu\nu}\rangle^{\text{ren}}$ does *not* vanish any more for a conformally invariant field (*conformal trace anomaly*). It is possible to show that the quantity $\langle \hat{T}^\mu_\mu\rangle^{\text{ren}}$ is independent of the choice of the state over which $\hat{T}_{\mu\nu}$ is averaged, and has the form

$$\langle \hat{T}^\mu_\mu\rangle^{\text{ren}} = \alpha\left(\mathcal{H} + \frac{2}{3}\Box R\right) + \beta\mathcal{K} + \gamma\,\Box R\,, \tag{11.1.10}$$

where

$$\mathcal{H} = C_{\mu\nu\rho\tau}\,C^{\mu\nu\rho\tau} = R_{\mu\nu\rho\tau}\,R^{\mu\nu\rho\tau} - 2R_{\mu\nu}\,R^{\mu\nu} + \frac{1}{3}\,R^2\,, \tag{11.1.11}$$

$$\mathcal{K} = {}^*R_{\mu\nu\rho\tau}\,{}^*R^{\mu\nu\rho\tau} = R_{\mu\nu\rho\tau}\,R^{\mu\nu\rho\tau} - 4R_{\mu\nu}\,R^{\mu\nu} + R^2\,. \tag{11.1.12}$$

Here ${}^*R_{\mu\nu\rho\tau} = e_{\mu\nu\alpha\beta}\,R^{\alpha\beta}{}_{\rho\tau}$. The coefficients α, β, and γ depend on the number of helicity states of each spin and on the renormalization scheme used. Point-separation, ζ-function, and dimensional renormalization all agree that[2]

$$\alpha = (2^9 \times 45 \times \pi^2)^{-1}\,[12\,h(0) + 18\,h(1/2) + 72\,h(1)]\,, \tag{11.1.13}$$

[2]The conformal trace anomaly was first noticed by Capper and Duff (1974) using a dimensional regularization scheme [see also Duff (1977)]. For the calculations in other regularization procedures see, e.g., Birrell and Davies (1982) and Fulling (1989).

$$\beta = (2^9 \times 45 \times \pi^2)^{-1} [-4\, h(0) - 11\, h(1/2) - 124\, h(1)] \,, \tag{11.1.14}$$

where $h(s)$ is the number of helicity states for fields of spin s. On the other hand, there is disagreement on the value of γ. Dimensional renormalization gives $\gamma = 0$, while point-separation and ζ-function renormalization give

$$\gamma = (2^9 \times 45 \times \pi^2)^{-1}[-120\, h(1)] \,. \tag{11.1.15}$$

Note that an additional term of type $\Box R$ that may appear in the case of the electromagnetic field vanishes in metrics satisfying the vacuum Einstein's equations ($R_{\mu\nu} = 0$), and in these spaces there is no ambiguity in the form of the conformal anomaly

$$\langle \hat{T}^{\mu}_{\mu} \rangle^{\text{ren}} = (\alpha + \beta)\, C_{\alpha\beta\gamma\delta}\, C^{\alpha\beta\gamma\delta} \,. \tag{11.1.16}$$

11.2 Choice of State and Boundary Conditions for Green's Functions

11.2.1 Unruh vacuum

Now we will describe the states and corresponding boundary conditions for Green's functions that are most frequently used in considering quantum effects in black holes.[3] For the sake of simplification, we limit the analysis to the case of massless fields. It is these fields that give the main contribution to the vacuum polarization.

The situation of obvious interest is that of a black hole arising as a result of collapse and the quantum field being originally in the ground, vacuum state. The corresponding Green's function,

$$G^{\text{in}}_{AB'}(x, x') = i\, \langle 0; \text{in} |\, T(\hat{\varphi}_A(x)\, \hat{\varphi}_{B'}(x')) | 0; \text{in} \rangle \,, \tag{11.2.1}$$

being symmetric as are all other Green's functions (11.1.3), is determined unambiguously by the following property: It coincides in the distant past (in the in-region) with the free causal Green's function in Minkowski space. Obviously, the behavior of G^{in} generally depends on details of the collapse forming the black hole. As a result, the determination of G^{in} is a complicated problem, difficult to solve. Recall, however, that after some time, the black hole becomes stationary.[4] The arguments used to prove the universality of the properties of the Hawking radiation at late times lead

[3] A general analysis of the problem of defining the vacuum in a spacetime with horizons can be found in the papers of Fulling (1977a,b), Sciama *et al.* (1981), Fulling and Ruijsenaars (1987), and the book by Wald (1994).

[4] Rather, it is nearly stationary because its parameters change to some extent as a result of quantum evaporation. But we have already mentioned that the rate of this change is negligibly small as long as the mass of the black hole is much greater than the Planck mass.

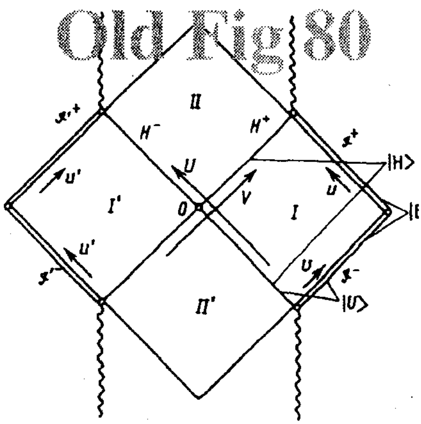


Figure 11.1: Penrose-Carter conformal diagram for eternal version of a rotating black hole. The straight lines point to surfaces bounding the external region I, where boundary conditions are fixed for Green's functions corresponding to the Hartle-Hawking ($|H\rangle$), Unruh ($|U\rangle$), and Boulware ($|B\rangle$) vacuum states.

to the conclusion that a sufficiently long time after the formation of the black hole, G^{in} is determined exclusively by the black hole parameters.

The following method is convenient for describing the properties of this "universal" Green's function. Consider the spacetime of an eternal black hole having the same parameters as the stationary black hole which forms in the collapse. In this spacetime, define a Green's function $G^U_{AB'}(x,x')$ as the solution of the equation (11.1.5) that at late times coincides with the asymptotic form of $G^{\text{in}}_{AB'}(x,x')$ and is obtained as its analytic continuation in the remaining part of the spacetime. Using the same line of arguments as in Section 10.2.2, one can prove that the following boundary conditions define the Green's function $G^U_{AB'}(x,x')$ unambiguously in the spacetime of an eternal black hole for values of the arguments x and x' that lie in the exterior region or at the event horizon: This function for a fixed value of x' in region I (Figure 11.1) is of negative frequency with respect to the affine parameter U on H^- for x on H^-, and of negative frequency in advanced time v on $\mathcal{J}^-$ for x on $\mathcal{J}^-$. The corresponding state $|U\rangle$ for which

$$G^U_{AB'}(x,x') = i\,\langle U|\,T(\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x'))|U\rangle \tag{11.2.2}$$

is called the *Unruh vacuum.*

In order to make this definition more accurate, let us consider the spacetime of an eternal version of the original black hole (see Figure 11.1). In the previous chapter we used modes φ^{in}, φ^{out}, φ^{up}, and φ^{down} for the description of quantum effects in the black hole exterior. These modes can be defined as solutions in the spacetime of the eternal black hole by assuming that they vanish in the region I'. We shall need another set of modes that vanish in the external region I and play a similar role in the black hole interior (region I'). In order to construct these modes, one can apply the discrete symmetry transformation J given by (H.53) to the *IN*-, *OUT*-, *UP*-, and *DOWN*-modes and accompany it by complex conjugation. We denote these new modes by $\varphi^{\rm ui}$, $\varphi^{\rm fno}$, $\varphi^{\rm dn}$, and $\varphi^{\rm umop}$. We already used modes $\varphi^{\rm dn}$ related to $\varphi^{\rm up}$ by (H.54). Similar relations connect $\varphi^{\rm in}$ with $\varphi^{\rm ui}$, $\varphi^{\rm out}$ with $\varphi^{\rm fno}$, and $\varphi^{\rm down}$ with $\varphi^{\rm umop}$.[5] For example, $\varphi^{\rm ui}$ is a solution which vanishes in the external region I and is determined in I' by the conditions that it vanishes on H^- and is of positive-frequency with respect to u' on $\mathcal{J}'^+$.

Solutions of massless field equations in the spacetime of the eternal black hole can be characterized by giving their values on $\mathcal{J}'^+ \cup H^- \cup \mathcal{J}^-$, and hence one can use a set of *IN*-, *UP*-, *DN*-, and *UI*-modes as the basis in the region $I \cup I' \cup II \cup II'$. For our purpose, instead of *UP*- and *DN*-modes it is more convenient to use their linear combinations (10.2.65) and (10.2.66) that are of positive frequency with the respect to the affine parameter U on H^-. We called them *D*- and *P*-modes. Using this basis, one can write

$$\hat{\varphi} = \sum_J \left[\varphi_J^{\rm in}\, \hat{a}_{{\rm in},J} + \varphi_J^{\rm p}\, \hat{a}_{{\rm p},J} + \varphi_J^{\rm d}\, \hat{a}_{{\rm d},J} + \varphi_J^{\rm ui}\, \hat{a}_{{\rm ui},J} + \text{Herm. Conj.}\right]. \tag{11.2.3}$$

Here "Herm. Conj." denotes the Hermitian conjugated expression. The operators of annihilation $\hat{a}_{\Lambda,J}$ and creation $\hat{a}^\dagger_{\Lambda,J}$ obey the standard canonical commutation relations (H.42). The Unruh vacuum is defined as a state $|U\rangle$ obeying the relations

$$\hat{a}_{{\rm in},J}|U\rangle = \hat{a}_{{\rm p},J}|U\rangle = \hat{a}_{{\rm d},J}|U\rangle = \hat{a}_{{\rm ui},J}|U\rangle = 0\,. \tag{11.2.4}$$

The Wightman function G^{U+}

$$\begin{aligned} G^{U+}_{AB'}(x,x') &= \sum_J \left[\varphi^{\rm in}_A(J;x)\,\bar{\varphi}^{\rm in}_{B'}(J;x') + \varphi^{\rm p}_A(J;x)\,\bar{\varphi}^{\rm p}_{B'}(J;x')\right.\\ &+ \left.\varphi^{\rm d}_A(J;x)\,\bar{\varphi}^{\rm d}_{B'}(J;x') + \varphi^{\rm ui}_A(J;x)\,\bar{\varphi}^{\rm ui}_{B'}(J;x')\right] \end{aligned} \tag{11.2.5}$$

can be used to construct the Green's function (11.2.2). If one of the points is located in the black hole exterior (region I), the last term in the square brackets vanishes. It is easy to show that the so-constructed Green's function meets the required boundary conditions.

11.2.2 Hartle-Hawking vacuum

Another case of interest is that of a black hole placed in a cavity in equilibrium with black-body radiation filling the cavity. Since this state is described by the density matrix $\hat{\rho}_\theta$ (10.4.22), the Green's function $G^H(x,x')$ corresponding to it is

$$G^H_{AB'}(x,x') = i\,\mathrm{Tr}[\,\hat{\rho}_\theta\, T(\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x'))]\,. \tag{11.2.6}$$

[5] See a footnote on page 371.

This Green's function can also be related to the spacetime of the eternal black hole.

For a non-rotating black hole (when superradiant modes do not exist), one can introduce a globally well-defined state $|H\rangle$ so that

$$G^{H}_{AB'}(x,x') = i\,\langle H|\,T(\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x'))|H\rangle \tag{11.2.7}$$

coincides with (11.2.6) provided x and x' are located in region I. As shown by Hartle and Hawking (1976), the corresponding Green's function satisfies the following boundary conditions: For a fixed value of x' it is a negative-frequency function with respect to the affine parameter U for x on H^- and a positive-frequency function with respect to the affine parameter V for x on H^+ (see Figure 11.1). The state $|H\rangle$ is called the *Hartle-Hawking vacuum.*

In order to provide the correct boundary conditions for $G^{H}_{AB'}(x,x')$, we use sets of modes *UP*, *DOWN*, *DN*, and *UMOP* that form a basis. The only problem is that this basis is not orthogonal. Namely, one has

$$< \varphi^{\rm up}_J\,, \varphi^{\rm down}_{J'} > = \bar{r}_J\,\delta_{JJ'}\,, \qquad < \varphi^{\rm dn}_J\,, \varphi^{\rm umop}_{J'} > = r_J\,\delta_{JJ'}\,. \tag{11.2.8}$$

One can easily verify that the following linear transformation produces an orthonormal basis

$$\varphi^{\rm up'}_J = a_J\,\varphi^{\rm up}_J + b_J\,\varphi^{\rm down}_J\,, \qquad \varphi^{\rm down'}_J = \bar{b}_J\,\varphi^{\rm up}_J + a_J\,\varphi^{\rm down}_J \tag{11.2.9}$$

$$\varphi^{\rm dn'}_J = a_J\,\varphi^{\rm dn}_J + \bar{b}_J\,\varphi^{\rm umop}_J\,, \qquad \varphi^{\rm umop'}_J = b_J\,\varphi^{\rm dn}_J + a_J\,\varphi^{\rm umop}_J \tag{11.2.10}$$

where

$$a_J = \frac{\sqrt{1+|\,t_J|}}{\sqrt{2}\,|\,t_J|}\,, \qquad b_J = -\,\frac{r_J}{\sqrt{2}\,|\,t_J|\sqrt{1+|\,t_J|}}\,. \tag{11.2.11}$$

We now define modes $\varphi^{\rm d'}_J$, $\varphi^{\rm p'}_J$, $\varphi^{\rm u'}_J$, and $\varphi^{\rm n'}_J$ by relations similar to (10.2.65) and (10.2.66)

$$\varphi^{\rm d'}_J = c_J\varphi^{\rm dn'}_J + s_J\,\bar{\varphi}^{\rm up'}_J\,, \qquad \varphi^{\rm p'}_J = c_J\,\varphi^{\rm up'}_J + s_J\,\bar{\varphi}^{\rm dn'}_J\,, \tag{11.2.12}$$

$$\varphi^{\rm u'}_J = c_J\,\varphi^{\rm umop'}_J + s_J\,\bar{\varphi}^{\rm down'}_J\,, \qquad \varphi^{\rm n'}_J = c_J\,\varphi^{\rm down'}_J + s_J\,\bar{\varphi}^{\rm umop'}_J\,, \tag{11.2.13}$$

with coefficients c_J and s_J given by (10.2.67). One can easily verify that the so-defined modes D', P', U', and N' are of positive frequency with respect to the affine parameters on H^+ and H^- and form an orthonormal basis. The decomposition of the field operator in this basis is

$$\hat{\varphi} = \sum_J \left[\varphi^{\rm d'}_J\,\hat{a}_{{\rm d'},J} + \varphi^{\rm p'}_J\,\hat{a}_{{\rm p'},J} + \varphi^{\rm n'}_J\,\hat{a}_{{\rm n'},J} + \varphi^{\rm u'}_J\,\hat{a}_{{\rm u'},J} + \text{Herm. Conj.}\right]. \tag{11.2.14}$$

The operators of annihilation $\hat{a}_{\Lambda,J}$ and creation $\hat{a}^\dagger_{\Lambda,J}$ obey the standard canonical commutation relations (H.42). The Hartle-Hawking vacuum is defined as the state $|H\rangle$ obeying the relations

$$\hat{a}_{\mathrm{p}',J}|H\rangle = \hat{a}_{\mathrm{d}',J}|H\rangle = \hat{a}_{\mathrm{n}',J}|H\rangle = \hat{a}_{\mathrm{u}',J}|H\rangle = 0\,. \tag{11.2.15}$$

The Wightman function G^{H+},

$$\begin{aligned} G^{H+}_{AB'}(x,x') &= \sum_J \Big[\varphi^{\mathrm{p}'}_A(J;x)\,\bar{\varphi}^{\mathrm{p}'}_{B'}(J;x') + \varphi^{\mathrm{d}'}_A(J;x)\,\bar{\varphi}^{\mathrm{d}'}_{B'}(J;x') \\ &+ \varphi^{\mathrm{n}'}_A(J;x)\,\bar{\varphi}^{\mathrm{n}'}_{B'}(J;x') + \varphi^{\mathrm{u}'}_A(J;x)\,\bar{\varphi}^{\mathrm{u}'}_{B'}(J;x')\Big]\,, \end{aligned} \tag{11.2.16}$$

can be used to construct the Green's function G^H. This Green's function evidently meets the required boundary conditions.

If the black hole is rotating, the situation is more complicated. As before, we can formally define G^H by imposing the same boundary conditions on H^+ and H^-, but due to the presence of superradiant modes we shall have problems. In order to suppress superradiant modes, we can assume that the black hole is surrounded by a mirror-like boundary. As we have mentioned in the preceding chapter, equilibrium of a rotating black hole in a cavity is possible only if the size of this cavity is sufficiently small. One can accordingly expect that if there is no cavity or if its boundary lies outside the "null cylinder", the Green's function $G^H(x,x')$ is, in a certain sense, pathological. Kay and Wald (1991) have shown that this is indeed so [see also Frolov and Thorne (1989)]. This is why we assume, when considering below the state $|H\rangle$ for a rotating black hole, that the black hole in question is placed in a cavity contained within the "null cylinder". For a slowly rotating black hole, the boundary of the cavity may lie very far from the black hole, so that its influence is negligible.[6] If the cavity size is small and the boundary effects are important, one needs to specify the boundary conditions at its surface Σ. Usually we assume that

$$\varphi|_\Sigma = 0\,. \tag{11.2.17}$$

The Green's function must also satisfy the appropriate boundary condition:

$$G^{H,\Sigma}(x,x')|_{x\in\Sigma} = 0\,. \tag{11.2.18}$$

If the role of the boundary surface Σ is important, we replace the symbol $|H\rangle$ with $|H,\Sigma\rangle$ and G^H with $G^{H,\Sigma}$.

In the presence of a boundary the sets of basis solutions must be modified. Namely, since we are interested in the field characteristics inside the cavity, we no longer need to consider *IN*- and *OUT*-modes. *UP*-modes are completely reflected by the boundary, and hence coincide on H^+ with *DOWN*-modes. For these reasons, it is sufficient to consider only one set of modes (we denote them by $\varphi^{\mathrm{up},\Sigma}_J$) that coincide

[6] We recall that for the large size of the cavity the equilibrium could be unstable.

with *UP*-modes on H^- and obey the boundary condition (11.2.17) on Σ. In order to define $|H,\Sigma\rangle$, we assume that besides the boundary Σ in the spacetime of the eternal black hole there exists another boundary Σ' located in region I' obtained from Σ by the action of the discrete symmetry $\mathbf{J}$ (H.53). We denote by $\varphi_J^{\mathrm{dn},\Sigma}$ a solution that coincides with the *DN*-mode on H^- and obeys the boundary condition (11.2.17) on Σ'. We use relations similar to (10.2.65) and (10.2.66) to define modes $\varphi_J^{\mathrm{d},\Sigma}$ and $\varphi_J^{\mathrm{p},\Sigma}$ that are of positive frequency with respect to the affine parameters on the horizon.

The field operator $\hat{\varphi}$ inside the cavity admits the following decomposition

$$\hat{\varphi} = \sum_J \left[\varphi_J^{\mathrm{d},\Sigma}\, \hat{a}_{\mathrm{d},J} + \varphi_J^{\mathrm{p},\Sigma}\, \hat{a}_{\mathrm{p},J} + \text{Herm. Conj.}\right], \tag{11.2.19}$$

and the state $|H,\Sigma\rangle$ is defined by the relations

$$\hat{a}_{\mathrm{p},J}|H,\Sigma\rangle = \hat{a}_{\mathrm{d},J}|H,\Sigma\rangle = 0\,. \tag{11.2.20}$$

The Wightman function $G^{H,\Sigma+}$,

$$G^{H,\Sigma+}_{AB'}(x,x') \;=\; \sum_J \left[\varphi_A^{\mathrm{p},\Sigma}(J;x)\,\bar{\varphi}_{B'}^{\mathrm{p},\Sigma}(J;x') + \varphi_A^{\mathrm{d},\Sigma}(J;x)\,\bar{\varphi}_{B'}^{\mathrm{d},\Sigma}(J;x')\right], \tag{11.2.21}$$

can be used to construct the Green's function (11.2.7). The so-constructed Green's function evidently meets all the required boundary conditions.

Hartle and Hawking (1976) demonstrated that the Green's function G^H (as well as $G^{H,\Sigma}$) has special analytical properties. In order to describe these properties, note that if a change of variables,

$$t = -\,i\,\tau, \qquad a = i\,b, \tag{11.2.22}$$

is carried out in the expression for the line element in the Kerr geometry (3.2.1), the resulting metric has signature $+\,+\,+\,+$. Moreover, one can show [Hartle and Hawking (1976), Hawking (1981)] that this metric is everywhere regular (including the surface $r = r_E = M + \sqrt{M^2+b^2}$ corresponding to the analytic continuation of the event horizon surface) provided the coordinate τ is cyclic, with the period $2\pi/\kappa_E$ ($\kappa_E = \kappa|_{a=ib}$). A regular manifold with this metric is known as the *Euclidean black hole.* On can also analytically continue the field equations and define the Euclidean Green's function G^E for these equations. The result obtained by Hartle and Hawking (1976) states that the Green's function $G^H(x,x')$ can be obtained by the analytic continuation (11.2.22) of the function $G^E(x,x')$

$$G^H(x,x') = \left[i\,G^E(x,x')\right]_{\substack{\tau=it\\ b=-ia}}\,. \tag{11.2.23}$$

In particular, for a massless scalar field, G^E is a symmetric solution of the equation

$$\Box_E\, G^E(x,x') = -\,\delta(x,x')\,, \quad \delta(x,x') = \frac{\delta(x-x')}{\sqrt{g_E}} \tag{11.2.24}$$

in the space of the Euclidean black hole, falling off at infinity and regular at the surface of the Euclidean horizon. Here

$$\Box_E = \Box|_{\substack{t=-i\tau \\ a=ib}} \,, \qquad g_E = g\,|_{\substack{t=-i\tau \\ a=ib}} \,. \tag{11.2.25}$$

The Euclidean Green's function $G^{E,\Sigma}$ is defined in a similar manner, but instead of the fall off at infinity one imposes boundary conditions on Σ. The analytical continuation of $G^{E,\Sigma}$ gives $G^{H,\Sigma}$.

Thus, it is possible to employ the following method for constructing $G^H(x,x')$: First find a Euclidean Green's function G_E and then obtain G^H using analytic continuation (11.2.23). This procedure is very similar to that of *Wick rotation* and transition to the Euclidean formulation that is frequently used in quantum field theory in flat spacetime. Quite often, this procedure considerably facilitates the calculation of G^H.

In the absence of walls the Hartle-Hawking state for a rotating black hole is badly defined. Because of the presence of superradiant modes the density matrix ρ_θ in (11.2.6) does not exist (see discussion in Section 10.4.5). One can still define G^H as a solution of (11.1.5) which is of positive frequency with respect to the affine parameter on H^+ and of negative frequency with respect to that on H^-. But Kay and Wald (1991) proved that this Green's function necessarily loses its Hadamard-type behavior in the black hole exterior. Moreover, there does not exist any Hadamard state which is invariant under the symmetries which generate the event horizon. In particular, there does not exist a pure state $|H\rangle$ in the spacetime of the eternal black hole for which relation (11.2.7) is valid. The problems that arise when one tries to introduce the notion of a particle connected with the positive frequency modes with respect to the affine parameters on both horizons for superradiant modes are discussed in [Frolov and Thorne (1989)]. All these problems are connected with superradiant modes.

11.2.3 Boulware vacuum

Let us briefly discuss one more choice of state which is useful for describing quantum effects in black holes and which was first suggested by Boulware (1975a,b, 1976). This state, denoted by $|B\rangle$, is known as the *Boulware vacuum.* For a non-rotating black hole the corresponding Green's function $G^B_{AB'}(x,x')$ satisfies the following boundary conditions: at a fixed value of $x' \in I$, it is a negative-frequency function of v for x on $\mathcal{J}^-$ and a positive-frequency function of u for x on $\mathcal{J}^+$.

To construct this Green's function we shall use the basis $\{\varphi_J^{\rm in}, \varphi_J^{\rm out}, \varphi_J^{\rm ui}, \varphi_J^{\rm fno}\}$. This basis is not orthogonal because

$$< \varphi_J^{\rm in}, \varphi_{J'}^{\rm out} > = \bar{R}_J\, \delta_{JJ'} \,, \qquad < \varphi_J^{\rm ui}, \varphi_{J'}^{\rm fno} > = R_J\, \delta_{JJ'} \,. \tag{11.2.26}$$

To make it orthogonal, we perform linear transformations similar to (11.2.9) and (11.2.10)

$$\varphi_J^{\rm in'} = A_J\, \varphi_J^{\rm in} + B_J\, \varphi_J^{\rm out} \,, \qquad \varphi_J^{\rm out'} = \bar{B}_J\, \varphi_J^{\rm in} + A_J\, \varphi_J^{\rm out} \tag{11.2.27}$$

$$\varphi_J^{\text{ui}'} = A_J\,\varphi_J^{\text{ui}} + \bar{B}_J\,\varphi_J^{\text{fno}}\,, \quad \varphi_J^{\text{fno}'} = B_J\,\varphi_J^{\text{ui}} + A_J\,\varphi_J^{\text{fno}}\,, \tag{11.2.28}$$

where

$$A_J = \frac{\sqrt{1+|\,T_J|}}{\sqrt{2}\,|\,T_J|}\,, \quad B_J = -\frac{R_J}{\sqrt{2}\,|\,T_J|\sqrt{1+|\,T_J|}}\,. \tag{11.2.29}$$

One can decompose the field operator $\hat{\varphi}$ as

$$\hat{\varphi} = \sum_J \Big[\varphi_J^{\text{in}'}\,\hat{a}_{\text{in}',J} + \varphi_J^{\text{out}'}\,\hat{a}_{\text{out}',J} + \varphi_J^{\text{ui}'}\,\hat{a}_{\text{ui}',J} + \varphi_J^{\text{fno}'}\,\hat{a}_{\text{fno}',J} + \text{Herm. Conj.}\Big]\,, \tag{11.2.30}$$

and define $|B\rangle$ by the relations

$$\hat{a}_{\text{in}',J}|B\rangle = \hat{a}_{\text{out}',J}|B\rangle = \hat{a}_{\text{ui}',J}|B\rangle = \hat{a}_{\text{fno}',J}|B\rangle = 0\,. \tag{11.2.31}$$

The Wightman function G^{B+},

$$\begin{aligned} G^{B+}_{AB'}(x,x') \;=\; & \sum_J \Big[\varphi_A^{\text{in}'}(J;x)\,\bar{\varphi}_{B'}^{\text{in}'}(J;x') + \varphi_A^{\text{out}'}(J;x)\,\bar{\varphi}_{B'}^{\text{out}'}(J;x') \\ & + \; \varphi_A^{\text{ui}'}(J;x)\,\bar{\varphi}_{B'}^{\text{ui}'}(J;x') + \varphi_A^{\text{fno}'}(J;x)\,\bar{\varphi}_{B'}^{\text{fno}'}(J;x')\Big]\,, \end{aligned} \tag{11.2.32}$$

can be used to construct the Green's function

$$G^{B}_{AB'}(x,x') = i\,\langle B|\,T(\hat{\varphi}_A(x)\,\hat{\varphi}_{B'}(x'))|B\rangle\,. \tag{11.2.33}$$

This Green's function evidently meets the required boundary conditions.

Consider a non-rotating spherical body of mass M and radius R_0 which is slightly larger than the gravitational radius $r_+ = 2M$ of this body. Since the corresponding Killing vector field $\xi^{\mu}_{(t)}\,\partial_\mu = \partial_t$ in such a static spacetime is everywhere timelike, every particle has positive energy, and particle creation is impossible. The corresponding vacuum $|B;\,R_0\rangle$ defined as the lowest energy state is stable. The Green's function G^B for a non-rotating black hole can be treated, in a certain sense, as the limit of $G_{B;R_0}$ as $R_0 \to r_+$. Obviously, no physical realization is possible of a static system whose size is arbitrarily close to the gravitational radius. In this limit, the particles on the surface of the body would move at infinitely large acceleration, and bringing them to rest requires infinitely large forces. As a result, the Green's function G^B which has simple, regular behavior far from the black hole and corresponds to the absence of quantum radiation both on $\mathcal{J}^+$ and $\mathcal{J}^-$ reveals "poor" analytical behavior close to the event horizon. The renormalized quantities $\langle B|\,\hat{T}_{\mu\nu}|B\rangle$ and $\langle B|\hat{\varphi}^2|B\rangle$ corresponding to G^B diverge both on H^+ and H^-.

Because of the presence of superradiant modes, G^B is not well defined for a rotating black hole.

11.2.4 Mode expansion for Hadamard's functions in the black hole exterior

To calculate the renormalized value of the energy-momentum tensor, one needs to know the values which the Green's function $G(x,x')$ assumes for neighbouring values of x and x'. However, this does not mean that the boundary conditions, imposed on the Green's function far from the point of interest for us, do not affect the behavior of $G(x,x')$ in the limit of coincident points. This is readily verified if one recalls that the wave equation determines the Green's function up to a solution of the homogeneous equation which is fixed unambiguously by the boundary conditions.

If the characteristic radius of curvature L of spacetime in the theory of a massive field (with mass m) is much greater than the Compton length $\lambda = \hbar/mc$, one can use the expansion in a small parameter $\varepsilon = (\lambda/L)^2$ and obtain a uniform approximation for the Green's function.[7] No such parameter exists if the field is massless. In the spacetime of a black hole the natural method of analyzing such fields is to expand them in their eigenmodes [Candelas (1980), Candelas *et al.* (1981)] because the wave equations for massless fields on a black hole background permit separation of variables.

We restrict ourselves by considering vacuum polarization effects in the exterior of a black hole. [For the calculation of these quantities inside the black hole, see Candelas and Jensen (1986)]. To calculate $\langle \hat{T}^\mu_\nu(x)\rangle^{\text{ren}}$ and $\langle\hat{\varphi}(x)\rangle^{\text{ren}}$ for points x lying outside the horizon, one needs to know Hadamard's function $G^{(1)}_{AB'}(x,x')$ with both points x and x' located in the black hole exterior. Using (11.1.7) and expressions (11.2.5), (11.2.16), (11.2.21), and (11.2.32) for the Wightman functions for different quantum states, one can obtain the required mode expansion for Hadamard's functions. We define

$$\varphi^{\Lambda}_{AB'}(J;x,x') = \bar{\varphi}^{\Lambda}_{A}(J;x)\,\varphi^{\Lambda}_{B'}(J;x') + \varphi^{\Lambda}_{A}(J;x)\,\bar{\varphi}^{\Lambda}_{B'}(J;x')\,. \tag{11.2.34}$$

Here the index Λ takes values "in", "out", and so on, specifying the type of the modes.

Let us consider at first $G^{U,(1)}_{AB'}(x,x')$. One has

$$\varphi^{\text{n}}_{AB'}(J;x,x') + \varphi^{\text{p}}_{AB'}(J;x,x') \stackrel{\text{ext}}{=} \coth(\pi\varpi_J\sigma_J/\kappa)\,\varphi^{\text{up}}_{AB'}(J;x,x')\,, \tag{11.2.35}$$

$$\varphi^{\text{ui}}_{AB'}(J;x,x') \stackrel{\text{ext}}{=} 0\,. \tag{11.2.36}$$

Here symbol $\stackrel{\text{ext}}{=}$ indicates that the relation is valid if both points x and x' are located in the black hole exterior. Using these relations, we get

$$G^{U,(1)}_{AB'}(x,x') \stackrel{\text{ext}}{=} \sum_{l,m,P}\left[\int_0^\infty d\omega\,\varphi^{\text{in}}_{AB'}(J;x,x')\right.$$

[7]Note that in the Riemannian space with the Euclidean signature of the metric, the Green's function of a massive field signature falls off exponentially as the distance between the points x and x' increases. If these points are far from the boundaries, the effect of boundary conditions is indeed negligible for these fields.

$$+ \int_{\Omega m}^{\infty} d\omega \, \coth \frac{\pi \varpi}{\kappa} \, \varphi^{\mathrm{up}}_{AB'}(J; x, x') \Big] . \tag{11.2.37}$$

In order to get $G^{H,\Sigma,(1)}_{AB'}(x, x')$, we note that

$$\varphi^{\mathrm{p},\Sigma}_{AB'}(J; x, x') + \varphi^{\mathrm{d},\Sigma}_{AB'}(J; x, x') \stackrel{\mathrm{ext}}{=} \coth(\pi \varpi_J \sigma_J / \kappa) \, \varphi^{\mathrm{up},\Sigma}_{AB'}(J; x, x') . \tag{11.2.38}$$

Using this relation and (11.2.21), one gets

$$G^{H,\Sigma,(1)}_{AB'}(x, x') = \sum_{l,m,P} \int_{\Omega m}^{\infty} d\omega \, \coth \frac{\pi \varpi}{\kappa} \, \varphi^{\mathrm{up},\Sigma}_{AB'}(J; x, x') . \tag{11.2.39}$$

The representation for $G^{H,(1)}_{AB'}(x, x')$ for a non-rotating black hole can be obtained in a similar manner. Note that

$$\varphi^{\mathrm{p}'}_{AB'}(J; x, x') + \varphi^{\mathrm{d}'}_{AB'}(J; x, x') \stackrel{\mathrm{ext}}{=} \coth(\pi \omega_J / \kappa) \, \varphi^{\mathrm{up}'}_{AB'}(J; x, x') , \tag{11.2.40}$$

$$\varphi^{\mathrm{u}'}_{AB'}(J; x, x') + \varphi^{\mathrm{n}'}_{AB'}(J; x, x') \stackrel{\mathrm{ext}}{=} \coth(\pi \omega_J / \kappa) \, \varphi^{\mathrm{down}'}_{AB'}(J; x, x') , \tag{11.2.41}$$

$$\varphi^{\mathrm{up}'}_{AB'}(J; x, x') + \varphi^{\mathrm{down}'}_{AB'}(J; x, x') = \varphi^{\mathrm{up}}_{AB'}(J; x, x') + \varphi^{\mathrm{in}}_{AB'}(J; x, x') . \tag{11.2.42}$$

Using these relations and (11.2.16), we have

$$G^{H,(1)}_{AB'}(x, x') = \sum_{l,m,P} \int_{0}^{\infty} d\omega \, \coth \frac{\pi \omega}{\kappa} \, \left[\varphi^{\mathrm{in}}_{AB'}(J; x, x') + \varphi^{\mathrm{up}}_{AB'}(J; x, x') \right] . \tag{11.2.43}$$

Finally, to construct $G^{B,(1)}_{AB'}(x, x')$ for a non-rotating black hole, we note that

$$\varphi^{\mathrm{in}'}_{AB'}(J; x, x') + \varphi^{\mathrm{out}'}_{AB'}(J; x, x') = \varphi^{\mathrm{in}}_{AB'}(J; x, x') + \varphi^{\mathrm{up}}_{AB'}(J; x, x') , \tag{11.2.44}$$

so that

$$G^{B,(1)}_{AB'}(x, x') = \sum_{l,m,P} \int_{0}^{\infty} d\omega \, \left[\varphi^{\mathrm{in}}_{AB'}(J; x, x') + \varphi^{\mathrm{up}}_{AB'}(J; x, x') \right] . \tag{11.2.45}$$

For generalization of expressions (11.2.43) and (11.2.45) to the case of a rotating black hole, see Candelas (1980), Candelas *et al.* (1981), and Frolov (1986).

11.3 $\langle \hat{T}^{\mu}_{\nu}\rangle^{\rm ren}$ and $\langle\hat{\varphi}^2\rangle^{\rm ren}$ in the Spacetime of a Black Hole

11.3.1 Christensen-Fulling representation

The mode expansion of Hadamard's functions constructed in the previous section can be used for the analysis of the behavior of $\langle \hat{T}^{\mu}_{\nu}\rangle^{\rm ren}$ and $\langle\hat{\varphi}^2\rangle^{\rm ren}$ in the spacetime of a black hole. Before presenting these results, we discuss some general properties of $\langle \hat{T}^{\mu}_{\nu}\rangle^{\rm ren}$ in the Schwarzschild spacetime. The symmetry of $\langle \hat{T}^{\mu}_{\nu}\rangle$ due to the symmetries of the background Schwarzschild gravitational field and the conservation law $\langle \hat{T}^{\mu}_{\nu}\rangle_{;\mu} = 0$ drastically reduce the number of its independent components. Namely, Christensen and Fulling (1977) demonstrated that any conserved $\langle \hat{T}^{\mu}_{\nu}\rangle$ in the spacetime of a non-rotating black hole admits the following representation:

$$\langle \hat{T}^{\mu}_{\nu}\rangle = \sum_{i=1}^{4} t_{(i)\nu}{}^{\mu}, \tag{11.3.1}$$

where $t_{(i)\nu}{}^{\mu}$ in the coordinates (t, r_*, θ, ϕ) have the form[8]

$$\begin{aligned}
t_{(1)\nu}{}^{\mu} &= \mathrm{diag}\left(-\frac{FH}{r^2}+\frac{1}{2}T,\ \frac{FH}{r^2},\ \frac{1}{4}T,\ \frac{1}{4}T\right),\\
t_{(2)\nu}{}^{\mu} &= \mathrm{diag}\left(-\frac{FG}{r^2}-2\Theta,\ \frac{FG}{r^2},\ \Theta,\ \Theta\right),\\
t_{(3)\nu}{}^{\mu} &= \frac{FW}{4\pi r^2}\,\tau^{\mu}_{\ \nu},\quad t_{(4)\nu}{}^{\mu} = \frac{NF}{4\pi r^2}\,\mathrm{diag}\,(-1,\,1,\,0,\,0)\,.
\end{aligned} \tag{11.3.2}$$

Here

$$\tau^{\mu}_{\ \nu} = \begin{pmatrix} -1 & -1 & 0 & 0\\ 1 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}, \tag{11.3.3}$$

and

$$F=(1-2M/r)^{-1},\quad T(r)=\langle \hat{T}^{\mu}_{\mu}(r)\rangle,\quad \Theta(r)=\langle \hat{T}^{\theta}_{\theta}(r)\rangle-\frac{1}{4}T(r), \tag{11.3.4}$$

$$H(r)=\frac{1}{2}\int_{2M}^{r}(r'-M)\,T(r')\,dr',\quad G(r)=2\int_{2M}^{r}(r'-3M)\,\Theta(r')\,dr'. \tag{11.3.5}$$

[8]The upper index of $t_{(i)\nu}{}^{\mu}$ enumerates columns, while the lower one enumerates rows of the matrices. Here and later diag(a,b,c,d) denotes a diagonal matrix with entries equal to a, b, c, d on the diagonal.

Each of the tensors ${t_{(i)}}_\nu^{\ \mu}$ satisfies the conservation law $({t_{(i)}}_\nu^{\ \mu})_{;\mu} = 0$. Only ${t_{(1)}}_\nu^{\ \mu}$ has nonzero trace, only ${t_{(2)}}_\nu^{\ \mu}$ has trace-free component ${t_{(2)}}_\nu^{\ \mu} - \frac{1}{4}\, t_{(2)}\, \delta^\mu_\nu$ whose $\theta\theta$-component is nonzero, only ${t_{(3)}}_\nu^{\ \mu}$ has non-diagonal components describing fluxes, and only ${t_{(4)}}_\nu^{\ \mu}$ is not regular on H^+.

In other words, an arbitrary energy-momentum tensor satisfying the conservation law and symmetry conditions imposed by the Schwarzschild metric is characterized unambiguously by fixing two functions $T(r)$ and $\Theta(r)$ (one of them, T, coincides with the trace of the tensor) and two constants W and N: W gives the intensity of radiation of the black hole at infinity ($W = -\, dM/dt$), and N vanishes if the energy-momentum tensor is regular on H^+. The radiation intensity W is not zero only in the Unruh vacuum. The coefficient N vanishes for the Unruh and the Hartle-Hawking vacuums.

11.3.2 Asymptotic values of $\langle \hat{T}^\mu_\nu \rangle^{\rm ren}$ and $\langle \hat{\varphi}^2 \rangle^{\rm ren}$ at the horizon and at infinity

The simplest case is the calculation of $\langle \hat{\varphi}^2(x) \rangle$ for a massless scalar field in the Schwarzschild metric. Note that for a general location of x the difference between any two of the above-described Hadamard's functions remains finite in the limit of coincident points because the divergences removed by renormalization have a universal form. Thus, for these finite differences we have

$$\langle U|\hat{\varphi}^2(x)|U\rangle - \langle H|\hat{\varphi}^2(x)|H\rangle = -2\sum_{l,m}\int_0^\infty \frac{d\omega}{e^{\omega/\theta}-1}\, |\varphi^{\rm in}(J;x)|^2\,, \tag{11.3.6}$$

$$\langle B|\hat{\varphi}^2(x)|B\rangle - \langle H|\hat{\varphi}^2(x)|H\rangle$$
$$= -2\sum_{l,m}\int_0^\infty \frac{d\omega}{e^{\omega/\theta}-1}\left[\,|\varphi^{\rm in}(J;x)|^2 + |\varphi^{\rm up}(J;x)|^2\right], \tag{11.3.7}$$

where $\theta = \kappa/2\pi$ is the black hole temperature.

Mode representations for Hadamard's functions make it possible to analyze the behavior of $\langle \hat{\varphi}^2 \rangle^{\rm ren}$ and $\langle \hat{T}^\mu_\nu \rangle^{\rm ren}$ close to $H^\pm$ and $\mathcal{J}^\pm$. We recall that the *IN*- and *UP*-modes for a massless scalar field in the Schwarzschild geometry have the following behavior (see Appendix G)

$$\varphi^\Lambda(\omega\ell m;x) = \frac{1}{\sqrt{4\pi\omega}}\, e^{-i\omega t}\, \frac{\chi^\Lambda_\ell(r,\omega)}{r}\, Y_{\ell m}(\theta,\phi)\,, \tag{11.3.8}$$

where

$$\chi^{\rm in}_\ell(r,\omega) \sim \begin{cases} T_{\ell\omega}\, e^{-i\omega r_*} & \text{as } r\to r_+\,,\\ e^{-i\omega r_*} + R_{\ell\omega}\, e^{+i\omega r_*} & \text{as } r\to\infty\,, \end{cases} \tag{11.3.9}$$

and

$$\chi_\ell^{\mathrm{up}}(r,\omega) \sim \begin{cases} r_{\ell\omega}\, \mathrm{e}^{-i\omega r_*} + \mathrm{e}^{+i\omega r_*} & \text{as } r \to r_+ , \\ t_{\ell\omega}\, \mathrm{e}^{+i\omega r_*} & \text{as } r \to \infty . \end{cases} \tag{11.3.10}$$

Using the asymptotic form of the solutions near the horizon and near infinity, one gets [Candelas (1980)]

$$\sum_{\ell=0}^{\infty} |\, \chi_\ell^{\mathrm{in}}(r,\omega)|^2 \sim \begin{cases} A(\omega) & \text{as } r \to r_+ , \\ 4\,\omega^2 r^2 & \text{as } r \to \infty , \end{cases} \tag{11.3.11}$$

$$\sum_{\ell=0}^{\infty} |\, \chi_\ell^{\mathrm{up}}(r,\omega)|^2 \sim \begin{cases} 16M^2\omega^2(1-2M/r)^{-1} & \text{as } r \to r_+ , \\ A(\omega) & \text{as } r \to \infty , \end{cases} \tag{11.3.12}$$

where

$$A(\omega) = \sum_{\ell=0}^{\infty} (2\ell+1)|\, t_{\ell\omega}|^2 . \tag{11.3.13}$$

Using these results, Candelas (1980) obtained[9]

$$\langle \hat{\varphi}^2(r) \rangle^U \sim \begin{cases} (192\pi^2 M^2)^{-1}\,[1-6B(M)] & \text{as } r \to r_+ , \\ (8\pi^2 r^2)^{-1} B(M) & \text{as } r \to \infty , \end{cases} \tag{11.3.14}$$

$$\langle \hat{\varphi}^2(r) \rangle^H \sim \begin{cases} (192\pi^2 M^2)^{-1} & \text{as } r \to r_+ , \\ (1/12)\theta^2 = (768\pi^2 M^2)^{-1} & \text{as } r \to \infty , \end{cases} \tag{11.3.15}$$

$$\langle \hat{\varphi}^2(r) \rangle^B \sim \begin{cases} -(768\pi^2 M^2)^{-1}(1-2M/r)^{-1} & \text{as } r \to r_+ , \\ o(r^{-2}) & \text{as } r \to \infty . \end{cases} \tag{11.3.16}$$

Here $\theta = (8\pi M)^{-1}$ is the temperature of the black hole, and

$$B(M) = \int_0^\infty \frac{d\omega\, A(\omega)}{\omega(\exp(8\pi M\omega)-1)} . \tag{11.3.17}$$

[9] In these formulas (and those that follow) the superscript "ren" is dropped in $\langle \hat{\varphi}^2 \rangle^{\mathrm{ren}}$ and $\langle \hat{T}_\nu^\mu \rangle^{\mathrm{ren}}$ because we will operate only with renormalized values of these quantities. The superscript after the angle brackets indicates the vacuum state for which the averaging is done. Thus, $\langle \hat{\varphi}^2 \rangle^B$ denotes $\langle B|\hat{\varphi}^2|B \rangle^{\mathrm{ren}}$.

In a similar manner, one can obtain the asymptotics of $\langle \hat{T}^\mu_\nu(r) \rangle$ for different vacuum states [Candelas (1980)]. In particular, for the leading terms of $\langle \hat{T}^\mu_\nu(r) \rangle^U$ at the horizon and at infinity one has

$$\langle \hat{T}^\mu_\nu(r) \rangle^U \sim \begin{cases} W(1 - r_+/r)^{-1}(4\pi r_+^2)^{-1}\, \tau^\mu_\nu & \text{as } r \to r_+ , \\ W(4\pi r_+^2)^{-1}\, \tau^\mu_\nu & \text{as } r \to \infty , \end{cases} \tag{11.3.18}$$

where τ^μ_ν in (t, r_*, θ, ϕ)-coordinates is the matrix given by (11.3.3), and

$$W = \frac{1}{2\pi} \int_0^\infty \frac{d\omega\, \omega\, A(\omega)}{\exp(8\pi M\omega) - 1} . \tag{11.3.19}$$

$\langle \hat{T}^\mu_\nu(r) \rangle$ has a similar behavior for other massless fields. The coefficients W for the massless scalar field ($s = 0$), two-component neutrino field ($s = 1/2$), electromagnetic field ($s = 1$), and gravitational field ($s = 2$) are [Page (1982), Elster (1983b)]

$$W_0 = 7.4 \times 10^{-5} M^{-2}, \qquad W_{1/2} = 8.2 \times 10^{-5} M^{-2},$$

$$W_1 = 3.3 \times 10^{-5} M^{-2}, \qquad W_2 = 0.4 \times 10^{-5} M^{-2}. \tag{11.3.20}$$

For the Boulware vacuum, one has in the neighborhood of a Schwarzschild black hole [Candelas (1980), Candelas *et al.* (1981), Sciama *et al.* (1981)]:

$$\langle \hat{T}^\mu_\nu(r) \rangle^B \sim -\frac{h_s}{2\pi^2(1 - r_+/r)^2} \int_0^\infty \frac{d\omega\, \omega\, (\omega^2 + s^2\kappa^2)}{e^{\omega/\theta} - (-1)^{2s}}\, \mathrm{diag}(-1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}) , \tag{11.3.21}$$

where h_s is the number of polarization states of the spin-s field. Quantities characterizing vacuum polarization in this state display this singular behavior because the state itself corresponds, as pointed out above, to a physically non-realizable situation. At large distances the components of $\langle \hat{T}^\mu_\nu(r) \rangle^B$ are expected to be of the order of the square of the Riemann tensor and to fall off like $O(r^{-6})$ [Candelas (1980)].

The renormalized stress-energy tensor in the Hartle-Hawking state remains finite at both horizons $H^\pm$, as well as at the bifurcation of horizons where the Killing vector $\xi_{(t)}$ vanishes. The symmetry condition $\mathcal{L}_{\xi_{(t)}} \langle \hat{T}^\mu_\nu \rangle^H = 0$ implies that at the bifurcation sphere one has

$$\langle \hat{T}^\mu_\nu \rangle^H = \mathrm{diag}(A, A, B, B) . \tag{11.3.22}$$

The sum $(A + B)$ is determined by the trace anomaly

$$A + B = \frac{1}{5760\pi^2}\, C^{\alpha\beta\gamma\delta}\, C_{\alpha\beta\gamma\delta} = \frac{1}{7680\pi^2 M^4} . \tag{11.3.23}$$

Define $A = (7680\pi^2 M^4)^{-1}\alpha$ and $B = (7680\pi^2 M^4)^{-1}\beta$. Using the mode expansions and the asymptotic form of the modes near the horizon, Candelas (1980) obtained by numerical calculations the following values for the coefficients α and β:

$$\alpha = 0.786, \qquad \beta = 0.214 . \tag{11.3.24}$$

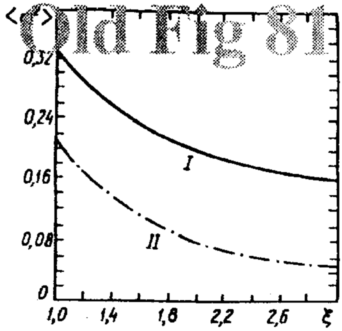


Figure 11.2: The values of $(8\pi M)^2\langle\hat{\varphi}^2\rangle$ as functions of $\xi = r/M - 1$: curve I – $\langle\hat{\varphi}^2\rangle^H$, curve II – $\langle\hat{\varphi}^2\rangle^U$.

At large distances one can neglect the effect of the gravitational field on the thermal bath surrounding the black hole. In accordance with this, $\langle\hat{T}^\mu_\nu\rangle^H$ at large distances has the following asymptotic form

$$\langle\hat{T}^\mu_\nu\rangle^H \approx \frac{\pi^2\theta^4}{90}\,\text{diag}(-3,1,1,1) = \frac{1}{8^4\times 90\times\pi^2 M^4}\,\text{diag}(-3,1,1,1)\,. \quad (11.3.25)$$

Here $\theta = (8\pi M)^{-1}$ is the temperature of the black hole.

11.3.3 Numerical results

Unfortunately, it is not possible to sum the series for $\langle\hat{\varphi}^2\rangle^{\text{ren}}$ and $\langle\hat{T}_{\mu\nu}\rangle^{\text{ren}}$ and to find an explicit final expression for these quantities for arbitrary radius r in the general case. Consequently, either numerical methods or methods of approximate summation of series are used to obtain a result. The work completed up to now has been limited mainly to the case of non-rotating black holes.

Numerical calculations published to date cover mainly the cases of the scalar and electromagnetic fields in the spacetime of a Schwarzschild black hole. First numerical calculations of $\langle\hat{\varphi}^2\rangle^H$ and $\langle\hat{\varphi}^2\rangle^U$ were performed by Fawcett and Whiting (1982). Their results are plotted in Figure 11.2. Howard and Candelas (1984) and Howard (1984) calculated the components $\langle\hat{T}^\mu_\nu\rangle^H$ for a conformal massless scalar field and found (and corrected) an error in the earlier calculations of this quantity by Fawcett (1983). Since there is no net flux of energy in the Hartle-Hawking state one can use the Christensen-Fulling representation to show that the tensor $\langle\hat{T}^\mu_\nu\rangle^H$ in the Schwarzschild geometry is diagonal in $(t-r)$-coordinates, and its value is completely determined by only one component (e.g., by $\langle\hat{T}^\theta_\theta\rangle^H$). To obtain $\langle\hat{T}^\theta_\theta\rangle^H$, one can use the mode expansion of the propagator G^H. It is convenient to make calculations in the Euclidean space obtained by Wick rotation $t = -i\tau$ [see (11.2.22)]. In particular, one has [Candelas (1980), Howard and Candelas (1984), Howard (1984)]

$$G^H(-i\tau, r, \theta, \phi; -i\tau', r', \theta', \phi') = \frac{i}{32\pi^2 M^2} \sum_{n=1}^{\infty} \cos[n\kappa(\tau - \tau')]$$

$$\times \sum_{\ell=0}^{\infty} (2\ell+1) P_\ell(\cos\gamma) \left[\frac{1}{n} p_\ell^n(\xi_<)\, q_\ell^n(\xi_>) - 2P_\ell(\xi_<)\, Q_\ell(\xi_>) \right], \quad (11.3.26)$$

where $\xi = r/M - 1$, $\kappa = 1/4M$,

$$\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi - \phi'), \quad (11.3.27)$$

and $\xi_<$ and $\xi_>$ denote the smaller and greater of ξ and ξ'. P_ℓ and Q_ℓ are Legendre functions, and p_ℓ^n and q_ℓ^n are solutions of the radial equation

$$\left[\frac{d}{d\xi}(\xi^2 - 1)\frac{d}{d\xi} - \ell(\ell+1) - \frac{n^2(1+\xi)^4}{16(\xi^2-1)} \right] R(\xi) = 0, \quad (11.3.28)$$

specified by the requirements that, for $n > 0$, $p_\ell^n(\xi)$ is the solution that remains bounded as $\xi \to 1$, and $q_\ell^n(\xi)$ is the solution that tends to zero as $\xi \to \infty$. First one applies the differential operator (11.1.8) to obtain $\langle \hat{T}^\theta_\theta \rangle^H$ for the split points. After this one puts $\tau = \tau' = \epsilon$, $r = r'$, $\theta = \theta'$, and $\phi = \phi'$.

By using the general expression obtained by Christensen (1976) for the subtracted term that is needed for the renormalization procedure, one can identically rewrite it in the form of a sum over n. After the subtraction one can take the limit $\epsilon \to 0$. In order to improve the convergence of the series in ℓ, Howard (1984) added to and subtracted from each term of the series a similar term in which the solutions p_ℓ^n and q_ℓ^n are replaced by similar solutions of (11.3.28) in the WKB approximation. As a result, one gets

$$\langle \hat{T}^\theta_\theta \rangle^H = \langle \hat{T}^\theta_\theta \rangle^H_{\rm analytic} + \langle \hat{T}^\theta_\theta \rangle^H_{\rm numeric}. \quad (11.3.29)$$

The first term $\langle \hat{T}^\theta_\theta \rangle^H_{\rm analytic}$, which requires for its calculation the solutions in the WKB approximation, can be obtained in explicit analytical form. The second term $\langle \hat{T}^\theta_\theta \rangle^H_{\rm numeric}$ requires numerical calculations. The above-described procedure guarantees that the series for $\langle \hat{T}^\theta_\theta \rangle^H_{\rm numeric}$ converges rapidly enough. Moreover, the calculations show that this term is a small correction to $\langle \hat{T}^\theta_\theta \rangle^H_{\rm analytic}$. Representation similar to (11.3.29) can be written for all the components of $\langle \hat{T}^\mu_\nu \rangle^H$:

$$\langle \hat{T}^\mu_\nu \rangle^H = \langle \hat{T}^\mu_\nu \rangle^H_{\rm analytic} + \langle \hat{T}^\mu_\nu \rangle^H_{\rm numeric}. \quad (11.3.30)$$

The calculations give [Howard and Candelas (1984), Howard (1984)]

$$\langle \hat{T}^\mu_\nu \rangle^H_{\rm analytic} = \frac{\pi^2 \theta^4}{90}\, t^\mu_\nu, \quad (11.3.31)$$

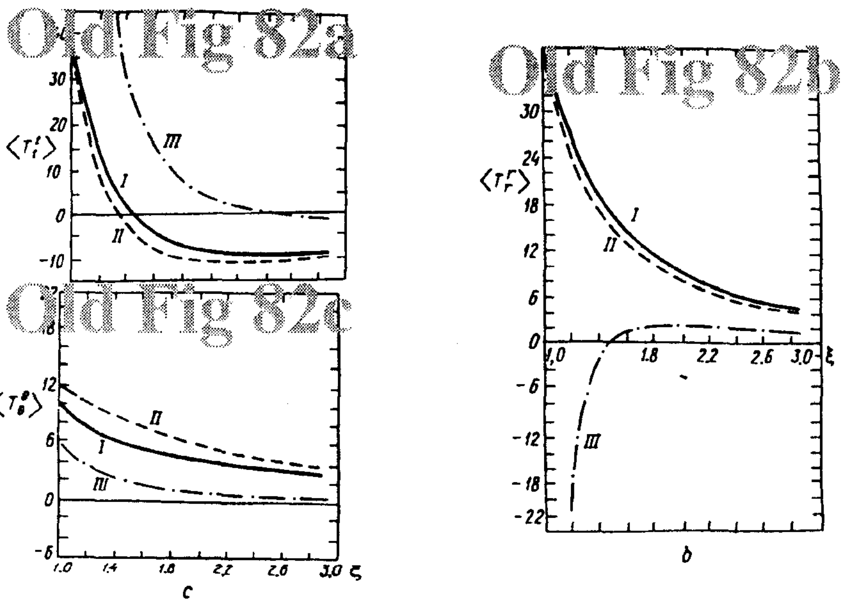


Figure 11.3: The values of the components $[90(8\pi M)^4/\pi^2]\langle \hat{T}^\mu_\nu \rangle$ for a massless scalar field in the Schwarzschild geometry as functions of $\xi = r/M - 1$. The figures show the behavior of the $\langle \hat{T}^t_t \rangle$ (a), $\langle \hat{T}^r_r \rangle$ (b), and $\langle \hat{T}^\theta_\theta \rangle = \langle \hat{T}^\phi_\phi \rangle$ (c) components of the energy-momentum tensor. Curves I – components $\langle \hat{T}^\mu_\nu \rangle^H$, curves II – components $\langle \hat{T}^\mu_\nu \rangle^H_{\text{analytic}}$, curves III – components $\langle \hat{T}^\mu_\nu \rangle^U$.

$$t^\mu_\nu = \frac{1 - z^6(4-3z)^2}{(1-z)^2}\,\text{diag}(-3,1,1,1) + 24\,z^6\,\text{diag}(3,1,0,0)\,, \tag{11.3.32}$$

where $z = 2M/r$.

The renormalized expectation value of quantum stress-energy tensor for a conformal massless scalar field in the Unruh vacuum state in the Schwarzschild geometry was calculated by Elster (1983b). This tensor has a non-vanishing non-diagonal component $\langle \hat{T}^r_t \rangle^U = -W/(4\pi r^2)$, where W is the black hole luminosity (11.3.19).

Figure 11.3 plots the nonzero components $\langle \hat{T}^\mu_\nu \rangle^H$ calculated by Howard and Candelas (1984) and Howard (1984), and the diagonal components of $\langle \hat{T}^\mu_\nu \rangle^U$ calculated by Elster (1983b).[10]

The renormalized expectation value of the stress-energy tensor in the Boulware vacuum state $\langle \hat{T}^\mu_\nu \rangle^B$ for a massless, conformally coupled scalar field in the Schwarzschil

[10]It is interesting to note that the components of $\langle \hat{T}^\mu_\nu \rangle^U$ describing the energy and angular momentum fluxes across the horizon of a rotating black hole can be calculated using a simple approach based on the equivalence principle [Frolov and Thorne (1989)]. For a general discussion of the equivalence principle in the quantum domain, see, e.g., Ginzburg and Frolov (1987).

spacetime was calculated by Jensen, McLaughlin, and Ottewill (1992). Numerical calculations of $\langle \hat{T}^{\mu}_{\nu} \rangle^{H}$ and $\langle \hat{\varphi}^2 \rangle^{H}$ in the Schwarzschild metric in the presence of boundaries were done by Elster (1982a,b, 1983a).

Jensen and Ottewill (1989) performed numerical calculations of the renormalized quantum stress-energy tensor for an electromagnetic field in the Hartle-Hawking state in the Schwarzschild geometry. They applied a similar procedure of extracting the semi-classical "analytic" part of $\langle \hat{T}^{\mu}_{\nu} \rangle^{H}_{\rm analytic}$ and numerically summing the series for $\langle \hat{T}^{\mu}_{\nu} \rangle^{H}_{\rm numeric}$ that has an improved convergence. Jensen and Ottewill (1989) obtained for the non-vanishing components of $\langle \hat{T}^{\mu}_{\nu} \rangle^{H}_{\rm analytic}$:

$$\langle \hat{T}^{\mu}_{\nu} \rangle^{H}_{\rm analytic} = \frac{\pi^2 \theta^4}{45}\, t^{\mu}_{\nu}\,, \tag{11.3.33}$$

$$t^{t}_{t} = -3 - 6z - 9z^2 - 12z^3 + 315z^4 - 78z^5 + 249z^6\,,$$

$$t^{r}_{r} = 1 + 2z + 3z^2 - 76z^3 + 295z^4 - 54z^5 + 285z^6\,, \tag{11.3.34}$$

$$t^{\theta}_{\theta} = t^{\phi}_{\phi} = 1 + 2z + 3z^2 + 44z^3 - 305z^4 + 66z^5 - 579z^6\,.$$

At large distances the renormalized quantum stress-energy tensor for the electromagnetic field has the same expected thermal behavior as for the massless scalar field

$$\langle \hat{T}^{\mu}_{\nu} \rangle^{H} \approx \frac{\pi^2 \theta^4}{90}\, h(s)\, \mathrm{diag}(-3, 1, 1, 1)\,. \tag{11.3.35}$$

The only difference is connected with the different number of polarization states ($h(0) = 1$, $h(1) = 2$). Jensen, McLaughlin, and Ottewill (1995) made similar calculations for gravitational perturbations in the Schwarzschild spacetime. The renormalized quantum stress-energy tensor for the electromagnetic field in the state of the Unruh vacuum was calculated by Jensen, McLaughlin, and Ottewill (1991).

Recently, Anderson, Hiscock, and Samuel (1995) generalized the results of Howard and Candelas (1984) and Howard (1984) and developed a method for computing $\langle \hat{\varphi}^2 \rangle$ and the stress-energy tensor $\langle \hat{T}^{\mu}_{\nu} \rangle$ of quantized scalar fields for the Boulware and Hartle-Hawking states in an arbitrary static spherically symmetric spacetime. The fields can be massless or massive with an arbitrary coupling ξ to the scalar curvature. By using the WKB approximation for the radial solutions, the authors developed a general scheme that allows one to represent $\langle \hat{T}^{\mu}_{\nu} \rangle$ in the form (11.3.29). A similar representation is valid for $\langle \hat{\varphi}^2 \rangle$:

$$\langle \hat{\varphi}^2 \rangle = \langle \hat{\varphi}^2 \rangle_{\rm analytic} + \langle \hat{\varphi}^2 \rangle_{\rm numeric}\,. \tag{11.3.36}$$

For the metric

$$ds^2 = -f(r)\, dt^2 + h(r)\, dr^2 + r^2\, d\omega^2\,, \tag{11.3.37}$$

one has [Anderson (1990), Anderson, Hiscock, and Samuel (1995)]

$$\langle\hat{\varphi}^2\rangle_{\text{analytic}} = \frac{\theta^2}{12f} + \frac{m^2}{16\pi^2} - \frac{1}{16\pi^2}\left[m^2 + \left(\xi - \frac{1}{6}\right)R\right]\ln\left(\frac{\mu^2 f}{4\lambda^2}\right) - \frac{f'^2}{96\pi^2 f^2 h} - \frac{f'h'}{192\pi^2 f h^2} + \frac{f''}{96\pi^2 f h} + \frac{f'}{48\pi^2 r f h}. \tag{11.3.38}$$

Here the prime denotes the derivative with respect to r, $\lambda = 2\pi\theta\exp(-C)$, and C is Euler constant. The parameter μ reflects the ambiguity in the renormalization procedure. For a conformal massless scalar field in the Schwarzschild spacetime relation (11.3.38) reproduces the expression obtained earlier by Fawcett and Whiting (1982)

$$\langle\hat{\varphi}^2\rangle^H = \frac{1}{768\,\pi^2 M^2}\,\frac{1-z^4}{1-z}. \tag{11.3.39}$$

The explicit expression for $\langle\hat{T}^\mu_\nu\rangle_{\text{analytic}}$ can be found in [Anderson, Hiscock, and Samuel (1995)]. This paper also contains the results of numerical calculations of the stress-energy tensor of massive and massless fields in Schwarzschild and Reissner-Nordström spacetimes.

11.3.4 Thermal atmosphere of black holes

The main feature of the tensor $\langle\hat{T}^\mu_\nu\rangle^H$ is that its components are finite on the event horizon. An observer at rest at a point r close to the event horizon records the local energy density $\epsilon = -\langle\hat{T}^t_t\rangle^H$. This quantity remains finite as $r \to r_+$. On the other hand, the temperature measured by the observer,

$$\theta_{\text{loc}} = \frac{\kappa}{2\pi}\left(1 - \frac{r_+}{r}\right)^{-1/2}, \tag{11.3.40}$$

grows infinitely near the horizon. This local temperature can be measured by using a two-level system as a thermometer: Transitions between levels are caused by the absorption and emission of quanta of the field (photons). After a sufficiently long exposure, the probability for a system to occupy the upper level will be less than that for the lower level by a factor $\exp(\Delta E/\theta_{\text{loc}})$ (ΔE is the energy difference between the levels). Other small-size detectors can be used in a similar manner [Unruh (1976b)]. It is not difficult to show that the temperature in the vicinity of r_+ is $\theta_{\text{loc}} \approx \theta_a = a/2\pi$, where a is the observer's acceleration; as $r \to r_+$, $\theta_{\text{loc}} \to \infty$.[11]

[11] Likewise, an observer moving at an acceleration a in a flat spacetime will also record a temperature $\theta_a = a/2\pi$ using the thermometers described above. This is so-called *Unruh effect*. From the standpoint of the accelerated observer, the standard Minkowski vacuum behaves, in a sense, just like thermal radiation at a temperature θ_a. Note that while recording the energy of these thermal-energy particles, the observer cannot measure their momentum with sufficient accuracy because the characteristic wavelength of this "radiation" is of the order of distance to the horizon. The same remark holds for the "particle" of thermal radiation recorded by an observer close to the black hole. For details, see Unruh (1976b), Unruh and Wald (1984).

The results plotted in Figure 11.3 imply that the radiation energy density ϵ in the neighborhood of such a point is $\epsilon \sim \sigma(\kappa/2\pi)^4 \ll \sigma\theta_a^4$. The law of $\sigma\theta^4$ is evidently violated. The reason of this violation is that close to the horizon the gravitational field changes greatly over distances of the order of the characteristic wavelength of thermal radiation. The fact that, in contrast to $\sigma\theta_a^4$, the quantity $\epsilon = -\langle \hat{T}^t_t\rangle^H$ is finite at the horizon, can be interpreted as follows: The contribution of vacuum polarization caused by a strong gravitational field close to the horizon compensates exactly for the divergence that would take place in the radiation energy density if the law $\sigma\theta_a^4$ were not violated.

To confirm this conclusion, let us consider a spacetime region close to the horizon. Denote by ρ the proper distance from the horizon

$$\rho = \int \frac{dr}{\sqrt{1-2m/r}} = \sqrt{r(r-2M)} + M\ln\frac{r-M+\sqrt{r(r-2M)}}{M}\,. \qquad (11.3.41)$$

For small ρ one has $r \approx 2M + \rho^2/8M$. Introduce new coordinates X, and Y

$$X^2 + Y^2 = 16M^2\tan^2(\theta/2)\,, \qquad Y/X = \tan\phi\,. \qquad (11.3.42)$$

If mass parameter M now tends to infinity with the values of the coordinates $(t, \rho, X, Y$ fixed, the Schwarzschild metric transforms into the metric

$$ds^2 = -\kappa^2\rho^2 dt^2 + d\rho^2 + dX^2 + dY^2\,. \qquad (11.3.43)$$

This is a so-called *Rindler metric* describing a homogeneous gravitational field (see Appendix C.2).

Consider a thermal gas of massless scalar particles in a homogeneous gravitational field. Since the local temperature depends on the distance ρ from the horizon, we choose the normalization point at $\rho = \kappa^{-1}$ and denote by θ the value of the local temperature at this point. The local temperature at other points is $\theta_{\rm loc}(\rho) = \theta/(\kappa\rho)$. The corresponding propagator $G(x,x')$ for the scalar field in this state can be obtained by analytical continuation of the Euclidean Green's function $G^E(x,x')$ which is a solution of the equation

$$\Box_E\, G^E(x,x') = -\,\delta(x,x')\,, \qquad (11.3.44)$$

in the space with the Euclidean metric

$$ds_E^2 = \kappa^2\,\rho^2\,d\tau^2 + d\rho^2 + dX^2 + dY^2\,. \qquad (11.3.45)$$

It is assumed that τ is periodic, and its period is θ^{-1}. For a value of the period different from $2\pi/\kappa$ there is a cone-like singularity at $\rho = 0$. The Euclidean Green's function $G^E(x,x')$ for such a space was obtained by Dowker (1978) and is of the form

$$G^E(x,x') = \frac{\theta}{4\pi\,\kappa\,\rho\,\rho'\sinh\lambda}\;\frac{\sinh(2\pi\,\theta\,\lambda/\kappa)}{\cosh(2\pi\,\theta\,\lambda/\kappa) - \cos[\,2\pi\,\theta\,(\tau-\tau')]}\,, \qquad (11.3.46)$$

where $x = (\tau, \rho, X, Y)$, $x' = (\tau', \rho', X', Y')$, and

$$\cosh\lambda = \frac{\rho^2 + \rho'^2 + (X - X')^2 + (Y - Y')^2}{2\rho\,\rho'} . \tag{11.3.47}$$

For renormalization it is sufficient to subtract the flat space Green's function, which in Rindler coordinates is

$$G_0^E(x, x') = \frac{1}{8\pi^2\rho\,\rho'}\,\frac{1}{\cosh\lambda - \cos[(\tau - \tau')\kappa]} . \tag{11.3.48}$$

The expression for $\langle\hat{\varphi}^2\rangle$ can be obtained by taking the coincidence limit

$$\langle\hat{\varphi}^2\rangle = \lim_{x\to x'} \left[G^E(x, x') - G_0^E(x, x')\right] .$$

The components of $\langle\,\hat{T}_\nu^\mu\rangle$ can be found applying the differential operator (11.1.8) to the difference $G^E - G_0^E$ and after that taking the coincidence limit. Straightforward calculations give

$$\langle\hat{\varphi}^2\rangle = \frac{1}{12(\kappa\,\rho)^2}\left[\theta^2 - \left(\frac{\kappa}{2\pi}\right)^2\right] , \tag{11.3.49}$$

$$\langle\,\hat{T}_\nu^\mu\rangle = \frac{\pi^2}{90(\rho\,\kappa)^4}\left[\theta^4 - \left(\frac{\kappa}{2\pi}\right)^4\right] \mathrm{diag}(-3, 1, 1, 1) . \tag{11.3.50}$$

Now note that $\theta_{\mathrm{loc}} = \theta/(\kappa\rho)$ is a local temperature of thermal radiation, and $\theta_U = w/2\pi = 1/(2\pi\rho)$ is the Unruh temperature, which is determined by the acceleration $w = 1/\rho$ of an observer who is at rest in the gravitational field. This allows one to rewrite (11.3.49) and (11.3.50) as follows

$$\langle\hat{\varphi}^2\rangle = \frac{1}{12}\left[\theta_{\mathrm{loc}}^2 - \theta_U^2\right] , \tag{11.3.51}$$

$$\langle\,\hat{T}_\nu^\mu\rangle = \frac{\pi^2}{90}\left[\theta_{\mathrm{loc}}^4 - \theta_U^4\right] \mathrm{diag}(-3, 1, 1, 1) . \tag{11.3.52}$$

In the more general case where beside massless scalar particles there are massless fermions (neutrinos) and photons, one has[12]

$$\langle\,\hat{T}_\nu^\mu\rangle = \frac{\pi^2}{90}\left(\theta_{\mathrm{loc}}^2 - \theta_U^2\right) \times \left[\left[h(0) + \frac{7}{8}\,h(1/2) + h(1)\right]\theta_{\mathrm{loc}}^2\right.$$

[12] This result follows from the expression for $\langle\,\hat{T}_\nu^\mu\rangle$ in a spacetime of a cosmic string obtained by Frolov and Serebriany (1987) [see also Dowker (1987)]. The metric of a cosmic string can be obtained from the Euclidean metric of the conical space (11.3.45) by making the analytical continuation $Y \to iT$. Since both metrics (for a gravitational field of a cosmic string and for the homogeneous gravitational field) can be obtained from the same Euclidean metric (11.3.45) by means of different analytical continuations, there exists a simple relation between $\langle\,\hat{T}_\nu^\mu\rangle$ in these two spacetimes.

$$- \left[h(0) + \frac{17}{8}\, h(1/2) + 11\, h(1) \right] \theta_U^2 \Big] \operatorname{diag}(-3,1,1,1)\,. \tag{11.3.53}$$

In the absence of the gravitational field (when $\theta_U = 0$), these expressions reproduce the standard results in flat spacetime. Terms modifying the standard relations are connected with the presence of a force acting on virtual particles which changes their state of motion. Hence, the terms containing the Unruh temperature describe the vacuum polarization effect. In a homogeneous gravitational field both of these terms (those connected with thermal effects and with the vacuum polarization) have the same structure and cancel one another when $\theta = \kappa/2\pi$. In the presence of curvature the cancellation is not exact but the corresponding curvature dependent corrections are finite. This explains why $\langle\hat\varphi^2\rangle^H$ and $\langle \hat T^\mu_\nu\rangle^H$ remain finite at the horizon.

11.3.5 Analytical approximations

Page's approximation

Page (1982) suggested a method of approximate computation of $\langle\hat\varphi^2\rangle^H$ and $\langle \hat T^\mu_\nu\rangle^H$ for a conformal invariant massless scalar field. The value of $\langle\hat\varphi^2\rangle^H$ in Page's approximation, $\langle\hat\varphi^2\rangle^H_P$, coincides with (11.3.39). Candelas and Howard (1984), Howard and Candelas (1984), and Howard (1984) have demonstrated that for the conformal massless scalar field the values of $\langle\hat\varphi^2\rangle^H_P$ and $\langle \hat T^\mu_\nu\rangle^H_P$ in Page's approximation fit very well the behavior of $\langle\hat\varphi^2\rangle^H$ and $\langle \hat T^\mu_\nu\rangle^H$ (the deviations of $\langle\hat\varphi^2\rangle^H_P$ from the true value of $\langle\hat\varphi^2\rangle^H$ do not exceed 1%, and those of the components $\langle \hat T^\mu_\nu\rangle^H_P$ do not exceed 20%).

Two propositions are basic for constructing Page's approximation:

(1) *Let there be two conformal spaces, and let the calculation of renormalized expectation values $\langle\hat\varphi^2\rangle$ and $\langle \hat T^\mu_\nu\rangle$ for a conformal massless scalar field be carried out in each of these spaces in the states obtained from each other by the same conformal transformation. Then the following combinations containing $\langle\hat\varphi^2\rangle$*

$$Z = g^{1/4} \left[\langle\hat\varphi^2\rangle + \frac{1}{288\,\pi^2} R \right] \tag{11.3.54}$$

and $\langle \hat T^\mu_\nu\rangle$

$$J^\mu_\nu = g^{1/2} \Big\{ \langle \hat T^\mu_\nu\rangle + \alpha \left[(C^{\alpha\mu}{}_{\beta\nu} \ln g)_{;\alpha}{}^{;\beta} + \frac{1}{2}\, R^\beta_\alpha\, C^{\alpha\mu}{}_{\beta\nu} \ln g \right]$$

$$+ \beta \left[2H^\mu_\nu - 4\, R^\beta_\alpha\, C^{\alpha\mu}{}_{\beta\nu} \right] + \frac{1}{6}\, \gamma\, I^\mu_\nu \Big\} \tag{11.3.55}$$

are invariant and do not depend on which conformal space they are calculated in.

Here

$$H_{\mu\nu} = -R^{\alpha}_{\mu}\,R_{\alpha\nu} + \frac{2}{3}\,R\,R_{\mu\nu} + \left(\frac{1}{2}\,R^{\alpha}_{\beta}\,R^{\beta}_{\alpha} - \frac{1}{4}\,R^2\right) g_{\mu\nu}\,, \tag{11.3.56}$$

$$I_{\mu\nu} = 2R_{;\mu\nu} - 2R\,R_{\mu\nu} + \left(\frac{1}{2}R^2 - 2R_{;\alpha}{}^{;\alpha}\right) g_{\mu\nu}\,. \tag{11.3.57}$$

The conformal trace-anomaly coefficients α, β, and γ are given by expressions (11.1.13) (11.1.15).

(2) *Let the metric*

$$ds^2 = -V^2dt^2 + h_{ij}\,dx^i\,dx^j \tag{11.3.58}$$

be a static solution of Einstein's vacuum equations ($V^2 = -\xi_{(t)\mu}\,\xi^{\mu}_{(t)}$, $\xi^{\mu}_{(t)}$ being the Killing vector field). Then the conformal trace anomaly vanishes in the space with the metric $d\tilde{s}^2 = V^{-2}\,ds^2$.

The metric $d\tilde{s}^2$ has the form

$$d\tilde{s}^2 = -dt^2 + \tilde{h}_{ij}\,dx^i\,dx^j$$

and is often called an *ultrastatic metric.* Page suggested to calculate $\langle\hat{\varphi}^2\rangle$ and $\langle\hat{T}^{\mu}_{\nu}\rangle$ first in the space with the ultrastatic metric $d\tilde{s}^2$, using the solution obtained by WKB approximation for a Green's function in it, and then return to the original physical space taking into account the invariance of the quantities (11.3.54) and (11.3.55). This approach gives for $\langle\hat{\varphi}^2\rangle^H$ of the scalar massless field in the Schwarzschild metric the expression (11.3.39), and for $\langle\hat{T}^{\mu}_{\nu}\rangle^H$ it gives the following approximate formula

$$\langle\hat{T}^{\mu}_{\nu}\rangle^H_P = \frac{\pi^2}{90}\,\theta^4_H\left[\frac{1-(4-3z)^2z^6}{(1-z)^2}(\delta^{\mu}_{\nu} - 4\,\delta^{\mu}_0\,\delta^0_{\nu}) + 24\,z^6(3\,\delta^{\mu}_0\,\delta^0_{\nu} + \delta^{\mu}_1\,\delta^1_{\nu})\right], \tag{11.3.59}$$

where $\theta_H = (8\pi M)^{-1}$ is the black hole temperature. The behavior of the nonzero components $\langle\hat{T}^{\mu}_{\nu}\rangle^H_P$ is shown in Figure 11.3 by dashed lines. This approximate expression was used by York (1985) for analyzing the back-reaction of vacuum polarization on the gravitational field of a black hole.

Page's method can also be used for determining the approximate values of $\langle\hat{\varphi}^2\rangle$ and $\langle\hat{T}^{\mu}_{\nu}\rangle$ in the Boulware vacuum. Calculations give [Frolov and Zel'nikov (1985a)]

$$\langle\hat{\varphi}^2\rangle^B_P = -\frac{M^2}{48\,\pi^2r^4(1-z)}\,, \tag{11.3.60}$$

$$\langle\hat{T}^{\mu}_{\nu}\rangle^B_P = \frac{M^2}{1440\,\pi^2r^6}\left[\left(\frac{4-3z}{2(1-z)}\right)^2(-\delta^{\mu}_{\nu} + 4\,\delta^{\mu}_0\,\delta^0_{\nu}) + 6(3\,\delta^{\mu}_0\,\delta^0_{\nu} + \delta^{\mu}_1\,\delta^1_{\nu})\right]. \tag{11.3.61}$$

Expressions (11.3.60)–(11.3.61) agree with asymptotic formulas (11.3.16) and (11.3.21) near the event horizon and manifest the correct asymptotic behavior at infinity. It seems that the accuracy with which (11.3.60) and (11.3.61) reproduce the exact values of $\langle \hat{\varphi}^2 \rangle^B$ and $\langle \hat{T}^\mu_\nu \rangle^B$ is of the same order as in the case of the Hartle-Hawking vacuum. Later it was shown that the analytic part which arise in the numerical calculations coincides with the result of Page's approximation [see Anderson (1990), Anderson, Hiscock, and Samuel (1995)]

Brown-Ottewill-Page approximation

Brown and Ottewill (1985) suggested a different method of determining $\langle \hat{\varphi}^2 \rangle$ and $\langle \hat{T}^\mu_\nu \rangle$ for conformally invariant fields. In the case of a conformal massless scalar field, the method again yields the expressions (11.3.39) and (11.3.59) found by Page's approximation. Brown and Ottewill noticed that conformal anomalies vanish not only in the spaces with the ultrastatic metric $d\tilde{s}^2 = V^{-2}\, ds^2$, but also in a wider class of spaces whose metrics are of the form $d\tilde{s}^2 = \exp(at)\, V^{-2}\, ds^2$. If we demand that the choice of state be such that not only the trace but all other components of $\langle \hat{T}^\mu_\nu \rangle$ in the space $d\tilde{s}^2$ vanish as well, then a quite definite value of $\langle \hat{T}^\mu_\nu \rangle$ is obtained after we return to the original space. Brown and Ottewill showed that for the Schwarzschild geometry the thus obtained expression for $a = 0$ coincides with $\langle \hat{T}^\mu_\nu \rangle^B_P$, and that for $a = -2\kappa = -(2M)^{-1}$, it correctly reproduces $\langle \hat{T}^\mu_\nu \rangle^H_P$. Similar approximate expressions can be obtained in the framework of this approach for the contributions of the neutrino and electromagnetic fields to $\langle \hat{T}^\mu_\nu \rangle$.

A refined version that unites this approach with Page's results was proposed by Brown, Ottewill, and Page (1986). The starting point of this work is an observation that $\langle \hat{T}^\mu_\nu \rangle^{\rm ren}$ can be obtained by the variation of the effective action

$$\langle \hat{T}^{\mu\nu} \rangle^{\rm ren} = \frac{2}{\sqrt{-g}} \frac{\delta W}{\delta g_{\mu\nu}} . \tag{11.3.62}$$

Under a conformal transformation the renormalized one-loop effective action $W[g_{\mu\nu}]$ for a massless conformal invariant field transforms according to the equation [Brown (1984), Brown and Ottewill (1985), Dowker (1986)]

$$W[e^{-2\omega} g_{\mu\nu}] = W[g_{\mu\nu}] - \alpha\, A[\omega; g] - \beta B[\omega; g] - \gamma\, C[\omega; g] , \tag{11.3.63}$$

where

$$A[\omega; g] = \int d^4x \sqrt{-g} \left\{ \omega \mathcal{H} + \frac{2}{3}[R + 3(\Box\omega - w^2)](\Box\omega - w^2) \right\} , \tag{11.3.64}$$

$$B[\omega; g] = \int d^4x \sqrt{-g} \left[\omega \mathcal{K} + 4R_{\mu\nu}\, \omega^{;\mu} \omega^{;\nu} - 2R\, w^2 + 2(w^2)^2 - 4w^2 \Box\omega \right] , \tag{11.3.65}$$

$$C[\omega; g] = \int d^4x \sqrt{-g} \left\{ [R + 3(\Box\omega - w^2)] \, (\Box\omega - w^2) \right\} . \tag{11.3.66}$$

Here $\mathcal{H} = C_{\mu\nu\rho\tau}C^{\mu\nu\rho\tau}$, $\mathcal{K} = {}^*R_{\mu\nu\rho\tau}{}^*R^{\mu\nu\rho\tau}$, $w^2 = \omega_{;\mu}\omega^{;\mu}$, and the coefficients α, β, and γ are the coefficients of the conformal trace anomaly given by (11.1.13)–(11.1.15).

Using the relation

$$\left[\frac{\delta}{\delta\omega}W[e^{-\omega}g_{\alpha\beta}]\right]_{\omega=0} = -2g_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}W[g_{\alpha\beta}] \equiv -\sqrt{-g}\,\langle\hat{T}^\mu_\mu\rangle^{\text{ren}}, \tag{11.3.67}$$

one gets

$$\langle\hat{T}^\mu_\mu\rangle^{\text{ren}} = \alpha\left(\mathcal{H} + \frac{2}{3}\Box R\right) + \beta\mathcal{K} + \gamma\Box R. \tag{11.3.68}$$

This relation reproduces the *conformal trace anomaly* (11.1.10).

Consider a static spacetime with a Killing vector ξ and choose $\omega = (1/2)\ln|\xi^2|$ in (11.3.63). Then we have

$$e^{-6\omega}(\tilde{T}^{\mu\nu} - \xi^\mu\xi^\nu\tilde{T}^\mu_\mu) = T^{\mu\nu} - \alpha T^{\mu\nu}_A - \beta T^{\mu\nu}_B - \gamma T^{\mu\nu}_C. \tag{11.3.69}$$

Here $\tilde{T}^{\mu\nu}$ is the renormalized stress-energy tensor in the *ultrastatic spacetime* with metric $|\xi^2|^{-1}g_{\mu\nu}$, and $T^{\mu\nu}_A$, $T^{\mu\nu}_B$, and $T^{\mu\nu}_C$ are stress-energy tensors obtained by the variation of $A(\omega;g)$, $B(\omega;g)$, and $C(\omega;g)$, respectively. In this variation procedure one keeps ξ^α fixed, i.e., one puts $\delta\xi^\alpha/\delta g_{\mu\nu} = 0$.

In the ultrastatic spacetime obtained by a conformal transformation from an Einstein spacetime ($R_{\mu\nu} = \Lambda g_{\mu\nu}$) the conformal trace anomaly vanishes.[13] The Gaussian approximation used by Page (1982) for vanishing temperature $\theta = 0$ is equivalent to the the choice $\tilde{T}^\mu_\nu = 0$. In the more general case when $\tilde{T}^\mu_\mu \neq 0$, Brown, Ottewill, and Page (1986) use the approximation ansatz [Zannias (1984)]

$$\tilde{T}^{\mu\nu} = \xi^\mu\xi^\nu\tilde{T}^\alpha_\alpha. \tag{11.3.70}$$

For this ansatz one has

$$\langle\hat{T}^{\mu\nu}\rangle_{BOP} = \alpha T^{\mu\nu}_A + \beta T^{\mu\nu}_B + \gamma T^{\mu\nu}_C. \tag{11.3.71}$$

It was demonstrated by Brown and Ottewill (1985) that the choice of conformal factor

$$\bar{\omega} = \kappa t + \frac{1}{2}\ln\left(|\xi^2|\right) \tag{11.3.72}$$

yields, under the assumptions $\gamma = 0$ and $\tilde{T}^{\mu\nu} = 0$, an approximation to the renormalized stress-energy tensor for a thermal state with temperature $\theta = \kappa/2\pi$. Alternatively, one can fix the temperature-dependent terms by demanding agreement with the values in the open Einstein universe, where one knows that the Gaussian

[13] This is not so for ζ-function or point-splitting renormalization in the case of spin 1.

approximation is exact. The latter procedure gives the following expression for the effective action in the BOP-approximation

$$W_{BOP} = \alpha\, A[\omega; g] + \beta B[\omega; g] + \gamma\, C[\omega; g]$$

$$+ q^{(2)}(2\pi)^2\theta^2 \int d^4x \sqrt{-g}\, \frac{1}{|\xi^2|^{1/2}} \left(\Box - \frac{1}{6}\right) \frac{1}{|\xi^2|^{1/2}} \tag{11.3.73}$$

$$+ q^{(4)}(2\pi)^4\theta^4 \int d^4x \sqrt{-g}\, \frac{1}{|\xi^2|^2},$$

where

$$q^{(2)} = \frac{1}{3}\alpha + \beta - \frac{1}{2}\gamma, \tag{11.3.74}$$

$$q^{(4)} = 2\alpha + 2\beta - \gamma. \tag{11.3.75}$$

A comparison of $\langle \hat{T}^\mu_\nu \rangle^H_{BOP}$ with the results of numerical study for the electromagnetic field [Jensen and Ottewill (1989)] shows that the difference becomes large in the vicinity of the horizon.

Frolov-Zel'nikov approximation

Frolov and Zel'nikov (1987, 1988) suggested a different approach which allows one to obtain an approximation for $\langle \hat{\varphi}^2 \rangle$ and $\langle \hat{T}^\mu_\nu \rangle$ in an arbitrary static spacetime. The basic idea of this approach is to use for the approximation scalars and tensors constructed from the curvature tensor, the Killing vector, and their covariant derivatives up to some given order. This requirement was called a *Killing ansatz.* By requiring that the so-constructed quantities possess known properties of $\langle \hat{\varphi}^2 \rangle$ and $\langle \hat{T}^\mu_\nu \rangle$, one gets expressions that can be used to approximate these quantities.

Consider a scalar $\langle \hat{\varphi}^2 \rangle_K$ constructed in the framework of the Killing ansatz that:

1. Has the same dimensionality as $\langle \hat{\varphi}^2 \rangle$, i.e., $(\text{length})^{-2}$
2. Is static, $\mathcal{L}_\xi \langle \hat{\varphi}^2 \rangle_K = 0$, and invariant under time reflection $t \to -t$
3. Is a polynomial in temperature θ and has high-temperature behavior $\langle \hat{\varphi}^2 \rangle_K \sim \theta^2/12$
4. Correctly reproduces the transformation law of $\langle \hat{\varphi}^2 \rangle$ under conformal transformations, i.e.,

$$|g|^{1/4} \left[\langle \hat{\varphi}^2 \rangle_K + \frac{R}{288\pi^2} \right] \tag{11.3.76}$$

is a conformal invariant.

Analysis shows [Frolov and Zel'nikov (1987, 1988)] that under these conditions $\langle\varphi^2\rangle_K$ necessarily has the form

$$\langle\hat\varphi^2\rangle_K = \frac{1}{12}\frac{\theta^2}{|\xi^2|} + q\left[w^2 + \frac{\xi^\alpha\xi^\beta}{\xi^2}R_{\alpha\beta} - \frac{1}{6}R\right] - \frac{1}{288\pi^2}R. \tag{11.3.77}$$

Here $w^2 = w_\mu w^\mu$, and $w_\mu = (1/2)\nabla_\mu \ln|\xi^2|$ is the acceleration of the Killing observer. An arbitrary coefficient q in this expression can be fixed if one requires that in the homogeneous gravitational field $\langle\hat\varphi^2\rangle_K$ reproduces the exact expression (11.3.51). This condition gives $q = -1/48\pi^2$. Substituting this value into (11.3.77), one gets

$$\langle\hat\varphi^2\rangle_{FZ} = \frac{1}{12}\left[\frac{\theta^2}{|\xi^2|} - \frac{w^2}{4\pi^2}\right] - \frac{1}{48\pi^2}\frac{\xi^\alpha\xi^\beta}{\xi^2}R_{\alpha\beta}. \tag{11.3.78}$$

Similarly, one can consider a tensor $\langle\hat T_{\mu\nu}\rangle_K$ constructed in the framework of the Killing ansatz that:

1. Has the same dimensionality as $\langle\hat T_{\mu\nu}\rangle$, i.e., (length)$^{-4}$
2. Is static, $\mathcal{L}_\xi\langle\hat T_{\mu\nu}\rangle_K = 0$, and invariant under the time reflection $t \to -t$
3. Is a polynomial in θ and has high-temperature behavior (11.3.35)
4. Is conserved $\langle\hat T^\mu_\nu\rangle_{;\mu} = 0$
5. Correctly reproduces the transformation law of $\langle\hat T_{\mu\nu}\rangle$ under the conformal transformations (11.3.55)

Frolov and Zel'nikov (1987, 1988) showed that such a tensor can be obtained by variation of the following effective action

$$W_{FZ} = W_{BOP} + \int dx^4\,|g|^{1/2}\left[q_1\mathcal{H} + q_2\left(w^2 - w^{;\alpha}_{;\alpha} - \frac{1}{6}R\right)^2\right]. \tag{11.3.79}$$

Here W_{BOP} is the effective action in the Brown-Ottewill-Page approximation, $\mathcal{H} = C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta}$, and q_1 and q_2 are two arbitrary constants.

In Einstein spacetimes the term $\mathcal{H}$ in the action does not contribute to the stress-energy tensor. On the other hand, there exist two more new traceless conserved tensors that can also be obtained by the variation procedure. In Einstein spacetime the most general stress-energy tensor in the Killing ansatz obeying the above conditions can be obtained from the following effective action:

$$\begin{aligned} W_{FZ} = W_{BOP} + \int dx^4\,|g|^{1/2}\Bigg[&q_2\left(w^2 - w^{;\alpha}_{;\alpha} - \frac{1}{6}R\right)^2 \\ &+ q_3\left[-(w^2)^2 + 2w^2 w^{;\alpha}_{;\alpha} + (w^{;\alpha}_{;\alpha})^2 + w^2 R - 2w_\alpha w_\beta R^{\alpha\beta}\right] \\ &+ q_4\left[(w^{;\alpha}_{;\alpha})^2 + \frac{1}{2}\ln|\xi^2|\left(-\frac{3}{4}R_{\alpha\beta}R^{\alpha\beta} + \frac{1}{4}R^2\right)\right]\Bigg]. \end{aligned} \tag{11.3.80}$$

Here q_2, q_3, and q_4 are arbitrary constants. In the general case two new tensors which arise as a result of the variation of the last two terms in the effective action give incorrect values for traces of the stress-energy tensor. However, in Einstein spacetimes this does not happen and their contribution to the trace vanishes.

One can show that if one applies the additional restriction that the corresponding stress-energy remains finite at the horizon, then the most general stress-energy tensor in the Frolov-Zel'nikov approximation in the spacetime of the Schwarzschild black hole is [Frolov and Zel'nikov (1987, 1988)]

$$\langle \hat{T}_{\mu\nu}\rangle_{FZ} = (\alpha+\beta)\,\tau^1_{\mu\nu} + b\,\tau^2_{\mu\nu} + c\,\tau^3_{\mu\nu}\,, \tag{11.3.81}$$

$$\tau^{1\,\mu}_{\ \ \nu} = 96\,\kappa^4\,z^6\,\mathrm{diag}(-1,-1,2,2)\,, \tag{11.3.82}$$

$$\begin{aligned}\tau^{2\,\mu}_{\ \ \nu} = {}&\left(\frac{\kappa}{2\pi}\right)^4\left[168\,z^6\,\mathrm{diag}(-1,1,0,0)\right.\\ &\left.+\left(1+2z+3z^2+4z^3+5z^4+6z^5-105z^6\right)\mathrm{diag}(-3,1,1,1)\right],\end{aligned} \tag{11.3.83}$$

$$\begin{aligned}\tau^{3\,\mu}_{\ \ \nu} = {}&\kappa^4 z^3\left[-12(1+z+z^2+2z^3)\,\mathrm{diag}(-1,1,0,0\right.\\ &\left.+(4+5z+6z^2+15z^3)\,\mathrm{diag}(-3,1,1,1)\right].\end{aligned} \tag{11.3.84}$$

Here α and β are coefficients of the conformal trace anomaly (11.1.13)–(11.1.14),

$$b = \frac{\pi^2}{90}\left[h(0) + \frac{7}{8}\,h(1/2) + h(1)\right], \tag{11.3.85}$$

and c is an arbitrary constant. For $c=0$ the stress-energy tensor $\langle \hat{T}_{\mu\nu}\rangle_{FZ}$ for a scalar massless field coincides with the result obtained in Page approximation. The freedom in the choice of constant c can be used to improve the approximation. In particular one can choose this constant so that $\langle \hat{T}_{\mu\nu}\rangle_{FZ}$ at the event horizon coincides with the exact value of $\langle \hat{T}_{\mu\nu}\rangle^H$. For this choice one has

$$c \approx \frac{1}{2880\pi^2}\left[-\frac{4}{35}h(0)\right], \qquad c = \frac{1}{2880\pi^2}\left[-4h(1)\right]. \tag{11.3.86}$$

The first value applies for the massless scalar field, and the second one, for the electromagnetic field. As far as we know, the exact value of the stress-energy tensor at the horizon for a neutrino field is not known.

We now briefly discuss the application of the Killing ansatz in a case when $R_{\mu\nu} \neq \Lambda\,g_{\mu\nu}$. The most important example is the Reissner-Nordström metric

$$ds^2 = -F\,dt^2 + F^{-1}\,dr^2 + r^2\,d\omega^2\,. \tag{11.3.87}$$

The quantity $\langle\hat{\varphi}^2\rangle_{FZ}$ given by (11.3.78) takes the form

$$\langle\hat{\varphi}^2\rangle_{FZ} = \frac{1}{12}\left(\frac{\kappa}{2\pi}\right)^2 \frac{r^4 - r_+^4}{r^2(r-r_+)(r-r_-)}\,. \tag{11.3.88}$$

It is interesting that at the horizon this gives the exact value for $\langle\hat{\varphi}^2\rangle^H$ obtained earlier by Frolov (1982). The stress-energy tensor calculated in the framework of the Killing ansatz contains a term proportional to $Z^{\mu\nu}\ln|\xi^2|$, where

$$Z^{\mu\nu}=\frac{\delta}{\delta g_{\mu\nu}}\int d^4x\,\sqrt{-g}\,C_{\alpha\beta\gamma\delta}\,C^{\alpha\beta\gamma\delta}\,.$$

In the general case, this term (which vanishes in Einstein spaces) is singular at the horizon. For the Reissner-Nordström metric one has

$$Z^{\mu}_{\nu}=16\,\frac{Q^2}{r^6}\,F(r)\,\mathrm{diag}(3,1,-2,-2)\,. \tag{11.3.89}$$

The components of this tensor in (t,r)-coordinates vanish on the horizon, but the regularity condition which requires that quantity $(T^t_t-T^r_r)/F$ remains finite is not satisfied. For this reason, $\langle\hat{T}^{\mu}_{\nu}\rangle_{FZ}$ is singular at the horizon of a charged black hole, while $\langle\hat{T}^{\mu}_{\nu}\rangle^H$ is finite. According to the results of Anderson, Hiscock, and Samuel (1995) similar singularities are present also in $\langle\hat{\varphi}^2\rangle_{\mathrm{analytic}}$ (when $R\neq 0$) and $\langle\hat{T}^{\mu}_{\nu}\rangle_{\mathrm{analytic}}$ (when $R_{\mu\nu}\neq 0$).

The results obtained up to now seem to indicate that the main contribution to the vacuum energy-momentum tensor of conformally invariant fields in the black hole background is connected with conformal anomalies. If these anomalies are properly taken into account, the resulting energy-momentum tensor reproduces sufficiently well the exact value of $\langle\hat{T}^{\mu}_{\nu}\rangle$ in the Ricci-flat spacetimes. It should be emphasized that in the general case both $\langle\hat{\varphi}^2\rangle$ and $\langle\hat{T}^{\mu}_{\nu}\rangle$ for a massless field are highly nonlocal functionals of the background geometry. The presented results indicate that when $R_{\mu\nu}\neq 0$ this nonlocality is really important.

It is interesting to note that in two-dimensional spacetime the conformal anomaly allows one to reconstruct the stress-energy tensor. The conformal anomaly in two dimensions is of the form

$$\langle\hat{T}^{\mu}_{\mu}\rangle=\frac{C_s}{24\,\pi}\,R\,,$$

where R is the scalar curvature, and C_s is a coefficient dependent on the spin s of the field ($C_0=1$). The total energy-momentum tensor $\langle\hat{T}^{\mu}_{\nu}\rangle$ is determined by its trace and conservation laws up to two functions of one variable, which correspond to the arbitrariness in the choice of the boundary conditions [Christensen and Fulling (1977)]. As a result, it is possible to calculate $\langle\hat{T}^{\mu}_{\nu}\rangle$ exactly in a number of two-dimensional models imitating the black hole [see Davies (1976), Davies, Fulling, and Unruh (1976), Unruh (1977), Fulling (1977b), Hiscock (1977), Frolov and Vilkovisky (1983), Balbinot and Brown (1984), Balbinot (1984), Kuroda (1984a), Birrell and Davies (1982)].

11.3.6 Exact results

We have already mentioned that exact values of $\langle\hat{\varphi}^2\rangle$ and $\langle \hat{T}^{\mu}_{\nu}\rangle$ were successfully calculated in a number of cases where the point of interest lies on the event horizon. This remarkable fact follows from the very special properties of spacetime near the event horizon. We will now discuss these properties in some detail.

The Kerr metric is symmetric under translation in time t and under rotations in ϕ. Let $\boldsymbol{\xi}_{(t)}$ and $\boldsymbol{\xi}_{(\phi)}$ be the corresponding Killing vector fields, and $\boldsymbol{\eta} = \boldsymbol{\xi}_{(t)} + \Omega^H \boldsymbol{\xi}_{(\phi)}$ be their linear combination which is tangent to the generators of the horizon. It is readily shown that $\boldsymbol{\xi}_{(\phi)}$ vanishes on the symmetry axis ($\theta = 0$, $\theta = \pi$), while $\boldsymbol{\eta}$ vanishes on the two-dimensional surface S of intersection of the horizons H^- and H^+ (the surface of horizon bifurcation). The antisymmetric tensors $\left(\xi_{(\phi)\mu;\nu}\right)_{\xi_{(\phi)}=0}$ and $\left(\eta_{\mu;\nu}\right)_{\eta=0}$ are non-degenerate.[14] Obviously, the polar ($\theta = 0$ and $\theta = \pi$) points x_0 of the surface S remain fixed under shifts both in t and ϕ.

If we stipulate that $\langle \hat{T}^{\mu}_{\nu}\rangle$ in the chosen state has the same symmetry properties as the background physical spacetime, this quantity must satisfy the equation

$$\mathcal{L}_{\xi}\langle \hat{T}_{\mu\nu}\rangle \equiv \xi^{\alpha}\,\langle \hat{T}_{\mu\nu}\rangle_{;\alpha} + \xi^{\alpha}{}_{;\mu}\,\langle \hat{T}_{\alpha\nu}\rangle + \xi^{\alpha}{}_{;\nu}\,\langle \hat{T}_{\mu\alpha}\rangle = 0\,, \tag{11.3.90}$$

where $\boldsymbol{\xi}$ is the Killing vector field and $\mathcal{L}_{\boldsymbol{\xi}}$ is the Lie derivative along it. At points where $\xi^{\alpha} = 0$, these equations turn into constraints on the algebraic structure of $\langle \hat{T}_{\mu\nu}\rangle$. It can be shown, when solving these equations, that such a regular energy-momentum tensor has the following form at the pole x_0 of the bifurcation surface S [Frolov and Zel'nikov (1985b)]:

$$\langle \hat{T}_{\mu\nu}\rangle = A(k_{\mu}\,l_{\nu} + l_{\mu}\,k_{\nu}) + B(m_{\mu}\,\bar{m}_{\nu} + \bar{m}_{\mu}\,m_{\nu})\,. \tag{11.3.91}$$

Here $(k, l, m, \bar{m})$ are the vectors of the complex null tetrad:

$$\begin{aligned} k_{\alpha}\,dx^{\alpha} &= -\,dt + \frac{\Sigma}{\Delta}\,dr + a\sin^2\theta\,d\phi\,, \\ l_{\alpha}\,dx^{\alpha} &= \frac{\Delta}{2\Sigma}\left(-\,dt - \frac{\Sigma}{\Delta}\,dr + a\sin^2\theta\,d\phi\right), \\ m_{\alpha}\,dx^{\alpha} &= \frac{1}{\sqrt{2}\,(r + ia\cos\theta)}\,[-\,ia\sin\theta\,dt + \Sigma\,d\theta + i(r^2 + a^2)\sin\theta\,d\phi\,]\,. \end{aligned} \tag{11.3.92}$$

Here $\Delta = r^2 - 2Mr + a^2$ and $\Sigma = r^2 + a^2\cos^2\theta$.

The difference between two constants A and B which enter (11.3.91) is fixed by the value of the conformal anomaly

$$2(B - A) = \langle \hat{T}^{\mu}_{\mu}\rangle = (\alpha + \beta)\,C_{\alpha\beta\gamma\delta}\,C^{\alpha\beta\gamma\delta}. \tag{11.3.93}$$

[14] A discussion of the general properties of surfaces formed by stable points under the action of a symmetry group can be found in Boyer (1969); see also Miller (1979).

In the case of a spherically symmetric black hole, any point of the surface S can be chosen for x_0. The nonzero components of the energy-momentum tensor then take the form

$$T_t^t = T_r^r = -A\,, \quad T_\theta^\theta = T_\phi^\phi = B. \tag{11.3.94}$$

Another factor also leads to substantial simplification of the problem of calculating $\langle\hat{\varphi}^2\rangle$ and $\langle\hat{T}_{\mu\nu}\rangle$ at the event horizon. We first demonstrate this for $\langle\hat{\varphi}^2\rangle$. In order to calculate this quantity, one needs to solve the equation

$$\Box\, G^H(x, x_0) = -\,\delta(x, x_0)\,,$$

which governs the Green's function $G^H(x, x_0)$. By making Wick rotation (11.2.13)–(11.2.14), one can reduce the problem to solving the equation for the Euclidean Green's function G_E

$$\Box_E\, G^E(x, x_0) = -\,\delta(x, x_0)\,. \tag{11.3.95}$$

Let $\Pi(x, x')$ be an arbitrary biscalar (e.g., Green's function $G^E(x, x')$ of a scalar field) with the symmetry properties of the background spacetime. Then

$$\mathcal{L}_\xi \Pi(x, x') + \mathcal{L}_{\xi'}\Pi(x, x') = \left(\xi^\alpha \frac{\partial}{\partial x^\alpha} + \xi^{\alpha'} \frac{\partial}{\partial x'^{\alpha'}}\right) \Pi(x, x') = 0\,. \tag{11.3.96}$$

Here $\mathcal{L}_\xi$ and $\mathcal{L}_{\xi'}$ are the Lie derivatives along the Killing vector field $\boldsymbol{\xi}$ with respect to the first and second argument, respectively. If the points x and x' do not coincide with the rotation axis and do not lie on the Euclidean horizon S, equation (11.3.96) shows that the biscalar $\Pi(x, x')$ is a function of the differences $\tau - \tau'$ and $\phi - \phi'$. If the point x' lies on the pole of the surface S, the function $\Pi(x, x_0)$ depends of neither τ nor ϕ.

This results, among other things, in a substantial simplification of the equation (11.3.95). In the coordinates $R = \Delta^{1/2}\sin\theta$, $z = (r - M)\cos\theta$ this equation in the Kerr metric becomes

$$\left[\frac{1}{R}\,\partial_R(R\,\partial_R) + \partial_z^2\right]\, G^E(R, z; x_0) = -\,\frac{\kappa}{4\pi^2 R}\,\delta(R)\,\delta(z - z_0)\,, \tag{11.3.97}$$

where $z_0 = \sqrt{M^2 - a^2}$. Note that the solution of this equation coincides with the potential of the field of a point-like charge $q = \kappa/8\pi^2$ placed at a point $z = z_0$ of the axis $x = y = 0$ ($R^2 = x^2 + y^2$) in a flat space. Using this solution, we obtain

$$G^H(x, x_0) = \frac{i\kappa}{8\pi^2(r - M - \sqrt{M^2 - a^2}\,\cos\theta)}\,. \tag{11.3.98}$$

Subtracting the divergent part from this expression [see (11.1.9)] and tending x to x_0, we obtain for $\langle\hat{\varphi}^2(x_0)\rangle^H$ at the pole of the horizon the following value [Frolov (1982, 1983c)]:

$$\langle\hat{\varphi}^2(x_0)\rangle^H = \frac{1}{48\,\pi^2}\,\frac{r_+^2 - 3a^2}{(r_+^2 + a^2)^2}\,. \tag{11.3.99}$$

(Candelas (1980) derived this expression for $a = 0$ by summing up the series representation of Green's function.)

This approach can be used to calculate $\langle\hat{\varphi}^2\rangle^H$ for a black hole placed in external axially symmetric static gravitational field [Frolov and Garcia (1983)] or enclosed in a mirror-walled cavity described by the equation $r = r_0$ $(\phi|_{r_0} = 0)$ [Frolov (1986)]. In the latter case, the problem reduces to calculating the field produced inside a conducting grounded ellipsoid of revolution,

$$\frac{R^2}{\Delta(r_0)} + \frac{z^2}{(r_0 - M)^2} = 1\,, \tag{11.3.100}$$

by a point-like charge placed at its focus. The corresponding value of $\langle\hat{\varphi}^2\rangle_{H;r_0}$ is

$$\langle\hat{\varphi}^2\rangle_{H;r_0} = \langle\hat{\varphi}^2\rangle^H - \delta_{r_0}\langle\hat{\varphi}^2\rangle\,, \tag{11.3.101}$$

where

$$\delta_{r_0}\langle\hat{\varphi}^2\rangle = \frac{1}{8\pi^2(r_+^2 + a^2)}\sum_{l=0}^{\infty}(2l+1)\,\frac{Q_l(b)}{P_l(b)}\,. \tag{11.3.102}$$

Here $b = (r_0 - M)/(M^2 - a^2)^{1/2}$, and P_l and Q_l are Legendre functions.

Frolov and Zel'nikov (1985b) used a similar method for calculating $\langle \hat{T}^{\mu}_{\nu}(x_0)\rangle^H$ for the electromagnetic field. (For calculations in the Schwarzschild spacetime see Elster (1984)). After correction of a technical mistake found by Jensen, McLaughlin, and Ottewill (1988), this method gives for the coefficients A and B in (11.3.91) the expressions

$$A = \frac{1}{1920\,\pi^2 M^4 r_+^3}[-15\,M^3 - 75\,M^2(M^2 - a^2)^{1/2} + 52\,(M^2 - a^2)^{3/2}]\,, \tag{11.3.103}$$

$$B = \frac{1}{1920\,\pi^2 M^4 r_+^3}[-15\,M^3 + 3\,M^2(M^2 - a^2)^{1/2} - 52\,(M^2 - a^2)^{3/2}]\,. \tag{11.3.104}$$

Let us look at one interesting property of expression (11.3.99). It is easily verified that

$$\langle\hat{\varphi}^2(x_0)\rangle^H = \frac{1}{48\,\pi^2}\,K\,, \tag{11.3.105}$$

where K is the Gaussian curvature of the two-dimensional surface of the black hole at its pole x_0. The explicit expression for K in the Kerr metric is given by (3.5.5). Expression (11.3.105) is also valid in the case of a black hole placed in an external static axially symmetric gravitational field.

It is easily proved that in Page's approximation, the quantity $\langle\hat{\varphi}^2\rangle^H_P$ at the event horizon of any static black hole (including black holes distorted by an external field) can always be written in the form (11.3.105). The quantity $\varepsilon = -\langle \hat{T}_t^t \rangle_P^H$, characterizing the energy density on the black hole surface, is given in this approximation by the expression [Frolov and Sánchez (1986)]

$$\varepsilon = -(7\alpha + 12\beta) K^2 + \alpha\, {}^{(2)}\!\Delta K, \tag{11.3.106}$$

where ${}^{(2)}\!\Delta$ is the two-dimensional Laplacian on the black hole surface, and α and β are coefficients (11.1.13)–(11.1.14) in expression (11.1.10) for conformal anomalies.

11.3.7 Vacuum polarization of massive fields

Let $L \sim M$ be the characteristic curvature radius near a black hole. Then the contribution of massive fields to vacuum polarization near black holes is smaller by a factor $\varepsilon = (\lambda/L)^2$, $\lambda = \hbar/mc$, than that of massless fields. An analysis of the contribution of massive fields is instructive for a number of reasons. Note, first of all, that if $\varepsilon \ll 1$, it is possible to separate the contributions of real and virtual particles to $\langle \hat{T}_\nu^\mu \rangle$. Thus, the contribution of real particles of the heat bath to $\langle \hat{T}_\nu^\mu \rangle^H$ in the Hartle-Hawking vacuum state contains a small factor $\exp(-\beta GMm/\hbar c)$, while the contribution of virtual particles is a power function of M^{-1}. The values of $\langle \hat{T}_\nu^\mu \rangle$ in different "vacua" differ by the contribution of real thermal radiation particles; hence, in the vicinity of a black hole, $\langle \hat{T}_\nu^\mu \rangle^B$ and $\langle \hat{T}_\nu^\mu \rangle^U$ for the massive field practically coincide with $\langle \hat{T}_\nu^\mu \rangle^H$ everywhere, except the exponentially narrow strip $[\sim r_+ \exp(-\beta GMm/\hbar c)]$ near the event horizon. The behavior of these quantities inside the strip is not identical: $\langle \hat{T}_\nu^\mu \rangle^B$ diverges on H^- and H^+, while $\langle \hat{T}_\nu^\mu \rangle^U$ diverges on H^- but remains finite on H^+.

A much more detailed study of the effect of vacuum polarization of massive fields is possible because one can make use of the expansion in a small dimensional parameter ε. If the background gravitational field satisfies Einstein's vacuum equations, the quantity $\langle \hat{T}_\nu^\mu \rangle^H$ for the massive scalar, spinor, and vector fields can be obtained to first non-vanishing order in ε by varying the following effective action [Frolov and Zel'nikov (1984)]:

$$\langle \hat{T}_{\mu\nu}(x) \rangle^H = -\frac{2}{\sqrt{-g}} \frac{\delta W}{\delta g^{\mu\nu}(x)}, \tag{11.3.107}$$

where

$$W = (96 \times 7!\pi^2 m^2)^{-1} \int d^4x \sqrt{-g} \left(A_s R_{\alpha\beta}{}^{\gamma\delta} R_{\gamma\delta}{}^{\varepsilon\zeta} R_{\varepsilon\zeta}{}^{\alpha\beta} \right.$$
$$\left. + B_s R\, R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} \right). \tag{11.3.108}$$

Here m is the mass of the field, and the coefficients A_s and B_s for the spin-s field are

$$A_0 = 1, \qquad A_{1/2} = -4, \qquad A_1 = 3,$$
$$B_0 = 18 - 84\xi, \qquad B_{1/2} = 12, \qquad B_1 = -30, \tag{11.3.109}$$

respectively (ξ is the coefficient of the term $R\hat{\varphi}^2$ in the action for the scalar field; the field is conformally invariant if $\xi = 1/6$ and $m = 0$). In this approximation, $\langle \hat{T}^\mu_\nu \rangle^H$ for a Schwarzschild black hole is of the form [Frolov and Zel'nikov (1982)]

$$\langle \hat{T}^\mu_\nu \rangle^H = \frac{M^2}{10080\,\pi^2\,m^2\,r^8}\,[A_s\,p^\mu_\nu + 3(2B_s + A_s)\,q^\mu_\nu\,], \tag{11.3.110}$$

where the nonzero components p^μ_ν and q^μ_ν (for $z = 2M/r$) are

$$\begin{aligned} p^t_t &= -15 + 16\,z\,, & p^r_r &= -3 + 4\,z\,, & p^\theta_\theta &= p^\phi_\phi = 9 - 11\,z\,, \\ q^t_t &= -10 + 11\,z\,, & q^r_r &= 4 - 3\,z\,, & q^\theta_\theta &= q^\phi_\phi = -12 + 14\,z\,. \end{aligned} \tag{11.3.111}$$

More recent numerical calculations by Anderson, Hiscock, and Samuel (1995) confirmed the accuracy of (11.3.110).

Note the qualitative similarity in the behavior of $\langle \hat{T}^\mu_\nu \rangle^H$ for massless and massive fields. The energy density $\varepsilon = -\langle \hat{T}^t_t \rangle^H$ of a massive scalar ($\xi = 1/6$) field is positive far from the black hole but reverses sign and becomes negative at the horizon. A similar effect occurs for a conformal scalar massless field. In the case of a massive vector field, the energy density ε is positive at the event horizon.

Rotation of the black hole produces a circular energy density flux in the surrounding space, due to vacuum polarization. This flux is described by the component $\langle \hat{T}^t_\phi \rangle \neq 0$. The expression for $\langle \hat{T}^\mu_\nu \rangle$ of massive fields in the Kerr metric was derived and analyzed by Frolov and Zel'nikov (1983, 1984).

In conclusion, it should be noted that not only has a qualitative understanding of the specifics of vacuum polarization effects in black holes been achieved, but so has a good quantitative description as well.

11.4 Quantum Mechanics of Black Holes

11.4.1 Introduction

We have described in detail the quantum effects of particle creation and vacuum polarization by black holes. These are important and best understood problems of *quantum black hole physics*. But certainly they do not exhaust the subject. Classical black holes are soliton-like solutions of gravitational equations. They are characterized by their position in space as well as by possible different "internal" states. One can expect that when quantum mechanics is applied to black holes, and a black hole is considered as a quantum object, its state is described by a wavefunction. This wavefunction contains information about a motion of the black hole in the surrounding spacetime and states of black hole excitations. The construction of a wavefunction of a black hole, derivation of dynamical equations which govern its time evolution, and development of the interpretation scheme and rules for the calculation of observables are subjects of *quantum mechanics of black holes*, a theory which has not been developed yet. There exist only some ideas and fragmentary results. We discuss them in this section.

We begin with remarks concerning general properties of a wavefunction describing a black hole. For a free motion of the black hole in an asymptotically flat spacetime one can expect that the wavefunction can be written as

$$\Psi(X, \bullet) = e^{-iP_\mu X^\mu}\, \Psi_{BH}(\bullet)\,. \tag{11.4.1}$$

Here P^μ is the four-momentum of the black hole, and X^μ are its "center-of-mass" coordinates. Both quantities, P^μ and X^μ, can be expressed in terms of data at null infinity. The part $\Psi_{BH}(\bullet)$ of the wavefunction describes black hole excitations, and $\bullet$ denotes the possible internal degrees of freedom of the black hole.

If a black hole is influenced by an external field, its motion is modified. As a result of this modification, the factor $\exp(-iP_\mu X^\mu)$ is changed by a wavefunction $\psi(X)$ corresponding to this motion. In the WKB-approximation $\psi(X) \sim \exp(iS)$, where S is the action for the motion of a black hole in the given field. An external field can also modify the part of the wavefunction describing internal states of the black hole.

If an external field is strong enough one can also expect a quantum creation of pairs of black holes, the process which is similar to charged particles creation by an electric field. The final state is a state with two black holes and its wavefunction can be written as a superposition of wavefunctions $\Psi(X_1, \bullet_1)\, \Psi(X_2, \bullet_2)$. The probability of the pair creation in the WKB-approximation is determined by the sub-barrier exponents $\exp(-W_E)$, where W_E is the effective action for a black hole calculated for its sub-barrier motion.

We consider examples of calculations of probability for a pair of black production by an external field in Section 11.4.4. But before this we shall make a few comments on the possible nature of the internal degrees of freedom of a black hole (Sections 11.4.2 and 11.4.3) and describe a no-boundary proposal for a wavefunction of a black hole Ψ_{BH} (Section 11.4.3).

11.4.2 Euclidean approach

The *no-boundary wavefunction of a black hole* was proposed by Barvinsky, Frolov, and Zel'nikov (1995). This construction, which is similar to the no-boundary wavefunction in quantum cosmology introduced by Hartle and Hawking (1983), essentially uses Euclidean methods. In this section we derive some of useful formulas and results connected with the Euclidean approach to quantum theory in the static black hole background.

Wavefunction of Hartle-Hawking vacuum in the "coordinate" representation

We begin by deriving a "coordinate" representation for the Hartle-Hawking state of a quantum field in the background of an eternal black hole. For the sake of simplicity, we restrict ourselves by considering a massless scalar field in the spacetime of a non-rotating black hole. The generalization to higher spins is straightforward.

Consider solutions of the field equation

$$\Box\varphi = 0\,.$$

For definition (11.2.15) of the Hartle-Hawking state in Section 11.2 we used modes D', P', N', and U' given by relations (11.2.12)–(11.2.13). In what follows we shall need a modification of these basic solutions. Namely, we consider solutions of the field equation in the black hole exterior (region R_+) which are of the form

$$\frac{e^{-i\omega_\lambda t}}{\sqrt{2\omega_\lambda}}\,\Phi_\lambda(\mathbf{x})\,, \tag{11.4.2}$$

where

$$\Phi_\lambda(\mathbf{x}) = \frac{1}{\sqrt{2\pi}}\,\frac{u_\ell^k(r;\omega)}{r}\,\tilde{Y}(\theta,\phi)\,. \tag{11.4.3}$$

Here $\lambda = \{k,\omega,\ell,m\}$, $k = 1,2$, and

$$\tilde{Y}(\theta,\phi) = \frac{1}{\sqrt{\pi}}\,P_\ell^m(\cos\theta)\begin{cases} \frac{1}{\sqrt{2}} & \text{if } m=0\,,\\ \sin(m\phi) & \text{if } -\ell\le m<0\,,\\ \cos(m\phi) & \text{if } 0<m\le\ell\,,\end{cases} \tag{11.4.4}$$

are *real* spherical harmonics. The radial functions $u_\ell^k(r;\omega)$

$$u_\ell^1 = \frac{1}{\sqrt{2}}\left(u_\ell^{\text{up}'} + u_\ell^{\text{down}'}\right), \quad u_\ell^2 = \frac{1}{i\sqrt{2}}\left(u_\ell^{\text{up}'} - u_\ell^{\text{down}'}\right), \tag{11.4.5}$$

are *real* solutions of equation (10.2.14)

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - V_\ell(r)\right] u_\ell^k(r;\omega) = 0\,, \tag{11.4.6}$$

$$V_\ell(r) = \left(1-\frac{2M}{r}\right)\left(\frac{\ell(\ell+1)}{r^2} + \frac{2M}{r^3}\right).$$

The so-defined modes $\Phi_\lambda(\mathbf{x})$ are *real* and form a complete set in the three-space (t = const) in the black hole exterior. They are normalized by conditions

$$\int d^3x\,|g^{tt}|\sqrt{|g|}\,\Phi_\lambda(\mathbf{x})\,\Phi_{\lambda'}(\mathbf{x}) = \delta_{\lambda\lambda'}\,, \quad \delta_{\lambda\lambda'} = \delta_{\omega\omega'}\,\delta_{\ell\ell'}\,\delta_{mm'}\,\delta_{kk'}\,. \tag{11.4.7}$$

Since the "internal" region R_- of the eternal black hole is isometric to R_+, a similar set of functions can be defined there. We add an extra index $\pm$ to distinguish these two different sets of modes. Thus, $\Phi_{\lambda+}(\mathbf{x})$ and $\Phi_{\lambda-}(\mathbf{x})$ denote functions in R_+ and R_-, respectively.[15]

Define $\varphi_{\lambda+}(x)$ as a solution of the field equation which vanishes in the left wedge R_- of the eternal black hole spacetime, and coincides with (11.4.2) in the right wedge R_+, and $\varphi_{\lambda-}(x)$ as a solution which vanishes in R_+ and coincides with (11.4.2) in R_-. The so-defined solutions $\varphi_{\lambda\pm}(x) \equiv \varphi_J^{k\pm}(x)$ $(J = \{\omega, \ell, m\})$ are related to the modes introduced in Section 11.2.2 as follows[16]

$$\varphi_J^{1+}(x) = \frac{1}{\sqrt{2}}\left[\varphi_J^{\mathrm{up}'}(x) + \varphi_J^{\mathrm{down}'}(x)\right],$$

$$\varphi_J^{2+}(x) = \frac{1}{i\sqrt{2}}\left[\varphi_J^{\mathrm{up}'}(x) - \varphi_J^{\mathrm{down}'}(x)\right], \tag{11.4.8}$$

$$\varphi_J^{1-}(x) = \frac{1}{\sqrt{2}}\left[\varphi_J^{\mathrm{dn}'}(x) + \varphi_J^{\mathrm{umop}'}(x)\right],$$

$$\varphi_J^{2-}(x) = \frac{i}{\sqrt{2}}\left[\varphi_J^{\mathrm{dn}'}(x) - \varphi_J^{\mathrm{umop}'}(x)\right]. \tag{11.4.9}$$

We also introduce another set of modes

$$\psi_{\lambda-}(x) = c_\lambda\, \varphi_{\lambda-}(x) + s_\lambda\, \bar{\varphi}_{\lambda+}(x)\,,$$

$$\psi_{\lambda+}(x) = c_\lambda\, \varphi_{\lambda+}(x) + s_\lambda\, \bar{\varphi}_{\lambda-}(x)\,, \tag{11.4.10}$$

related to D'-, P'-, U'-, and N'-modes defined by (11.2.12) and (11.2.13) by the relations similar to (11.4.8) and (11.4.9). The field operator $\hat{\varphi}(x)$ decomposition in modes $\varphi_{\lambda\pm}(x)$ and modes $\psi_{\lambda\pm}(x)$ is

$$\hat{\varphi}(x) = \sum_{\lambda\pm}\left[\hat{b}_{\lambda\pm}\, \varphi_{\lambda\pm}(x) + \hat{b}^\dagger_{\lambda\pm}\, \bar{\varphi}_{\lambda\pm}(x)\right]$$

$$= \sum_{\lambda\pm}\left[\hat{a}_{\lambda\pm}\, \psi_{\lambda\pm}(x) + \hat{a}^\dagger_{\lambda\pm}\, \bar{\psi}_{\lambda\pm}(x)\right]. \tag{11.4.11}$$

Operators $\hat{a}_{\lambda\pm}$, $\hat{a}^\dagger_{\lambda\pm}$, $\hat{b}_{\lambda\pm}$, and $\hat{b}^\dagger_{\lambda\pm}$ are related as follows

$$\hat{b}_{\lambda-} = c_\lambda\, \hat{a}_{\lambda-} + s_\lambda\, \hat{a}^\dagger_{\lambda+}\,, \quad \hat{b}_{\lambda+} = c_\lambda\, \hat{a}_{\lambda+} + s_\lambda\, \hat{a}^\dagger_{\lambda-}\,. \tag{11.4.12}$$

[15] We denote here by R_+ and R_- external and internal regions of the eternal black hole, which earlier were denoted as I and I', respectively. This helps us to simplify relations since the same index $\pm$ can be used to distinguish the quantities related to $R_\pm$.

[16] It is assumed that in the definition of these modes we also use "real angular harmonics" $\bar{Y}_{\ell m}$.

These relations coincide with the special type of canonical transformation (I.43) for a pair of coupled oscillators considered in Appendix I.4. Hence, the Hartle-Hawking vacuum state $|H\rangle$ defined by relations

$$\hat{a}_{\lambda\pm}|H\rangle = 0 \tag{11.4.13}$$

can be presented in the form

$$|H\rangle = \exp\left[\sum_\lambda \left[-\ln c_\lambda + \frac{s_\lambda}{c_\lambda}\hat{b}^\dagger_{\lambda-}\,\hat{b}^\dagger_{\lambda+}\right]\right]|0\rangle\,, \tag{11.4.14}$$

where $\hat{b}_{\lambda\pm}|0\rangle = 0$.

The first of relations (11.4.11) can be rewritten as

$$\hat{\varphi}(x) = \sum_{\lambda\pm}\hat{\varphi}_{\lambda\pm}(t)\,\Phi_{\lambda\pm}(\mathbf{x})\,, \tag{11.4.15}$$

where

$$\hat{\varphi}_{\lambda\pm}(t) = \frac{1}{\sqrt{2\omega_\lambda}}\left(e^{-i\omega_\lambda t}\,\hat{b}_{\lambda\pm} + e^{i\omega_\lambda t}\,\hat{b}^\dagger_{\lambda\pm}\right).$$

In other words, the quantum field $\hat{\varphi}(x)$ can be decomposed into an infinite set of quantum oscillators, $\hat{\varphi}_{\lambda\pm}(t)$ being position operators of these oscillators. Using this fact and relation (I.52), we can write the following expression for the Hartle-Hawking state $|H\rangle$ in the "coordinate"-representation [Barvinsky, Frolov, and Zel'nikov (1995)]

$$\begin{aligned}\Psi_H[\varphi_-,\varphi_+] &= \langle\varphi_-,\varphi_+|H\rangle \\ &= \Psi_0\exp\left\{-\sum_\lambda\left[\frac{\omega_\lambda}{2}\left(\frac{\cosh(\frac{\omega_\lambda\beta_0}{2})}{\sinh(\frac{\omega_\lambda\beta_0}{2})}(\varphi^2_{\lambda-}+\varphi^2_{\lambda+}) - \frac{2\varphi_{\lambda-}\varphi_{\lambda+}}{\sinh(\frac{\omega_\lambda\beta_0}{2})}\right)\right]\right\},\end{aligned} \tag{11.4.16}$$

where $\beta_0 = 8\pi M$ is the inverse Hawking temperature of the black hole, and Ψ_0 is a normalization constant.

This wavefunction has a quite simple interpretation. Consider the section $t = 0$ of the eternal black hole. This section is the Einstein-Rosen bridge, which consists of two sheets Σ_- and Σ_+ connected at the bifurcation two-sphere of the horizons $r = r_+$ (see Figure 11.4). The field $\varphi(x)$ at $t = 0$ can be written as

$$\varphi(t=0,\mathbf{x}) = \varphi_+(\mathbf{x}) + \varphi_-(\mathbf{x})\,, \tag{11.4.17}$$

where

$$\varphi_\pm(\mathbf{x}) = \sum_\lambda \varphi_{\lambda\pm}\,\Phi_{\lambda\pm}(\mathbf{x})\,. \tag{11.4.18}$$

The wavefunction $\Psi_H[\varphi_-,\varphi_+]$ is the probability amplitude to have in a chosen state $|H\rangle$ the field configurations $\varphi_\pm(\mathbf{x})$ on $\Sigma_\pm$.

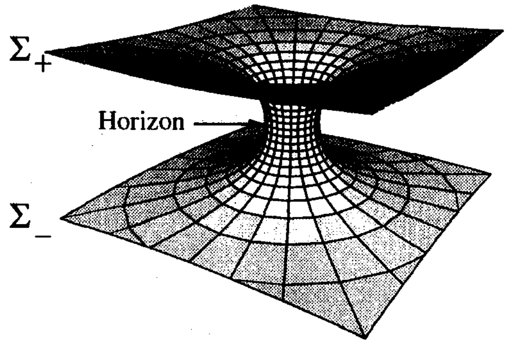


Figure 11.4: The embedding diagram of the global Cauchy surface Σ; $t = 0$. This is an Einstein-Rosen bridge connecting two asymptotically flat three-spaces. The bifurcation sphere of the future and past horizons separates the interior Σ_- of a black hole from its exterior Σ_+.

Wavefunction of Hartle-Hawking vacuum and an Euclidean action

By comparing (11.4.16) with expression (I.7), one can make the important observation that the wavefunction for the Hartle-Hawking vacuum in the "coordinate"-representation can be written in the form

$$\Psi_H[\varphi_-, \varphi_+] = \Psi_0 \exp\left[-\sum_\lambda S_E(\varphi_{\lambda-}, 0; \varphi_{\lambda+}, \beta_0/2)\right], \tag{11.4.19}$$

where S_E is the Euclidean action of a quantum oscillator; $\varphi_{\lambda-}$ is the "initial" position of the oscillator; $\varphi_{\lambda+}$ is its "final" position, and $\beta_0/2$ is the interval of the Euclidean time. We shall use this form of the wavefunction to establish its relation with the Euclidean action for the field $\varphi(x)$.

For this purpose, we note that the Kruskal-Schwarzschild metric (2.7.17)–(2.7.18) is a real Lorentzian-signature section of the complex manifold parameterized by the real radial r, $2M \le r < \infty$, and complex time $z = \tau + i\,t$ coordinates:

$$U = -\left(\frac{r}{2M} - 1\right)^{1/2} \exp\left\{\frac{1}{2}\left(\frac{r}{2M} - 1\right) + i\,\frac{z - 2\pi M}{4M}\right\}, \tag{11.4.20}$$

$$V = \left(\frac{r}{2M} - 1\right)^{1/2} \exp\left\{\frac{1}{2}\left(\frac{r}{2M} - 1\right) - i\,\frac{z - 2\pi M}{4M}\right\}. \tag{11.4.21}$$

The sectors R_+ and R_- of the Kruskal spacetime with asymptotically flat infinities are generated by the following segments in the complex plane of z

$$R_\pm : \quad z = \pm 2\pi M + i\,t \quad \text{for} \quad -\infty < t < \infty, \tag{11.4.22}$$

and analytically joined by the real Euclidean section R_E

$$R_E: \quad z = \tau \quad \text{for} \quad -2\pi M \leq \tau \leq 2\pi M . \tag{11.4.23}$$

Here t is the usual Killing time coordinate in the Schwarzschild metric, while τ is its Euclidean analogue. The *Euclidean black hole* metric, i.e., the metric on the Euclidean section R_E, is

$$ds_E^2 = \left(1 - \frac{2M}{r}\right) d\tau^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2\, d\omega^2 . \tag{11.4.24}$$

In the vicinity of $r = r_+ = 2M$ this metric is

$$ds_E^2 \approx \kappa^2\, l^2\, d\tau^2 + dl^2 + r_+^2\, d\omega^2 , \tag{11.4.25}$$

where l is the proper distance from the *Euclidean horizon* r_+, and $\kappa = 1/4M$ is the surface gravity. This form of the metric implies that in the general case there exists a *cone-like singularity* at $r = r_+$. This singularity is absent only if the coordinate τ plays the role of an angular coordinate and is periodic with the period $\beta = 8\pi M$. The corresponding regular Euclidean space is known as the *Gibbons-Hawking instanton*[17] [Gibbons and Hawking (1977)] (see Figure 11.5).

The Euclidean section (11.4.23) represents a half-period part of this instanton (see Figure 11.6). The boundary $\Sigma_+ \cup \Sigma_-$ of this part at $\tau_\pm = \pm 2\pi M$ is the Einstein-Rosen bridge shown in Figure 11.4. At this boundary the Euclidean section analytically matches to the Lorentzian sectors R_+ and R_- on the Penrose diagram of the Kruskal metric.

Consider now a scalar field $\varphi(\tau, \mathbf{x})$ on the Euclidean manifold R_E described by the action[18]

$$W_E[\varphi] = \frac{1}{2} \int d^4x\, g_E^{1/2}\, g_E^{\mu\nu}\, \nabla_\mu \varphi\, \nabla_\nu \varphi . \tag{11.4.26}$$

In what follows we denote the boundary fields on the two asymptotically flat components of the Einstein-Rosen bridge $\Sigma_\pm$ by $\varphi_\pm(\mathbf{x})$:

$$\varphi(x)\Big|_{\Sigma_\pm} \equiv \varphi(\pm \beta_0/4, \boldsymbol{x}) = \varphi_\pm(\boldsymbol{x}), \quad \beta_0 \equiv 8\pi M. \tag{11.4.27}$$

With this notation the solution to the Euclidean field equation

$$\left\{ g^{1/2}\, g^{\tau\tau}\, \frac{d^2}{d\tau^2} + \partial_a (g^{1/2}\, g^{ab}\, \partial_b) \right\} \varphi(\tau, \boldsymbol{x}) = 0 \tag{11.4.28}$$

[17]The gravitational action calculated on this Euclidean solution is finite, as it is required for an instanton.

[18]We do not include a term $\xi R \varphi^2$ in the action since for the Schwarzschild metric $R = 0$, and this term contributes neither to the field equation nor to the action.

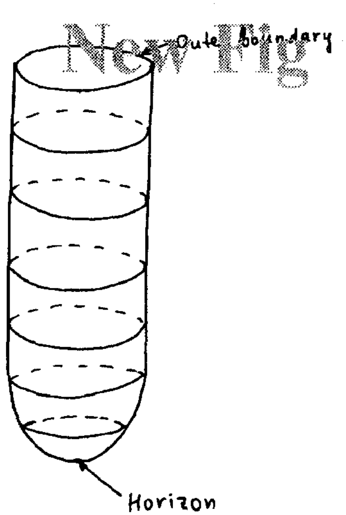


Figure 11.5: The embedding diagram of the $(\tau - r)$-section of the Gibbons-Hawking instanton.

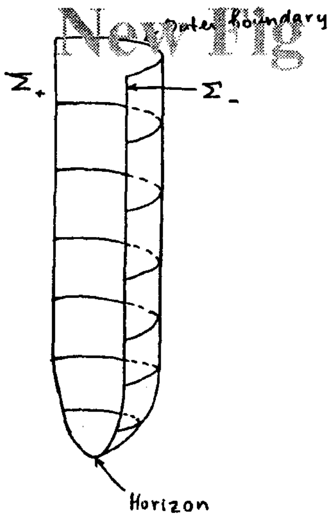


Figure 11.6: Here one half of the Gibbons-Hawking gravitational instanton is depicted. The global Cauchy surface $\Sigma = \Sigma_+ \cup \Sigma_-$ is one boundary of this Euclidean manifold, and spatial infinity ∂M_∞ is another. The arguments of the wavefunction $\Psi_H[\varphi_-, \varphi_+]$ are the boundary values of quantum fields on the two asymptotically flat parts of the Einstein-Rosen bridge $\Sigma_\pm$, respectively.

can be written as a decomposition

$$\varphi(\tau, \boldsymbol{x}) = \sum_\lambda \varphi_\lambda(\tau)\, \Phi_\lambda(\boldsymbol{x}), \tag{11.4.29}$$

where $\varphi_\lambda(\tau)$ is a solution of the equation

$$\frac{d^2\varphi_\lambda}{d\tau^2} - \omega_\lambda^2 \varphi_\lambda = 0, \tag{11.4.30}$$

with the boundary conditions

$$\varphi_\lambda(-\beta_0/4) = \varphi_{\lambda-}, \quad \varphi_\lambda(+\beta_0/4) = \varphi_{\lambda+}. \tag{11.4.31}$$

Relations (11.4.29)–(11.4.30) show that after the decomposition into spatial modes the field is represented as a set of Euclidean oscillators enumerated by the collective index λ. Moreover, using the normalization condition (11.4.7) and the field equation (11.4.28), one can rewrite the action (11.4.26) in the form

$$W_E[\varphi] = \frac{1}{2}\int d\tau \sum_\lambda \left[\left(\frac{d\varphi_\lambda}{d\tau}\right)^2 + \omega_\lambda^2 \varphi_\lambda^2\right]. \tag{11.4.32}$$

That is, the Euclidean action $W_E^{(\beta_0/2)}[\varphi_-(\mathbf{x}), \varphi_+(\mathbf{x})]$ calculated for the field on a half of the Gibbons-Hawking instanton is represented as a "sum" of the Euclidean actions for oscillators. Using (I.7), one has

$$W_E^{(\beta_0/2)}[\varphi_-(\mathbf{x}), \varphi_+(\mathbf{x})]$$

$$= \sum_\lambda \left[\frac{\omega_\lambda}{2}\left(\frac{\cosh(\omega_\lambda \beta_0/2)}{\sinh(\omega_\lambda \beta_0/2)}(\varphi_{\lambda-}^2 + \varphi_{\lambda+}^2) - \frac{2\varphi_{\lambda-}\,\varphi_{\lambda+}}{\sinh(\omega_\lambda \beta_0/2)}\right)\right], \tag{11.4.33}$$

where $\varphi_{\lambda\pm}$ are defined by (11.4.18).

To summarize, we showed that the wavefunction of the Hartle-Hawking vacuum state in the "coordinate"-representation can be written as

$$\Psi_H[\varphi_-, \varphi_+] = \Psi_0 \exp\left\{-W_E^{(\beta_0/2)}[\varphi_-(\mathbf{x}), \varphi_+(\mathbf{x})]\right\}, \tag{11.4.34}$$

where $W_E^{(\beta_0/2)}$ is the value of the Euclidean action for a scalar field solution calculated on half of the Gibbons-Hawking instanton.[19]

---

[19]Similarly, the Euclidean action for particles is used in the path-integral derivation of black hole radiance by Hartle and Hawking (1976). Note that the Euclidean Green's function for a field $\varphi(x)$ can be written as follows

$$G^E(x,x') = \int \mathcal{D}[\varphi]\,\varphi(x)\,\varphi(x')\,e^{-W_E[\varphi]},$$

where the integration is performed over functions φ on the Gibbons-Hawking instanton. On the other hand, one can represent the same quantity in terms of the Feynman path integral over the particle trajectories connecting points x and x' [Hartle and Hawking (1976)]

$$G^E(x,x') = \int_0^\infty dT\,e^{-m^2 T}\int \mathcal{D}[x]\,e^{-S_E[x]},$$

where m is the mass of the particle, and

$$S_E[x] = \frac{1}{4}\int_0^T d\tau\, g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau},$$

is the Euclidean action for a particle which starts at a spacetime point x' at $\tau = 0$ and arrives at x at $\tau = T$. The equivalence of the two representations for the Green's function is connected with the particle-field duality.

This result implies in particular that the wavefunction $\Psi_H[\varphi_-,\varphi_+]$ allows the following functional integral representation

$$\Psi_H[\varphi_-,\varphi_+] = C \int\limits_{(\varphi|_{\Sigma_\pm}=\varphi_\pm)} \mathcal{D}[\varphi]\, e^{-W_E[\varphi]} . \tag{11.4.35}$$

Here $W_E[\varphi]$ is the Euclidean action, and C is a constant which can be found from the normalization condition. The integration is performed over functions $\varphi(x)$ on half of the black hole instanton which satisfy given boundary conditions $\varphi_\pm(\mathbf{x})$ on the boundary of the region. In order to prove (11.4.35), we represent $\varphi(x)$ as $\varphi(x) = \varphi_0(x) + \tilde{\varphi}(x)$, where $\varphi_0(x)$ is a solution of the field equation $\Box\varphi_0 = 0$ with boundary conditions $\varphi_0|_{\Sigma_\pm} = \varphi_\pm(\mathbf{x})$. One has

$$W_E[\varphi] = W_E[\varphi_0] + W_E[\tilde{\varphi}] , \tag{11.4.36}$$

and hence

$$\int\limits_{(\varphi|_{\Sigma_\pm}=\varphi_\pm)} \mathcal{D}[\varphi]\, e^{-W_E[\varphi]} = \tilde{C}\, e^{-W_E[\varphi_0]} , \tag{11.4.37}$$

where

$$\tilde{C} = \int\limits_{(\tilde{\varphi}|_{\Sigma_\pm}=0)} \mathcal{D}[\tilde{\varphi}]\, e^{-W_E[\tilde{\varphi}]} \tag{11.4.38}$$

is an unimportant normalization constant independent of the boundary conditions. Relations (11.4.33)–(11.4.35) were derived for the massless scalar field, but it is evident that similar relations are valid for massive fields and for the fields of higher spin as well. They also allow a generalization to the case of Fermi fields. The representation (11.4.35) will be our starting point in the discussion of the no-boundary ansatz for the wavefunction of a black hole in the next section. But before coming to this discussion we make a few more comments on other properties of the Hartle-Hawking state.

Density matrix in the "coordinate" representation and thermofield dynamics

A state of the quantum field in the black hole exterior is described by the density matrix obtained by tracing $|H\rangle\langle H|$ over the "internal variables"

$$\hat{\rho}[\varphi'_+,\varphi_+] = \mathrm{Tr}_-\left(|H\rangle\langle H|\right) = \int \mathcal{D}[\varphi_-]\, \overline{\Psi}_H[\varphi'_+,\varphi_-]\, \Psi_H[\varphi_-,\varphi_+] . \tag{11.4.39}$$

The factorization of the wavefunction implies that the problem of finding $\hat{\rho}[\varphi_+,\varphi'_+]$ can be reduced to the calculation for a single oscillator. The measure $\mathcal{D}[\varphi_-]$ of integration in the integral over states $\varphi_-(\mathbf{x})$ on Σ_- is to be understood as $\prod_\lambda d\varphi_{\lambda -}$. Simple calculations give

$$\hat{\rho}[\varphi'_+, \varphi_+] = \rho_0 \exp\left\{-W_E^{(\beta_0)}[\varphi'_+(\mathbf{x}), \varphi_+(\mathbf{x})]\right\}. \tag{11.4.40}$$

That is, $\hat{\rho}[\varphi'_+, \varphi_+]$ has a form similar to (11.4.34), with two important differences: (i) The Euclidean action is taken for the interval β_0 (not $\beta_0/2$) of the Euclidean time τ, and (ii) The normalization factor ρ_0 differs from Ψ_0.

Comparing (11.4.40) with the thermal density matrix for an oscillator at the temperature β_0^{-1} in the "coordinate" representation, (I.40), we may conclude that this density matrix describes thermal excitation of the field φ in the black hole exterior, the temperature of the excitations (as measured at infinity) being equal to the Hawking temperature of the black hole. We already obtained this result earlier in Section 10.4.5 using the representation of the density matrix in terms of operators of creation and annihilation of particles propagating in the black hole exterior.

Israel (1976) noticed that there is a close similarity between the origin of the thermal density matrix for a black hole and the approach of Umezawa and Takahashi [Umezawa and Takahashi (1975), Umezawa, Matsumoto, and Takahashi (1981)]. Umezawa and Takahashi proposed a formulation of a quantum field theory at finite temperature called *thermo-field dynamics.* Their aim was to use the wavefunctions of quantum field theory instead of the mixed states of the usual finite-temperature theories. They succeeded in doing this by doubling the number of degrees of freedom and introducing a new additional "fictitious" system, identical to the original one. They assume a special choice of the "vacuum" state for the combined system and derive the thermal density matrix for the physical observables by averaging over "fictitious" states. Their result agrees with the usual one obtained by calculating the trace of the statistical operator and the operator of an observable. Laflamme (1989a,b) developed this analogy and demonstrated that the introduction of a fictitious field and the vacuum in thermo field dynamics can be derived from Euclidean path integrals. In particular, in the case of a black hole one obtains expression (11.4.34) for the wavefunction of the enlarged system with a doubled set of variables.

11.4.3 The no-boundary wavefunction of a black hole

The no-boundary ansatz

As we already mentioned, a study of the propagation of perturbations in a real black hole can be reduced to the analogous problem in its "eternal version". For the latter, one can specify the initial data on the section $t = 0$ which has the topology of the Einstein-Rosen bridge. We denote it by Σ, and denote by Σ_+ and Σ_- its external and internal parts, respectively. It is evident that the data on Σ_- do not influence the black hole exterior. In the spacetime of an "eternal version" of a black hole, such perturbations propagate to the future entirely inside the horizon. It is evident that the corresponding perturbations in a "physical" black hole also always remain beneath the horizon. That is why these data should be identified with the internal degrees of freedom of a black hole [Frolov and Novikov (1993b), Barvinsky, Frolov, and Zel'nikov (1995)]. Such an approach can be generalized to the case when the perturbations are not small.

Among all perturbations describing the propagation of physical fields, a special role is played by gravitational perturbations. The corresponding initial data on the Einstein-Rosen bridge can be considered either as a tensor field, or, equivalently, as small deformations of the initially spherical geometry of the Einstein-Rosen bridge. In this sense one can relate different configurations of the gravitational fields in the system with deformations of the Einstein-Rosen bridge, obeying the necessary constraints, existing in the theory. For physical excitations with finite energy the condition of asymptotic flatness at both infinities of the Einstein-Rosen bridge should be preserved. To summarize, in classical physics the space of initial physical configurations of a system including a black hole can be related to the space of "deformations" of the Einstein-Rosen bridge of the eternal black hole and possible configurations of other fields on it (besides the gravitational one), which obey the constraints and preserve asymptotic flatness. One can use now this configuration space to define the wavefunction of a black hole.

The main idea can be described as follows. Fix a three-dimensional manifold with a wormhole topology $R \times S^2$ and consider all three-dimensional metrics on it which possess two asymptotically flat regions. Consider also all configurations of matter fields on this manifold. The space of 3-geometries and matter fields will be considered as the configuration space for our problem. The wavefunction of a black hole is defined as a functional on this configuration space. It should be stressed that the metric and fields on the "internal" part Σ_- of space are to be considered as defining the internal state of a black hole.

Certainly, there exists a wide ambiguity in the choice of the wavefunction. As in standard quantum mechanics, this ambiguity reflects different possibilities for the "preparation" of the system. Moreover, after the choice of the initial state is made, one must apply the dynamical equations in order to obtain the future evolution of the state in time. In our case such an evolution equation is known as *Wheeler-DeWitt equation*. We shall not discuss this subject here, but confine ourselves to describing the (in some sense) simplest possible quantum state of a black hole. This state is the so-called *no-boundary wavefunction* proposed by Barvinsky, Frolov, and Zel'nikov (1995).

The no-boundary wavefunction of a black hole is defined as a functional of the 3-geometry and matter fields on a surface $\Sigma = R \times S^2$. It is given by the Euclidean path integral over 4-geometries and spacetime matter fields

$$\Psi\left[{}^3g(\boldsymbol{x}), \varphi(\boldsymbol{x})\right] = \int \mathcal{D}g\; \mathcal{D}\varphi\; e^{-W[g,\varphi]} \tag{11.4.41}$$

Here, $W[g, \varphi]$ is the Euclidean action. The integration is performed over Euclidean

4-geometries g and matter-field configurations φ on a spacetime M with a boundary ∂M. The integration variables are subject to the conditions $({}^3g(\boldsymbol{x}),\ \varphi(\boldsymbol{x})),\ \ \boldsymbol{x} \in \partial M$ – the collection of 3-geometry and boundary matter fields on

$$\partial M = \Sigma \cup \partial M_\infty \tag{11.4.42}$$

where $\Sigma = R \times S^2$ is the hypersurface with the topology of the Einstein-Rosen bridge connecting two asymptotically flat three-dimensional spaces. These boundary values are just the argument of the wavefunction (11.4.41). This proposal for the wavefunction is similar to the famous no-boundary wavefunction in quantum cosmology, suggested by Hartle and Hawking (1983).

The no-boundary wavefunction in the semi-classical approximation

Let us calculate the no-boundary wavefunction in the semi-classical approximation. If (g_0, ϕ_0) is a point of the extremum of the action W_E, then we can write

$$g = g_0 + \tilde{g}, \qquad \varphi = \varphi_0 + \tilde{\varphi}, \tag{11.4.43}$$

$$W_E[g, \phi] = W_0[g_0, \varphi_0] + W_2[\tilde{g}, \tilde{\varphi}] + \dots . \tag{11.4.44}$$

In accordance with this decomposition, the no-boundary wavefunction (11.4.41) in the semi-classical approximation is

$$\Psi_\beta\left[{}^3g(\boldsymbol{x}), \varphi(\boldsymbol{x})\right] = \Psi^0_\beta\left[{}^3g_0(\boldsymbol{x}), \varphi_0(\boldsymbol{x})\right] \times \Psi^1_\beta\left[{}^3\tilde{g}(\boldsymbol{x}), \tilde{\varphi}(\boldsymbol{x})\right], \tag{11.4.45}$$

where

$$\Psi^0_\beta\left[{}^3g_0(\boldsymbol{x}), \varphi_0(\boldsymbol{x})\right] = \mathrm{e}^{-W_0[g_0,\varphi_0]}, \tag{11.4.46}$$

is a classical (tree-level) contribution, and

$$\Psi^1_\beta\left[{}^3\tilde{g}(\boldsymbol{x}), \tilde{\varphi}(\boldsymbol{x})\right] = \int \mathcal{D}\tilde{g}\ \mathcal{D}\tilde{\varphi}\ e^{-W_2[\tilde{g},\tilde{\varphi}]} \tag{11.4.47}$$

is a one-loop part.

We consider a theory for which $\phi_0 = 0$ so that g_0 is a solution of the vacuum Einstein equations. The corresponding Euclidean solution is a part M_β of the Gibbons-Hawking instanton (see Figure 11.7), i.e., the Euclidean Schwarzschild solution

$$ds^2 = F\, d\tau^2 + F^{-1}\, dr^2 + r^2\, d\omega^2, \qquad F = 1 - r_+/r, \tag{11.4.48}$$

with $\tau \in (-\frac{1}{4}\beta_\infty, \frac{1}{4}\beta_\infty)$. (For the special choice $\beta_\infty = 8\pi M$ this part is half of the instanton.)

The Euclidean Einstein-Hilbert action is

$$W_0[g_0] = -\frac{1}{16\pi}\int_{M_\beta} R_E \sqrt{g_E}\, d^4x + \frac{1}{8\pi}\int_{\partial M_\beta} K\sqrt{h}\, d^3x$$

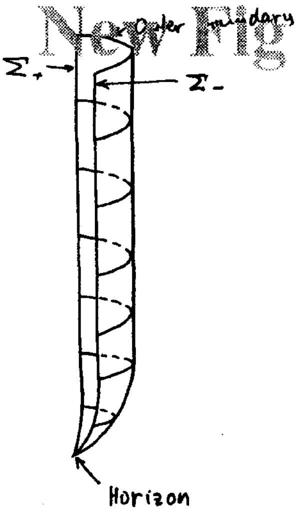


Figure 11.7: Part M_β of the Gibbons-Hawking instanton, which gives the main contribution into the no-boundary wavefunction Ψ_β in the semi-classical approximation.

$$+\frac{1}{8\pi}\int_S k\,\sqrt{\sigma}\,d^2x\,. \tag{11.4.49}$$

The first two terms on the right-hand side are just the Euclidean version of (A.96). The last term is included because of the presence of corners S_+, S_B^-, and S_B^+. In the last integral $\sqrt{\sigma}\,d^2x$ is the surface area element on S, and $k = \arccos(n_1 \cdot n_2)$ is the intersection angle between normals n_1 and n_2 to the three-hypersurfaces Σ_1 and Σ_2 intersecting at S. The corner term is necessary in order to ensure that the variational procedure is self-consistent [Hayward (1993), Hawking and Hunter (1996)].

Since $R = 0$ for solution (11.4.48), only the surface and corner terms contribute to $W_0[g_0]$. We first assume that the boundary Σ_B is located at a finite radius $r = r_B$. In the absence of perturbations the intrinsic metrics of the boundaries Σ_B and $\Sigma_\pm$ are

$$ds_B^2 = F(r_B)\,d\tau^2 + r_B^2\,d\omega^2\,, \qquad \tau \in \left(-\frac{\beta_\infty}{4}, \frac{\beta_\infty}{4}\right),$$

$$ds_\pm^2 = F(r)^{-1}\,dr^2 + r^2\,d\omega^2\,, \qquad r \in (r_+, r_B)\,. \tag{11.4.50}$$

The external boundary Σ_B is a part of the cylinder $R \times S^2$ with the proper length of the generator equal to $\beta/2 = \sqrt{F_B}\,\beta_\infty/2$. Hence, the wavefunction we are looking for is parameterized by two parameters r_B (radius of the external boundary) and β (the inverse temperature on it).

The extrinsic curvature vanishes identically on $\Sigma_\pm$. The calculation of the trace of the extrinsic curvature K for Σ_B is straightforward and gives

$$-K_B = \frac{1}{2}\frac{r_+}{r_B^2}\frac{1}{\sqrt{F(r_B)}} + \frac{2\sqrt{F(r_B)}}{r_B}. \tag{11.4.51}$$

By using this expression, one gets

$$-\frac{1}{8\pi}\int_{\Sigma_B} K\sqrt{h}\, d^3x = \frac{1}{2}\beta\, r_B\sqrt{F(r_B)} + \frac{\beta\, r_+}{8\sqrt{F(r_B)}}. \tag{11.4.52}$$

For corners $S_B^\pm$ the intersection angle is $k_{S_B^\pm} = \pi/2$, while the intersection angle for S_+ is $k_{S_+} = \pi - \left[\beta/(2r_+ F_B^{1/2})\right]$. Hence,

$$-\frac{1}{8\pi}\int_S k\sqrt{\sigma}\, d^2x = \frac{\pi}{2}r_+^2\left(1 - \frac{\beta}{4r_+\sqrt{F(r_B)}}\right) + \frac{\pi\, r_B^2}{2}. \tag{11.4.53}$$

By combining these results, one gets

$$W_0[g_0] = \frac{1}{2}\beta\, r_B\left(1 - \sqrt{F(r_B)}\right) - \frac{\pi}{2}r_+^2 - \frac{1}{2}\beta\, r_B - \frac{\pi\, r_B^2}{2}. \tag{11.4.54}$$

Hence, we have $\Psi_\beta^0[{}^3g_0(\boldsymbol{x}), \varphi_0(\boldsymbol{x})] = \Psi_\beta(M)$

$$\Psi_\beta(M) = N_0\, e^{-\frac{1}{2}\beta r_B[1-(1-2M/r_B)^{1/2}]+2\pi M^2}. \tag{11.4.55}$$

The last two terms of (11.4.54) do not depend on the mass M and are included in the normalization factor N_0.

Now we calculate the one-loop contribution to the no-boundary wavefunction of a black hole. First of all, we note that each of the fields (including the gravitational perturbations) gives an independent contribution to W_2. As a result, Ψ_β^1 is a product of wavefunctions $\Psi_\beta^1[\varphi]$ depending on only one particular type of field φ. (We recall that $\phi_0 = 0$, and hence the value φ on the boundary coincides with its perturbation $\tilde{\phi}$.) Note that Gaussian integrals (11.4.47) in the definition of Ψ_β^1 coincide with (11.4.35), and hence $\Psi_\beta^1 = \Psi_\beta[\varphi_-, \varphi_+]$

$$\Psi_\beta^1[\varphi_-, \varphi_+] = \Psi_0 e^{-W_E^\beta[\varphi_-, \varphi_+]}, \tag{11.4.56}$$

where

$$W_E^\beta[\varphi_-, \varphi_+] = \sum_\lambda \left\{ \frac{\omega_\lambda \cosh(\frac{\omega_\lambda\beta_\infty}{2})}{2\sinh(\frac{\omega_\lambda\beta_\infty}{2})}(\varphi_{\lambda,+}^2 + \varphi_{\lambda,-}^2) - \frac{\omega_\lambda}{\sinh(\frac{\omega_\lambda\beta_\infty}{2})}\varphi_{\lambda,+}\varphi_{\lambda,-}\right\}. \tag{11.4.57}$$

This action is a sum of Euclidean actions for quantum oscillators of frequency ω_λ for the interval $\beta_\infty/2$ of the Euclidean time with the initial "position" of the oscillator φ_- and the its final "position" φ_+. It is also assumed that an additional summation over all physical fields must be done.

To summarize, we obtain the following expression for the no-boundary wavefunction of a black hole in the semi-classical approximation [Barvinsky, Frolov, and Zel'nikov (1995), Frolov (1996)]

$$\Psi_\beta[M, \varphi_+, \varphi_-] = \Psi_\beta(M) \times \Psi_\beta[\varphi_-, \varphi_+] , \tag{11.4.58}$$

where $\Psi_\beta(M)$ and $\Psi_\beta[\varphi_-, \varphi_+]$ are defined by (11.4.55) and (11.4.56), respectively.[20]

The no-boundary wavefunction depends on the parameter β. The form of (11.4.58) suggests the following interpretation. $\Psi_\beta(M)$ is the probability amplitude to find a black hole of mass M in a state Ψ_β, while $\Psi_\beta[\varphi_-, \varphi_+]$ gives the probability amplitude for different excitations $\{\varphi_-, \varphi_+\}$ of the black hole.

11.4.4 Creation of black hole pairs by an external field

Quantum creation of black holes

We already discussed the particular process of black hole formation by the spherical collapse of a scalar field. One can also expect formation of black holes when there are two colliding waves of scalar or other (e.g., gravitational) fields provided that at the moment of collision, the energy E is concentrated in a region of dimension less than GE/c^4. For example, consider the head-on collision of two almost plane gravitational waves. Denote their transverse size by L, their thicknesses by a, their typical amplitudes by h, and their frequency by ω. At the moment of collision, in a volume $L^2 a$ there is a concentration of energy density of order $h^2\omega^2$. Hence, the total mass in this region is $M \sim h^2\omega^2L^2a$. One can expect formation of a black hole of mass M provided $M \gg L$ and $M \gg a$ [for the discussion of this process see Yurtsever (1988)]. The quantum analogue of this effect is quantum black hole creation in the scattering of gravitons of super-high energy [Veneziano (1993), Fabbrichesi *et al.* (1994), Aref'eva, Viswanathan, and Volovich (1995)].

Recently, another type of quantum effect involving black holes attracted interest. This effect is analogous to the effects of charged particles creation by an external field. The creation of a pair of black holes from vacuum implies that there exists an external source providing the necessary energy to make this process possible. The various concrete mechanisms which have been discussed include black hole pair production by a constant electromagnetic field [Gibbons (1986), Garfinkle and Strominger

[20] Kuchař (1994) developed the scheme of canonical quantization of spherically symmetric vacuum spacetimes and demonstrated that $\Psi(M)$ is a wavefunction describing an unchanging superposition of black holes of different masses. The quantization of a collapsing thin dust shell was performed in papers [Berezin (1990, 1997), Hájíček, Kay, and Kuchař (1992), Dolgov and Khriplovich (1997)].

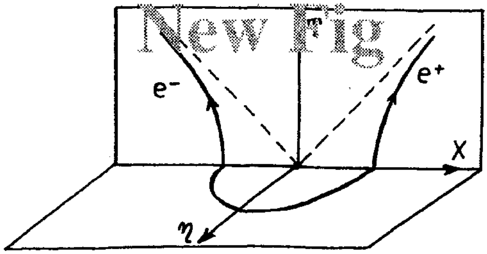


Figure 11.8: Charged particles creation by an external electric field.

(1991), Dowker *et al.* (1994), Garfinkle, Giddings, and Strominger (1994), Hawking, Horowitz, and Ross (1995), Yi (1995)], by snapping of cosmic strings [Hawking and Ross (1995), Emparan (1995), Eardley *et al.* (1995)], by domain walls [Caldwell, Chamblin, and Gibbons (1996)], and by a cosmological constant [Mann and Ross (1995), Bousso and Hawking (1996)].

Instantons and the probability of pairs creation in a static field

Let us consider first the creation of charged black hole pairs by a constant electromagnetic field. This process is similar to the effect of electron-positron pair creation by a constant electric field discussed in Section 10.1.2. Particles of mass m and with charge e are created at a separation $l = m/(eE)$, where E is the strength of the electric field. After creation they are moving with a constant acceleration and their world lines in the Minkowski spacetime are given by the equations

$$T = l\sinh(\tau/l), \qquad X = \pm l\cosh(\tau/l)\,. \tag{11.4.59}$$

We choose the axis X to coincide with the direction of $\mathbf{E}$, and put $Y = Z = 0$. Allow the time t to take complex values $t = T + i\,\eta$. Then trajectory (11.4.59) can be analytically continued to imaginary values of time $t = i\,\eta$

$$\eta = l\sin(\tilde{\tau}/l), \qquad X = l\cos(\tilde{\tau}/l)\,, \tag{11.4.60}$$

where $\tilde{\tau} = i\,\tau$. The semi-circle $0 < \tilde{\tau} < \pi$ connects the points $T = 0$, $X = \pm l$ (initial points of real particle trajectories) (see Figure 11.8).

This solution of the Euclidean equations of motion is an *instanton* connecting a spacetime without charged particles at $t = T + i\infty$ and a spacetime with a couple of created particles at $t = T + i\,0$. The Euclidean action for the semi-circle

$$S_E = m\int_0^\pi d\tilde{\tau} - e\int_0^\pi A_\mu \frac{dX^\mu}{d\tilde{\tau}}\,d\tilde{\tau}\,, \tag{11.4.61}$$

is $S_E = \pi m\,l/2$. (The second integral is equal to half of the value $\pi m\,l$ of the first one.) The probability of charged pair production is proportional to $\exp(-2S_E) = \exp(-\pi m^2/eE)$. This result we already cited in Section 10.1.2.

A similar expression can be easily obtained for the semi-classical probability rate of monopole pair production by a magnetic field

$$\exp(-2S_E) = \exp(-\pi m^2/qB)\,, \tag{11.4.62}$$

where q is the magnetic charge of the monopole and B is the strength of the magnetic field. To leading order in qB this expression reproduces the result of calculations in grand unified theories (GUTs). The monopole in GUTs is a nonlocal object and the rate for the creation of oppositely charged monopoles was calculated by estimating the action for the corresponding field instanton (see Affleck and Manton (1982), Affleck, Alvarez, and Manton (1982) where also the derivation of the semi-classical expression (11.4.62) can be found).

Ernst's metric

The probability of production of a pair of charged black holes can be calculated in a similar way. One needs to find a solution to the Euclidean gravity equations (a *gravitational instanton*) that can be sewed onto a solution describing the motion of a pair of black holes in the physical spacetime. The semi-classical probability rate is $\exp(-W_E)$, where W_E is the Euclidean action for the instanton.

For this purpose, let us consider the following solution of the Einstein-Maxwell equations obtained by Ernst (1976b)

$$ds^2 = \frac{Q^2}{(x-y)^2A^2}\left[G(y)\,dt^2 - \frac{dy^2}{G(y)} + \frac{dx^2}{G(x)} + \frac{G(x)}{Q^4}\,d\phi^2\right], \tag{11.4.63}$$

$$A_\phi = -\frac{2}{BQ}\left(1+\frac{1}{2}Bqx\right) + k\,, \tag{11.4.64}$$

where the functions $Q = Q(x,y)$ and $G(z)$ are

$$G(z) = (1 + r_- Az)(1 - z^2 - r_+ Az^3)\,, \tag{11.4.65}$$

$$Q = \left(1+\frac{1}{2}Bqx\right)^2 + \frac{B^2}{4A^2(x-y)^2}\,G(x)\,. \tag{11.4.66}$$

Here $q^2 = r_- r_+$ and k is an arbitrary constant. We define also $m = (r_- + r_+)/2$.

We show now that this solution describes the uniformly accelerated motion of a pair of magnetically charged black holes.[21] The parameters m and q are connected with the mass and magnetic charge of the black holes; B determines the strength of the magnetic field, and A defines the acceleration of the black holes.

[21] Relations for *electrically* charged black holes moving in an electric field have a similar form.

In order to provide a physical interpretation of the *Ernst metric*, we consider its various limits. If we set the black hole parameters m and q (or equivalently r_- and r_+) to zero in (11.4.63) and (11.4.64), we get

$$ds^2 = \frac{Q^2}{(x-y)^2A^2}\left[(1-y^2)\,dt^2 - \frac{dy^2}{1-y^2} + \frac{dx^2}{1-x^2} + \frac{1-x^2}{Q^4}\,d\phi^2\right], \qquad (11.4.67)$$

$$A_\phi = -\frac{2}{BQ} + k\,, \quad Q = 1 + \frac{B^2}{4A^2(x-y)^2}(1-x^2)\,. \qquad (11.4.68)$$

Let us introduce new coordinates

$$\rho = \frac{\sqrt{1-x^2}}{(x-y)A}\,, \quad \eta = \frac{\sqrt{y^2-1}}{(x-y)A}\,. \qquad (11.4.69)$$

Since

$$d\rho^2 + d\eta^2 = \frac{1}{A^2(x-y)^2}\left(\frac{dx^2}{1-x^2} + \frac{dy^2}{y^2-1}\right), \qquad (11.4.70)$$

one can rewrite (11.4.67) as

$$ds^2 = Q^2\left[-\eta^2\,dt^2 + d\rho^2 + d\eta^2\right] + \frac{\rho^2}{Q^2}\,d\phi^2\,. \qquad (11.4.71)$$

By making the further coordinate transformation

$$T = \eta\sinh t\,, \quad X = \eta\cosh t\,, \qquad (11.4.72)$$

we get

$$ds^2 = Q^2\left[-\,dT^2 + dX^2 + d\rho^2\right] + \frac{\rho^2}{Q^2}\,d\phi^2\,, \qquad (11.4.73)$$

where

$$Q = 1 + \frac{B^2\rho^2}{4}\,. \qquad (11.4.74)$$

This is the metric of *Melvin magnetic universe* [Bonnor (1953), Melvin (1964, 1965)]. The magnetic field is directed along the z-axis, and its field strength $\sqrt{(F_{\mu\nu}F^{\mu\nu}/2)} = B/Q^2$ has a maximum B on the axis $\rho = 0$ and decreases to zero at infinity.

In the absence of a magnetic field the Ernst metric (11.4.63) takes the form

$$ds^2 = \frac{1}{(x-y)^2A^2}\left[G(y)\,dt^2 - \frac{dy^2}{G(y)} + \frac{dx^2}{G(x)} + G(x)\,d\phi^2\right], \qquad (11.4.75)$$

with $G(z)$ given by (11.4.65). This is the so-called *C-metric*. Its physical meaning is discussed in detail by Kinnersley and Walker (1970). By introducing new coordinates

$$u = A^{-1}\left(t + \int \frac{dy}{G(y)}\right), \quad r = \frac{1}{A(x-y)}, \tag{11.4.76}$$

the C-metric can be presented in the form

$$ds^2 = -H du^2 - 2\, du\, dr - 2A\, r^2\, du\, dx + r^2 \left[\frac{dx^2}{G(x)} + G(x)\, d\phi^2\right], \tag{11.4.77}$$

where $H = -A^2 r^2 G(x - 1/(Ar))$. For $m = q = 0$, $G(z) = 1 - z^2$, and one has

$$\begin{aligned} ds^2 = &- (1 - 2A\, r\cos\theta + A^2 r^2 \sin^2\theta)\, du^2 - 2\, du\, dr \\ &- 2A\, r^2 \sin\theta\, du\, d\theta + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,. \end{aligned} \tag{11.4.78}$$

This is a flat-space metric written in Newman-Unti (1963) coordinates. The transformation required to cast (11.4.78) into a Minkowski coordinate system is

$$\begin{aligned} T &= (A^{-1} - r\cos\theta)\sinh(Au) + r\cosh(Au)\,, \\ X &= (A^{-1} - r\cos\theta)\cosh(Au) + r\sinh(Au)\,, \\ Y &= r\sin\theta\cos\phi\,, \quad Z = r\sin\theta\sin\phi\,. \end{aligned} \tag{11.4.79}$$

Provided $A > 0$ and $r > 0$, the (u, r, θ, ϕ)-coordinates cover only the half-space $T + X > 0$. The locus $r = 0$ is one branch of a hyperbola with constant acceleration A, parameterized in terms of its proper time u:

$$T = A^{-1}\sinh(Au)\,, \quad X = A^{-1}\cosh(Au)\,. \tag{11.4.80}$$

In the zero-acceleration limit, metric (11.4.77) takes the form

$$ds^2 = -\left(1 - \frac{2m}{r} + \frac{q^2}{r^2}\right) du^2 - 2\, du\, dr + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,. \tag{11.4.81}$$

This is the *Reissner-Nordström metric* in the retarded Eddington-Finkelstein coordinates. The vector potential $A_\phi = -q\cos\theta$ describes the magnetic field created by the monopole magnetic charge q.

Combining the information from these various limiting cases, one can conclude that the C-metric describes the combined gravitational and electromagnetic field of a pair of uniformly accelerating charged black holes [Kinnersley and Walker (1970)] (see Figure 11.9).

We consider now the geometry

$$dl^2 = r^2 \left[\frac{dx^2}{G(x)} + G(x)\, d\phi^2\right], \tag{11.4.82}$$

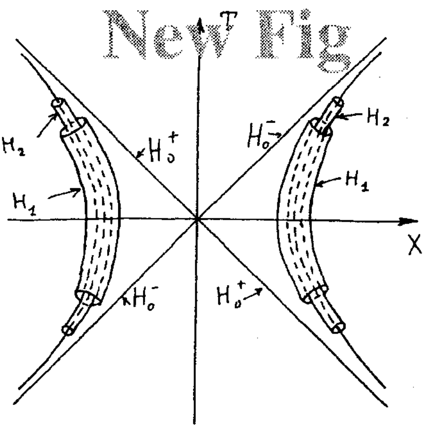


Figure 11.9: Qualitative picture of the charged C-metric. The geometry is extended to include both magnetically charged black holes. In the absence of the "left" black hole the solution is singular on the horizon H^-. The electromagnetic field has a δ-like singularity on H^- similar to the one discussed in Section 7.2.3. The outer horizons $H_0^\pm$ have the topology of two intersecting planes. The horizons $H_{1,2}$ have the topology $R^1 \times S^2$. They are the event horizon, H_1, and the Cauchy horizons, H_2, for each of the black holes.

of two-dimensional sections $u =$ const, $r =$ const of C-metric (11.4.77). The polynomial of fourth order $G(z)$, (11.4.65), has four roots. It is convenient to set $z_1 = -1/(Ar_-)$ and let $z_2 \le z_3 < z_4$, where z_2, z_3 and z_4 are the roots of the polynomial $f(z) = 1 - z^2 - r_+ A z^3$. At $z = \pm\infty$ $f(\pm\infty) = \mp\infty$. The function $f(z)$ has a local maximum $f(0) = 1$ at $z = 0$ and a local minimum at $z = -2/(3r_+ A)$. Hence, $z_3 < 0$ and $z_4 > 0$, and $f(z) \ge 0$ in the interval $z_3 \le z \le z_4$. Since $r \le r_+$, one also has $z_1 < z_3$. Thus, $G(z) \ge 0$ in the interval $z_3 \le z \le z_4$, and metric (11.4.82) has the Euclidean signature there and is conformal to the metric of a two-sphere S^2. We call z_3 the south pole and z_4 the north pole of the sphere. For this choice the south pole points towards spatial infinity, and the north pole points towards the other black hole. Metric (11.4.82) is regular at its pole z_i if ϕ is an angular coordinate and has period $\Delta\phi_i = 4\pi/G'(z_i)$. The metric is free from a conical singularity if and only if $G'(z_3) = -G'(z_4)$. Simple analysis shows that, save for the two exceptions $mA = 0$ and $eA = mA > (12)^{-1/2}$, the spacetime of the C-metric always contains a nodal timelike two-surface [Kinnersley and Walker (1970)]. If these singularities are chosen to run from the north poles of black holes to infinity, each of the two corresponding conical singularities can be related with a *cosmic string* which provides the force required for the acceleration of the black holes.

After these remarks we return to the discussion of the general Ernst solution. We restrict $z_3 \le x \le z_4$ in order for the metric to have Lorentz signature. The range of y is $-\infty < y < x$. The surface $y = z_1$ is the inner (Cauchy) horizon of the black hole;

$y = z_2$ is the black hole event horizon, and $y = z_3$ is the acceleration horizon. For $z_1 < z_2$ the black holes are nonextremal, and for $z_1 = z_2$ they are extremal.

The function $G(z)$ for the Ernst solution has the same form as for the C-metric. That is why one can directly apply the above analysis to the case of the motion of black holes in a magnetic field. The form of the metric (11.4.63) implies that the condition for the absence of node singularities at both (north and south) poles is

$$\frac{G'(z_3)}{Q^2(z_3)} = -\frac{G'(z_4)}{Q^2(z_4)}. \tag{11.4.83}$$

We recall that

$$Q = \left(1 + \frac{1}{2}Bqx\right)^2 + \frac{B^2r^2}{4}G(x) \tag{11.4.84}$$

in general depends on both coordinates, r and x. However, the value of Q for $x = z_i$, where $G(x)$ vanishes is independent of r. We define $L = Q(x = z_3)$; then when (11.4.83) is satisfied, the spheres are regular as long as ϕ has the period

$$\Delta\phi = \frac{4\pi L^2}{G'(z_3)}. \tag{11.4.85}$$

In what follows we assume that the acceleration A is small and expand z_i in terms of $r_\pm A$

$$z_1 = -\frac{1}{r_-A}, \quad z_2 = -\frac{1}{r_-A} + r_+A + \dots,$$
$$z_3 = -1 - \frac{r_+A}{2} + \dots, \quad z_4 = 1 - \frac{r_+A}{2} + \dots. \tag{11.4.86}$$

Using these relations, one gets

$$-\frac{G'(z_3)}{G'(z_4)} = 1 - 2(r_- + r_+)A + \dots,$$
$$\frac{Q^2(z_3)}{Q^2(z_3)} = 1 - 4Bq + \dots. \tag{11.4.87}$$

Hence, in the presence of the external field, condition (11.4.83) takes the form

$$qB = mA + \dots. \tag{11.4.88}$$

This relation has an obvious interpretation: Node singularities are absent if the acceleration of the black holes is provided entirely by the action of the external magnetic field.

Ernst's instanton and the probability of creation of pairs of charged black holes

We now consider the Euclidean section $\tau = it$ of the Ernst solution. In the non-extremal case, the corresponding metric has the Euclidean signature for $z_2 \leq y \leq z_3$. In the general case, the Euclidean metric possesses conical singularities at $y = z_2$ (the black hole horizon) and $y = z_3$ (the acceleration horizon). The solution is regular if

$$G'(z_2) = -G'(z_3), \tag{11.4.89}$$

and if the Euclidean time τ is periodic with period

$$\beta = \Delta\tau = \frac{4\pi}{G'(z_3)}. \tag{11.4.90}$$

For small acceleration $Ar_+ \ll 1$ condition (11.4.89) can be written in the form

$$\frac{r_+ - r_-}{4\pi r_+^2} = \frac{A}{2\pi} + \dots . \tag{11.4.91}$$

The two relations, (11.4.88) and (11.4.91), restrict the possible parameters q, m, B, and A, specifying a solution, so that only two of them can be fixed independently. In particular, since A is small, $r_+ - r_-$ must also be small, so that the black holes are almost extremal.

Consider now the section $t = 0$ of the Ernst solution (11.4.63). This section represent a space with two charged black holes which are at rest at this moment of time. The Ernst solution is time-symmetric with respect to the moment of closest approach of the two black holes. We identify the surfaces of their horizons at this moment, so that topologically the configuration is similar to the one presented in Figure 7.5. The topology of this configuration is $S^2 \times S^1 - \{p\}$, where the missing point is $x = y$ (which can occur only at z_2) and corresponds to infinity. The Euclidean section $\tau = it$ of the Ernst solution with $0 \leq \tau \leq \Delta\tau/2$ interpolates between the Melvin magnetic universe and a similar universe with a pair of black holes. Technically, it is much simpler to describe the "double" instanton obtained by gluing this instanton to a copy of itself along the $S^2 \times S^1 - \{p\}$ boundary [Garfinkle and Strominger (1991)]. The "double" instanton has topology $S^2 \times S^2 - \{p\}$, (x, ϕ) and (y, τ) being angular coordinates on the spheres and p being the missing point of Euclidean infinity. The Euclidean time τ for the "double" instanton varies from 0 to $\Delta\tau$. The Euclidean action calculated for the "double" instanton is [Dowker *et al.* (1994), Hawking, Horowitz, and Ross (1995)]

$$W_E = \frac{2\pi L^2}{A^2 G'(z_3)(z_3 - z_1)}. \tag{11.4.92}$$

In the limit of small acceleration this expression takes the form $W_E \approx \pi r_-/A \approx \pi q/B$. Hence, for black holes close to extremal $m \approx q$, the expression $\exp(-W_E)$ in the limit

$A \to 0$ correctly reproduces (11.4.62) [Garfinkle and Strominger (1991)]. Hawking, Horowitz, and Ross (1995) proved that W_E allows another representation:

$$W_E = -\frac{1}{4}(\Delta\mathcal{A} + \mathcal{A}_{BH}) \,. \tag{11.4.93}$$

Here $\Delta\mathcal{A}$ is the change of the area of the acceleration horizon due to the presence of a black hole, and $\mathcal{A}_{BH}$ is the black hole area. The presence of the additional factor $\mathcal{A}_{BH}$ can be interpreted as the increase in the probability of black hole pair creation because of the presence of the internal degrees of freedom of the black holes. We return to this interpretation later in Chapter 12, when we shall be discussing the statistical mechanical aspects of black hole physics.

Black hole pair creation by snapping of cosmic strings

A small modification of the above described formalism allows one to find the semiclassical probability of another possible under-barrier process: the creation of black hole pairs by the breaking of a cosmic string [Hawking and Ross (1995), Emparan (1995), Eardley *et al.* (1995)]. Let us consider the C-metric (11.4.75) but assume now that condition (11.4.83) is violated. Then the Lorentzian charged C-metric describes a pair of charged black holes accelerating away from one another. The acceleration is caused by the conical singularities along the X-axis. To avoid a conical singularity *between* the two black holes the following periodicity condition must be satisfied

$$\Delta\phi = \frac{4\pi}{|G'(z_4)|} \,. \tag{11.4.94}$$

We recall that z_4 corresponds to the north poles of the black holes that point towards each other. Conical singularities directed along the X-axis from the south poles of the black holes to infinity can be interpreted as the presence of cosmic strings that provide the acceleration of the black holes. The conical deficit angle

$$\delta = 2\pi\left(1 - \left|\frac{G'(z_3)}{G'(z_3)}\right|\right) \tag{11.4.95}$$

is related to the mass per unit length $\mu = \delta/8\pi$ of the cosmic string. This interpretation is supported by more accurate considerations. In particular, it was shown that a wide class of cosmic strings can really end in a black hole [Hawking and Ross (1995), Eardley *et al.* (1995)]. A cosmic string with a black hole at its end is in many aspects similar to a cosmic string with a monopole at one end. The Euclidean section of the C-metric thus gives an instanton describing the breaking of the cosmic string and creation of the pair of black holes at the ends of the broken cosmic string. As before, the condition (11.4.90) removes additional unphysical conical singularities of the Euclidean solution. Calculations give for the Euclidean action [Hawking and Ross (1995), Emparan (1995), Eardley *et al.* (1995)]

$$W_E = -\frac{1}{4}(\Delta\mathcal{A} + \mathcal{A}_{BH}) = \frac{2\pi}{A^2|G'(z_4)|(z_3 - z_1)} \,. \tag{11.4.96}$$

In the limit of small acceleration $r_+ A \ll 1$, the black holes are close to extremal; the mass per unit length of the string is

$$\mu \approx r_+ A\,, \tag{11.4.97}$$

and the Euclidean action takes the form

$$W_E \approx \frac{\pi m^2}{\mu}\,, \tag{11.4.98}$$

where m is the black hole mass. This expression is in agreement with the expression for the action describing breaking of a cosmic string and creation of a pair of monopoles at the end [Vilenkin (1982)]

$$W_E = \frac{\pi M^2}{\mu}\,, \tag{11.4.99}$$

where M is the mass of the monopole.

In a similar way, by finding the corresponding instantons and calculating the Euclidean action, one can describe other effects of black hole pair creation. For example, black hole pairs can be created by domain walls [Caldwell, Chamblin, and Gibbons (1996)], and by the cosmological constant[Mann and Ross (1995) and Bousso and Hawking (1996)]. It should be emphasized that the probability of these tunneling processes is negligible. For example, for realistic strings $\mu \sim 10^{-6} m_{\rm Pl}/l_{\rm Pl}$, so that for black holes of mass greater than the Planck mass the rate of black hole pair formation is less than $\exp(-10^6)$. So far the main importance of these results is connected with a demonstration of the possibility of the quantum effects that result in changes of the spatial topology. Another important aspect is the demonstration that the probability is usually greater than the expected quantity by the factor $\exp \mathcal{A}_{BH}/4$, which might be connected with an existence of different internal degrees of freedom of the created black holes.

Chapter 12

Thermodynamics and Statistical Mechanics of Black Holes

12.1 Black Holes and Thermodynamics

Hawking's discovery of thermal radiation from black holes was a complete surprise to most specialists, even though quite a few indications of a close relationship between black hole physics and thermodynamics had emerged before this discovery.

Wheeler seems to have been the first to notice that the very existence of a black hole in the classical theory of gravitation contradicts the law of non-decreasing entropy. Indeed, imagine that a black hole swallows a hot body possessing a certain amount of entropy. Then the observer outside it finds that the total entropy in the part of the world accessible to his observation has decreased. This disappearance of entropy could be avoided in a purely formal way if we simply assigned the entropy of the ingested body to the inner region of the black hole. In fact, this "solution" is patently unsatisfactory because any attempt by an "outside" observer to determine the amount of entropy "absorbed" by the black hole is doomed to failure. Quite soon after the absorption, the black hole becomes stationary and completely "forgets", as a result of "balding", such "details" as the structure of the ingested body and its entropy.

If we are not inclined to forgo the law of non-decreasing entropy because a black hole has formed somewhere in the Universe, we have to conclude that any black hole *by itself* possesses a certain amount of entropy and that a hot body falling into it not only transfers its mass, angular momentum and electric charge to the black hole, but also its entropy S as well. As a result, the entropy of the black hole increases by at least S. Bekenstein (1972, 1973a) noticed that the properties of one of the black hole characteristics – its area $\mathcal{A}$ – resemble those of entropy. Indeed, the Hawking's area theorem implies that the area $\mathcal{A}$ does not decrease in any classical processes; that is, it behaves as entropy does. It was found, in fact, that the analogy of black hole physics to thermodynamics is quite far-reaching. It covers both gedanken experiments with specific thermodynamic devices (like the heat engine) and the general laws of thermodynamics, each of which has an analogue in black hole physics.

An arbitrary black hole, like a thermodynamic system, reaches an equilibrium (stationary) state after the relaxation processes are completed. In this state, it is completely described by fixing a small number of parameters: M, J, Q. The area $\mathcal{A}$ of a stationary black hole is a function of these parameters:

$$\mathcal{A} = 4\pi\left(2M^2 - Q^2 + 2M\sqrt{M^2 - Q^2 - J^2/M^2}\right). \tag{12.1.1}$$

This relation can be inverted, yielding a formula for the internal energy of a black hole:

$$M \equiv M(\mathcal{A}, J, Q) = \left[\frac{\pi[(Q^2 + \mathcal{A}/4\pi)^2 + 4J^2]}{\mathcal{A}}\right]^{1/2}. \tag{12.1.2}$$

The internal energies of two stationary black holes with slightly differing values of area (by $d\mathcal{A}$), angular momentum (by dJ), and electric charge (by dQ), differ by an amount

$$dM = \frac{\kappa}{8\pi}\,d\mathcal{A} + \Omega^H\,dJ + \Phi^H\,dQ, \tag{12.1.3}$$

where $\kappa = 4\pi\sqrt{M^2 - Q^2 - J^2/M^2}\,/\mathcal{A}$ is the surface gravity; $\Omega^H = 4\pi J/M\mathcal{A}$ is the angular velocity, and $\Phi^H = 4\pi Q\,r_+/\mathcal{A}$ is the electric potential of the black hole. The second and third terms of this formula describe the changes in rotational energy and electric energy, respectively.

This relation is similar to the first law of thermodynamics. An analogue of temperature (the variable conjugate to entropy) is a quantity proportional to surface gravity κ. Hawking's result on the thermal nature of the radiation of stationary black holes not only supports this analogy, but also supplies the coefficient relating temperature T_H to surface gravity κ:

$$T_H = \hbar\,\kappa/2\pi c\,k. \tag{12.1.4}$$

Note that relation (12.1.3) is identical to the first law of thermodynamics[1]

$$dE = \theta\,dS^H + \Omega^H\,dJ + \Phi^H\,dQ, \tag{12.1.5}$$

if one assumes the following expression for the entropy of the black hole

$$S^H = \mathcal{A}/4l_{\mathrm{Pl}}^2, \qquad l_{\mathrm{Pl}}^2 = \hbar G/c^3. \tag{12.1.6}$$

This quantity is known as the Bekenstein-Hawking entropy.

These arguments give sufficient ground for taking the the analogy between black hole physics and thermodynamics quite seriously. In this chapter we discuss the

[1] $\theta = k\,T_H$ is the temperature T_H measured in energy units, and it differs from the latter by the factor k. The entropy in these units is dimensionless.

thermodynamics and statistical mechanics of systems containing black holes. We shall demonstrate the fruitfulness of thermodynamical ideas in the application to black hole physics. However, as we shall see, black holes considered as thermodynamical systems possess a number of peculiar properties. We shall discuss these properties in detail later. Now we just make a few comments.

1. Since gravity is a long-range force, self-gravitating systems in a state of thermal equilibrium are usually macroscopically inhomogeneous. This inhomogeneity is extremely important in the presence of a black hole.

2. The decrease of the size (the gravitational radius) of a black hole results in a decrease of its internal energy (mass). This property (characteristic for self-gravitating systems) results in a negative heat capacity for a non-rotating black hole and possible instability of a system containing such a black hole.

3. Though relation (12.1.5) can be used to the study of a black hole transition from one stationary configuration to another, it cannot by itself guarantee that similar relations are valid for the total system including the black hole and surrounding matter. The mechanism which makes possible the equilibrium of the black hole with surrounding thermal radiation is based on the existence of the Hawking radiation, i.e., a quantum gravity effect.

4. For fixed angular momentum and charge of the black hole its temperature (surface gravity) and size (surface area) are defined by one parameter (mass), and hence are not independent. For the standard thermodynamical systems the temperature (intensive parameter) and size (extensive parameter) are independent.

5. Usually thermodynamics gives us a description of systems which consist of a large number of particles (constituents) and can be derived from the microscopic theory of the constituents by means of statistical mechanical methods. A black hole (at least as a classical object) is nothing but an empty space with a very strong gravitational field. It is a highly non-trivial question whether the similarity between black holes and the ordinary thermodynamical system goes so far as to include also the possibility of a microscopical statistical-mechanical foundation of black hole thermodynamics.

We return to these problems later, but first we describe the thermodynamical analogy in black hole physics in more detail.

12.2 Mass Formulas

12.2.1 Integral mass formula

We begin by deriving a useful representation for the mass of a stationary system containing a black hole. We recall that the angular velocity Ω^H, surface gravity

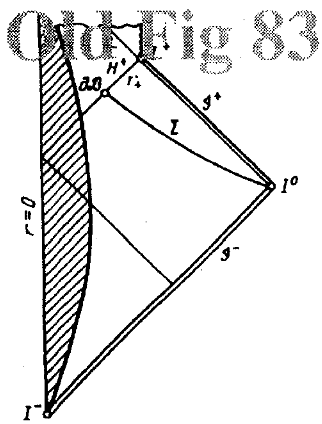


Figure 12.1: Spacetime of a stationary black hole (illustration to the derivation of the mass formula).

κ, and electric potential Φ^H of a stationary black hole are constant on the event horizon (see Section 6.3). This property is essential for the derivation of the so-called *mass formula.* This formula establishes a relationship between the black hole mass observed at infinity and the geometric characteristics of the event horizon. Specifically, Bardeen, Carter, and Hawking (1973) proved that the black hole mass observed at infinity, M^∞, in a stationary axisymmetric asymptotically flat spacetime can be written in the form

$$M^\infty = -\int_\Sigma (2\,T^\mu_\nu - T\,\delta^\mu_\nu)\,\xi^\nu_{(t)}\,d\sigma_\mu + 2\,\Omega^H J^H + \frac{\kappa}{4\pi}\mathcal{A}\,. \tag{12.2.1}$$

Here, Ω^H is the angular velocity; J^H the angular momentum; κ the surface gravity; $\mathcal{A}$ is the area of the black hole, and T^μ_ν the total energy-momentum tensor of the stationary axisymmetric distribution of matter and fields outside the black hole. The integration is carried out on a spacelike asymptotically flat surface Σ that intersects the event horizon along a two-dimensional surface $\partial\mathcal{B}$. The surface Σ is chosen in such a manner that $\xi^\mu_{(\phi)}$ is tangent to $\partial\mathcal{B}$ and Σ is orthogonal to $\xi^\mu_{(t)}$ at asymptotic infinity.

Formula (12.2.1) is proved as follows. For an arbitrary Killing vector field ξ^μ

$$\xi^{\mu;\nu}{}_{;\nu} = -R^\mu{}_\nu\,\xi^\nu. \tag{12.2.2}$$

This relation can be obtained by contracting (A.30) in α and β. Choosing $\xi^\mu_{(t)}$ for ξ^μ, integrating (12.2.2) over the surface Σ, and using Stokes' theorem [see (A.87)], we find

$$\int_{\partial\Sigma} \xi^{\mu;\nu}_{(t)}\,d\sigma_{\mu\nu} = -\int_\Sigma R^\mu{}_\nu\,\xi^\nu_{(t)}\,d\sigma_\mu\,, \tag{12.2.3}$$

where $d\sigma_{\mu\nu}$ and $d\sigma_\mu$ are elements of the surface Σ and its boundary $\partial\Sigma$. If the surface Σ is chosen as described above, its boundary $\partial\Sigma$ consists of the black hole boundary $\partial\mathcal{B}$ and a two-dimensional surface $\partial\Sigma_\infty$ at spatial infinity (Figure 12.1).

We now show that the black hole mass M^∞, measured by a distant observer in the asymptotically flat region, is given by the expression

$$M^\infty = \frac{1}{4\pi}\int_{\partial\Sigma_\infty} \xi^{\mu;\nu}_{(t)}\, d\sigma_{\mu\nu}\,. \tag{12.2.4}$$

To do this, we consider the effect which the black hole exerts on test particles. We assume that a test body is at rest far from the black hole. The four-acceleration of this body is then

$$a^\mu = -\frac{\xi^\nu_{(t)}\,\xi^\mu_{(t);\nu}}{\xi^\alpha_{(t)}\,\xi_{(t)\alpha}}\,. \tag{12.2.5}$$

Let Σ be a spacelike surface orthogonal to the four-velocity of the body, $u^\mu = \xi^\mu/|\,\xi_\alpha\,\xi^\alpha\,|^{1/2}$, in the asymptotic region. The gravitational field in this region is weak so that the relation of its invariant four-dimensional characteristics to the Newtonian description is easily established. Thus, the vector $a^\nu \approx u^\nu\,\xi^\mu_{(t);\nu}$, which lies in Σ, has only three nonzero components. In the Newtonian theory, this three-dimensional vector characterizes the strength of the gravitational field and is related to the gravitational potential U by the formula $a_i = U_{,i}$. By Gauss' theorem, the flux of this vector across any closed two-dimensional surface $\partial\Sigma_\infty$ (that lies in Σ) enclosing the gravitating body equals $4\pi\, M^\infty$, where M^∞ is the mass of the body. Let $\tilde{n}^\mu$ be the unit vector of the outward normal to $\partial\Sigma_\infty$ that lies in Σ. Then

$$M^\infty = \frac{1}{4\pi}\int_{\partial\Sigma_\infty} \xi^{\mu;\nu}_{(t)}\,\tilde{n}_\mu\, u_\nu\, d^2\sigma\,, \tag{12.2.6}$$

where $d^2\sigma$ is an element of the area $\partial\Sigma_\infty$. Since $\xi^{\mu;\nu}$ is an antisymmetric tensor, we can replace the expression $\tilde{n}_\mu\, u_\nu\, d^2\sigma$ with $d\sigma_{\mu\nu} = \tilde{n}_{[\mu}\, u_{\nu]}\, d^2\sigma$. Expression (12.2.6) then transforms to (12.2.4).

Likewise, it can be shown that the total angular momentum J^∞ of the system, measured by a distant observer (e.g., by recording the Lense-Thirring drag), is given by the following formula:[2]

$$J^\infty = -\frac{1}{8\pi}\int_{\partial\Sigma_\infty} \xi^{\mu;\nu}_{(\phi)}\, d\sigma_{\mu\nu}\,. \tag{12.2.7}$$

[2]Direct checking will demonstrate that formulas (12.2.4) and (12.2.7) hold for the mass M and angular momentum J if one takes into account that the metric far from the rotating body can be written in the following form:

$$ds^2 = -\left[1 - \frac{2M}{r} + O(r^{-2})\right] dt^2 - \left[\frac{4J}{r}\sin^2\theta + O(r^{-2})\right] dt\, d\phi +$$
$$+\ (1 + O(r^{-1}))\left[dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\right].$$

Using the relation (12.2.3) and a similar relation for $\xi_{(\phi)}$, we can obtain, taking into account (12.2.4) and (12.2.7), that

$$M^{\infty} = -\int_{\Sigma} (2\,T^{\mu}_{\nu} - T^{\alpha}_{\alpha}\,\delta^{\mu}_{\nu})\,\xi^{\nu}_{(t)}\,d\sigma_{\mu} + M^{H}, \tag{12.2.8}$$

$$J^{\infty} = \int_{\Sigma} T^{\mu}_{\nu}\,\xi^{\nu}_{(\phi)}\,d\sigma_{\mu} + J^{H}, \tag{12.2.9}$$

where the integrals on the right-hand sides describe the contributions of matter and fields outside the black hole to the total mass M^{∞} and angular momentum J^{∞} of the system, while

$$M^{H} = \frac{1}{4\pi}\int_{\partial\mathcal{B}} \xi^{\mu;\nu}_{(t)}\,d\sigma_{\mu\nu} \tag{12.2.10}$$

and

$$J^{H} = -\frac{1}{8\pi}\int_{\partial\mathcal{B}} \xi^{\mu;\nu}_{(\phi)}\,d\sigma_{\mu\nu} \tag{12.2.11}$$

are the contributions of the mass and angular momentum of the black hole itself to M^{∞} and J^{∞}.[3]

The expression for M^{H} can be transformed as follows. We recall that $l^{\mu} = \xi^{\mu}_{(t)} + \Omega^{H}\xi^{\mu}_{(\phi)}$ is a null generator of the event horizon of the rotating black hole. Expressing $\xi^{\mu}_{(t)}$ in terms of l^{μ} and $\xi^{\mu}_{(\phi)}$, we find

$$\int_{\partial\mathcal{B}} \xi^{\mu;\nu}_{(t)}\,d\sigma_{\mu\nu} = 8\pi\,J^{H}\,\Omega^{H} + \int_{\partial\mathcal{B}} l^{\mu;\nu}\,d\sigma_{\mu\nu}\,. \tag{12.2.12}$$

We choose a complex null tetrad $[m_{\mu}\,\bar{m}^{\mu} = -\,l_{\mu}\,n^{\nu} = 1]$ so that the vectors m^{μ} and $\bar{m}^{\mu}$ lie in the plane tangent to $\partial\mathcal{B}$. Then $d\sigma_{\mu\nu} = l_{[\mu}\,n_{\nu]}\,d\mathcal{A}$, where $d\mathcal{A}$ is an area element of the two-dimensional surface $\partial\mathcal{B}$. Using the definition (6.3.20) of the surface gravity κ and also its constancy, we can write the integral on the right-hand side of (12.2.12) in the form

$$\int_{\partial\mathcal{B}} l^{\mu;\nu}\,d\sigma_{\mu\nu} = \kappa\mathcal{A}\,, \tag{12.2.13}$$

[3]It must be emphasized that the expression (A.82) which we use for the surface element $d\sigma_{\mu\nu}$ agrees with that employed by Bardeen, Carter, and Hawking (1973) and Carter (1973a), but is twice as small as that in Carter (1979) and Damour (1982). Note also that the orientation of $d\sigma_{\mu\nu}$ on the surfaces $\partial\Sigma_{\infty}$ and $\partial\mathcal{B}$ was chosen so as to have

$$\int_{\partial\Sigma} \varphi^{\mu\nu}\,d\sigma_{\mu\nu} = \int_{\partial\Sigma_{\infty}} \varphi^{\mu\nu}\,d\sigma_{\mu\nu} - \int_{\partial\mathcal{B}} \varphi^{\mu\nu}\,d\sigma_{\mu\nu}\,.$$

where

$$\mathcal{A} = \int_{\partial\mathcal{B}} d\mathcal{A} \tag{12.2.14}$$

is the total area of the black hole surface. Using this equality and substituting (12.2.12) into (12.2.8), we arrive at the mass formula (12.2.1).

Recall that T_ν^μ in this formula is the total energy-momentum tensor of matter and fields outside the black hole. If an electromagnetic field is present, the tensor T_ν^μ is composed of two parts: The energy-momentum tensor of matter, $T^{(\mathrm{m})\mu}_{\ \nu}$, and the energy-momentum tensor of the electromagnetic field, $T^{(\mathrm{em})\mu}_{\ \nu}$. Using the expression (F.16) for $T^{(\mathrm{em})\mu}_{\ \nu}$, we can transform formula (12.2.1) into the following form [Carter 1973a, 1979)]:

$$M^\infty = -\int_\Sigma \left(2\,T^{(\mathrm{m})\mu}_{\ \nu} - T^{(\mathrm{m})\alpha}_{\ \alpha}\,\delta^\mu_\nu\right)\xi^\nu_{(t)}\,d\sigma_\mu + \widetilde{M}^H + M^{(\mathrm{em})}_{\mathrm{ext}}, \tag{12.2.15}$$

where $M^{(\mathrm{em})}_{\mathrm{ext}}$ is the contribution to the total mass due to the presence of electric current j^μ outside the black hole,

$$M^{(\mathrm{em})}_{\mathrm{ext}} = \int_\Sigma \left[-2\,\xi^\nu_{(t)}\,A_\nu\,j^\mu + j^\nu A_\nu\,\xi^\mu_{(t)}\right] d\sigma_\mu\,, \tag{12.2.16}$$

and $\widetilde{M}^H$ is the black hole mass including the energy of its electromagnetic field:

$$\widetilde{M}^H = 2\,\Omega^H \tilde{J}^H + \Phi^H Q + \frac{\kappa}{4\pi}\mathcal{A}. \tag{12.2.17}$$

Here

$$\tilde{J}^H = J^H + J^{(\mathrm{em})H}\,, \qquad J^{(\mathrm{em})H} = \frac{1}{4\pi}\int_{\partial\mathcal{B}} \xi^\alpha_{(\phi)}\,A_\alpha\,F^{\mu\nu}\,d\sigma_{\mu\nu}\,, \tag{12.2.18}$$

and

$$Q = \frac{1}{4\pi}\int_{\partial\mathcal{B}} F^{\mu\nu}\,d\sigma_{\mu\nu} \tag{12.2.19}$$

is the electric charge of the black hole. In the above formulas, A_μ denotes the vector potential of the electromagnetic field, vanishing at infinity and satisfying the relations $\mathcal{L}_{\xi_{(t)}} A_\mu = \mathcal{L}_{\xi_{(\phi)}} A_\mu = 0$.[4] If there is no matter or current outside the black hole, its mass M^∞ measured at infinity coincides with $\widetilde{M}^H$. Relation (12.2.17) for an isolated black hole was derived by Smarr (1973a).

[4]The existence of a gauge in which these relations are satisfied can be verified if one uses the conditions $\mathcal{L}_{\xi_{(t)}} F_{\mu\nu} = \mathcal{L}_{\xi_{(\phi)}} F_{\mu\nu} = 0$ of stationarity and axial symmetry of the field $F_{\mu\nu}$.

12.2.2 Differential mass formula

The integral mass formula (12.2.1) makes it possible to calculate the mass difference dM^{∞} for two static or stationary axisymmetric configurations containing a black hole. We denote the appropriate metric variation by $\delta g_{\mu\nu}$ and specify the gauge arbitrariness $\delta g_{\mu\nu} \to \delta g_{\mu\nu} + \delta\lambda_{(\mu;\nu)}$ (inherent in the choice of this quantity) in such a way that $\delta\xi^{\mu}_{(t)} = \delta\xi^{\mu}_{(\phi)} = 0$, with unchanged position of the event horizon. Then the general expression for the mass variation is written in the following form [Bardeen, Carter, and Hawking (1973)]:

$$dM^{\infty} = \Omega^H \, dJ^H + \frac{\kappa}{8\pi}\, dA - \delta \int_{\Sigma} T^{\mu}_{\nu}\, \xi^{\nu}_{(t)}\, d\sigma_{\mu} - \frac{1}{2}\int_{\Sigma} T^{\mu\nu}\, \delta g_{\mu\nu}\, \xi^{\alpha}_{(t)}\, d\sigma_{\alpha} \, . \qquad (12.2.20)$$

In the case when an ideal fluid, described by the energy-momentum tensor $T^{(\mathrm{m})\mu}_{\nu}$, revolves around the black hole at a local angular velocity Ω, and the system is stationary and axially symmetric both before and after its parameters undergo a change, this general relation takes the following form [Carter (1973a), Damour (1982)]:

$$dM^{\infty} = \Omega^H \, d\tilde{J}^H + \Phi^H \, dQ + \frac{\kappa}{8\pi}\, \delta A + \delta q \, , \qquad (12.2.21)$$

where

$$\delta q = \int \Omega\, \delta \left(d^3 J^{(\mathrm{m})} + \xi^{\alpha}_{(\phi)} A_{\alpha}\, d^3 Q\right) - \int \left(\xi^{\alpha}_{(t)} + \Omega\, \xi^{\alpha}_{(\phi)}\right) A_{\alpha}\, \delta(d^3 Q)$$

$$+ \int \overline{T}\, \delta \left(d^3 S^{(\mathrm{m})}\right) + \int \bar{\mu}\, \delta(d^3 N) \, . \qquad (12.2.22)$$

Here, $d^3 J^{(\mathrm{m})} = T^{(\mathrm{m})\mu}_{\nu}\, \xi^{\nu}_{(\phi)}\, d\sigma_{\mu}$ is an element of local angular momentum of matter; $d^3 Q = j^{\mu}\, d\sigma_{\mu}$ is the local charge distribution; j^{μ} is the electric current; $\overline{\theta} = [-(\boldsymbol{\xi}_{(t)} + \Omega\, \boldsymbol{\xi}_{(\phi)})^2]^{1/2}\, \theta$; θ is the local temperature; $d^3 S^{(\mathrm{m})}$ is the local distribution of the matter entropy; $\bar{\mu} = \mu\overline{\theta}/\theta$, μ is the local chemical potential, and $d^3 N$ is the local distribution of the number of particles.

The integral and differential expressions for the mass formula given in this section are convenient for analyzing a numerous aspects of the processes that change the parameters of black holes. These expressions are among the main relations employed in the description of the analogues of the laws of thermodynamics in black hole physics. If matter and fields are absent outside the black hole, the last term in formula (12.2.21) vanish, and it becomes identical to formula (12.1.3) for the Kerr-Newman black hole.

12.3 Four Laws of Black Hole Physics

12.3.1 Four laws of black hole thermodynamics

According to the thermodynamic analogy in black hole physics, the quantities

$$\theta = \frac{\hbar \kappa}{2\pi k c}, \qquad S^H = \frac{\mathcal{A}}{4l_{\rm Pl}^2}, \qquad E = Mc^2 \tag{12.3.1}$$

(κ is the surface gravity, $\mathcal{A}$ is the area and M is the mass of the black hole) play the role of temperature, entropy, and internal energy of the black hole, respectively. Bardeen, Carter, and Hawking (1973) formulated the four laws of black hole physics, which are similar to the four laws of thermodynamics.

Zeroth law. *The surface gravity κ of a stationary black hole is constant everywhere on the surface of the event horizon.*

Thermodynamics does not permit equilibrium when different parts of a system are at different temperatures. The existence of a state of thermodynamic equilibrium and temperature is postulated by the zeroth law of thermodynamics. The zeroth law of black hole physics plays a similar role. This proposition was proved in Section 6.3.3 under the assumption that the energy dominance condition is satisfied.

First law. *When the system incorporating a black hole switches from one stationary state to another, its mass changes by*

$$dM = \theta\, dS^H + \Omega^H d\tilde{J}^H + \Phi^H dQ + \delta q, \tag{12.3.2}$$

where $d\tilde{J}^H$ and dQ are the changes in the total angular momentum and electric charge of the black hole, respectively, and δq is the contribution to the change in the total mass due to the change in the stationary matter distribution outside the black hole.

This is nothing more but the differential mass formula (12.2.21), and δq for an ideal fluid has the form (12.2.22).

Second law. *In any classical process, the area of a black hole, $\mathcal{A}$, and hence its entropy S^H, do not decrease:*

$$\Delta S^H \geq 0. \tag{12.3.3}$$

This analogue of the second law of thermodynamics is a consequence of Hawking's area theorem, which is valid provided the weak energy condition is satisfied (Section 5.4). In both cases (thermodynamics and black hole physics), the second law signals the irreversibility inherent in the system as a whole, and thus singles out the direction of the time arrow. In thermodynamics, the law of non-decreasing entropy signifies that the part of the internal energy that cannot be transformed into work grows with time. Quite similarly, the law of non-decreasing area of a black hole signifies that the fraction of a black hole's internal energy that cannot be extracted grows with time. As in thermodynamics, the quantity S^H stems from the impossibility of extracting any information about the structure of the system (in this case, the structure of the black hole).

Third law. In thermodynamics, the third law has been formulated in a variety of ways[5]. Two (essentially equivalent) formulations, due to Nernst, state that: (1) Isothermal reversible processes become isentropic in the limit of zero temperature, and (2) It is impossible to reduce the temperature of any system to the absolute zero in a finite number of operations. A stronger version, proposed by Planck, states that: The entropy of any system tends, as $\theta \to 0$, to an absolute constant, which may be taken as zero.

Bardeen, Carter, and Hawking (1973) formulated the analogue of the **third law** for black holes in the following form: *It is impossible by any procedure, no matter how idealized, to reduce the black hole temperature to zero by a finite sequence of operations.*

Since the black hole temperature θ vanishes simultaneously with κ, this is only possible if an isolated stationary black hole is extremal: $M^2 = a^2 + Q^2$. The impossibility of transforming a black hole into an extremal one is closely related to the impossibility of realizing a state with $M^2 < a^2 + Q^2$ in which a naked singularity would appear and the "cosmic censor" principle would be violated. An analysis of specific examples [see, e.g., Wald (1974a)] show that the closer the state of the black hole comes to the extremal one, the more restrictive are the conditions on the possibility of making the next step.

Israel (1986b) emphasized that it is difficult to define the meaning of "finite sequence of operations" considering only quasi-static processes which were analyzed by Bardeen, Carter, and Hawking (1973). He proposed and proved the following version of the **third law**: *A non-extremal black hole cannot become extremal at a finite advanced time in any continuous process in which the stress-energy tensor of accreated matter stays bounded and satisfies the weak energy condition in the a neighborhood of the outer apparent horizon.*

The **proof** of this law is based on the earlier result of Israel (1986a) according to which: *A trapped surface can be extended, in a locally-preserving fashion, to a three-cylinder that is everywhere spacelike provided it encounters no singularities in its development and weak energy condition is satisfied.* This result implies that initially trapped surface is permanently sealed off inside the black hole. This makes impossible a process of transition of a black hole to its extremal state during a finite time interval because in such a process the trapped surface would disappear.

[5]The third law of thermodynamics was discovered by Nernst (1906). It was extended and reformulated by Planck (1911). The precise conditions under which this law is valid are quite subtle, and there were many attempts to clarify them (see Wilson (1957), Callen (1960), Lewis and Randal (1961), Kestin (1968)). For a discussion of the third law in connection with black hole physics, see Israel (1986b)

It must be specially emphasized that another formulation of the third law of thermodynamics, stating that the entropy of a system vanishes at zero absolute temperature, is not valid for black holes because the area $\mathcal{A}$ remains finite as $\kappa \to 0$.[6] Wald (1997) discussed the possibility of violation of the third law in this formulation for a class of simple thermodynamical systems.

12.3.2 Generalized second law

Formulation of the generalized second law

Quantum effects violate the condition for the applicability of Hawking's area theorem. Thus, quantum evaporation reduces the area of black holes, and inequality (12.3.3) is violated. On the other hand, black hole radiation is thermal in nature and the black hole evaporation is accompanied by a rise in entropy in the surrounding space. One may expect that the so-called *generalized entropy* $\tilde{S}$, defined as the sum of the black hole entropy S^H and the entropy of the radiation and matter outside the black hole, S^m,

$$\tilde{S} = S^H + S^m , \tag{12.3.4}$$

does not decrease. In fact, we note that the rate of increase (by the clock of a distant observer) of mass and entropy of matter in the black hole exterior, because of Hawking radiation of a massless spin-s field, can be written in the form

$$\frac{dM^m}{dt} = -\frac{dM^H}{dt} = \frac{1}{4}\sigma_s\, h_s\, \Sigma_s\, \theta^4 , \qquad \frac{dS^m}{dt} = \frac{1}{3}\sigma_s\, B_s\, h_s\, \Sigma_s\, \theta^3 , \tag{12.3.5}$$

where h_s is the number of polarizations of the field; $\sigma_s = \pi^2/30$ for bosons and $7\pi^2/240$ for fermions; Σ_s is the effective cross section of the black hole; θ is its temperature, and B_s is a dimensionless coefficient of order unity. On the other hand, the change in the entropy S^H of a non-rotating black hole is related to the change in its mass M by the formula

$$dS^H = \theta^{-1}\, dM^H . \tag{12.3.6}$$

Comparing (12.3.5) and (12.3.6), we find [Zurek (1982)]

$$R \equiv -\frac{dS^m}{dS^H} = \frac{4}{3} B_s . \tag{12.3.7}$$

Numerical results obtained using formulas (10.5.14) and (10.5.31) [Zurek (1982), Page (1983)] demonstrate that the coefficient B_s is always greater than 3/4; hence, the generalized entropy $\tilde{S}$ increases when an isolated black hole emits radiation. It can be shown [Zurek (1982)] that if there is black body radiation at a temperature $\tilde{\theta}$

[6] Gibbons and Kallosh (1995) and Hawking, Horowitz, and Ross (1995) drew attention to the fact that under special conditions the entropy of an extremal black hole might vanish.

outside the black hole, the generalized entropy again increases, except in the case when $\tilde{\theta} = \theta$. In this special case, the increase in entropy in the space around the black hole due to its evaporation is exactly compensated for by a decrease in the entropy in this space, due to the accretion of thermal radiation onto the black hole.

These arguments give sufficient grounds for assuming that the following law holds:[7]

Generalized second law. *In any physical process involving black holes, the generalized entropy $\tilde{S}$ does not decrease.*

$$\Delta\tilde{S} = \Delta S^H + \Delta S^m \geq 0 . \tag{12.3.8}$$

The fact that the generalized second law includes, on an equal footing, the seemingly very different quantities, S^m (which characterizes the "degree of chaos" in the structure of the physical matter) and S^H (which is a geometric characteristic of the black hole), is a new indication of their profound similarity. In fact, the very possibility of this relation is rooted in Einstein's equations, which relate the physical characteristics of matter with the geometric characteristics of spacetime.

Proof of the generalized second law

A simple direct proof of the generalized second law for a quasi-stationary black holes was given by Frolov and Page (1993). The idea of the proof is following. A quasi-stationary black hole may be considered to emit thermal radiation in *UP*-modes described by a density matrix $\hat{\rho}_0$.[8] Suppose there is also radiation with density matrix $\hat{\rho}_1$ incident on the hole from far away (e.g., past null infinity, $\mathcal{J}^-$) in the corresponding *IN*-modes (i.e., incoming modes which are of positive frequency at $\mathcal{J}^-$). We use the semiclassical approximation and assume that the radiation in these two sets of modes will be quantum mechanically uncorrelated, giving an initial product state

$$\hat{\rho}_{\text{initial}} \equiv \hat{\rho}_{01} = \hat{\rho}_0 \otimes \hat{\rho}_1 . \tag{12.3.9}$$

After the initial state $\hat{\rho}_{01}$ interacts with the potential barrier separating the horizon from infinity, and possibly interacts with itself as well, it will have evolved unitarily into a (generally) correlated final state

$$\hat{\rho}_{\text{final}} \equiv \hat{\rho}_{23} \neq \hat{\rho}_2 \otimes \hat{\rho}_3 , \tag{12.3.10}$$

[7] The generalized second law in this formulation was first suggested by Bekenstein (1972, 1973b, 1974) before the quantum radiation of black holes was discovered.

[8] For the definition of *UP*-, *IN*-, *DOWN*-, and *OUT*-modes, see Section 10.2. Note that, according to the standard definition, these modes are defined for the eternal black hole. In the case in which the black hole arises from gravitational collapse and becomes quasi-stationary, the *UP*-modes are defined to be the same in the future stationary region as the *UP*-modes of the eternal black hole with the same future stationary region. As before, even when speaking about a black hole that was formed as a result of collapse, we use terminology developed for the eternal black hole. The possibility and convenience of using the eternal version of a black hole was discussed in Section 10.2.

where $\hat{\rho}_2 = \mathrm{tr}_3\,\hat{\rho}_{23}$ is the density matrix of the radiation in the *OUT*-modes escaping to future null infinity $\mathcal{J}^+$, and $\hat{\rho}_3 = \mathrm{tr}_2\,\hat{\rho}_{23}$ is that in the *DOWN*-modes that are swallowed by the future horizon H^+.

The entropy of each of these states is $S_i = -\,\mathrm{tr}(\hat{\rho}_i \ln \hat{\rho}_i)$. Because the evolution from $\hat{\rho}_{01}$ to $\hat{\rho}_{23}$ is unitary, $S_{23} = S_{01}$. Furthermore, since $\hat{\rho}_{01}$ is uncorrelated, but $\hat{\rho}_{23}$ is generically partially correlated, the entropies of these states obey the inequality [Araki and Lieb (1970)]

$$S_2 + S_3 \geq S_{23} = S_{01} = S_0 + S_1. \tag{12.3.11}$$

The first law of black hole physics implies that

$$dS^H = \theta^{-1}\left(dM - \Omega^H dJ - \Phi^H dQ\right) = \tilde{\theta}^{-1}\, dE\,, \tag{12.3.12}$$

where $\tilde{\theta}$ and E are the local temperature and energy as measured by an observer corotating with the hole near the horizon. If E_0 and E_3 are the expectation values of the local energies of the emitted state $\hat{\rho}_0$ and the absorbed state $\hat{\rho}_3$, respectively, then the semiclassical approximation, combined with the first law, says that

$$\Delta S^H = \tilde{\theta}^{-1}(E_3 - E_0)\,, \tag{12.3.13}$$

assuming that the changes to the hole are sufficiently small so that $\tilde{\theta}$ stays nearly constant throughout the process (the quasi-stationary approximation).

Equation (12.3.11) and inequality (12.3.12) imply that the change in the generalized entropy is

$$\begin{aligned}\Delta\tilde{S} &= \Delta S_{BH} + \Delta S_{rad} = \tilde{\theta}^{-1}(E_3 - E_0) + S_2 - S_1 \\ &\geq \left(S_0 - \tilde{\theta}^{-1}E_0\right) - \left(S_3 - \tilde{\theta}^{-1}E_3\right).\end{aligned} \tag{12.3.14}$$

Now for fixed $\tilde{\theta}$ and for equivalent quantum systems (as the *UP*-modes of $\hat{\rho}_0$ and the corresponding *DOWN*-modes of $\hat{\rho}_3$ are by *CPT* reversal, for a quasi-stationary black hole), $S - \tilde{\theta}^{-1}E$ is a Massieu function [Massieu (1869), Callen (1960)] (essentially the negative of the local free energy divided by the temperature) which is maximized for the thermal state $\hat{\rho}_0$. Therefore,

$$\Delta\tilde{S} \geq 0\,, \tag{12.3.15}$$

which is the generalized second law. This is an explicit mathematical demonstration of what Zurek and Thorne (1985) argued verbally, that the generalized second law is a special case of the ordinary second law, with the black hole as a hot, rotating, charged body that emits thermal radiation uncorrelated with what is incident upon it.

This proof applies to any emission, scattering, or absorption process, even for interacting fields, so long as the semiclassical approximation applies, and so long as the changes take place in a quasi-stationary manner.

Role of quantum effects

It is instructive to discuss a gedanken experiment that has been proposed to show the validity of the generalized second law [Bekenstein (1973b, 1974)]. Imagine that radiation (or matter) possessing energy E^m and entropy S^m is placed in a box with impenetrable (mirror) walls, and let this box be slowly lowered toward a non-rotating black hole. When the bottom of the box is in the immediate vicinity of the black hole surface, the bottom lid is opened and the contents of the box fall into the black hole. The box is then hoisted up.

Let us evaluate the change in generalized entropy in this process. Let all dimensions of the box, including its height l, be small in comparison with the black hole radius r_g. When the lower face of the box reaches the black hole, the upper face is at $r = r_g + \Delta r$, where $\Delta r \approx l^2/4r_g$. The energy transferred to the black hole (taking into account the redshift) does not exceed the amount

$$\varepsilon = \left(1 - \frac{r_g}{r}\right)^{1/2} E^m \approx \frac{l}{2r_g} E^m, \tag{12.3.16}$$

and its entropy increases by

$$\delta S^H \sim \theta^{-1}\, \varepsilon \approx 2\pi\, l\, E^m. \tag{12.3.17}$$

These arguments imply that the generalized entropy is indeed non-decreasing in this process provided the inequality $S^m \leq 2\pi l\, E^m$ is satisfied.[9]

For example, consider a massless particle in the box. Its energy can be estimated by using the uncertainty relation as $E^m \sim hl^{-1} = 2\pi l^{-1}$ (for $\hbar = 1$), and hence for this particle $2\pi l E^m \sim 1$. On the other hand, the decrease S^m of the entropy because of the absorption of one particle is $S^m \leq 1$. Thus, one can expect that relation $S^m \leq 2\pi\, l\, E^m$ is satisfied for such a process.

But if we slightly modify this gedanken experiment and choose l to be sufficiently small and much smaller than the other two dimensions of the box, L, we may conclude, at first glance, that the generalized second law can be violated since the energy $E^m \sim h/L$ of a quantum can be made arbitrary small. Unruh and Wald (1982, 1983a) proved that this conclusion is too hasty and wrong. Indeed, the analysis ignores effects due to vacuum polarization. When these effects are taken into account, the situation is changed drastically.

Let us assume that the box has mirror walls. As the box slowly approaches the black hole, the walls undergo progressively higher acceleration, and the accelerated motion of the mirror boundaries generates energy fluxes (due to quantum effects). This behavior is well known in flat spacetime [see DeWitt (1975), Birrell and Davies (1982)]. Its physical cause is the following. When a mirror boundary reflects radiation, charges and currents are induced on it. A similar phenomenon takes place in the

[9] On the validity of this constraint for real physical systems, see Bekenstein (1982, 1983), Unruh and Wald (1983a).

vacuum. The charges and currents "induced" by zero-point oscillations fluctuate and their expectation values are zero. In the presence of the boundary the renormalized expectation value of the stress-energy tensor is modified (the *Casimir effect*). For a stationary boundary the net flux of quantum radiation vanishes. If such a boundary begins to move with an acceleration, the "induced" charges and currents emit radiation. When a flat mirror moves at an increasing acceleration in flat spacetime, energy fluxes in the direction of the acceleration vector are generated on both sides of the mirror.[10]

Unruh and Wald (1982, 1983a) noticed that a similar effect takes place when a box with mirror walls is lowered into a black hole. Slow lowering of the box produces an additional positive energy flux into the black hole due to vacuum polarization by the walls. As a result, the energy ε transferred to the black hole and the related amount of the change of the entropy S^H are increased, and the generalized entropy does not decrease.

If the box is rigid, then the acceleration of the bottom is always greater than that of the upper wall. The net radiation of the mirror walls into the box reduces the energy inside. If an empty box is lowered, the energy inside it may become negative. At a point where the acceleration is a ($a \gg \kappa$), the energy density in the box is of order $-\sigma\,\theta_a^4$, where $\theta_a = a/2\pi$ is the local temperature measured by an observer at rest in this point (see Section 11.3.4). If we now open this box, a flux of negative energy into the black hole will be formed. This will stop when the energy-momentum density of radiation in the box becomes comparable to the energy-momentum density $\langle \hat{T}^\mu_\nu \rangle$ in the neighborhood of the point. If the lid of the box is then closed and the box is pulled out, the observer finds that it is filled with thermal radiation at a temperature $\theta_a = a/2\pi$ [for more details, see e.g., Anderson (1994)]. As a result of this cyclic process, the mass of the black hole slightly decreases and the energy equal to the difference between the energy of extracted radiation and the work done in the cycle can be utilized for doing work [Unruh and Wald (1982, 1983a)]. This additional energy properly taken into account allows one to demonstrate the validity of the generalized second law for this process without using the inequality $S^m < 2\pi l E^m$ [Unruh and Wald (1982, 1983a)].

By repeating the cycle, it is possible to continuously extract energy, even from cold (massive) black holes. Unruh and Wald (1983b) obtained a theoretical restriction on the admissible power produced in such a process: $|\,dE/dt\,| \le c^5/G \approx 3.6 \times 10^{59}$ erg/s. This process also obeys the generalized second law of black hole physics.

[10]For a scalar massless field in a two-dimensional spacetime this flux, $T_{\mu\nu}\, v^\mu\, (a^\nu/a)$ (v^μ is the four-velocity, and a^μ is the four-acceleration), equals $(12\,\pi)^{-1}\, da/d\tau$ (τ is the proper time of the observer on the surface of the mirror) [DeWitt (1975), Birrell and Davies (1982)]. For the generalization to four dimensions, see Anderson and Israel (1997).

12.3.3 Entropy as the Noether charge

The differential mass formula which plays the key role in establishing the four laws of black hole physics requires the Einstein equations for its derivation. One may suspect that the established relations between properties of black holes and thermodynamical systems have a deeper origin and remain valid for a more general class of theories of gravity which still allow black holes. We shall discuss some of these generalizations in Chapter 13. But even if we preserve the Einstein equations at the classical level, quantum corrections inevitably modify them. In particular, as we discussed earlier, higher-order curvature corrections are present in the effective action of quantum gravity. Though these corrections are small for black hole masses $M \gg m_{\rm Pl}$, their very existence requires generalization of the black hole laws. This generalization was recently proposed by Wald (1993). Namely, Wald demonstrated that the first law of thermodynamics can be formulated in any diffeomorphism-invariant theory in any spacetime that contains a stationary black hole with bifurcate Killing horizon, and the black hole entropy is related to the *Noether charge* [Wald (1993), Iyer and Wald (1994)].

The key role in this construction is played by the *Noether current* connected with the general covariance of the theory. To define it, let us consider a theory with the action

$$W = \int_\Omega dx\ \Lambda(x, \varphi_A, \varphi_{A,\alpha}, \varphi_{A,\alpha\beta}, \ldots)\,, \tag{12.3.18}$$

where φ_A is the complete set of dynamical variables (including the spacetime metric) enumerated by the collective index A. Dots indicate that the Lagrangian density Λ may depend also on higher derivatives of φ_A. The integration is performed over the part Ω of the spacetime, and dx is the element of coordinate volume. Consider the infinitesimal transformations

$$x^\mu \to x^\mu + \xi^\mu(x)\,, \quad \varphi_A(x) \to \varphi_A(x) + \delta\varphi_A(x)\,. \tag{12.3.19}$$

The functional W remains invariant under transformations (12.3.19) for an arbitrary volume Ω if and only if the following relation is valid

$$\frac{\delta\Lambda}{\delta\varphi_A}\delta\varphi_A + \frac{\partial \mathcal{J}^\beta}{\partial x^\beta} = 0\,, \tag{12.3.20}$$

where

$$\frac{\delta\Lambda}{\delta\varphi_A} = \frac{\partial\Lambda}{\partial\varphi_A} - \frac{\partial}{\partial x^\alpha}\left(\frac{\partial\Lambda}{\partial\varphi_{A,\alpha}}\right) + \frac{\partial^2}{\partial x^\alpha \partial x^\beta}\left(\frac{\partial\Lambda}{\partial\varphi_{A,\alpha\beta}}\right) + \ldots\,, \tag{12.3.21}$$

$$\mathcal{J}^\alpha = \Lambda\,\xi^\alpha + \frac{\partial\Lambda}{\partial\varphi_{A,\alpha}}\,\delta\varphi_A + \frac{\partial\Lambda}{\partial\varphi_{A,\alpha\beta}}\,\frac{\partial\,\delta\varphi_A}{\partial x^\beta} - \frac{\partial}{\partial x^\beta}\left(\frac{\partial\Lambda}{\partial\varphi_{A,\alpha\beta}}\right)\delta\varphi_A + \ldots\,. \tag{12.3.22}$$

Here (...) denotes additional terms that are present in the case when the Lagrangian density Λ depends on higher than second derivatives of φ_A. If the equations of motion $\delta\Lambda/\delta\varphi_A = 0$ are satisfied, the Noether current density $\mathcal{J}^\alpha$ is conserved: $\mathcal{J}^\alpha_{,\alpha} = 0$.

Consider a covariant action, i.e., an action invariant under coordinate transformations $x^\mu \rightarrow x^\mu + \xi^\mu(x)$. In this case, the transformation of the field induced by coordinate transformations is $\delta\varphi_A = \mathcal{L}_\xi\varphi_A$, where $\mathcal{L}_\xi$ is the Lie derivative in the direction of ξ^α. We define the Noether current density corresponding to this transformation as $\mathcal{J}^\alpha(\varphi,\xi)$. The conservation of $\mathcal{J}^\alpha(\varphi,\xi)$ implies that (at least locally)

$$\mathcal{J}^\alpha(\varphi,\xi) = \mathcal{N}^{\alpha\beta}{}_{,\beta}\,, \qquad \mathcal{N}^{\alpha\beta} = \mathcal{N}^{[\alpha\beta]}\,. \tag{12.3.23}$$

Since $\mathcal{J}^\alpha$ depends linearly on ξ^α and is conserved for an arbitrary ξ^α, it is possible to uniquely and globally define $\mathcal{N}^{\alpha\beta}(\varphi,\xi)$, which depends on no more than $(k-1)$ derivatives of ξ^α, where k denotes the highest derivative of any dynamical field occurring in the Lagrangian density Λ [Wald (1990)].

If the metric $g_{\alpha\beta}$ is among the dynamical variables (it is just this case which is interesting for our discussion), it is convenient to work not with densities Λ, $\mathcal{J}^\alpha$, and $\mathcal{N}^{\alpha\beta}$ but with ordinary scalars and tensors

$$L = (-g)^{-1/2}\Lambda\,, \qquad J^\alpha = (-g)^{-1/2}\mathcal{J}^\alpha\,, \qquad N^{\alpha\beta} = (-g)^{-1/2}\mathcal{N}^{\alpha\beta}\,. \tag{12.3.24}$$

In this notation equation (12.3.23) takes the form

$$J^\alpha = N^{\alpha\beta}{}_{;\beta}\,. \tag{12.3.25}$$

By integrating this relation over a surface Σ and using Stokes' theorem (A.87), one gets

$$\int_\Sigma J^\alpha\, d\Sigma_\alpha = \int_{\partial\Sigma} N^{\alpha\beta}\, d\sigma_{\alpha\beta}\,. \tag{12.3.26}$$

The integral of $N^{\alpha\beta}$ over a closed two-dimensional surface σ

$$N(\sigma,\xi) = \int_\sigma N^{\alpha\beta}\, d\sigma_{\alpha\beta}\,, \tag{12.3.27}$$

is called the *Noether charge* of σ relative to ξ^α.

Consider a stationary, axisymmetric, asymptotically flat spacetime. Define by $\chi^\alpha = \xi^\alpha_{(t)} + \Omega^H\,\xi^\alpha_{(\phi)}$ the Killing vector which is the generator of the Killing horizon. Let us apply (12.3.26) for $\xi^\alpha = \chi^\alpha$ and choose for Σ a three-hypersurface extending from asymptotic infinity to the bifurcation surface. Upon considering a variation of the fields away from the background solution, Wald (1993) found an identity relating surface term at infinity to another on the horizon, in the form expected for the first law

$$\delta N(\sigma_H,\chi) = \delta M - \Omega^H\,\delta J\,. \tag{12.3.28}$$

Here δM and δJ are the variations of the mass and angular momentum of the system measured at infinity, and $\delta N(\sigma_H, \chi)$ is the variation of the Noether charge of the horizon. In the process of variation the contravariant components of χ^α are chosen to be fixed.

It is easy to show that for vacuum black holes in the Einstein theory relation (12.3.28) coincides with (12.2.21) (with $\delta q = 0$). In fact, for the Einstein theory one has [Iyer and Wald (1994)]

$$L = \frac{1}{8\pi} R, \tag{12.3.29}$$

$$J^\alpha = \frac{1}{8\pi} \left[\chi^{[\alpha;\beta]}{}_{;\beta} + G^\alpha{}_\beta \, \chi^\beta \right] . \tag{12.3.30}$$

Here $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R$ is the Einstein tensor, which vanishes for vacuum spaces. Thus, we have

$$N^{\alpha\beta} = \frac{1}{8\pi} \chi^{[\alpha;\beta]} . \tag{12.3.31}$$

Relation (12.2.13) shows that

$$N(\sigma_H, \chi) = \frac{1}{8\pi} \int_\sigma d\sigma_{\alpha\beta} \, \chi^{[\alpha;\beta]} = \frac{1}{8\pi} \kappa \mathcal{A} . \tag{12.3.32}$$

Since $\delta\chi^\alpha = 0$ everywhere, and $\chi^\alpha = 0$ on the two-surface of bifurcation of horizons, one has $\delta(\chi^\alpha{}_{;\beta}) = 0$ and $\delta\kappa = 0$. Hence, (12.3.28) takes the form

$$\frac{\kappa}{2\pi} dS = dM - \Omega^H \, dJ , \tag{12.3.33}$$

where $S = \mathcal{A}/4$.

The remarkable fact is that relation (12.3.28) is valid for an arbitrary covariant theory, and hence it allows a natural generalization of the definition of the black hole entropy. If a theory does not contain higher than second derivatives of the dynamical variables, the Noether charge $N(\sigma, \chi)$ is a linear function of χ^α and its first derivative. On the bifurcation sphere $\chi^\alpha = 0$, and hence $N^{\alpha\beta} \sim \chi^{\alpha;\beta}$. The value of $N(\sigma, \chi)$ depends on the normalization of χ. One can remove this dependence by dividing $N(\sigma, \chi)$ by the surface gravity κ. The following quantity, which depends only on the geometry and physical fields near the horizon was proposed by Wald (1993) as the natural generalization of the black hole entropy :

$$S = \frac{2\pi}{\kappa} N(\sigma, \chi) = \frac{2\pi}{\kappa} \int_\sigma N^{\alpha\beta} \, d\sigma_{\alpha\beta} . \tag{12.3.34}$$

In the Einstein theory this expression evidently reproduces the Bekenstein-Hawking entropy of the black hole. Relation (12.3.33) (the first law) remains valid in a more general theory with the entropy of a black hole defined by (12.3.34). If a theory contains higher than the second derivatives, then $N(\sigma, \chi)$ formally contains higher than first order derivatives of χ. Using (A.30), these higher derivatives can always be eliminated. Expression (12.3.34) defines the entropy of a black hole in this case provided the corresponding procedure has been performed before the substitution of $N^{\alpha\beta}$ into the relation.

The advantages of Wald's definition of black hole entropy are:

1. One works directly with an arbitrary Lagrangian, and there is no need to construct the corresponding Hamiltonian, as in [Sudarsky and Wald (1992), Brown and York (1993a,b), Brown (1995)].
2. Possible ambiguities in the definition of the Lagrangian (adding a total divergence) and Noether charge do not affect the value of the entropy.
3. The method does not require "Euclideanization" of the black hole metric, the procedure which is often used in other approaches. The comparison of the Noether charge and Euclidean methods can be found in [Iyer and Wald (1995)].
4. Though in the original definition of the black hole entropy the integration is performed over the bifurcation surface, the entropy in fact can be calculated by (12.3.34) with the integral taken over an arbitrary slice of the horizon of a stationary black hole [Jacobson, Kang, and Myers (1994)].
5. Though the expression (12.3.34) was derived for a stationary black hole it allows natural generalizations to the time-dependent case.

Explicit expressions for the black hole entropy in some of the non-Einsteinian theories can be found in [Iyer and Wald (1994), Jacobson, Gungwon, and Myers (1994)].

12.4 Black Hole as a Thermodynamic System

12.4.1 Equilibrium of a black hole and thermal radiation in a box

Let us look more closely at the case of a non-rotating uncharged black hole surrounded by black body radiation at a temperature θ_{rad}.[11] We have already mentioned that if this temperature coincides with that of the black hole, θ, we deal with equilibrium in which the accretion of radiation onto the black hole is balanced for by the Hawking radiation of the hole.[12]

[11] Moss (1985) pointed out that when a black hole undergoes quantum evaporation, a "bubble" of the new phase may form around it as a result of the phase transition in the surrounding space. Under certain conditions, the particles emitted by the black hole are reflected at the bubble walls, accumulate inside it, and produce a high-temperature medium outside the black hole [see also Hiscock (1987)].

[12] Gibbons and Perry (1976, 1978) demonstrated that this conclusion remains unaltered when the interaction among thermal quanta is taken into account.

It is not difficult to verify that if the size of the thermal bath is infinitely large, then the equilibrium is unstable. Indeed, let a black hole absorb, in response to a random fluctuation, less energy during some time interval than it has radiated away. In this case, its mass slightly diminishes, but its temperature θ increases, which results in a further increase of the radiation rate and a further reduction of the black hole mass. On the other hand, a fluctuation increasing the mass of the black hole reduces its temperature and also the rate of the Hawking radiation. In this case, the accretion of radiation onto the black hole becomes the dominating process. In other words, two situations are possible if there is a sufficient amount of radiation around the black hole: either complete evaporation of the black hole or unlimited growth of its size.

This characteristic of the behavior of non-rotating, uncharged black holes is directly traced to their heat capacity,

$$C = \theta \left(\frac{\partial S^H}{\partial \theta} \right)_{J=Q=0} , \tag{12.4.1}$$

being negative ($C = -8\pi M^2$). The negative heat capacity signifies that a reduction in the system's energy increases its temperature ($dE = C\, d\theta$). This property is characteristic for systems with long-range attractive forces, and thus for systems with gravitational self-action. It is readily shown (using, e.g., the virial theorem) that a reduction in the dimensions of such a system, resulting in a reduction of its potential and total energies, produces a simultaneous increase in the kinetic energy of its particles (the temperature of the body).

The situation is quite different if a black hole is part of a thermodynamical system which has a finite size. For example, if a black hole is placed in a reservoir of radiation and the total energy of the system is fixed, a stable equilibrium configuration can exist [Davies (1977), Gibbons and Perry (1978)]. Let the temperature of radiation be θ_{rad}; then its energy E^m and entropy S^m are

$$E^m = \sigma\, V\, \theta_{\text{rad}}^4, \qquad S^m = \frac{4}{3} \sigma\, V\, \theta_{\text{rad}}^3 , \tag{12.4.2}$$

$$\sigma = \frac{\pi^2}{30} \left(h_b + \frac{7}{8} h_f \right) . \tag{12.4.3}$$

Here V is the volume of the reservoir, and h_b and h_f are the number of polarization of massless bosons and fermions, respectively. For simplicity, we treat only massless fields. The condition of stable equilibrium in a system consisting of a reservoir of radiation with a black hole inside it is that the generalized entropy,

$$\tilde{S} = S^H + S^m = 4\pi M^2 + \frac{4}{3} \sigma\, V\, \theta_{\text{rad}}^3 , \tag{12.4.4}$$

is maximal for a fixed value of total energy,

$$E = M + \sigma\, V\, \theta_{\text{rad}}^4 . \tag{12.4.5}$$

Using the relation $dM/d\theta_{\rm rad} = -4\,\sigma\,V\,\theta_{\rm rad}^3$, implied by (12.4.5), we can show that the extremum of S is achieved if $\theta_{\rm rad} = \theta = 1/(8\pi M)$ which means that the radiation and black hole temperatures coincide. This equilibrium is stable if $d^2S/d\theta^2 < 0$; this is equivalent to the inequality

$$E^m < \frac{1}{4} M. \tag{12.4.6}$$

This equilibrium is stabilized by the following mechanism. Assume, as before, that a black hole absorbs more energy in a fluctuation than it has radiated away. Its temperature, and hence its rate of radiation, are then reduced. However, the rate of accretion onto the black hole has also decreased because of the diminished amount of radiation outside it. If condition (12.4.6) is satisfied, the second of these effects is more important so that the black hole returns to the original state after its mass diminishes owing to the excess of radiation over accretion. The situation with fluctuations due to the reduction of black hole mass is analogous.

Condition (12.4.6) can be reformulated as a restriction on the volume V. We define

$$V_0 = \frac{2^{20}\,\pi^4}{5^5\,\sigma}\,E_0^5\,. \tag{12.4.7}$$

If $V > V_0$, then for the given energy E_0 the most probable state is thermal radiation without a black hole. In the case of the reversed inequality ($V < V_0$), the system contains a black hole surrounded by thermal radiation at a temperature $\theta_{\rm rad} = \theta$ [Hawking (1976a)]. The process of black hole formation at $V = V_0$ when V decreases resembles a first-order phase transition and is similar to the process of nucleation of a liquid droplet in the course of vapor cooling.

We can also consider a system with a fixed volume V_0. Then for $E > E_0$ the most probable state is a black hole in thermal equilibrium with the radiation. For $E < E_0$ the black hole is absent. For this problem the energy E plays the role of the control parameter and the total entropy is the corresponding thermodynamical potential. Different states of a system in the volume V with fixed energy E form a *microcanonical ensemble.*

The system (a black hole in the box) possesses a simple scaling property. Namely, if we make the transformation $M \to \lambda M$, we can maintain consistency by requiring $V \to \lambda^5 V$ and $\theta_{\rm rad} \to \lambda^{-1}\theta_{\rm rad}$. The transformation of the temperature $\theta_{\rm rad}$ of the radiation is consistent with the transformation of the black hole temperature $\theta \sim M^{-1}$. Under these scaling transformations, both energy (12.4.5) and generalized entropy (12.4.4) transform homogeneously: $E \to \lambda E$ and $\tilde{S} \to \lambda^2 \tilde{S}$. Qualitative properties of the system (such as the stability of the equilibrium) can depend only on the scale invariant combination of the external parameters VE^{-5} [Schumacher, Miller, and Zurek (1992)]. Condition (12.4.7) is in agreement with this observation. One of the important consequences of the scaling is that the size of the system $L \sim V^{1/3}$ containing in the equilibrium a black hole can be made much larger than the gravitational radius $r_g = 2M$. Indeed, scaling $L/r_g \to \lambda^{2/3}L/r_g$ allows one to make this ratio arbitrarily large.

Another possible type of gedanken experiment with a black hole is the case when the external control parameter is the temperature at the boundary (the *canonical ensemble*). The corresponding thermodynamical potential is the free energy $F = F(\theta, V) = E - \theta\, S$. In Section 12.5 we shall show that a stable thermal equilibrium is possible only if the size of the cavity is comparable with the size of the black hole. For example, for a non-rotating black hole the stability condition is $L < 3M$, where L is the cavity radius. One can also arrive at this conclusion by using the following simple argument. Consider a cavity of small size so that it is possible to neglect the contribution of the radiation to the energy and entropy. If r_B is the radius of the cavity, then the black hole mass M and the temperature on the boundary θ are related by

$$\theta^{-1} = 8\pi M \left(1 - \frac{2M}{r_B}\right)^{1/2} . \tag{12.4.8}$$

It follows from this relation that

$$-\frac{1}{6\pi^2\theta^3 M}\,\frac{d\theta}{dM} = 1 - \frac{3M}{r_B}, \tag{12.4.9}$$

and $d\theta/dM$ changes sign at $r_B = 3M$. Since the entropy of the black hole is a monotonically increasing function of its mass, relation (12.4.1) shows that its heat capacity is positive, and the equilibrium of the black hole in the cavity is stable if $r_B < 3M$.[13] A simple scaling law is not valid in this case since the action of the gravitational field on the radiation cannot be neglected.

Different aspects of the thermodynamics of black holes in a cavity, and in particular problems connected with rotation of the black hole are discussed by Kaburaki and Okamoto (1991), Schumacher, Miller, and Zurek (1992), Kaburaki, Okamoto, and Katz (1993), Parentani, Katz, and Okamoto (1995).

12.4.2 Heat capacity. Thermal fluctuations

The heat capacity of an isolated charged rotating black hole, calculated with a formula similar to (12.4.1), is

$$C_{J,Q} = \theta\,\frac{\partial S}{\partial \theta}\bigg|_{J,Q} = \frac{MTS^3}{\pi J^2 + (\pi/4)Q^4 - T^2S^3} . \tag{12.4.10}$$

[13]The question whether it is possible to contruct a "real" physical system which can play the role of such a thermostat is not trivial. In the usual thermodynamics, a system which is considered as the thermostat has arbitrarily large heat capacity, and the emission and absorption of energy does not change its temperature. Page (1993b) argued that in systems containing a black hole, a finite size of the cavity may create problems.

If we write $J^2 = \alpha M^4$ and $Q^2 = \beta M^2$, it is readily shown that C undergoes a sign reversal for values of the parameters α and β that satisfy the relation [Davies (1977)]

$$\alpha^2 + 6\alpha + 4\beta - 3 = 0. \tag{12.4.11}$$

At this point $C_{J,Q}$ becomes infinite. This property of heat capacity is typical, to a certain extent, of the properties of heat capacity of ordinary matter in second-order phase transitions. For an uncharged ($\beta = 0$) black hole $\alpha = 2\sqrt{3} - 3$, while for a non-rotating (α=0) black hole $\beta = 3/4$. The transition from negative to positive values in these (and other) heat capacities can be characterized by critical exponents obeying the scaling laws [Lousto (1993, 1996)]. It should be noted that the creation of charged particles and the quantum analogue of superradiance make it difficult to treat rigorously the physical peculiarities caused by the above-described behavior of heat capacity $C_{J,Q}$ [Davies (1977), Hut (1977), Sokolowski and Mazur (1980)].

Local deviations of macroscopic observables from their average equilibrium value are known as fluctuations. For instance, in a canonical ensemble of total volume V, the mean quadratic fluctuation $\overline{(\Delta E)^2}$ of the energy is

$$\overline{(\Delta E)^2} = T^2 \left.\frac{\partial E}{\partial T}\right|_V = T^2\, C_V\,. \tag{12.4.12}$$

Similarly, for a fixed energy of the system (in the microcanonical ensemble) the local temperature fluctuates in accordence with

$$\overline{(\Delta T)^2} = T^{-2} \left.\frac{\partial T}{\partial E}\right|_V = (T^2\, C_V)^{-1}\,. \tag{12.4.13}$$

However, gravitation is a long-range force. That is why for gravitating systems, where one cannot treat the remaining part of the system as a passive reservoir, the relations (12.4.12) and (12.4.13) should be modified. The discussion of thermal fluctuations for a black hole in a cavity can be found in [Parentani, Katz, and Okamoto (1995)].

The thermodynamic analogy in black hole physics, which we have discussed in this chapter, has so far been confined to *equilibrium thermodynamics* (i.e., to considering equilibrium states and various relations between the characteristics of these states). Actuelly, this analogy is broader. It can also be traced in *nonequilibrium* thermodynamics which describes irreversible transitions of a system between its states and the processes taking place when the system passes to the state of thermodynamic equilibrium [Damour (1979)]. A general analysis of the problems of nonequilibrium thermodynamics of black holes can be found in Sciama (1981). On the change of a black hole entropy in nonequilibrium processes, see Hawking and Hartle (1972), Bekenstein (1974), Carter (1979).

Candelas and Sciama (1997) showed that black hole radiation obeys the principles of irreversible thermodynamics in the form of a *fluctuation-dissipation theorem*. The dissipation is connected with the absorption of the "ordered" energy by a black hole and its subsequent re-radiation in the Hawking process.

12.5 Euclidean Approach to Black Hole Thermodynamics

12.5.1 Euclidean formulation

In quantum statistical mechanics the temperature is naturally connected with the property of periodicity of thermal Green's functions (and other quantities decribing the system) in imaginary time, the period being the inverse temperature of the system. This makes it useful to consider an Euclidean space approach. By making Wick rotation of the time $t \to -i\tau$ in a static spacetime, one gets a metric with a Euclidean signature. The usual quantum field theory at zero temperature and also quantum field statistical theory can be obtained by analytic continuation from the Euclidean version of the theory. Nonzero temperature arises if the topology of the Euclidean time is not R^1 but S^1. This idea can be easily realised in the Feynmann path-integral approach by specifying that the integration is to be taken over configurations which are periodic (for bosons) and antiperiodic (for fermions) in Euclidean time.

This approach, which relates thermal properties to the geometry and topology of a spacetime, is very attractive for quantum gravity and especially for considering thermodynamical systems including black holes. Gibbons and Hawking (1977) and Hawking (1979) were the first to describe the thermodynamics of black holes by exploiting the properties of the Euclidean action and its relation to the partition function of a canonical ensemble [Feynman and Hibbs (1965)]. For this purpose, they used the path integral approach.

In the path integral approach, the amplitude of the propagation of a quantized field Φ from some initial configuration $|\Phi_1, t_1\rangle$ to some final configuration $|\Phi_2, t_2\rangle$ is given by the path integral

$$\langle \Phi_2, t_2 | \Phi_1, t_1 \rangle = \int \mathcal{D}[\Phi] \exp(iW[\Phi]) , \tag{12.5.1}$$

where W is a classical action and the integration is taken over all the configurations of the field Φ with the initial value Φ_1 at t_1 and the final value Φ_2 at t_2. On the other hand, one has

$$\langle \Phi_2, t_2 | \Phi_1, t_1 \rangle = \langle \Phi_1 | \exp(-iH(t_2 - t_1)) | \Phi_2 \rangle, \tag{12.5.2}$$

where H is the Hamiltonian. If one puts $t_2 - t_1 = -i\beta$ and $\Phi_2 = \Phi_1$ and sums over Φ_1, then one gets

$$Z \equiv \mathrm{Tr} \exp(-\beta H) = \int \mathcal{D}[\Phi] \exp(iW[\Phi]). \tag{12.5.3}$$

The path integral is taken over all the configurations of the field Φ which are periodic (or antiperiodic) in the imagine time with the period β. The quantity Z is the partition function for the canonical emsemble for the field Φ at temperature $\theta = \beta^{-1}$.

The generalization to the case of the gravitational field is straightforward. One needs only to include in the path integral the integration over metrics and to take into account in (12.5.3) the Faddeev-Popov ghost contribution. The dominant contribution to the path integral will come from the metric g and matter fields φ which are near to the extremum $\Phi_0 = \{g_0, \varphi_0\}$ of the classical action. In this semi-classical approximation, one has

$$\ln Z \approx iW[g_0, \varphi_0] . \tag{12.5.4}$$

For a Euclidean black hole which is an extremum of the (Euclidean) gravitational action the inverse themperature β is related to the mass M as $\beta = 8\pi M$, and $iW = -W_E = 4\pi M^2$. The Euclidean action is related to the free energy F as $W_E = \beta F$, and the corresponding entropy is [Gibbons and Hawking (1977)]

$$S = -\frac{\partial F}{\partial \theta} \equiv \beta^2 \frac{\partial F}{\partial \beta} = 4\pi M^2 .$$

That is, it coincides with the Bekenstein-Hawking entropy (12.1.6) of the black hole.

The Euclidean approach gives the correct answer. Nevertheless, the general validity of this method of calculation of the thermodynamical characteristics of a black hole remains unclear. The following two obstacles apparently limit the success of the idea:

1. The heat capacity of an isolated non-rotating black hole is negative and such a black hole cannot be in a stable equilibrium with the heat bath. This makes the canonical ensemble of an infinitely large system including a black hole ill-defined.

2. If one assumes the usual interpretation of the entropy, then the density of states of a system containing a black hole must grow as $\exp S_H = \exp(4\pi M^2)$ so that the sum defining the partition function $\mathrm{Tr}\exp(-\beta H)$ would not converge.

12.5.2 Boundary conditions

Both difficulties were overcome by York and collaborators [York (1986), Whiting and York (1988), Braden, Whiting, and York (1987), Braden *et al.* (1990), Brown *et al.* (1990), Brown, Martinez, and York (1991)]. They stressed the importance of the finite size of the system and propose to consider a black hole located in a cavity with the boundary data fixed on the boundary of the cavity. The boundary conditions are to be chosen in a proper way in order:

1. To guarantee the existence and stability of the equilibrium state
2. To be essentially the same canonical data which are needed for the thermodynamical description as well as for the dynamical Euclidean field theory description. Different boundary conditions single out different thermodynamic ensembles.

For simplicity, we restrict ourselves to consideration of spherically symmetric black holes in thermal equilibrium with thermal radiation. (Delicacies which arise in the case of a rotating black hole are discussed in [Brown, Martinez, and York (1991)].) In the case of a canonical ensemble, the thermodynamic data that define the ensemble are the size and temperature. The static gravitating systems are homogeneous in time but not in space. As a result of the redshift effect, the local temperature measured by static observers at different points is different. That is why, when speaking about the temperature, we need to specify where it is measured. In our case, that is the temperature at the boundary. We assume that the boundary is also spherically symmetric, and the temperature over it is constant. We denote the inverse value of the temperature at the boundary by β. In order to fix the size of the system, we fix the area $\mathcal{A}_B = 4\pi r_B^2$ of the boundary. The topology of the Euclidean black hole is $R^2 \times S^2$, and the topology of the boundary is $S^1 \times S^2$. The above boundary data have a well-defined geometric meaning. The inverse temperature β is the proper length of S^1, and $\mathcal{A}_B$ is the area of S^2. This means that the boundary data appear as the boundary values of dynamical variables; namely, components of the metric tensor. The quantities E (energy) and p_B (pressure), thermodynamically conjugate to β and $\mathcal{A}_B$, are at the same time canonically conjugate to the 3-metric with respect to the radial foliation. The partition functions for the canonical $(\beta, \mathcal{A}_B)$ and the microcanonical $(E, \mathcal{A}_B)$ ensembles (as well as to their conjugates with $\mathcal{A}_B \to p_B$) can be directly described as path integrals [Braden *et al.* (1990)]. The grand canonical ensemble differs from the canonical one only by the way in which the matter variables are described.

12.5.3 Calculation of the Euclidean action

We describe now the main steps of the derivation of the thermodynamical properties of charged non-rotating black holes in the framework of the grand canonical ensemble approach following the paper by Braden *et al.* (1990). The Euclidean metric which is the result of the Wick rotation of a static spherically symmetric metric can be written in the form

$$ds_E^2 = V^2\, d\tau^2 + N^2\, dx^2 + r^2\, d\omega^2 . \tag{12.5.5}$$

Here the radial coordinate $x \in [0, 1]$, the Euclidean time τ has period 2π. The metric coefficients V, N and r are functions of x, and $d\omega^2$ is the line element on the unit sphere. We write

$$V_B = V(1), \quad r_B = r(1), \quad \mathcal{A}_B = 4\pi r_B^2, \quad r_H = r(0). \tag{12.5.6}$$

The inverse temperature $\beta \equiv \theta^{-1}(r_B)$ is related to the proper length of the round S^1 of the boundary, and it is of the form

$$\beta = 2\pi V_B . \tag{12.5.7}$$

The regularity conditions at the Euclidean horizon ($x = 0$) give

$$V(0) = 0, \quad \left[\frac{V'}{N}\right]_{x=0} = 1, \quad \left[\frac{r'}{N}\right]_{x=0} = 0. \tag{12.5.8}$$

Here and later $(\ )' = d(\ )/dx$.

By suitable gauge transformations, the vector-potential of the electromagnetic field can be reduced to the form

$$A_\mu \, dx^\mu = A_\tau(x) \, d\tau \,, \qquad A_\tau(0) = 0\,. \tag{12.5.9}$$

The latter condition guarantees the regularity of the potential one-form at the Euclidean horizon. The Wick rotation $t = -i\,\tau$ generates the transformation $A_\tau(x) = -iA_t(x)$ of the zero-component of the potential one-form. It means that for a charged black hole solution A_τ is pure imaginary. We write

$$A_\tau = -i\,\Phi(x)\,. \tag{12.5.10}$$

We fix the boundary condition for the electromagnetic field by putting

$$\Phi(1) \equiv V_B \, \Phi_B = \frac{\beta \Phi_B}{2\pi}\,. \tag{12.5.11}$$

The factor V_B which appears in the above formula allows the more direct interpretation of Φ_B as the electric potential measured by an observer located at the boundary. (Φ_B coincides with the value of A_μ in the proper orthonormal frame.) We fix the boundary value for the electromagnetic field by fixing Φ_B. The set $\{\beta, A_B, \Phi_B\}$ forms the necessary data for the grand canonical ensemble.

Denote by $W_E = -iW$ the Euclidean action corresponding to the Lorentzian action W. The Euclidean action for the gravitational and electromagnetic fields is

$$W_E = -\frac{1}{16\pi}\int_M R_E \, g_E^{1/2} \, d^4x + \frac{1}{16\pi}\int_M d^4x \, g_E^{1/2} \, F_{\mu\nu} \, F^{\mu\nu}$$
$$+ \frac{1}{8\pi}\int_{\partial M} (K_E - K_E^0) \, h_E^{1/2} \, d^3y\,, \tag{12.5.12}$$

where the subscript E indicates that the corresponding quantity is to be calculated for the Euclidean metric. Here h_{ij} is the metric induced on ∂M; K is the mean extrinsic curvature on ∂M, and K_0 is chosen to make the action with the given boundary ∂M vanish in flat space.

The partition function for the grand canonical ensemble under consideration can be approximated as

$$Z \approx \exp(-I_*)\,, \tag{12.5.13}$$

where I_* is the Euclidean action (12.5.12) for the interior of the cavity evaluated for metric (12.5.5) and the electromagnetic field (12.5.10) with fixed values of the boundary data $\{\beta, \mathcal{A}_B, \Phi_B\}$. Direct substitution gives

$$I_* = -\pi \int_0^1 dx \left[\frac{2r\, r'\, V' + V r'^2 + r^2\, \Phi'^2\, V^{-1}}{N} + NV\right] + 2\pi\, V(1)\, r(1) - \pi\, r^2(0)\,. \tag{12.5.14}$$

It is easy to verify that the variation of I_* with respect to N, V, r, and Φ gives the complete set of equations equivalent to the Einstein equations for spherically symmetric static spacetimes.

An important characteristic property of general relativity is that the energy on a spacelike hypersurface can be evaluated as a surface integral. Hence, in a static geometry the action can be expressed by a boundary term provided the Hamiltonian constraints ($G^\tau_\tau = 0$ and $F^{\tau\mu}_{\ ;\mu} = 0$) are satisfied. The constraints can be integrated explicitly. The electromagnetic field constraint has the solution

$$\Phi' = \frac{Q}{r^2}\left(V^2 N^2\right)', \tag{12.5.15}$$

where Q is a constant (charge of the black hole). The gravitational constraint has the solution

$$V^2\, r'^2 = 1 - \frac{C}{r} + \frac{Q^2}{r^2}\,. \tag{12.5.16}$$

The constant C can be obtained by using the regularity conditions (12.5.8), and it is of the form

$$C = r_H + \frac{Q^2}{r_H}\,. \tag{12.5.17}$$

The reduced action after substitution of the constraints solution into it takes the form [Whiting and York (1988), Braden *et al.* (1990)]

$$I_* = I_*(\beta, \Phi_B, r_B; Q, r_H) = \beta\left[r_B - (r_B - r_H)^{1/2}\left(r_B - \frac{Q^2}{r_H}\right)^{1/2} - \Phi_B\, Q\right] - \pi\, r_H^2\,. \tag{12.5.18}$$

This action is exact for smooth, static, spherically symmetric Euclidean geometries. The imposed constraints (12.5.15) and (12.5.16) can be used to specify two of the four arbitrary functions. The remarkable fact is that the reduced action I_* does not depend explicitly on the remaining two functions. It depends only on the charge Q and the radius r_H of the horizon (which is related to the mass M of the black hole by $r_H = M + (M^2 - Q^2)^{1/2}$) as parameters, and hence I_* is not a functional but a function. The stationary points of I_* with respect to these parameters are defined by the equations

$$\left[\frac{\partial I_*}{\partial Q}\right]_{\beta,\Phi_B,r_B} = 0\,, \qquad \left[\frac{\partial I_*}{\partial r_H}\right]_{\beta,\Phi_B,r_B} = 0\,. \tag{12.5.19}$$

These equations allows one to relate the charge Q and the radius of the horizon r_H with Φ_B, r_B and β

$$\Phi_b = \frac{Q}{r_H}\left(1-\frac{r_H}{r_B}\right)^{1/2}\left(1-\frac{Q^2}{r_H r_B}\right)^{-1/2}, \tag{12.5.20}$$

$$\beta = 4\pi r_H\left(1-\frac{Q^2}{r_H^2}\right)^{-1}\left(1-\frac{r_H}{r_B}\right)^{1/2}\left(1-\frac{Q^2}{r_H r_B}\right)^{1/2}. \tag{12.5.21}$$

Denote by I the reduced action I_* evaluated at the extremum point. Then the approximate value of $\ln Z$ (the *Massieu function*) is

$$\ln Z \approx -I. \tag{12.5.22}$$

12.5.4 Thermodynamical parameters

By using the obtained Massieu function I, one can easily get the mean values of different physical observables. For the mean value of the charge one has

$$\langle Q\rangle = \beta^{-1}\left[\frac{\partial I}{\partial \Phi_B}\right]_{\beta,r_B} = Q\,. \tag{12.5.23}$$

The average energy $\langle E\rangle$ is

$$\begin{aligned}\langle E\rangle &= \left[\frac{\partial I}{\partial \beta}\right]_{\Phi_B,r_B} - \Phi_B\,\beta^{-1}\left[\frac{\partial I}{\partial \Phi_B}\right]_{\beta,r_B} \\ &= r_B - r_B\left(1-\frac{r_H}{r_B}\right)^{1/2}\left(1-\frac{Q^2}{r_H r_B}\right)^{1/2}.\end{aligned} \tag{12.5.24}$$

And finally, the entropy S^H is

$$S^H = \beta\left[\frac{\partial I}{\partial \beta}\right]_{\Phi_B,r_B} - I = \pi\, r_H^2\,. \tag{12.5.25}$$

This expression reproduces the Bekenstein-Hawking entropy (12.1.6)

It should be stressed that the parameters $\{\beta, r_B, \Phi_B\}$ which enter the reduced action I are independent variables. A quite remarkable fact is that the same action I

evaluated for the Euclidean Schwarzschild metric (i.e., for $\beta = 4\pi r_H(1 - r_H/r_B)^{1/2}$) coincides in the limit $r_B \to \infty$ with the Bekenstein-Hawking entropy.

The thermal stability of a system is determined by its heat capacity. For an uncharged black hole in a cavity the heat capacity is

$$C_B = -\beta \left[\frac{\partial S_H}{\partial \beta} \right]_{r_B} = \frac{8\pi M^2 \left(1 - \dfrac{2M}{r_B}\right)}{\dfrac{3M}{r_B} - 1}. \tag{12.5.26}$$

The heat capacity is negative, and the system is unstable for $r_B > 3M$. For $2M < r_B < 3M$ the heat capacity is positive and the system is stable. This result shows that a properly chosen finite size of the system containing a black hole does really guarantee the thermal stability of the system. [For the stability conditions for the other ensembles and for ensembles containing a charged black hole see, Braden *et al.* (1990), Brown *et al.* (1990)].

The other problem mentioned above and connected with divergence of the density of states can be also solved by considering systems with a finite size [Braden, Whiting, and York (1987)]. The number of energy levels between E and $E + dE$ for the gravitational field in the black hole topological sector can be estimated as follows :

$$dN(E) = d\left[e^{S^H(E)}\right] = e^{S^H} \frac{dS^H}{dE} dE. \tag{12.5.27}$$

By using (12.5.24), we get for an uncharged black hole

$$E = r_B \left(1 - \sqrt{1 - \frac{r_H}{r_B}}\right) \tag{12.5.28}$$

and

$$S^H = S^H(E) = \pi E^2 \left(2 - \frac{E}{r_B}\right)^2. \tag{12.5.29}$$

The density of states

$$\frac{dN}{dE} = 4\pi E \left(1 - \frac{E}{r_B}\right)\left(2 - \frac{E}{r_B}\right) e^{S^H(E)} \tag{12.5.30}$$

grows with energy E and after reaching a maximum value drops steeply to zero at $E = r_B$. For this maximal value of the energy, the size of the black hole coincides with the size of the cavity r_B. For $E > r_B$ the density of states vanishes ($dN/dE = 0$).

12.5.5 Conical singularity method

In the conclusion of this section, we describe a modification of the Euclidean approach to the black hole thermodynamics known as the *conical singularity method.*

To describe the method, let us again consider the static spherically symmetric Euclidean metric (12.5.5):

$$ds_E^2 = V^2\, d\tau^2 + N^2\, dx^2 + r^2\, d\omega^2 . \tag{12.5.31}$$

As before, we impose conditions (12.5.6) and (12.5.8), but we do not require the period of τ to be 2π. Instead of this we impose the periodicity $0 \leq \tau \leq \alpha$, where α is an arbitrary positive parameter. Expression (12.5.7) for the inverse temperature $\beta = \theta^{-1}(r_B)$ (defined as the proper length of the circumference S^1 of the boundary) is also modified to

$$\beta = \alpha\, V_B . \tag{12.5.32}$$

Near the horizon $x = 0$ the metric (12.5.31) has the following asymptotic form :

$$ds_E^2 = \left(\frac{\alpha}{2\pi}\right)^2 l^2\, d\tilde{\tau}^2 + dl^2 + r_H^2\, d\omega^2 . \tag{12.5.33}$$

Here $\tilde{\tau} = 2\pi\tau/\alpha$, $0 \leq \tilde{\tau} \leq 2\pi$, $r_H = 2M$, and $l = \int_0^x dx' N(x')$ is the proper distance from the horizon. This metric is a direct sum of two two-dimensional metrics, one being locally flat, and the other being a metric on a sphere of radius r_H. If α differs from 2π, the locally flat two-dimensional metric is not globally flat and has a cone-like singularity. Thus, metric (12.5.33) describes a space $\mathcal{M}_\alpha$ which near the horizon has a structure $\mathcal{C}_\alpha \times S^2$, where $\mathcal{C}_\alpha$ is a cone with the deficit of the polar angle $2\pi - \alpha$. Therefore, if $\alpha \neq 2\pi$, the space $\mathcal{M}_\alpha$ is regular everywhere except at the surface $x = 0$, where it has a non-differentiable singularity. As a result, the scalar curvature and components of the Riemann tensor at the surface $x = 0$ can be defined only as distributions [Sokolov and Starobinsky (1977), Geroch and Traschen (1987), Hayward and Louko (1990), Bañados, Teitelboim, and Zanelli (1994), Fursaev and Solodukhin (1995)].

It is possible to calculate integral geometric characteristics on manifolds with conical singularities by smearing the singularity. For instance, for the space (12.5.31) the function $V(x)$ can be replaced by a function $\tilde{V}(x,a)$ where a is a "regularization" parameter. The regularized function $\tilde{V}(x,a)$ coincides with $V(x)$ for $x \gg a$ and obeys the condition

$$\left(\frac{\alpha}{2\pi}\right)\left[\frac{\tilde{V}'}{N}\right]_{x=0} = 1 . \tag{12.5.34}$$

Because of the last condition, the regularized space has the structure $R^2 \times S^2$, and its Riemann tensor is well defined at $x = 0$. By using this smearing, the integral of the scalar curvature R over the space with a conical singularity can be defined as the limit

$$\int_{\mathcal{M}_\alpha} R\, g_E^{1/2}\, d^4x = \lim_{a\to 0} \int \tilde{R}\, \tilde{g}_E^{1/2}\, d^4x\,, \tag{12.5.35}$$

where the quantities $\tilde{R}$ and $\tilde{g}_E^{1/2}$ are calculated for the regularized space. In this way one obtains (for details, see Fursaev and Solodukhin (1995))

$$\int_{\mathcal{M}_\alpha} R\, g_E^{1/2}\, d^4x = 2\,(2\pi - \alpha)\mathcal{A} + \int_{\mathcal{M}_\alpha - S^2} R\, g_E^{1/2}\, d^4x\,. \tag{12.5.36}$$

Here $\mathcal{A} = 4\pi\, r_H^2$ and integration in the last term is performed over the regular domain of $\mathcal{M}_\alpha$.

Using equation (12.5.36), one gets for the Euclidean action (12.5.12) on the singular space $\mathcal{M}_\alpha$ the following expression

$$W_E^{\mathrm{CS}}[\mathcal{M}_\alpha] = -\frac{1}{8\pi}(2\pi - \alpha)\mathcal{A} + W_E[\mathcal{M}_\alpha - S^2]\,. \tag{12.5.37}$$

The volume part of the action $W_E[\mathcal{M}_\alpha - S^2]$ calculated for metric (12.5.31) is $\alpha I_*/2\pi$, where I_* is defined by equation (12.5.14). After imposing the constraints (12.5.15) and (12.5.16), the conical singularity action (12.5.37) takes the form

$$I_*^{\mathrm{CS}} = \left(\frac{\alpha}{2\pi} - 1\right) r_H^2 + \frac{\alpha}{2\pi}\, I_*\left(\frac{2\pi}{\alpha}\beta, \Phi_B, r_B; Q, r_H\right). \tag{12.5.38}$$

The remarkable fact is that when expressed in terms of the boundary parameters (β, Φ_B, r_B), the action I_*^{CS} has a form identical to (12.5.18) :

$$I_*^{\mathrm{CS}}(\beta, \Phi_B, r_B; Q, r_H) = \beta\left[r_B - (r_B - r_H)^{1/2}\left(r_B - \frac{Q^2}{r_H}\right)^{1/2} - \Phi_B\, Q\right] - \pi\, r_H^2\,. \tag{12.5.39}$$

In other words, the form of the functional I_* calculated for a regular manifold with a smeared singularity remains invariant in the limit when the smearing is removed. Besides the boundary conditions, the function I_*^{CS} depends also on two additional parameters r_H and Q. In particular, if I_*^{CS} is calculated for a static solution of Einstein-Maxwell equations, r_H is the radius of the event horizon, and Q is the charge of the black hole. It should be emphasized that the parameters r_H and Q which specify the solution and enter the arguments of I_*^{CS} can be chosen arbitrarily. The black hole entropy is defined as

$$S^H = \beta\frac{\partial I_*^{\mathrm{CS}}}{\partial\beta} - I_*^{\mathrm{CS}} = \pi\, r_H^2\,. \tag{12.5.40}$$

To derive the entropy in the conical-singularity method, one first makes off-shell calculations and only afterward takes the on-shell limit ($\alpha = 2\pi$). The latter means that r_H and Q in (12.5.40) must be expressed as functions of the boundary parameters (β, Φ_B, r_B) by solving equations (12.5.20) and (12.5.21).

The conical-singularity approach allows one to consider a wider class of metrics. This gives more flexibility and sometimes simplifies calculations of thermodynamical parameters of black holes.

12.6 Statistical-Mechanical Foundation of Black Hole Thermodynamics

12.6.1 Introduction

The relationship of the thermal properties of black holes to the loss of information about the spacetime region inside the hole is in accordance with the general informational approach to thermodynamics that was formulated by Szillard (1929) and later elaborated on by a number of physicists and mathematicians [see, e.g., the monographs of Brillouin (1956) and Poplavsky (1981)]. This approach essentially consists of postulating a direct relationship between the lack of information about the physical system and its entropy.

The information on the state of the collapsed matter is "cut off" by a strong gravitational force. The black hole "forgets" its history, remembering only the "macroscopic" characteristics: mass, charge, and angular momentum. Correspondingly, the entropy S^H of a black hole is a measure of the amount of information lost as a result of collapse. The expected number of distinct ("microscopic") states of a black hole with prescribed parameters M, J, Q is proportional to $\exp[S(M, J, Q,)]$ [Bekenstein (1973b, 1980), Hawking (1976a), Wald (1979b)]. The direct calculation from first principles of this number of states is a very complicated problem.

In order to explain the problem, let us compare two identical gedanken experiments. One of them is made with a *black body*, and the other with a *black hole*. Consider a spherical box filled with radiation. We assume that it has the same temperature as the heated boundary. In the first box we place a black body, and in the second one we place a black hole. We assume that the temperature of the black body and the black hole coincide with the temperature of the radiation so that we have thermal equilibrium. Now let us change the temperature of the boundaries in both systems in the same manner. In response to this change the free energies also change.

We know well what happens with the black body. Being formed of atoms, it has an internal structure and internal degrees of freedom. In the process of heating the system, a part of the added energy is absorbed by the black body, and this portion of energy increases the internal energy of the body. This results in an additional excitation of the internal degrees of freedom. Statistical mechanics relates the observable thermodynamical parameters of the black body with its structure. In particular, it relates the entropy of the black body to the number of its internal degrees of freedom.

Qualitatively, the behavior of a black hole in the thermal bath is similar. During heating a part of the energy is absorbed by the black hole. However, in the case of the black hole we are dealing with an empty spacetime. So if we wish to develop the statistical-mechanical foundations of black hole thermodynamics, we must be prepared to relate the black hole's degrees of freedom to properties of an empty space.

The problem seems to be even more intriguing if the entropy of a black hole is compared with the entropy of other forms of the matter in the Universe. The entropy of a supermassive black hole of mass $10^9 M_\odot$ is 10^{95}, while the entropy of all other matter in the visible part of the Universe is of order 10^{88}. In other words, there is simply not enough matter in the Universe to explain the huge value of the entropy of a single black hole.

To solve this puzzle, two different ideas have been proposed:

1. To relate black hole entropy to the information lost in the process of black hole formation
2. To relate black hole entropy to properties of the vacuum in a strong gravitational field

12.6.2 Black hole entropy

Information-theoretical approach

In the case of a black body, one can define its entropy by counting different possibilities of preparing the system in a final state with given macroscopic parameters from microscopically different initial states :

$$S^{\mathrm{I}} = -\sum_n p_n \ln p_n \,, \tag{12.6.1}$$

with p_n being the probabilities of different initial states. This definition directly relates the entropy of the body to the information lost in the process of its formation. For this reason, one can call the so-defined entropy the *informational entropy*.

Historically, the first attempts at a statistical-mechanical foundation of the entropy of a black hole were connected with the informational approach [Bekenstein (1973)., Zurek and Thorne (1985)]. According to this approach, the black hole entropy is interpreted as *"the logarithm of the number of quantum-mechanically distinct ways that the hole could have been made"* [Zurek and Thorne (1985), see also Chapter VIII by Thorne, Zurek, and Price in the book by Thorne, Price, and Macdonald (1986)].

In order to estimate this number, one assumes that the black hole is made in a total time not longer than the time of black hole evaporation $t_H \sim M^3$ [see relation

(10.1.20)]. The vast majority of the ways to prepare a black hole with given mass M involves sending in one quantum after another with their energy ϵ kept as small as possible, i.e., $\epsilon \sim M^{-1}$. The total number of quanta required for the formation of a black hole of mass M is $N_B \sim M/\epsilon \sim M^2$. Using the expression $27\pi M^2$ for the effective cross-section of the hole, one can estimate the total phase-space volume that the particles occupied before their injection into the black hole as $\sim 27\pi M^2 t_H \epsilon^3$. The density of single-particle states in the phase space is $\sim (2\pi)^{-3}$. Thus, $N \sim M^2 \sim N_B$ is the total number of single-particle states in which N_B quanta can be injected into the black hole. The number of ways to make the black hole is $\mathcal{N} \sim (N-1+N_B)!/[(N-1)!N_B!]$ and its logarithm agrees with the Bekenstein-Hawking entropy

$$S^{BH} \sim \log \mathcal{N} \sim N_B . \tag{12.6.2}$$

One might expect that if initially there were not one but several species of particles, the number of possibilities to create the black hole would increase. This does not happen since t_H then becomes smaller so that N remains the same. The accreting particle configurations which provide the vast majority of the ways to prepare the black hole are simply an approximate "time reversal" of the products of the black hole evaporation. Zurek and Thorne (1985) showed that the so-defined informational entropy of a black hole is simply related to the amount of information lost by stretching the horizon.

Quite interesting results were obtained for a special model proposed by Bekenstein and Mukhanov [Mukhanov (1986), Bekenstein and Mukhanov (1995)]. In this model, the mass of a black hole is assumed to be quantized. The mass spectrum follows from the assumption that the horizon area $\mathcal{A}$ should be quantized in integers $\mathcal{A} = \alpha\hbar n$. The black hole of mass M is identified with a system at the level $n \sim M^{1/2}$, so that an absorption or emission of a particle by a black hole is accompanied by the transition from one energy level n to another $n \pm 1$. The entropy is related to the degeneracy $g(n)$ of the level n: $S = \ln g(n)$. The degeneracy is identified with the number of different ways to reach the level n by injecting individual quanta into the black hole. It is easy to show that the number of ways in which the black hole at level n can be formed from $n = 0$ by going up the staircase of levels is 2^{n-1}. So that for the value $\alpha = 4\ln 2$, one gets the correct value of the Bekenstein-Hawking entropy S^{BH}.

It should be emphasized that in this consideration one focuses on the behavior black hole alone. The number of species of particles that exist in nature and participate formation of the black hole does not enter at all. In other words, Bekenstein and Mukhanov (1995) count different trajectories in the space of the black hole parameters, while Zurek and Thorne (1985) count different initial states of the matter that form a black hole. Since one can specify a theory describing the matter, the method of Thorne and Zurek allows (at least in principle) an exact formulation based on first principles. In the model of Bekenstein and Mukhanov, the existence of internal discrete levels of a black hole and their degeneracy are postulated and not derived.

General relations connecting the black hole entropy with the possible number of black hole "preparations" are very important, but they do not give the complete solution of the problem. First of all, it is important to bear in mind the following difference between a black hole and an ordinary physical body. The particles that were used to build the body are always there and can in principle be re-extracted. The particles that enter a black hole have fallen into the singularity and cannot be recovered. We cannot say that a black hole at some late moment of time "consists" of all the particles that had fallen into it earlier. Gedanken experiments with the black hole "performed" at late times do not change the state of these particles. For this reason, in order to have a closer similarity to standard physical thermodynamics, one would like to define the dynamical degrees existing at the moment of the experiment.

There exists another problem with the explanation of black hole entropy by counting the different ways in which the black hole can formed. The time of evaporation of a black hole of mass M is of order

$$t_H \sim 10^{10}\text{years}\left(\frac{M}{10^{15}g}\right) . \tag{12.6.3}$$

It is this time that is required for processes of black hole formation that give the vast majority of the ways in the Thorne-Zurek argument. On the other hand, we know that the stellar and massive black holes which exist in nature were certainly formed after the Big Bang, not earlier than say 10–20 billion years ago. This additional information drastically reduces the number of possibilities and the logarithm of this number is much smaller than the expected entropy $4\pi M^2$.

Bekenstein and Mukhanov (1995) do not consider the time required for black hole formation in this process and assume that "*quantum transitions do not take time*". However, one can expect that in order to be able to distinguish the different ways, one should send the next quantum with a delay $\Delta t \sim 1/\Delta E$ after the previous one generated a "jump" by the absorption of the energy ΔE. In this case, the time required to prepare the black hole by "trying" the vast majority of the ways is again t_H. So for black holes for which the known time of formation is much smaller than the evaporation time t_H, there exists the same potential problem.

York (1983) proposed an approach which is another version of the "sum over possibilities". York noticed that quantum evaporation of a black hole thermally excites its gravitational quasinormal modes. He tried to connect the black hole entropy with these modes. Since the logarithm of the number of thermally excited quasinormal modes at given moment of time is much less than M^2, in order to obtain the correct answer for the entropy, York proposed to define it as the logarithm of the number of distinct excitation states of these modes in the process of black hole evaporation.

Black hole entropy and the properties of the vacuum

Another important class of attempts to explain black hole entropy is to relate the Bekenstein-Hawking entropy to the properties of the vacuum in the gravitational field of a black hole. We recall that a vacuum is a physical medium which possesses a number of non-trivial properties. Even in empty space, in the absence of a long-lived particles, there exist zero-point fluctuations of physical fields. An observer at rest near the black hole can register these zero-point fluctuations in the form of thermal atmosphere of the black hole (see Section 11.3.4). Zurek and Thorne (1985) argued that the informational entropy S^I they obtained by counting the ways of the black hole formation can be related to "*the logarithm of the number of configurations that the hole 'atmosphere', as measured by stationary observers, could assume in the presence of its background noise of acceleration radiation*".

Earlier, Gerlach (1976) tried to relate the black hole entropy to the logarithm of "*total number of vacuum fluctuation modes responsible for the emission of black body radiation*" by the black hole.

't Hooft (1985) proposed that "*the particle wavefunctions extremely close to the horizon must be modified in a complicated way by gravitational interactions between ingoing and outgoing particles*". He considered a so-called *brick wall model* in which it is assumed that field modes simply vanish within some fixed distance from the horizon. This boundary condition corresponds to the situation where there exists a mirror-like boundary outside the horizon, and one restricts oneself to considering a region in the exterior to the boundary. 't Hooft proposed to relate the black hole entropy to the entropy of thermal radiation at the Hawking temperature located outside the mirror-like boundary. Simple calculations give

$$S \sim \alpha \frac{\mathcal{A}}{l^2}, \tag{12.6.4}$$

where $\mathcal{A}$ is the surface area of the black hole; l is the proper distance of the mirror from the horizon, and α is a numerical coefficient depending on the field parameters. For $l \sim l_{\text{Pl}}$ one gets the same order of magnitude as the Bekenstein-Hawking entropy $S^{BH} = \mathcal{A}/(4l_{\text{Pl}}^2)$. This model does not explain how the thermal properties of fields outside the black hole are connected with the loss of information concerning the states in the black hole interior. The author himself stressed that "this model should be seen as an elementary exercise, rather than an attempt to describe physical black holes accurately".

Bombelli *et al.* (1986) demonstrated that a non-trivial entropy arises already in a flat spacetime for a system in the vacuum state provided one restricts oneself by considering a subsystem located in a spatially restricted part of the space. Let Ω be such a region and $\partial\Omega$ be its boundary. A free field φ can be decomposed into modes $\varphi_J(x)$, so that the amplitude of a single mode can be identified with a quantum oscillator. The wavefunction $\varphi_J(x)$ defines the amplitude of probability to find a quantum near point x. Among all the modes there always exists a subset of modes that do not vanish both inside and outside $\partial\Omega$. If the information on the states of the oscillators inside Ω is unavailable, one can obtain a reduced density matrix $\hat{\rho}$ for the oscillators outside Ω by integrating out over the states in the region Ω. Bombelli *et al.* (1986) showed that the entropy $S = -\text{Tr}(\hat{\rho} \ln \hat{\rho})$ for this density matrix is divergent, and the leading divergence is of the same form as (12.6.4) with $\mathcal{A}$ being the surface area of $\partial\Omega$, l being the proper distance cutoff parameter, and α being the numerical coefficient depending on the field parameters. Bombelli *et al.* (1986) proposed to use this so-called *entanglement entropy* in order to explain the origin of black hole entropy. A similar idea was also proposed later by Srednicki (1993).

Model of a black hole with dynamical interior

A refined version of this approach which makes clear the thermal origin of S for a black hole was proposed by Frolov and Novikov (1993b). Their model is based on the fact that the black hole interior is a dynamical system with non-trivial properties.

At first sight, this looks like a paradox. We know that (at least in the classical theory of general relativity) the state of a black hole at late times is completely specified by a finite number of parameters. For a non-rotating uncharged black hole one needs to know only one parameter (the mass M) to describe all its properties. This is true not only for the exterior where this property is a consequence of the no hair theorems, but also for the black hole interior [we shall discuss this in more detail in Chapter 14].

The reason why it is impossible for an isolated black hole at late times to have non-trivial excitations in its interior in the vicinity of the horizon is basically the same as for the exterior regions. These states could be excited only if a collapsing body emits a pulse of fields or particles immediately after it crosses the event horizon. Due to the redshift effect the energy of the emitted pulse must be exponentially large in order to reach the late time region with any reasonable energy. A short time after the formation of a black hole (say $\Delta t \sim 100M$) this is virtually impossible because it requires an energy of emission much greater than the black hole mass M.

In quantum physics the situation is quite different due to the presence of zero-point fluctuations. To analyze the states of a quantum field, it is convenient to use its decomposition into modes. Besides the positive-frequency modes which have positive energy, there exist also positive-frequency modes with negative energy (we called then DN-modes, see Section 10.1.2). In a non-rotating uncharged black holes such modes can propagate only inside the horizon where the Killing vector used to define the energy is spacelike. It is possible to show that at late times these states are thermally excited for any regular initial state of the field and the corresponding temperature coincides with the black hole temperature $\theta_H = (8\pi M)^{-1}$. It is natural to identify the dynamical degrees of freedom of a black hole with the states of all physical fields inside the black hole [Frolov and Novikov (1993b)]. The set of the fields must include the gravitational field.

Figure 12.2 gives a schematic picture of a black hole at some "instant of time". Quanta of different fields are created by the Hawking process. The quanta are created in pairs. Energy conservation does not allow for both components of the pair to be created outside the black hole. Different possibilities are shown in Figure 12.2.

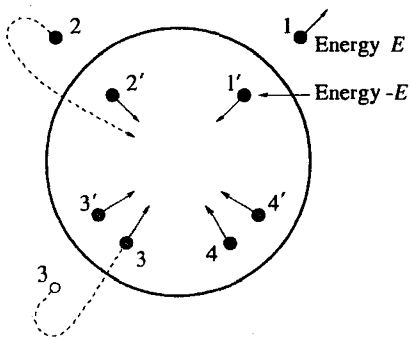


Figure 12.2: Illustration of the Hawking process.

The particle 1 has enough energy to penetrate the potential barrier and to escape to infinity. Such particles form the flux of Hawking radiation from the black hole. Particle 2 has not enough energy to escape to infinity and will fall back into the black hole later, but at the moment of consideration it is located outside the black hole. Particle 3 was created earlier outside the horizon, but after reflection by the potential barrier it has fallen into the black hole. Particle 4 as well as its companion 4′ are created inside the black hole.

After averaging of the given initial state over the states of the particles located outside the black hole at the given moment of time, one obtains the density matrix describing the state of the black hole interior. It can be identified with the density matrix of the black hole itself. In accordance with this definition the density matrix of a black hole is

$$\hat{\rho}_H = \underbrace{Tr}_{\text{"visible"}} \left(\hat{\rho}_{\text{total}} \right), \tag{12.6.5}$$

where $\hat{\rho}_{\text{total}}$ is the total density matrix for quantum fields, and the trace operation is performed over the modes located outside the horizon (which are "visible" to a distant observer).

The corresponding statistical-mechanical entropy is

$$S^{SM} = - \underbrace{Tr}_{\text{"invisible"}} \left(\hat{\rho}_H \ln \hat{\rho}_H \right), \tag{12.6.6}$$

where the trace operation is performed over the modes located inside the horizon (which are "invisible" to a distant observer). In fact, pairs 3-3′ and 4-4′ do not contribute to the entropy because both of the components are located inside the black hole, and their state is pure. This means that only pairs 1–1′ and 2–2′ are responsible for the entropy of the black hole. A particle belongs to the type 1 if its momentum ℓ is small: $\ell \lesssim \sqrt{27}\, ME$. Since particles with energies greater than

$\theta_H = (8\pi M)^{-1}$ are practically not created, one can conclude that particles of type 2 (trapped by the potential barrier) give the main contribution to the statistical-mechanical entropy of the black hole. (For more details, see [Frolov and Novikov (1993b)].)

By using the mode decomposition and the fact that the internal modes are thermally populated with respect to the Killing energy, one gets

$$S^{SM} = \underbrace{\Sigma}_{\text{"visible" J}} s_J. \tag{12.6.7}$$

Here

$$s_J = \frac{\beta\omega}{\exp(\beta\omega) - 1} - \ln\left[1 - \exp(-\beta\omega)\right] \tag{12.6.8}$$

is the entropy of a quantum oscillator of frequency ω at the temperature $\theta = \beta^{-1}$ (see Appendix I). It is important to stress that in this model the dynamical degrees of freedom of a black hole are connected with "invisible" modes; nevertheless, the summation in (12.6.7) is performed over the "visible" states. The reason is simple: Only those pairs for which one of the components is "visible" actually contribute to the entropy.

The fact that the summation in (12.6.7) is performed over the "visible" states is in an agreement with 't Hooft's (1985) brick wall proposal. It should also be emphasized that the entropy S^{SM} can be considered as a special case of the entanglement entropy [Bombelli *et al.* (1986), Srednicki (1993)]. The remarkable feature of the statistical-mechanical entropy (12.6.7) for a black hole is that it is related to the entropy of "thermal excitations". This property is connected with the fact that the boundary of the black hole coincides with the Killing horizon. There is an elegant possibility [Barvinsky, Frolov, and Zel'nikov (1995)] of describing the internal degrees of freedom above introduced in the framework of the *no-boundary wavefunction of the black hole* described in Section 11.4.3.

It can be shown that for any chosen field number of modes grows without bound as one considers regions located closer and closer to the horizon. For this reason, the contribution of a field to the statistical-mechanical entropy of a black hole calculated by counting the internal modes of a black hole is formally divergent. In order to make it finite, one could restrict oneself by considering only those modes which are located at a proper distance from the horizon greater than some chosen value l. For this choice of cutoff the contribution of a field to the statistical-mechanical entropy of a black hole is

$$S^{SM} = \alpha\,\frac{\mathcal{A}}{l^2}, \tag{12.6.9}$$

where $\mathcal{A}$ is the surface area of the black hole, and a dimensionless parameter α depends on the type of field.

The statistical-mechanical entropy S^{SM} in this as well as other "dynamical" approaches possesses the following main properties:

1. $S^{SM} \sim \mathcal{A}$, where $\mathcal{A}$ is the surface area of a black hole.
2. S^{SM} is divergent and requires regularization $S^{SM} \sim \mathcal{A}/l^2$, where l is the cutoff parameter.
3. S^{SM} depends on the number of fields which exist in nature.
4. For $l \sim l_{Pl}$ one has $S^{SM} \sim S^{BH}$.

One-loop contribution to the black hole entropy

It might be suggestive to use a special choice of the cutoff parameter l and to identify S^{SM} with S^{BH}. Unfortunately, the situation is not so simple. There are at least the following three problems which must be taken into account in a discussion of the relation between S^{SM} and S^{BH}:

1. S^{SM} is generated by quantum fields, but the same fields in the general case contribute to the ultraviolet divergences which modify the gravitational action.
2. A black hole really is a thermodynamical system, but it also possesses a very special property which singles it out; namely, its size and temperature are not independent parameters. Under these conditions the standard proof of the equality S^{SM} and S^{BH} does not work [Frolov (1995)].
3. York's derivation of S^{BH} uses only a tree-level gravitational action, while the derivation of S^{SM} requires one-loop calculations.

These problems are closely related to one another. In order to discuss them in more detail, we need to develop further the formalism presented in Section 12.5.1.

Complete information concerning the canonical ensemble of black holes with a given inverse temperature β at the boundary is contained in the partition function $Z(\beta)$ given by the Euclidean path integral [Hawking (1979)]

$$Z(\beta) = \int \mathcal{D}[\Phi] \exp(-W_E[\Phi]). \tag{12.6.10}$$

Here the integration is taken over all fields (including the gravitational field) that are real on the Euclidean section and are periodic in the imaginary time coordinate τ with period β. The quantity Φ is understood as a collective variable describing the fields. In particular, it contains the gravitational field. Here $\mathcal{D}[\Phi]$ is the measure on the space of fields Φ, and W_E is the Euclidean action of the field configuration. The action W_E includes the Euclidean Einstein action. The state of the system is determined by the choice of boundary conditions on the metrics that one integrates over. For the canonical ensemble and for the gravitational fields inside a spherical box of radius r_B at temperature θ, one must integrate over all metrics inside r_B which are periodically identified in the imaginary time direction with period $\beta = \theta^{-1}$. Such a partition function must describe in particular a thermal ensemble of black holes.

The partition function Z is related to the effective action $\Gamma = -\ln Z$ and to the free energy $F = \beta^{-1}\Gamma = -\beta^{-1}\ln Z$.

By using the stationary-phase approximation, one gets

$$\beta F \equiv -\ln Z = W_E[\Phi_0] - \ln Z_1 + \dots . \tag{12.6.11}$$

Here Φ_0 is the (generally speaking, complex) solution of the classical field equations for the action $W_E[\Phi]$ obeying the required periodicity and boundary conditions. Besides the tree-level contribution $W_E[\Phi_0]$, the expression (12.6.11) includes also one-loop corrections $\ln Z_1$, connected with the contributions of the field perturbations on the background Φ_0, as well as higher-order terms in the loop expansion, denoted by $(\dots)$. The one-loop contributions can be written as follows : $\ln Z_1 = -\frac{1}{2}\operatorname{Tr}\ln(-D)$, where D is the field operator for the field perturbations inside the box r_B.

The one-loop contribution contains divergences and requires renormalization. In order to be able to absorb these divergences into the renormalization of the coefficients of the initial classical action, we choose the latter in the form

$$W_{\text{cl}} = \int d^4x\,\sqrt{g}\,L, \tag{12.6.12}$$

$$L = \left[-\frac{\Lambda_{\text{bare}}}{8\pi G_{\text{bare}}} - \frac{R}{16\pi G_{\text{bare}}} + c^1_{\text{bare}}\,R^2 + c^2_{\text{bare}}\,R^2_{\mu\nu} + c^3_{\text{bare}}\,R^2_{\alpha\beta\mu\nu}\right]. \tag{12.6.13}$$

The subscript "bare" indicates that corresponding coupling constant is used in the initial action, and later it will be renormalized by quantum corrections.

The divergent part of the one-loop effective action has the same structure as the initial classical action (12.6.12), and hence one can write

$$\Gamma = \Gamma^{\text{ren}}_{\text{cl}} + \Gamma^{\text{ren}}_{1}, \tag{12.6.14}$$

$$\Gamma^{\text{ren}}_{\text{cl}} = \Gamma_{\text{cl}} + \Gamma^{\text{div}}_{1}, \qquad \Gamma^{\text{ren}}_{1} = \Gamma_1 - \Gamma^{\text{div}}_{1}. \tag{12.6.15}$$

Here $\Gamma^{\text{ren}}_{\text{cl}}$ is identical to the initial classical action with the sole change that all the bare coefficients Λ_{bare}, G_{bare}, and c^i_{bare} are replaced by their renormalized versions Λ_{ren}, G_{ren}, and c^i_{ren}

$$\frac{\Lambda_{\text{ren}}}{G_{\text{ren}}} = \frac{\Lambda_{\text{bare}}}{G_{\text{bare}}} + \alpha_{(4)}\,, \tag{12.6.16}$$

$$\frac{1}{G_{\text{ren}}} = \frac{1}{G_{\text{bare}}} + \alpha_{(2)}\,, \tag{12.6.17}$$

$$c^i_{\text{ren}} = c^i_{\text{bare}} + \alpha^i_{(0)}\,. \tag{12.6.18}$$

The concrete form of the divergent coefficients $\alpha_{(n)}$ depends on the chosen regularization scheme.

We shall refer to (12.6.14) as the loop expansion of the renormalized effective action. After multiplying the renormalized effective action by β^{-1}, we get the expansion for the renormalized free energy.

The effective action Γ contains complete information about the system under consideration. In particular, the variation of Γ with respect to the metric provides the equations for the quantum average metric $\bar{g} = \langle g \rangle$:

$$\frac{\delta \Gamma}{\delta \bar{g}} = 0 . \tag{12.6.19}$$

One usually assumes that quantum corrections are small and solves this equation perturbatively:

$$\bar{g} = g_{\text{cl}} + \delta g , \tag{12.6.20}$$

where g_{cl} is a solution of the classical equations.

One also usually assumes that the renormalized values of Λ_{ren} and c^i_{ren} vanish: $\Lambda_{\text{ren}} = c^i_{\text{ren}} = 0$. The "classical" field g_{cl} is a solution of the classical Einstein equations. In our case g_{cl} is the Euclidean black hole metric, while the metric $\bar{g}$ describes the Gibbons-Hawking instanton, deformed due to the presence of quantum corrections to the metric. The quantity $\Gamma[\bar{g}]$, expressed as a function of the boundary conditions (β and r_0), specifies the thermodynamical properties of the black hole.

By a simple repetition of the calculations of Section 12.5 starting from the action $\Gamma^{\text{ren}}_{\text{cl}}$, one can show that the Bekenstein-Hawking entropy $S^{BH}_{\text{ren}} \equiv \mathcal{A}/4G_{\text{ren}}$ coincides with $\beta^2\, d(\beta \Gamma^{\text{ren}}_{\text{cl}})/d\beta$. The quantum correction Γ^{ren}_1, calculated on the regular Euclidean space of the Gibbons-Hawking instanton, is finite. It contains an additional factor $\hbar$, so that for dimensional reasons the corresponding contribution S^{ren}_1 to the entropy is smaller than S^{BH} by the factor m^2_{Pl}/M^2.[14] S^{ren}_1 can be related to the contribution to the entropy of the thermal radiation outside the black hole. To summarize, the thermodynamical entropy (that is, the entropy which enters the laws of thermodynamics) of the black hole is

$$S^{TD} = S^{BH}_{\text{ren}} + S^1_{\text{ren}} . \tag{12.6.21}$$

Let us now discuss the relation between S^1_{ren} and S^{SM}. First we note that the derivation of the thermodynamical entropy of a black hole requires *on-shell* calculations. This means that in order to get S^{TD}, one compares two different Euclidean black hole *equilibrium* configurations with slightly different temperatures. For each of them $M = M(\beta)$. S^{TD} is given by

$$S^{TD} = \beta^2 \frac{dF}{d\beta} , \tag{12.6.22}$$

[14]Due to the presence of the conformal anomalies one might expect that the leading corrections to S^{ren}_1 are of order $\ln M$ (see, e.g., [Fursaev (1995)]).

where $d/d\beta$ is the *total derivative* of the free energy with respect to the inverse temperature β.

On the other hand, the statistical-mechanical entropy of a black hole S^{SM} requires *off-shell* calculations [Frolov, Fursaev, and Zel'nikov (1996a,b)]. The words *off-shell* mean that in the calculations one must consider the inverse temperature β and the black hole mass M as independent parameters. Strictly speaking, for $\beta_\infty \equiv \beta(1 - 2M/r_B)^{-1/2} \neq \beta_H \equiv 8\pi M$ there is no regular vacuum Euclidean solution with the Euclidean black hole topology $R^2 \times S^2$. Such a solution can be obtained only if one excludes a horizon sphere S^2. If we denote the off-shell free energy by $\tilde{F}(\beta, M)$, then

$$S^{SM} = \left.\frac{\partial \tilde{F}}{\partial \beta}\right|_M . \tag{12.6.23}$$

For $\beta_\infty = \beta_H$ the singularity disappears and $F(\beta) = \tilde{F}(\beta, M)$. Hence, we have

$$\frac{dF}{d\beta} = \left.\frac{\partial \tilde{F}}{\partial \beta}\right|_M + \left.\frac{\partial \tilde{F}}{\partial M}\right|_\beta \frac{d\tilde{M}}{d\beta} . \tag{12.6.24}$$

As a result of this relation, we have [Frolov (1995)]

$$S^{TD} = S^{BH}_{\text{ren}} + S^{SM} - \Delta S . \tag{12.6.25}$$

The term

$$\Delta S = - \left.\frac{\partial \tilde{F}}{\partial M}\right|_\beta \frac{d\tilde{M}}{d\beta} \tag{12.6.26}$$

cancels the divergence of S^{SM} so that $S^1_{\text{ren}} = S^{SM} - \Delta S$ is always finite.[15]

The relation (12.6.25) can be used to explain of the entropy subtraction procedure proposed earlier by Zurek and Thorne (1985).

[15]The above general formulas can be illustrated by explicit calculations in a simplified two-dimensional model [Frolov, Fursaev, and Zel'nikov (1996b)]. In this model one considers a special version of two-dimensional dilaton gravity, which can be obtained from the Einstein action by its reduction to spherical spacetimes, and uses a massless quantum field to describe the corresponding quantum degrees of freedom. The main result is that the entropy of a black hole can be written in the form

$$S^{TD} = \bar{S}^{BH} + S^{SM} - S^{SM}_{\text{Rindler}} .$$

The quantities which enter this relation have the following meaning. $\bar{S}^{BH}$ is the Bekenstein-Hawking entropy with the only change that for its calculation one must use instead of the classical gravitational radius r_+ its quantum-corrected version. S^{SM} is the statistical-mechanical entropy calculated either by the volume cutoff or by the brick wall method. S^{SM}_{Rindler} is the analogous quantity calculated by the same method in the Rindler space (that is, in flat spacetime) provided that in both calculations (for the black hole and in the Rindler space) one uses identical values of the proper-distance cutoff parameter l.

To summarize, quantum excitations of a black hole are responsible for S^{SM}. There are two effects which are important when one establishes the relation of S^{SM} to the thermodynamical properties of black hole: One-loop divergences which renormalize the gravitational coupling constant in the definition of S^{BH} and the renormalization of S^{SM} due to subtraction ΔS in accordance with (12.6.25). We analyze now a relation between these two renormalizations.

Black hole entropy and renormalization

The origin of the divergence of S^{SM} is connected with the fact that in the close vicinity of the horizon there exist an infinite number of modes which contribute to S^{SM}. In expression (12.6.9) this infinity is suppressed by assuming that the summation over modes is performed only over the modes located at a proper distance larger than l from the horizon. This procedure is known as the *volume cutoff method.* A similar result can be obtained by using other schemes of regularization. For the discussion of the relation between S^{BH} and S^{SM} it is more convenient to use a *covariant regularization.*

Pauli-Villars regularization is an example of such a covariant regularization which is often used for calculations [see e.g., Demers, Lafrance, and Myers (1995)]. In this method, besides the given field of mass m, one introduces additional fields. In the simplest case, the number of these fields is five. Two of them, with masses M_k have the same statistics as the field m, and the other three have masses M'_r have the "wrong" statistics (i.e., they are fermions for scalars and bosons for spinors). The masses are restricted by the relations

$$m^2+\sum_{k=1}^{2} M_k^2-\sum_{r=1}^{3} M'^2_r=0\,,\quad m^4+\sum_{k=1}^{2} M_k^4-\sum_{r=1}^{3} M'^4_r=0\,. \tag{12.6.27}$$

To be concrete, we choose a special one-parameter solution of these equations

$$M_1=M_2=\sqrt{3\mu^2+m^2}\,,$$

$$M'_1=M'_2=\sqrt{\mu^2+m^2}\,,\qquad M'_3=\sqrt{4\mu^2+m^2}\,. \tag{12.6.28}$$

The statistical-mechanical entropy calculated in the Pauli-Villars regularization is

$$S^{SM}=\frac{\mathcal{A}}{48\pi}\sum_i h(s_i)\,B(m_i^2,\mu^2)\,, \tag{12.6.29}$$

where $h(s_i)$ is the number of polarizations of the field i of spin s_i, and

$$B(m^2,\mu^2)=m^2\ln m^2+\sum_k M_k^2\ln M_k^2-\sum_r M'^2_r\ln M'^2_r\,. \tag{12.6.30}$$

In the limit $\mu \to \infty$, the quantity B diverges as μ^2. This quadratic divergence μ^2 corresponds to the quadratic divergence l^{-2} in the volume cutoff method of regularization (cf. (12.6.9)). One can interpret this by saying that the covariant cutoff which removes divergences at high energies $E \sim \mu$ at the same time gives a cutoff at the proper distance $l \sim \mu^{-1}$.

It was suggested by Susskind and Uglum (1994) and Callan and Wilczek (1995) that the ultraviolet divergence of the Newton coupling constant might be related to the divergence of S^{SM}. In order to analyze this possibility, we establish the relation between S^{SM} and S^{BH}. To compare these quantities, we first need to take into account that G_{ren}, which enters $S^{BH}_{\text{ren}} = \mathcal{A}/G_{\text{ren}}$, is the observable value of the Newtonian coupling constant. It differs from the bare coupling G_{bare} by contributions due to the quantum fields

$$\frac{1}{G_{\text{ren}}} = \frac{1}{G_{\text{bare}}} + \frac{1}{G_{\text{div}}}, \tag{12.6.31}$$

where for the same Pauli-Villars regularization used above one has

$$\frac{1}{G_{\text{div}}} = \frac{1}{12\pi}\sum_i h(s_i)\, B(m_i^2,\mu^2) - \frac{1}{2\pi}\sum_s \xi_s\, B(m_s^2,\mu^2)\,. \tag{12.6.32}$$

Here ξ_s is the parameter of non-minimal coupling of the scalar field ϕ_s with the scalar curvature R. By comparing $\mathcal{A}/4G_{\text{ren}}$ with (12.6.29), one can conclude that S^{BH} can be equal to S^{SM} only if $G^{-1}_{\text{bare}} = 0$ and $\xi = 0$. The first relation indicates that there is no initial (or bare) gravity, so that the entire gravitational action is induced as a result of quantum (one-loop) effect of a system of quantized fields. But as soon as the initial bare parameter G^{-1}_{bare} does not vanish, there would always be an inevitable non-statistical (geometrical) contribution to the entropy.

The second relation, $\xi = 0$, implies that the scalar fields which might be present among the quantum fields must necessarily be minimally coupled with $\xi = 0$. Under this assumption and in the absence of an initially infinite (and negative) coupling constant G^{-1}_{bare}, it is impossible to have a finite value of G^{-1}_{ren} because the contribution of all (boson and fermion) fields to G^{-1}_{div} is positive and divergent. In other words, one cannot have $S^{BH} = S^{SM}$ in a theory where G^{-1}_{ren} is one-loop finite.

One can say that the attempts to explain the Bekenstein-Hawking entropy S^{BH} by counting quantum excitations of a black hole meet an evident difficulty because S^{BH} arises at the tree level, while the entropy of quantum excitations is a one-loop quantity.

Does this mean that it is impossible to explain the Bekenstein-Hawking entropy by counting quantum excitations of a black hole in the scheme described above? This conclusion is too hasty. There exist a class of theories where S^{BH} is directly related to the statistical-mechanical entropy. They are the so-called theories of *induced gravity.*

12.6.3 Sakharov's induced gravity and black hole entropy

A theory of induced gravity was proposed by Sakharov in 1968 [see also Sakharov (1976) and the review by Adler (1982)]. According to Sakharov's idea, general relativity can be considered as a low-energy effective theory where the metric $g_{\mu\nu}$ becomes a dynamical variable as a result of quantum effects in a system of heavy constituents propagating in the external gravitational background. Gravitons in this picture are to some extent analogous to the phonon field describing collective excitations of a lattice in the low-temperature limit of the theory.

According to Sakharov's idea, the background fundamental theory is described by an action $I[\varphi_i, g_{\mu\nu}]$ of fields φ_i propagating in an external geometry with the metric $g_{\mu\nu}$. The metric originally is not dynamical so that the original gravitational action is simply put equal to zero. By averaging over the constituent fields φ_i, one gets a dynamical effective action for the metric $g_{\mu\nu}$

$$\exp(-W[\,g_{\mu\nu}]) = \int \mathcal{D}[\varphi_i]\exp(-I[\varphi_i, g_{\mu\nu}]\,)\,. \qquad (12.6.33)$$

An observable gravitational field is an extremum of the effective action $W[g_{\mu\nu}]$. The idea is to relate the Bekenstein-Hawking entropy to the statistical mechanics of the ultra-heavy constituents which induce gravity in the low energy limit of the theory [Jacobson (1994)].

This idea was recently developed by Frolov, Fursaev, and Zel'nikov (1996c). The model proposed in this paper consists of Dirac fermion and scalar boson field constituents. It was demonstrated that one can obtain a vanishing induced cosmological constant and a finite induced Newton coupling constant G for a special choice of spectrum of the masses of the constituents provided at least some of the scalar fields are non-minimally coupled. The Newton constant G is a function of the parameters of the constituents

$$\frac{1}{G} = \frac{1}{12\pi}\left(\sum_s (1-6\xi_s)\; m_s^2\,\ln m_s^2 + 2\sum_d m_d^2\,\ln m_d^2\right), \qquad (12.6.34)$$

where ξ is the parameter of the non-minimal coupling $\xi R\varphi^2$ in the action. The non-minimality of the scalar fields plays an important role since it allows one to make finite both G and the entropy of constituents [Frolov, Fursaev, and Zel'nikov (1996c), Frolov and Fursaev (1997a,b)]. The value of the induced Newton constant is dominated by the masses of the heaviest constituents and at least some of the constituents must have Planck mass $m \sim m_{Pl}$.

Relation (12.6.33) allows one to express the black hole entropy as a functional integral over the configurations of the constituent fields in a given gravitational background. Considering the states of constituents near the horizon one basically returns to the original model discussed by Frolov and Novikov (1993b). Pairs of constituents are created in the vicinity of the horizon, at least one component of the pair being created inside the black hole. The only (but very important) difference is that instead of counting the degrees of freedom of the light (physically observable) particles one repeats similar calculations for the heavy (unobservable at low energies) constituents. The heavy constituents can effectively be excited only in a narrow region of Planckian size close to the horizon. The remarkable fact is that, since the entropy of the black hole is generated by the same fields that generate the low energy gravity itself, the statistical-mechanical entropy of the constituents in such models always coincides with the Bekenstein-Hawking entropy. This model allows one also to relate the number of states of a black hole of fixed mass M to the number of physical states of the constituents with fixed total energy [Frolov and Fursaev (1997a,b)].

The proposed class of induced-gravity models certainly cannot be considered as a fundamental theory of quantum gravity. They are simply phenomenological models which reproduce correctly the low energy behavior of the gravitational field. But since the results do not depend on the particular choice of the model from this class, one might hope that the same property, independence of the entropy of macroscopic black holes from the particular realization of the fundamental theory at high energies, is of more general significance. Frolov, Fursaev, and Zel'nikov (1996c) proposed to consider this consistency of the statistical mechanics of constituents in fundamental theory with the standard low energy gravitational calculations as a general principle which was called the *low-energy censorship conjecture.*

12.6.4 Black hole entropy in superstring theory

Superstring theory is a promising candidate for quantum gravity. One of the most interesting developments of superstring theory are recent results relating the entropy of a special class of black holes to the counting of string quantum excitation states. This subject goes far beyond the scope of the book, so we restrict ourselves to brief comments. More details can be found e.g., in the reviews by Horowitz (1996, 1997), Maldacena (1996), and Akhmedov (1997), and references therein.

According to superstring theory, all particles are excitations of a one-dimensional object (string). String theory is characterized by two fundamental parameters: the string scale, l_s, and the string coupling constant g. The string tension is $T \sim l_s^{-2}$. Besides a finite number of massless fields, which include spin 2 gravitons, scalar dilatons, and some gauge fields, there exist an infinite number of massive string excitations, enumerated by an integer N, with masses $m^2 = N/l_s^2$. These massive states have a large degeneracy $\sim \exp N^{1/2}$.

When strings are quantized in a curved spacetime, there are a number of consistency conditions which impose constraints on the possible gravitational and other massless fields configurations and on the number of spacetime dimensions. The constraints on the massless fields can be rewritten in the form of dynamical equations for a specially chosen effective action. In the low-energy limit, when the curvature $\mathcal{R}$ is much less than l_s^{-2}, the consistency equations take the form of the Einstein equations with matter consisting of the other massless fields (see Section 13.1.2). The gravitational coupling constant is defined as $G \sim g^2 l_s^2$. In the simplest case, the number of spacetime dimensions required by the consistency conditions is 10. So initially one has a gravity theory in ten dimensions. In order to describe a physical four-dimensional world, one assumes that the extra dimensions are compactified. As a result of compactification, besides the gravitational field, the theory in four dimensions has an extra number of gauge and scalar fields. For example, $N = 8$ supergravity in four dimensions is the low energy limit of type II string theory compactified on a six-torus T^6. The difference from Einstein gravity is that $N = 8$ *supergravity* contains many extra massless fields: $U(1)$ gauge fields, scalar fields and various fermionic fields. Because of the presence of these extra fields, supergravity possesses an additional symmetry relating bosons and fermions, called *supersymmetry*. Despite this difference the theory has the same solutions describing a charged black hole. Extremally charged black holes with $\sqrt{G}\, M = Q$ are of the most interest. For such a black hole the Bekenstein-Hawking entropy $S^{BH} = \mathcal{A}/(4G) = (\pi/4)Q^2$ does not depend on the gravitational constant and is defined by the charge only.

Counting of the number of string excitations corresponding to a black hole is based on the following observation. Let us fix the string scale l_s and consider the coupling constant g as a quantity that (at least formally) can be changed. The entropy $\mathcal{A}/(4G)$ is independent of g. At weak coupling constant the horizon size $\sim g\, l_s Q$ is much smaller than the string size, so that one can use the flat spacetime description. As one increases the coupling constant, the horizon size becomes comparable with the string scale l_s and a black hole forms. In supersymmetric theory there exists a general inequality on the mass and charge: $M \geq kQ$, where k is a constant depending on the theory. States that saturate this bound are called Bogomol'nyi-Prasad-Sommerfield states, or *BPS-states*. They have the property that their mass cannot receive quantum corrections. The number of these states is a topological invariant. In string theory an extremal black hole can be considered as formed by specially chosen BPS-states, and the number of ways in which it can be formed is an invariant, independent of the string coupling constant g. That is why instead of making calculations for the black hole entropy in the regime of strong coupling, one can perform the calculations for weak coupling, that is, in flat spacetime. Technically this problem is much simpler.

In string theory there exist different charges. There are two types of gauge fields called NS (for Neveu-Schwarz) and R (for Ramond), and they can carry either electric or magnetic charge. A black hole with a single charge has an extremal limit with a singular horizon. In order to form an extremal black hole with a regular horizon, one needs to combine several different types of charges. The Reissner-Nordström metric is a solution to string theory, but several gauge fields must be nonzero, and the charge Q is a function of these charge parameters.

Besides simple string fluctuations, there are also soliton-like states called *D-branes*. An extremal black hole with nonzero surface area is considered as being built from elementary constituents; that is, solitons with a single charge. For weak coupling constant one must calculate the number of string soliton states in flat spacetime that have the same total charge Q as the black hole. For large values of Q the answer is $\exp(\pi Q^2/4)$, and hence the result reproduces the expected value of the Bekenstein-Hawking entropy $S^{BH} = \mathcal{A}/(4G)$. This result was first obtained for five dimensional extremal black holes [Strominger and Vafa (1996)]. Later it was confirmed for five dimensional extremal rotating black holes [Breckenridge *et al.* (1997)] and slightly non-extremal black holes [Callan and Maldacena (1996), Horowitz and Strominger (1996), Breckenridge *et al.* (1996)]. It is also valid for four-dimensional extremal [Maldacena and Strominger (1996), Johnson, Khuri, and Myers (1996)] and slightly non-extremal [Horowitz, Lowe, and Maldacena (1996)] black holes. It is important that the surface area is a function of several parameters, which include different string charges, so that the proved equality of the entropy of a black hole to the logarithm of the number of different string configurations is a statement about a function, not a number.

In spite of the impressive results obtained in this approach, it should be emphasized that the mechanism of the universality of the low energy expression for the black hole entropy is still unclear in string theory. Each type of black hole solution requires separate calculations. Moreover, supersymmetry plays a key role in the string calculations of the entropy, and it is not clear how far these calculations can be generalized, so that they would be able to include the case of a neutral Schwarzschild or Kerr black hole, which are certainly of most interest.

Chapter 13

Black Holes in Unified Theories

13.1 Non-Einsteinian Black Holes

13.1.1 Introduction

The black holes we have considered up to now are solutions of the Einstein or Einstein-Maxwell equations. Only these black holes are important for astrophysical applications. Recently, there has been an enormous increase of interest in black hole physics on the part of scientists traditionally working in high energy physics. This is largely connected with attempts to construct a unified theory of all interactions, including gravity.

The idea that the various forces of nature might be unifiable has a long history. The unification of electric and magnetic fields by Faraday and the development of a theory of the electromagnetic field by Maxwell was the beginning of this history. Einstein made many attempts to unify gravity and electromagnetism. Recently, this idea was revived and has become one of the most promising in theoretical physics after the success in the unification of the electromagnetic and weak interactions and the construction of various models of Grand Unified Theories in which the strong interaction (quantum chromodynamics) is unified with the electroweak one. Finding a theory which unifies gravity with the other forces in nature is presently considered as one of the most elusive goals of theoretical physics. It is strongly believed that without such a unification it will be impossible to develop a theory which reconciles quantum mechanics and gravity, and to build a quantum theory of gravity, this Holy Grail of theoretical physics.

In the attempts to unify different interactions several ideas and principles were proposed, which have proved to be very deep and powerful.

One such principle is *gauge invariance.* Coordinate transformations in general relativity are a well-known example of *gauge transformations.* The first Noether theorem relates a global symmetry with conservation laws. Local symmetries differ greatly from global ones. They reduce the number of independent degrees of freedom of the system. Only invariants with respect to the local symmetry transformations (gauge invariants) enter the description of physical observables. On the other hand, in the general case one cannot solve the constraints imposed by the gauge invariance and work purely in terms of the "internal" gauge independent quantities. The reason is that dynamical equations in such variables are nonlocal and cannot be obtained by variation of an action. Both gravity and electromagnetism are gauge fields, the fields that exhibit the symmetry of the theory with respect to the corresponding gauge group of transformations. The gravitational field in general relativity is described by a massless field of spin two. Its nonlinear interactions are determined by a non-Abelian local group of coordinate transformations. The Yang-Mills field is a massless field of spin one. For a non-Abelian gauge group its equations are also nonlinear. The Maxwell field is an Abelian Yang-Mills fields with a $U(1)$ gauge group.

Another important idea proposed in the seventies is *supersymmetry*, that is, the symmetry between bosons and fermions. Within a few years the global supersymmetry of the early work was extended to local supersymmetry or *supergravity*, a remarkably rich extension of the Einstein and Einstein-Maxwell theory.

Most recent efforts have been directed at studying theories (e.g., supergravity) in which the number of spacetime dimensions is greater than four of the world which we observe. This idea goes back to the work of Kaluza (1921) and Klein (1926). The basic idea of this approach is that the extra dimensions are compactified and due to the small value of the compactification radius, excitation of the modes connected with the internal space requires very high energy. That is why under normal conditions these dimensions are not observable. On the other hand, the metric tensor in higher dimensions has more components. After projecting it into the physical four-dimensional spacetime, these extra components may be used to describe electromagnetic and/or other physical fields.

A refined version of the theory which naturally incorporates these ideas is a *superstring theory*. In string theory world lines representing trajectories of particles are replaced by two-dimensional world sheets, strings orbits. Such strings can be of any size, but under ordinary conditions they are tiny, about 10^{-32} cm. This size is deduced by comparing the predictions of string theory for Newton constant and the fine structure constant. Since strings are nonlocal objects, string theory does not contain quantum ultraviolet divergences, which are characteristic of standard local quantum field theories. The condition for the self-consistency of quantum string theory appears to be very restrictive. That such theories do exist was established in the early 1980's as a result of almost fifteen years of investigation. Moreover, there are only a few such consistent theories. The very latest study suggests that, in fact, all of them are equivalent. String theory produces Einstein gravity in the low-energy limit, that is, for energies much smaller than the Planck energy $10^{28}eV$. The (boson) string theory makes sense only in 26 dimensions, while its supersymmetric version is viable in 10 dimensions. The Kaluza-Klein reduction of these extra dimensions to the four physical dimensions unifies gravity with matter and makes important predictions concerning the nature of the "matter". For this reason, string theory is now considered by many physicists as the most promising unified theory of everything. It is also very often used as the modern version of quantum gravity.

One may expect that black holes, which allow one to relate gravity, quantum theory and thermodynamics in the four-dimensional case and which play an important role in understanding non-perturbative effects in a quantum theory of gravity may be important in the unified theories. The low energy limit of these theories contains gravity and some other massless or low-mass fields. How far can the results of classical black hole physics be extended? What happens to black holes when additional fields and extra-dimensions are allowed? These are the questions that have been analyzed in numerous recent works. It should be emphasized that the expected unification of the Einstein gravity occurs at Planck energies.[1] For this reason, the corresponding modification of Einstein gravity is practically unimportant for astrophysical black holes. The only expected situations when such effects can play a significant role are final states of evaporation of primordial black holes (Chapter 15) and spacetime regions near singularities inside black holes (Chapter 14).

Since there already exist more than a hundred publications on different aspects of black holes in unified theories, this topic itself could be a subject of a separate book. In this chapter we collect only some of the main results and certainly do not pretend to completeness. Before beginning the discussion of black holes in unified theories, we make a few additional remarks on string theory and the low-energy effective action for gravity in it.

13.1.2 Low-energy effective action in string theory

Consider a closed bosonic string $X^\mu(\xi)$ propagating in a non-trivial background in the target space, and write its action as

$$S[X] = \frac{1}{4\pi\alpha'}\int d^2\xi\,\sqrt{h}\,\big[h^{AB}\,g_{\mu\nu}(X)\,\partial_A X^\mu\,\partial_B X^\nu$$

$$+\,e^{AB}B_{\mu\nu}(X)\,\partial_A X^\mu\,\partial_B X^\nu + \alpha'\,R^{(2)}\,\varphi(X)\big]\,. \tag{13.1.1}$$

Here h_{AB} and $R^{(2)}$ are a metric and a curvature of a two-dimensional manifold; X^μ are coordinates in the target space, and $g_{\mu\nu}$, $B_{\mu\nu}$, and φ are a metric, an antisymmetric tensor, and a scalar function on it. The dimensional coupling constant α' is the inverse string tension. The action has the form of the nonlinear σ-model. For consistency of the quantization, the model should be conformally invariant, so that the longitudinal

[1]We already mentioned in Chapter 11 that quantum corrections to the Einstein equations become important in this regime. Even if one starts with the Einstein theory, these corrections modify it. The corrections include, in particular, higher in curvature terms. On black hole solutions in theories of gravitation with higher-order curvature corrections, see e.g., Wiltshire (1986, 1988), Myers (1987), Whitt (1988), and references cited therein.

modes of the string decouple from physical amplitudes. The first two terms of the classical action are explicitly conformally invariant. The third term, which violates the classical conformal invariance, contains the same additional factor α' as the one-loop quantities, and it is required to improve the quantum properties of the model.

The local conformal invariance implies that the trace of the two-dimensional world-sheet stress-energy tensor of the theory vanishes. In the model (13.1.1) this invariance is broken by the term containing the dilaton field φ and by quantum conformal anomalies. The general structure of the trace is

$$2\pi\, T^A_A = \sqrt{h}\left[\beta^g_{\mu\nu}\, h^{AB}\,\partial_A X^\mu\,\partial_B X^\nu + \beta^B_{\mu\nu}\, e^{AB}\,\partial_A X^\mu\,\partial_B X^\nu + \beta^\varphi\, R^{(2)}\right]. \quad (13.1.2)$$

The quantities $\beta^g_{\mu\nu}$, $\beta^B_{\mu\nu}$, and β^φ are local functionals of the coupling functions $g_{\mu\nu}$, $B_{\mu\nu}$, and φ. The self-consistency of the quantum theory requires that

$$\beta^g_{\mu\nu} = \beta^B_{\mu\nu} = \beta^\varphi = 0\,. \quad (13.1.3)$$

It can be shown that in the one-loop approximation these conditions follow from the following action [Callan *et al.* (1985)]

$$W = \int d^D x\,\sqrt{-g}\,e^{-2\varphi}\left[\Lambda + R + 4(\nabla\varphi)^2 - \frac{1}{12}\,H_{\mu\nu\lambda}\,H^{\mu\nu\lambda}\right]. \quad (13.1.4)$$

Here $H_{\mu\nu\lambda} = 3\nabla_{[\mu}\,B_{\nu\lambda]}$, and Λ is a constant which is related to the spacetime dimension D. For a bosonic string it vanishes when $D = 26$. For supersymmetric strings with fermions the critical number of dimensions is $D = 10$. For these critical dimensions $D = 10$ or 26 the theory is free from divergences and anomalies. One can also consider other number of dimensions by introducing additional conformal fields. Λ also depends on the central charge of possible "internal" conformal fields if they are present. The low-energy effective action (13.1.4) can also contain an additional term $(-F^2)$ for a gauge field A and some massless fermions.

It is clear from the form of action (13.1.4) that e^φ plays the role of a coupling constant which reflects the strength of quantum corrections. The complete string action includes higher-curvature corrections. In the application to the black hole problem one usually discusses solutions of the low-energy effective theory (13.1.4) where the fermions are omitted. One also include the Maxwell field associated with a $U(1)$ subgroup of the gauge group.

13.2 Four-Dimensional Black Holes

13.2.1 Dilaton black holes

The simplest generalization of Einstein-Maxwell theory including the dilaton field is described by the action

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,e^{-2\varphi}\left[R + 4(\nabla\varphi)^2 - F_{\mu\nu}F^{\mu\nu}\right]. \quad (13.2.1)$$

It can be obtained from the low energy action of the string theory (13.1.4) by dropping all the fields except for the metric $g_{\mu\nu}$, a dilaton φ, and a Maxwell field $F_{\mu\nu}$. As before, we put the Newtonian coupling constant $G = 1$. Restoring G in (13.2.1), one can see that the quantity $G\exp(2\varphi)$ plays the role of an effective gravitational coupling constant which might depend on x. Before discussing black hole solutions for the action (13.2.1) we note that by making a conformal transformation $g^E_{\mu\nu} = e^{-2\varphi}\, g_{\mu\nu}$ the action can be rewritten in a more familiar form

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g_E}\,\left[R_E - 2(\nabla\varphi)^2 - e^{-2\varphi}\,F_{\mu\nu}\,F^{\mu\nu}\right]. \tag{13.2.2}$$

In the absence of the Maxwell field the action reduces to the standard Einstein theory with a scalar massless field as the matter. According to the no hair theorem (see Section 6.7), the only spherically symmetric black hole solution is the Schwarzschild one, with $\varphi = 0$. For this solution $g_{\mu\nu} = g^E_{\mu\nu}$. In a general case when $\varphi \neq 0$, this equality is violated. In string theory $g_{\mu\nu}$ plays a more important role since strings are propagated in this metric. On the other hand, in terms of $g^E_{\mu\nu}$ the action has the Einsteinian form. In should be emphasized that global geometrical properties of spacetimes with these two metrics may differ. In order to distinguish between the metrics, one usually refer to the quantities calculated for $g^E_{\mu\nu}$ and for $g_{\mu\nu}$ as quantities in the *Einstein frame* and in the *string frame*, respectively.

There exists a remarkable symmetry property that allows one to obtain a one-parameter family of solutions starting with any static solution of (13.2.1) [Sen (1991, 1992a), Hassan and Sen (1992)]. If one starts with the Schwarzschild solution ($g_{\mu\nu}, \varphi = 0, A_\mu = 0$), one gets ($\tilde{g}_{\mu\nu}, \tilde{\varphi}, \tilde{A}_\mu$), where g_{ij} is unchanged, $\tilde{A}_\mu = \tilde{A}_t\,\delta^t_\mu$, and

$$\begin{aligned}
\tilde{g}_{tt} &= \frac{g_{tt}}{[1 + (1 + g_{tt})\sinh^2\alpha]^2}\,,\\
\tilde{A}_t &= -\frac{(1 + g_{tt})\sinh 2\alpha}{2\sqrt{2}\,[1 + (1 + g_{tt})\sinh^2\alpha]}\,,\\
\exp(-2\tilde{\varphi}) &= [1 + (1 + g_{tt})\sinh^2\alpha]\exp(-2\varphi)\,.
\end{aligned} \tag{13.2.3}$$

The solution in the string frame obtained from the Schwarzschild metric is

$$ds^2 = -\left(1 - \frac{2m}{\tilde{r}}\right)\left(1 + \frac{2m}{\tilde{r}}\sinh^2\alpha\right)^{-2} dt^2 + \left(1 - \frac{2m}{\tilde{r}}\right)^{-1} d\tilde{r}^2 + \tilde{r}^2\,d\omega^2, \tag{13.2.4}$$

$$A_t = -\frac{m\sinh 2\alpha}{\sqrt{2}\,[\,\tilde{r} + 2m\sinh^2\alpha]}\,, \qquad \exp(-2\varphi) = 1 + \frac{2m}{\tilde{r}}\sinh^2\alpha\,. \tag{13.2.5}$$

The metric has a horizon at $\tilde{r} = 2m$ and a curvature singularity at $\tilde{r} = 0$, where $F_{\mu\nu}\,F^{\mu\nu}$ is also divergent. Since $\exp\varphi \to 0$ near $r = 0$, the effective string coupling becomes weak. At large distances $\varphi \to 0$.

Rescaling the metric and introducing a new radial coordinate $r = \tilde{r} + 2m\sinh^2\alpha$, one obtains the solution in the Einstein frame

$$ds_E^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r\left(r - \frac{Q^2}{M}\right)d\omega^2, \qquad (13.2.6)$$

$$F_{rt} = \frac{Q}{r^2}, \qquad \exp(2\varphi) = 1 - \frac{Q^2}{Mr}, \qquad (13.2.7)$$

where the physical mass M and charge Q are

$$M = m\cosh^2\alpha, \qquad Q = \frac{m}{\sqrt{2}}\sinh 2\alpha. \qquad (13.2.8)$$

The metric (13.2.6) in the (r,t)-plane coincides with the Schwarzschild metric, while the area of the sphere of radius r is smaller, and vanishes for $r = Q^2/M$.

The magnetically charged black holes can be obtained from the electrically charged solution by means of the duality transformation

$$g_{\mu\nu}^E \to g_{\mu\nu}^E, \qquad \varphi \to -\varphi, \qquad F_{\mu\nu} \to \exp(-2\varphi)\,\frac{1}{2}\,e_{\mu\nu}{}^{\alpha\beta}\,F_{\alpha\beta}. \qquad (13.2.9)$$

The metric in the Einstein frame remains unchanged, but because the dilaton fields is transformed, one has the following form of the metric for a magnetically charged black hole in the string frame:

$$ds^2 = -\frac{\left(1 - \frac{2M}{r}\right)}{\left(1 - \frac{Q^2}{Mr}\right)}dt^2 + \frac{dr^2}{\left(1 - \frac{2M}{r}\right)\left(1 - \frac{Q^2}{Mr}\right)} + r^2\,d\omega^2. \qquad (13.2.10)$$

As one approaches the singularity $r = Q^2/M$, the surface area of a sphere $r = \text{const}$ remains finite.

The solutions for dilaton black holes were obtained by Gibbons (1982) [see also Gibbons and Maeda (1988), Garfinkle, Horowitz, and Strominger (1991)].

There exists also an exact solution for a black hole with both electric (Q_E) and magnetic (Q_M) charges [Gibbons and Maeda (1988), Kallosh *et al.* (1992)]. In the Einstein frame it takes the form

$$ds_E^2 = -\frac{(r - r_-)(r - r_+)}{r^2 - r_0^2}dt^2 + \frac{r^2 - r_0^2}{(r - r_-)(r - r_+)}dr^2 + (r^2 - r_0^2)\,d\omega^2, \quad (13.2.11)$$

$$\exp(2\varphi) = \frac{r + r_0}{r - r_0}. \qquad (13.2.12)$$

Here

$$r_0 = \frac{Q_M^2 - Q_E^2}{2M}, \qquad r_\pm = M \pm \sqrt{M^2 + r_0^2 - Q_E^2 - Q_M^2}. \qquad (13.2.13)$$

An uncharged rotating black hole in the dilaton gravity has the Kerr metric. Its charged version can be obtained by applying a generalization of the transformation (13.2.3) to the Kerr metric [Sen (1992b)]. For more detailed discussion of dilaton black holes and their properties, see the review by Horowitz (1992). On the uniqueness theorem for dilaton black holes see Masood-ul-Alam (1993).

13.2.2 Black holes with non-Abelian hair

Colored black holes

It is quite easy to obtain solutions of the Einstein-Yang-Mills equations with Abelian hair (see e.g., Frolov (1973)). Yasskin (1975) proved that for every solution of the Einstein-Maxwell equations outside the sources, an $(N-1)$-parameter family of exact solutions of the Einstein-Yang-Mills equations can be constructed for the N-parameter gauge group. This family with Abelian hair has the same metric as the original solution.

After Bartnik and Mckinnon (1988) discovered a non-trivial self-consistent solution of the Einstein-Yang-Mills equations, a variety of black hole solutions with non-Abelian hair have been found. In this section we describe some of these solutions.

We begin with the simplest case of Einstein-Yang-Mills equations with the $SU(2)$ gauge group described by the action

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,\left[R - \frac{1}{g^2}\,\mathrm{Tr}(\mathbf{F}^2)\right]. \tag{13.2.14}$$

Here $\mathbf{F}$ is the Yang-Mills curvature (field strength), $\mathbf{F} = d\mathbf{A} + \mathbf{A}\wedge\mathbf{A}$, and g is a self-coupling constant of the Yang-Mills field. Consider a spherically symmetric metric

$$ds^2 = -\left(1 - \frac{2m(r)}{r}\right)B^2(r)\,dt^2 + \left(1 - \frac{2m(r)}{r}\right)^{-1} dr^2 + r^2\,d\omega^2\,, \tag{13.2.15}$$

and choose the following ansatz for the spherically symmetric Yang-Mills field

$$\mathbf{A} = (1 + w(r))[-\boldsymbol{\tau}_\phi\,d\theta + \boldsymbol{\tau}_\theta \sin\theta\,d\phi]\,. \tag{13.2.16}$$

Here

$$\begin{aligned}
\boldsymbol{\tau}_r &= \frac{1}{2i}\,[\boldsymbol{\sigma}_1 \sin\theta\cos\phi + \boldsymbol{\sigma}_2 \sin\theta\sin\phi + \boldsymbol{\sigma}_3\cos\theta]\,,\\
\boldsymbol{\tau}_\theta &= \frac{1}{2i}\,[\boldsymbol{\sigma}_1 \cos\theta\cos\phi + \boldsymbol{\sigma}_2 \cos\theta\sin\phi - \boldsymbol{\sigma}_3\sin\theta]\,,\\
\boldsymbol{\tau}_\phi &= \frac{1}{2i}\,[-\boldsymbol{\sigma}_1 \sin\phi + \boldsymbol{\sigma}_2 \cos\phi]\,,
\end{aligned} \tag{13.2.17}$$

and $\boldsymbol{\sigma}_i$ are the Pauli spin matrices. The field strength $\boldsymbol{F}$ for this ansatz is

$$\boldsymbol{F} = -w'\,\boldsymbol{\tau}_\phi\,dr\wedge d\theta + w'\,\boldsymbol{\tau}_\theta\,dr\wedge\sin\theta\,d\phi - (1-w^2)\,\boldsymbol{\tau}_r\,d\theta\wedge\sin\theta\,d\phi\,. \quad (13.2.18)$$

One chooses the boundary condition $w \to \pm 1$ at infinity which guarantees boundedness of the energy of the system. Besides this, one imposes the conditions of asymptotic flatness: $m(r) \to M = \text{finite}$, $B(r) \to 1$ as $r \to \infty$, and the condition for existence of a horizon $2m(r_H) = r_H$ and $B(r_H) < \infty$.

There exist trivial solutions of the system of Einstein-Yang-Mills equations such as the Schwarzschild black hole ($w = \pm 1$) or the magnetic Reissner-Nordström black hole ($w = 0$). For the latter solution the structure of the Yang-Mills hair is effectively Abelian.

Besides these solutions, there exists a discrete family of non-trivial (definitely non-Abelian) solutions (*colored black holes*) [Volkov and Gal'tsov (1989, 1990), Künzle and Masood-ul-Alam (1990), Bizon (1990)]. For these solutions $w(r)$ is a non-trivial function that has an asymptotic behavior $|w| \sim 1 - c/r,\ c > 0$. The solutions of this family are parametrized by an integer number n: the number of nodes of the Yang-Mills potential. The mass of the solution increases with increasing n. All solutions are unstable with respect to radial linear perturbations. The number of unstable modes increases as the node number increases [Sudarsky and Wald (1992)]. Einstein-Yang-Mills black holes with the gauge group $SU(N)$ were discussed by Künzle (1991, 1994) and Kleihaus, Kunz, and Sood (1997) (see also references therein). Solutions for slowly rotating colored black holes were obtained by Volkov and Straumann (1997).

There are known dilatonic modifications of colored black holes [Donetz and Gal'tsov (1993), Torii and Maeda (1993)]. They are solutions for the action

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\left[R - (\nabla\sigma)^2 - \frac{1}{g^2}\,e^{-\alpha\sigma}\,\mathrm{Tr}(\mathbf{F}^2)\right]. \quad (13.2.19)$$

This type of action arises from different types of unified theories, including a superstring theory ($\alpha = 1$), and 5-dimensional Kaluza-Klein theory ($\alpha = \sqrt{3}$).

The spacetime metric and the Yang-Mills potential ansatz are the same as earlier. In addition to the above described boundary conditions it is required that $\sigma \to 0$ as $r \to \infty$, and $\sigma(r_H)$ is finite. The properties of *dilatonic colored black holes* are similar to the properties of the ordinary colored black holes [see Kleihaus, Kunz and Sood (1996) and references therein]. A paper by Lavrelashvili (1997) contains review of black hole solutions with non-Abelian hair.

Skyrme black holes

Another interesting class of solutions for black holes with non-Abelian hair was obtained by considering the nonlinear sigma model interacting with the gravitational field. The simplest form of the nonlinear sigma model which has stable solitonic solutions was given by Skyrme (1961), who used this model to study low energy behavior of pions. The basic field variable of the nonlinear sigma model of Skyrme is an $SU(2)$ valued function $U(x)$. The corresponding Einstein-Skyrme action is

$$W = \int d^4x \sqrt{-g} \left[\frac{1}{16\pi} R - \frac{1}{32g_S^2} \operatorname{Tr}(\boldsymbol{F}^2) + \frac{1}{4} f_S^2 \operatorname{Tr}(\boldsymbol{A}^2) \right], \tag{13.2.20}$$

where $\boldsymbol{F} = \boldsymbol{A} \wedge \boldsymbol{A}$, and $\boldsymbol{A} = \boldsymbol{U}^\dagger \nabla \boldsymbol{U}$. The metric for the *Skyrme black hole* is of the form (13.2.15), while for U one uses the so-called *hedgehog ansatz*

$$U = \exp[\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}\, \chi(r)], \tag{13.2.21}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices, and $\hat{\mathbf{r}}$ is a unit radial vector. Skyrme black hole solutions were obtained by Luckock and Moss (1986), Droz, Heusler, and Straumann (1991), and Bizon and Chmaj (1992). There exist two particle-like solutions of the Skyrme model. One can exist in the absence of gravity, and the other is a soliton-like solution stabilized by gravity. The two branches of black hole solutions, which start from these particles, merge at some critical radius of the horizon. Beyond this critical point where the mass and the horizon radius reach their maximal value, there are no non-trivial black holes. Below the critical point, for each value of the mass their exist two different Skyrme black holes. The one that possesses the smaller value of r_H is unstable with respect to radial perturbations [Luckock and Moss (1986), Heusler, Droz and Straumann (1991)].

Proca black holes

The action for the massive Yang-Mills field (Proca field) interacting with gravity is

$$W = \int d^4x \sqrt{-g} \left[\frac{1}{16\pi} R - \frac{1}{4\pi g^2} \left(\frac{1}{4} \operatorname{Tr}(\boldsymbol{F}^2) - \frac{\mu^2}{2} \operatorname{Tr}(\boldsymbol{A}^2) \right) \right]. \tag{13.2.22}$$

Here μ is the mass of the Yang-Mills field, and $\mathbf{F} = d\mathbf{A} + \mathbf{A} \wedge \mathbf{A}$. *Proca black hole* is a solution with metric (13.2.15), vector potential (13.2.16), and boundary conditions the same as for a colored black hole. There exist two branches of Proca black hole solutions. One gives a particle-like solution in the limit when $r_H \to 0$ which can exist even in the absence of gravity, while the other gives a self-gravitating particle-like solution. The two branches of black hole solutions that start from these particles, merge at some critical gravitational radius r_H, where a Proca black hole has a maximal mass M and maximal size of the horizon r_H. Beyond this point there are no non-trivial solutions [Greene, Mathur, and O'Neill (1993), Torii, Maeda, and Tachizawa (1995)]. On the stability properties of Proca black holes see [Torii, Maeda, and Tachizawa (1995)].

Sphaleron black holes

A *sphaleron black hole* is a black-hole-like solution of the theory

$$W = \int d^4x \sqrt{-g} \left[\frac{1}{16\pi} R - \frac{1}{16\pi g^2} \operatorname{Tr}(\boldsymbol{F}^2) - \frac{1}{4\pi} (D_\mu \Phi)^\dagger (D_\mu \Phi) - \frac{1}{4\pi} V(\Phi) \right], \tag{13.2.23}$$

where

$$D_\mu = \partial_\mu + \boldsymbol{\tau} \cdot \boldsymbol{A}_\mu , \qquad \boldsymbol{F} = d\boldsymbol{A} + \boldsymbol{A} \wedge \boldsymbol{A} , \tag{13.2.24}$$

$$V(\Phi) = \lambda \left(\Phi^\dagger \Phi - \Phi_0^2 \right)^2 . \tag{13.2.25}$$

$\boldsymbol{A}$ is the $SU(2)$ Yang-Mills potential; $\boldsymbol{F}$ is the field strength, and $\boldsymbol{\Phi}$ is the Higgs field. For the metric and the Yang-Mills field one uses the same ansatz and boundary conditions as for the colored black hole solutions. For the complex doublet of Higgs fields one uses the ansatz

$$\boldsymbol{\Phi}(x) = \frac{1}{\sqrt{2}} e^{-2\pi \hat{\boldsymbol{r}} \cdot \boldsymbol{\tau}} \begin{pmatrix} 0 \\ r^{-1} \Phi(r) \end{pmatrix} . \tag{13.2.26}$$

The boundary condition at infinity for the Higgs field is $\Phi(\infty) = \Phi_0$. Sphaleron black holes are always unstable.

Black holes in monopoles

Magnetic monopoles were first discussed by Dirac (1931) in the framework of Maxwell electrodynamics. 't Hooft (1974) and Polyakov (1974) discovered a monopole as a solution of the $SO(3)$ Yang-Mills-Higgs theory. Solutions describing gravitating monopoles were obtained by Cho and Freund (1975) and van Nieuwenhuizen *et al.* (1976). Later, Breitenlohner, Forgács, and Maison (1992), Lee, Nair, and Weinberg (1992a-c), and Ortiz (1992) found solutions describing a *monopole black hole.*

The $SO(3)$ Einstein-Yang-Mills-Higgs theory is described by the action

$$W = \int d^4x \sqrt{-g} \left[\frac{1}{16\pi} R - \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} - \frac{1}{2} D_\mu \Phi^a D^\mu \Phi^a - \frac{\lambda}{4} \left(\Phi^a \Phi^a - v^2 \right)^2 \right], \tag{13.2.27}$$

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - e\, \epsilon^{abc} A^b_\mu A^c_\nu , \tag{13.2.28}$$

$$D_\mu \Phi^a = \partial_\mu \Phi^a + e\, \epsilon^{abc} A^b_\mu \Phi^c . \tag{13.2.29}$$

Here A^a_μ and $F^a_{\mu\nu}$ $(a = 1,2,3)$ are the $SO(3)$ Yang-Mills field potential and field strength, respectively. Φ^a is a triplet of real Higgs fields. For monopole black hole solutions one uses the metric (13.2.15) and the hedgehog ansatz for the Yang-Mills and Higgs fields

$$\Phi^a = v\,\hat{r}^a\,h(r)\,, \qquad A^a_i = \omega^c_i\,\epsilon^{cab}\,\hat{r}^b\,\frac{1-w(r)}{e\,r}\,, \qquad A^a_0 = 0\,. \tag{13.2.30}$$

Here $\hat{r}^a$ is a unit radial vector in the internal space, and ω^c_i is a triad in this space. The boundary conditions imposed at infinity $r \to \infty$ are

$$m(\infty) = \text{finite}\,, \quad B(\infty) = 1\,, \quad h(\infty) = 1\,, \quad w(\infty) = 0\,. \tag{13.2.31}$$

These boundary conditions are the same as for a regular gravitating monopole solution. The black hole monopole is singled out as the solution that has a *regular* horizon $2m(r_H) = r_H$ at r_H.

For arbitrary values of v and λ there always exists a trivial Reissner-Nordström solution

$$m(r) = M - \frac{2\pi}{e^2\,r}\,, \quad B = 1\,, \quad w(r) = 0\,, \quad h(r) = 1\,. \tag{13.2.32}$$

Non-Abelian black holes can exist when the gravitational radius $r_g \sim M$ is less than the monopole size $1/(e\,v)$. For given parameters v and λ, there may be one or more (up to three) non-Abelian black hole solutions in addition to the Reissner-Nordström one. On the stability of the monopole black hole solutions, see Tachizawa, Maeda, and Torii (1995). It is quite interesting that for some range of the parameters a trivial Reissner-Nordström solution is unstable with respect to linear radial perturbations. As a result of this instability, a non-Abelian monopole black hole is formed [Lee, Nair, and Weinberg (1992a-c)].

For a general review of black holes with non-Abelian hair and the analysis of their stability, see [Torii, Maeda, and Tachizawa (1995) and Tachizawa, Maeda, and Torii (1995)].

13.2.3 Quantum hair

Axionic black holes

The simplest action for the axionic black hole is

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,\left[R + H_{\mu\nu\lambda}\,H^{\mu\nu\lambda}\right], \tag{13.2.33}$$

where H is the three-form strength of the Kalb-Ramond axion field, and two-form B is its potential.

In general, axions have axial fermionic couplings that induce anomalous coupling to gauge bosons which are of the Chern-Simons form. It means that the axion field strength H is of the form

$$H = dB + \omega_L - \omega_Y , \tag{13.2.34}$$

where ω_L and ω_Y are the Chern-Simons forms for the Lorentz spin connection and Yang-Mills gauge connection, respectively,

$$\omega_L = \mathrm{Tr}(\omega \wedge d\omega + \frac{2}{3}\omega \wedge \omega \wedge \omega) , \quad \omega_Y = \mathrm{Tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A) , \tag{13.2.35}$$

with $\omega_{ab\mu} = -e^{\alpha}_{b}\, e_{a\alpha;\mu}$ the spin connection and e^{μ}_{a} the orthonormal tetrad. The following Bianchi identity for H follows from (13.2.34) and (13.2.35)

$$dH = \mathrm{Tr}(R \wedge R) - \mathrm{Tr}(F \wedge F) . \tag{13.2.36}$$

Let us define a one-form ${}^*H$ dual to H as

$$({}^*H)_\mu = e_{\mu\alpha\beta\gamma}\, H^{\alpha\beta\gamma} . \tag{13.2.37}$$

Then the equation of motion $H^{\alpha\beta\gamma}{}_{;\gamma} = 0$ implies that $({}^*H)_\mu = a_{;\mu}$, while the Bianchi identity (13.2.36) gives

$$\Box a = J , \tag{13.2.38}$$

where the source term J is proportional to $e^{\alpha\beta\mu\nu}\,(R_{\rho\lambda\alpha\beta}\, R_{\mu\nu}{}^{\rho\lambda} - F_{\alpha\beta}\, F_{\mu\nu})$.

Campbell *et al.* (1990) attracted attention to the fact that for a rotating black hole the *Hirzebruch signature density* $\mathrm{Tr}(R \wedge R) \equiv e^{\alpha\beta\mu\nu}\, R_{\rho\lambda\alpha\beta}\, R_{\mu\nu}{}^{\rho\lambda}$ does not vanish, so that in the theory with a non-minimally coupled Kalb-Ramond axion field there exists *axion hair*. This was the first known example of stable, dynamical, non-gauge hair exhibited by a black hole. Campbell *et al.* (1990) calculated the axion field for a slowly rotating black hole and showed that it has the following asymptotic behavior

$$a(r, \theta) \sim \frac{5a\cos\theta}{12Mr^2} , \tag{13.2.39}$$

which exhibits the r^{-2} long-range fall off one would expect of a dipole.

In a similar way, if a black hole is non-rotating, but it posesses both electric and magnetic charges (a so-called *dyon black hole*), then the non-vanishing anomalous term $F_{\alpha\beta}\,{}^*F^{\alpha\beta} \equiv e^{\alpha\beta\mu\nu}\, F_{\alpha\beta}\, F_{\mu\nu}$ plays the role of a source for the axion field, so that the black hole again has axion hair [Campbell, Kaloper, and Olive (1991)]. On the generalization of these results to the case of black holes in string theory, see [Campbell, Kaloper, and Olive (1992)].

The axion hair connected with the rotation and charges of a black hole can be generated by other mechanisms. For example, this happens in the theory with the action

$$W = \frac{1}{16\pi} \int d^4x\, \sqrt{-g}\, \left[R - F_{\mu\nu}\, F^{\mu\nu} - 2(\nabla a)^2 + \lambda\, a\, F_{\alpha\beta}\,{}^*F^{\alpha\beta} \right] . \tag{13.2.40}$$

If a static black hole has both electric and magnetic charges, so that $F_{\alpha\beta}\,{}^*F^{\alpha\beta} \neq 0$, then the last term in the action generates a source for the axion field a. As a result, the black hole has axion hair [Lee and Weinberg (1991)]. Quite interesting that a rotating black hole without charges can also have axion hair in the theory (13.2.40) [Reuter (1992)]. This result follows from the existence of the bosonic chiral anomaly discovered by Dolgov, Khriplovich, and Zakharov (1987) [see also Dolgov *et al.* (1989)]. Namely, when an Abelian gauge field is quantized in the background with a non-vanishing Hirzebruch signature density, then the vacuum expectation of the chiral current

$$K^\mu = e^{\alpha\beta\mu\nu}\, A_\beta\, F_{\mu\nu} \tag{13.2.41}$$

is not conserved, and the pseudoscalar $F_{\alpha\beta}\,{}^*F^{\alpha\beta}$ acquires a vacuum expectation value:

$$\langle K^\mu\rangle_{;\mu} = \frac{1}{2}\langle F_{\alpha\beta}\,{}^*F^{\alpha\beta}\rangle = \frac{1}{192\pi^2}\, e^{\alpha\beta\mu\nu}\, R_{\rho\lambda\alpha\beta}\, R_{\mu\nu}^{\ \ \rho\lambda}\,. \tag{13.2.42}$$

One can describe the axion field generation as the two-step process. The black hole rotation generates $\langle F_{\alpha\beta}\,{}^*F^{\alpha\beta}\rangle$ by means of the bosonic chiral anomaly, while the latter plays the role of the source for the axion field [Reuter (1992)].

Quantum axion hair

Axion black holes as well as dilaton black holes and black holes with non-Abelian fields can be considered as examples of black holes with non-trivial hair. Matter fields associated with hair have a non-vanishing stress-energy tensor which modifies the vacuum metric. Besides this type of hair, a black hole might have additional charges that do not modify the metric, and hence cannot be discovered by studying the gravitational interaction of the black hole. An *axion quantum hair* is an example first considered by Bowick *et al.* (1988).

The starting point is again the action for the axionic black hole

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\,\left[R + H_{\mu\nu\lambda}\,H^{\mu\nu\lambda}\right]. \tag{13.2.43}$$

But now we assume that the axions are minimally coupled so that $H = dB$, where the two-form B is a potential. A black hole of mass M and axion charge q is a solution

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{dr^2}{1 - \dfrac{2M}{r}} + r^2\,d\omega^2\,, \tag{13.2.44}$$

$$B_{\mu\nu}\,dx^\mu \wedge dx^\nu = \frac{q}{4\pi}\,\sin\theta\,d\theta \wedge d\phi\,. \tag{13.2.45}$$

It is easy to verify that for this axion potential B the field H vanishes, and hence the axion field equations are trivially satisfied. Moreover, the stress-energy tensor of the axion field vanishes so that the Schwarzschild metric (13.2.44) is evidently a solution of the Einstein equations with the ansatz (13.2.45) for the axion field. It is possible to show that this solution is the only static solution to the Einstein-axion field equations which has a regular event horizon, is asymptotically flat, and has finite charge q [Bowick *et al.* (1988)].

The axion charge

$$q = \int_{S^2} B_{\mu\nu}\, d\sigma^{\mu\nu} \tag{13.2.46}$$

can be measured neither by gravitational experiments nor by any *local* measurements of the axion field. Nevertheless, this new additional black hole characteristic which is of pure topological origin can be determined by nonlocal Aharonov-Bohm type experiments. For example, a string interacting with the axion field would "feel" the axion charge if its world sheet encircles the black hole. In particular, one can measure the charge, by creating a pair of strings at some point, allowing one of them to lasso the hole and the other to avoid the hole, and then interfering with one another. Since this Aharonov-Bohm effect is quantum in nature, one can call the axion field *quantum hair.*

Black holes with discrete gauge charges

A black hole can carry quantum hair associated with discrete gauge charges. It may happen when the local continuous gauge symmetry breakdown leaves unbroken a residual discrete subgroup of the gauge group. Charges connected with the unbroken discrete symmetry can reside on a black hole providing it with new physical characteristics [Krauss and Wilczek (1989), Coleman, Preskill, and Wilczek (1991, 1992)].

Consider a $U(1)$ gauge theory that contains two scalar fields η and ξ carrying charge pe and e, respectively, and described by the action

$$W = -\int d^4x\,\sqrt{-g}\left[\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + |(\partial_\mu + i\,p\,e A_\mu)\,\eta\,|^2 + |(\partial_\mu + i\,e\,A_\mu)\,\xi\,|^2 \right.$$
$$\left. + \frac{\lambda}{4}(\,|\,\eta\,|^2 - v^2)^2 + V_{\text{int}}(\,|\,\eta\,|, |\,\xi\,|\,)\right], \tag{13.2.47}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{13.2.48}$$

Field η undergoes a condensation

$$\langle\eta(x)\rangle = v\,. \tag{13.2.49}$$

Before the condensation the theory is invariant under the local gauge transformations

$$\eta(x) \to e^{ipe\lambda(x)}\,\eta(x)\,, \quad \xi(x) \to e^{ie\lambda(x)}\,\xi(x)\,, \quad A_\mu \to A_\mu - \partial_\mu\lambda\,. \tag{13.2.50}$$

After the condensation this continuous symmetry is broken since the homogeneous ground state is invariant only when λ is an integer multiple of $2\pi/pe$. Nevertheless, residual transformations still act non-trivially on ξ, and their actions generate the discrete group $\mathbb{Z}_p$. Since the Higgs field η can only screen charges which are a multiple of pe, there exists a conserved charge mod(pe) associated with the observable [Krauss and Wilczek (1989)]

$$\exp\left(\frac{2\pi i\, Q}{pe}\right) = \exp\left(\frac{2\pi i}{pe}\int \boldsymbol{E}\cdot d\boldsymbol{s}\right). \tag{13.2.51}$$

If a two-dimensional integral is taken over the black hole surface, one obtains the discrete gauge charge of the black hole. This charge can be measured (at least in principle) in the following gedanken experiment. The broken symmetry theory of field η contains topologically stable strings with magnetic flux $2\pi/pe$. A particle ξ with a charge e can test this magnetic flux by means of the Aharonov-Bohm type scattering. Similarly, the gauge discrete charge of a black hole can be detected by the Aharonov-Bohm interaction of a cosmic string if its world sheet encircles the black hole.

Even in the absence of actual strings, this effect changes the phases in the amplitudes for virtual processes in which a string world sheet envelops the charge. Suppose a virtual string is nucleated at a point of the event horizon, sweeps around it, and shrinks and annihilates after this. The motion of a string carrying the magnetic flux generates an electric field in the radial direction. Averaging over all possible points of nucleation cancels the magnetic field, but the electric field survives. As a result of such virtual processes, there is a non-vanishing electric field in the vicinity of the horizon [Coleman, Preskill, and Wilczek (1991, 1992)]. Its amplitude is proportional to $\sin(2\pi Q/pe\hbar)\exp(-S_{\text{vortex}}/\hbar)$, where S_{vortex} is the Euclidean action for a string sweeping around the horizon. The field strength falls off exponentially as $\exp(-r/l)/r$, where l is the thickness of the string. Likewise, in the theories that allow magnetic monopoles a black hole might have a discrete magnetic charge and a short-range magnetic field [Coleman, Preskill, and Wilczek (1991, 1992)].

Primary and secondary hair

All black hole solutions discussed above can be considered as different counterexamples to the original no hair conjecture. A strong interpretation of this conjecture would say that black holes settle down to the Kerr-Newman one, characterized completely by mass, angular momentum, and electric charge.

The conjecture in this form is certainly not correct. By modifying the theory and adding extra fields, one can obtain black holes which possess additional physical characteristics besides gravitational and electromagnetic fields. As we have seen, in the presence of continuous gauge fields a black hole can have fields generated by the corresponding conserved charges. This observation implies the following natural generalization of the strong no hair conjecture: There can be no non-trivial fields outside the black hole aside from the gravitational, electromagnetic, and more general continuous gauge fields.

It is easy to see that even this modified form of the no hair conjecture cannot be correct. The dilaton black hole (Section 13.2.1) is an evident counter-example. The dilaton scalar field (which is not a gauge field) is present outside the horizon and the metric is modified by the presence of this field. The mechanism of generation of dilaton hair is quite simple. Because of the term $\exp\varphi F^2$ in the action, in the presence of an electric field F^2 plays the role of the source for the dilaton. As a result, the dilaton field is uniquely specified by giving the mass M and the electric charge Q of the black holes. Such hair can be called *induced* or *secondary hair* [Coleman, Preskill, and Wilczek (1991, 1992)]. Secondary hair of black holes is easily generated e.g., by perturbative quantum corrections to the matter effective action. These corrections contain a direct interaction of the curvature with matter fields. The spacetime curvature plays the role of source for matter fields even in the absence of real charges.

Another, weaker interpretation of the no hair conjecture is the statement that the properties of black holes are completely determined, within any given theory, by the values of its mass, angular momentum, and continuous gauge charges. The weaker form of the no hair conjecture is also violated. Black holes with quantum hair, considered in the present section, give counter-examples. Axionic charge or discrete gauge charge are additional parameters which expand the space of states of black holes. The electric or magnetic short-range quantum hair associated with discrete gauge charges are so-called *primary hair* [Coleman, Preskill, and Wilczek (1991, 1992), Krauss and Liu (1997)].

13.3 Lower-Dimensional Black Holes

Till now we have considered a gravitational theory and black holes in a four-dimensional spacetime since it is what can be (at least in principle) directly related to observations. Nevertheless, it is instructive to study what might be physical properties of the world in the case if the number of spacetime dimensions were different from four. There are quite a lot of papers where black hole solutions were obtained and studied in the spacetimes with the number of dimensions either lower or higher than four. We briefly describe obtained results for two and three dimensional black holes in this section. Properties of higher dimensional black holes are considered in the next section.

13.3.1 Three-dimensional black holes

The Einstein-Hilbert action in $(2+1)$-dimensional spacetime is

$$W[g] = \frac{1}{16\pi G}\int d^3x\,\sqrt{-g}\,(R - 2\Lambda + L_m)\,. \tag{13.3.1}$$

Λ is a cosmological constant, and L_m is the matter Lagrangian. We omit the surface term which is not important for our considerations and put as before $c = 1$. The gravitational constant G has the dimension of *length*. In this section we use units in which $8G = 1$. The gravitational equations have the form $(a, b = 0, 1, 2)$

$$R_{ab} - \frac{1}{2}\,g_{ab}\,(R - 2\Lambda) = \pi\,T_{ab}. \tag{13.3.2}$$

In a $(2+1)$-dimensional spacetime the number of independent components of the Riemann tensor (six) is the same as for the Ricci tensor R_{ab}, so that the Riemann tensor is completely determined by R_{ab}:

$$R_{abcd} = R_{ac}\,g_{bd} - R_{ad}\,g_{bc} + R_{bd}\,g_{ac} - R_{bc}\,g_{ad} - \frac{1}{2}R\,(g_{ac}\,g_{bd} - g_{ad}\,g_{bc})\,. \tag{13.3.3}$$

This means that by specifying the matter distribution T_{ab}, one specifies also the complete curvature of spacetime, and there are no propagating degrees of freedom in $(2+1)$-gravity. In particular, outside the matter where $R_{ab} = 0$, the spacetime is locally flat, or has constant curvature Λ if there is a cosmological constant.

These properties of three-dimensional general relativity imply many simplifications, and for this reason it is often used as a toy model for study the problems of classical and quantum gravity [Deser, Jackiw, and 't Hooft (1984), Deser and Jackiw (1988), 't Hooft (1988), Witten (1988,1989)]. It might be surprising that this toy model is of any interest, especially because 3-gravity has no Newtonian limit [Barrow, Burd, and Lancaster (1986)]. Nevertheless, study of three-dimensional gravity has brought a number of interesting results. One of them is the unexpected existence of black hole solutions [Bañados, Teitelboim, and Zanelli (1992, 1993)].

The BTZ black hole is a solution with $T_{ab} = 0$ and $\Lambda = -1/l^2$, and it is described by the metric

$$ds^2 = -\Phi\,dt^2 + \frac{dr^2}{\Phi} + r^2(d\phi - \omega\,dt)^2\,, \tag{13.3.4}$$

where

$$\Phi = -M + \frac{r^2}{l^2} + \frac{J^2}{4r^2}\,, \qquad \omega = -\frac{J}{2r^2}\,, \qquad |\,J\,| \le Ml\,. \tag{13.3.5}$$

$M > 0$ plays the role of the black hole mass, and J is the angular momentum. The spacetime has two horizons at $r = r_\pm$, where

$$r_\pm^2 = \frac{Ml^2}{2}\left\{1 \pm \left[1 - \left(1 - \frac{J}{Ml}\right)^2\right]^{1/2}\right\}. \tag{13.3.6}$$

Metric (13.3.4) allows two Killing vectors $\xi^a_{(t)}\,\partial_a = \partial_t$ and $\xi^a_{(\phi)}\,\partial_a = \partial_\phi$. The vector $\xi^a_{(t)}$ becomes null at the static-limit surface

$$r_{\rm stat} = M^{1/2}\,l \tag{13.3.7}$$

located outside the event horizon r_+. The surface gravity of the event horizon is $\kappa = (r_+^2 - r_-^2)/(l^2\,r_+)$.

It is remarkable that a BTZ black hole, being a solution of the field equations with negative cosmological constant, is locally isometric to anti-de Sitter space. One important difference between the BTZ solution and anti-de Sitter space is the topology. Anti-de Sitter spacetime can be embedded in a flat space of one higher dimension with the metric

$$ds^2 = -\,dT^2 - dU^2 + dX^2 + dY^2\,. \tag{13.3.8}$$

The induced metric on the hypersurface

$$-T^2 - U^2 + X^2 + Y^2 = -\,l^2 \tag{13.3.9}$$

has constant negative curvature, and hence it is a solution of the vacuum Einstein equations with negative cosmological constant $\Lambda = -1/l^2$. The embedded spacetime is periodic in the time direction. To obtain anti-de Sitter space, one uses its universal cover, obtained by "unwinding" about the XY-axis.

The embedded surface (13.3.9) can be locally parametrized as

$$\begin{aligned} T &= \sqrt{\frac{r^2}{M} - l^2}\,\sinh\frac{\sqrt{M}}{l}\,t\,, \quad U = \frac{r}{\sqrt{M}}\cosh\left(\sqrt{M}\,\phi\right)\,, \\ X &= \sqrt{\frac{r^2}{M} - l^2}\,\cosh\frac{\sqrt{M}}{l}\,t\,, \quad Y = \frac{r}{\sqrt{M}}\sinh\left(\sqrt{M}\,\phi\right)\,. \end{aligned} \tag{13.3.10}$$

In these coordinates the induced metric is

$$ds^2 = -\left(-M + \frac{r^2}{l^2}\right)dt^2 + \frac{dr^2}{-M + \dfrac{r^2}{l^2}} + r^2\,d\phi^2\,. \tag{13.3.11}$$

This is the metric of a non-rotating BTZ black hole provided one makes the identification $\phi \equiv \phi + 2\pi$. In fact, this identification, which makes the topology of the spacetime different from the topology of the anti-de Sitter space, distinguishes the BTZ metric globally. Though the mass parameter M can be excluded from the metric (13.3.11) by rescaling the coordinates t and r, this rescaling changes the period of ϕ and makes it dependent on M. In other words, appropriate identification not only changes the topology but also introduces a scale parameter. The metric of a rotating BTZ black hole can be obtained in a similar way by choosing a proper identification in anti-de Sitter space [Bañados, Teitelboim, and Zanelli (1992,1993), Carlip (1995)]. One can also generate multi-black-hole solutions by using more general types of identifications [Brill (1996)]. For a review of (2 + 1)-dimensional black holes and their properties, see [Carlip (1995)].

13.3.2 Two-dimensional black holes

In two dimensions the Einstein action is a topological invariant, and hence it has no dynamical content. The simplest way to obtain dynamics is to include a dilaton field. An example of such a theory naturally arises in the framework of the usual four-dimensional Einstein gravity if we restrict ourselves by considering spherically symmetric spacetimes. The reduced action for the spherically symmetric ansatz is (see Appendix B.2)

$$W[g,r] = \frac{c^3}{4G}\int d^2x\,\sqrt{-g}\,\left[r^2R + 2(\nabla r)^2 + 2\right]. \tag{13.3.12}$$

Here g_{AB} and R are two-dimensional metric and curvature, respectively, $A, B = 0, 1$. We restored c and G in the formula and omit the surface term. Let us choose some length scale parameter l and define $r = l\exp(-\varphi)$. Then the action takes the form

$$W[g,\varphi] = \frac{1}{4G'}\int d^2x\,\sqrt{-g}\,e^{-2\varphi}\left[R + 2(\nabla\varphi)^2 + l^{-2}\,V(\varphi)\right], \tag{13.3.13}$$

where $V(\varphi) = 2\exp(2\varphi)$. $G' = G/(l^2c^3)$ plays the role of the two-dimensional Newton constant.

Since the curvature interacts directly with the dilaton field φ, its variation with respect to the metric g_{AB} does not vanish identically. The corresponding equations are of the form $\varphi_{|AB} \sim g_{AB}$. According to the generalized Birkhoff's theorem (see Appendix B.3), these equations guarantee the existence of a Killing vector $\xi^A = e^{AB}\varphi_{|B}$. It is evident that the only solution to the "dilaton gravity" (13.3.13) is the two-dimensional sector of the Schwarzschild metric

$$ds^2 = -\left(1 - \frac{r_+}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{r_+}{r}}, \quad \varphi = -\ln(r/l). \tag{13.3.14}$$

The most general action for two-dimensional dilaton gravity is

$$W[g,\varphi] = \frac{1}{4G'}\int d^2x\,\sqrt{-g}\left[W(\varphi)\,R + U(\varphi)\,(\nabla\varphi)^2 + l^{-2}\,V(\varphi)\right]. \tag{13.3.15}$$

The ambiguity in the dilaton field reparametrization $\varphi \to \tilde{\varphi}(\varphi)$ and conformal transformation $g_{AB} \to \Omega^2(\varphi)g_{AB}$ can be used to eliminate two of the three arbitrary functions W, U, V. In particular, one can present the action (13.3.13) in the form

$$W[g,\Phi] = \frac{1}{4}\int d^2x\,\sqrt{-g}\,[\Phi R + V(\Phi)]. \tag{13.3.16}$$

The equations of motion take the simple form

$$R = -\frac{dV}{d\Phi}, \quad \Phi_{|AB} = \frac{1}{2l^2}V g_{AB}. \tag{13.3.17}$$

The latter equation implies that $\xi^A = e^{AB}\,\Phi_{|B}$ is a Killing vector. For the coordinate choice $x = \Phi$, $g_{tx} = 0$, the most general solution is

$$ds^2 = -A\,dt^2 + \frac{dx^2}{A}, \tag{13.3.18}$$

where $dA(\Phi)/d\Phi = V(\Phi)$. The solution contains an arbitrary integration constant. If the function $A(x)$ vanishes at some point x_0, then the solution possesses a Killing horizon at x_0.

The following special choice of the two-dimensional dilaton action is motivated by string theory [Witten (1991), Mandal, Sengupta, and Wadia (1991)]

$$W[g,\varphi] = \frac{1}{4G'}\int d^2x\,\sqrt{-g}\,e^{-2\varphi}\left[R + 4(\nabla\varphi)^2 + 4\lambda^2\right]. \tag{13.3.19}$$

A general solution is parametrized by one constant and is of the form

$$ds^2 = -A\,dt^2 + \frac{dr^2}{A}, \tag{13.3.20}$$

$$A = 1 - \frac{M}{\lambda}\,e^{-2\lambda r}, \quad \varphi = -\lambda r. \tag{13.3.21}$$

The metric is static, asymptotically flat and describes a two-dimensional black hole. The horizon is at $r = r_+ \equiv \frac{1}{2\lambda}\ln(M/\lambda)$.

One can easily obtain a charged black hole solution by adding an $F_{AB}\,F^{AB}$ term into action (13.3.19). In two dimensions $F_{AB} = F e_{AB}$. The solution contains two arbitrary constants "mass" M and "charge" Q [Frolov (1992)]. The metric has the form of (13.3.21) and

$$A = 1 - \frac{M}{\lambda}\,e^{-2\lambda r} + \frac{Q^2}{4\lambda^2}\,e^{-4\lambda r}, \quad \varphi = -\lambda r, \quad F = Qe^{-2\lambda r}. \tag{13.3.22}$$

The global spacetime structure for this solution is similar to the structure of the Reissner-Nordström metric. It has two horizons at

$$r = r_\pm \equiv \frac{1}{2\lambda}\ln\left[\left(M \pm /\sqrt{M^2 - Q^2}\right)/2\lambda\right]. \tag{13.3.23}$$

On the properties of two-dimensional black holes in string-motivated dilaton gravity see [Frolov (1992), Strominger (1995)]. On black hole solutions in a more general class of two-dimensional dilaton gravity theories[2], see [Louis-Martinez and Kunstatter (1994), Chan and Mann (1995), Gegenberg, Kunstatter, and Louis-Martinez (1995)].

[2]It is interesting to note that the internal geometry of an infinitely long stationary cosmic string trapped by a stationary black hole coincides with a geometry of a two-dimensional black hole [Frolov, Hendy, and Larsen (1996)]. The "event horizon" of such a "string hole" is at the line of intersection of the string with an infinite-redshift surface. Perturbations ("stringons") propagating along the string are "living" in the spacetime of this two-dimensional "string hole". In particular, the "string hole" emits "stringons" with thermal spectrum in the process similar to the Hawking radiation.

13.4 Multi-dimensional Black Holes

13.4.1 Einsteinian multi-dimensional black holes

D-dimensional gravity

We begin by considering the gravitational theory in D-dimensional spacetime [Myers and Perry (1986)]. The signature of the metric g_{AB} $(A, B = 0, 1, \ldots D-1)$ is $(-, +, +, \ldots +)$. The natural generalization of the Einstein action is

$$W[g] = \int d^D x\, \sqrt{-g} \left(\frac{1}{16\pi} R + L_m \right) . \tag{13.4.1}$$

We omit the surface term which is not important for our considerations and put $c = G = 1$ (G has dimensions $(length)^{D-2}$). The above L_m is the Lagrangian density for any other fields one may wish to consider.

We assume that the spacetime is asymptotically flat, i.e., its spatial infinity has the topology of S^{D-2}, and that suitable coordinates can be chosen in which the metric has the asymptotic form

$$g_{AB} = \eta_{AB} + O\left(1/r^{D-3}\right) .$$

The mass and angular momentum of an isolated system can be defined by comparing the asymptotic behavior of the gravitational field with the gravitational field in the Newtonian limit (i.e., when the field is weak and it is created by a non-relativistic distribution of matter). The linearization

$$g_{AB} = \eta_{AB} + h_{AB} \tag{13.4.2}$$

of the Einstein equations

$$R_{AB} - \frac{1}{2} g_{AB} R = 8\pi T_{AB} \tag{13.4.3}$$

with the gauge fixing condition

$$\left(h^{AB} - \frac{1}{2} \eta^{AB} h^C_C \right)_{;B} = 0 \tag{13.4.4}$$

gives

$$\Box h_{AB} = 16\pi \left(T_{AB} - \frac{1}{D-2} \eta_{AB} T^C_C \right) . \tag{13.4.5}$$

A solution of these equations for a non-relativistic matter (when $|T_{00}| \gg |T_{0I}| \gg |T_{IJ}|$, $I, J = 1, 2, \ldots D-1$) has the following asymptotic form:

$$h_{00} \simeq \frac{16\pi}{(D-2) A_{D-2}} \frac{M}{r^{D-3}} , \tag{13.4.6}$$

$$h_{IJ} \simeq \frac{16\pi}{(D-2)(D-3)A_{D-2}} \frac{M}{r^{D-3}} \delta_{IJ} , \tag{13.4.7}$$

$$h_{0I} \simeq -\frac{8\pi}{A_{D-2}} \frac{x^J}{r^{D-1}} J_{JI} . \tag{13.4.8}$$

Here x^K are Cartesian coordinates with origin at the center of mass of the system in which the matter is at rest as a whole and

$$A_{D-2} = \frac{2\pi^{(D-1)/2}}{\Gamma(\frac{D-1}{2})}$$

is the area of a unit $(D-2)$-sphere S^{D-2}. The constants M and J_{IJ}, which are defined as

$$M = \int d^{D-1}x \, T_{00} , \tag{13.4.9}$$

$$J^{IJ} = \int d^{D-1}x \left(x^I T^{J0} - x^J T^{I0}\right) , \tag{13.4.10}$$

can be identified with the mass and angular momentum of an isolated gravitating system. In $(3+1)$-dimensional spacetime the three-dimensional antisymmetric tensor J^{ij} has only one component. In higher dimensions by suitable rotations of spatial coordinates it can be transformed into the following standard form

$$J = \begin{pmatrix} 0 & J_1 & 0 & 0 & \dots \\ -J_1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & J_2 & \dots \\ 0 & 0 & -J_2 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} . \tag{13.4.11}$$

For even D the last row and column of the matrix vanish. Thus, in higher dimensions there are $[(D-1)/2]$ independent components of the angular momentum. (The expression $[(D-1)/2]$ denotes the integer part of $(D-1)/2$). This number is related to the existence of $[(D-1)/2]$ Casimirs of the group $SO(D-1)$ of spatial rotations. The parameters J_k $(k = 1, 2, \dots [(D-1)/2])$ correspond to the rotations in the planes $x^{2k-1} - x^{2k}$, i.e., the planes in which the system is spinning. The angular rotations in these planes commute with one another and the generators of the corresponding rotations form the Cartan subalgebra.

A natural higher-dimensional generalization of a stationary axially symmetric spacetime is a spacetime in which besides a Killing vector $\xi_{(0)}$ which is timelike at spatial infinity possesses $[(D-1)/2]$ spacelike Killing vectors ξ_k which correspond to the mutually commutative rotations in two-dimensional planes. The subscript $k =$

$1, 2, \ldots [(D-1)/2]$ enumerates these two-planes. In a stationary axially symmetric spacetime the parameters M and J_k can be written in the following covariant form:

$$M = -\frac{D-2}{D-3}\frac{1}{8\pi}\int \nabla^A \xi^B_{(t)}\, d\sigma_{AB}\,, \tag{13.4.12}$$

$$J_k = \frac{1}{8\pi}\int \nabla^A \xi^B_{(k)}\, d\sigma_{AB}\,, \tag{13.4.13}$$

where the integration is taken over the $(D-2)$-sphere at spatial infinity. In four dimensions these expressions coincide with relations (12.2.4) and (12.2.7).

D-dimensional non-rotating black holes

Particular vacuum spherically-symmetric solutions of the Einstein equations in higher dimensions were obtained by Tangherlini (1963). They have the form

$$ds^2 = -B\,dt^2 + \frac{dr^2}{B} + r^2\, d\omega^2_{D-2}\,, \tag{13.4.14}$$

where

$$B = 1 - \frac{16\pi M}{(D-2)A_{D-2}}\frac{1}{r^{D-3}}\,, \tag{13.4.15}$$

and $d\omega^2_{D-2}$ is a line element on the unit sphere S^{D-2}. For $D = 4$ this solution evidently reproduces the Schwarzschild metric.

The solution (13.4.14) can be easily extended to the solution of the Einstein-Maxwell equations. The analog of the Reissner-Nordström solution reads [Myers and Perry (1986)]

$$B = 1 - \frac{16\pi M}{(D-2)A_{D-2}}\frac{1}{r^{D-3}} + \frac{2Q^2}{(D-2)(D-3)}\frac{1}{r^{2(D-3)}}\,, \tag{13.4.16}$$

$$A_B\, dx^B = \frac{Q}{D-3}\frac{1}{r^{D-3}}\, dt\,, \tag{13.4.17}$$

where

$$Q = \frac{1}{A_{D-2}}\int F^{AB}\, d\sigma_{AB} \tag{13.4.18}$$

is the electric charge of a black hole, and A_B is the electromagnetic potential.

The global structure of the spacetime is the same as for the Reissner-Nordström metric. A black hole exists when $Q \le M$, and there is naked singularity in the case when $M < Q$. Unlike the electrically charged case, there is no higher dimensional generalization of magnetically charged black hole for the simple reason that there is no magnetic charge in higher dimensions.

D-dimensional rotating black holes

The higher dimensional generalization of the Kerr solution was obtained by Myers and Perry (1986). It contains $1+[(D-1)/2]$ parameters. Define $D = 2d+1+\epsilon$, where $\epsilon = 0$ for odd D and $\epsilon = 1$ for even D. The metric in Boyer-Lindquist coordinates can be written in the following form which is valid for both odd and even D

$$ds^2 = -dt^2 + \sum_{k=1}^{d} [(r^2+a_k^2)(d\mu_k^2 + \mu_k^2\, d\phi_k^2)]$$

$$+ \frac{m\, r^{2-\epsilon}}{\Pi F}\left[dt + \sum_{k=1}^{d} a_k\, \mu_k^2\, d\phi_k\right]^2 + \frac{\Pi F}{\Pi - m r^{2-\epsilon}}\, dr^2 + \epsilon\, r^2\, d\alpha^2\,. \qquad (13.4.19)$$

For these solutions

$$m = \frac{16\pi}{(D-2)A_{D-2}}\, M\,, \quad a_k = \frac{D-2}{2}\frac{J_k}{M}\,, \quad \sum_{k=1}^{d}\mu_k^2 + \epsilon\,\alpha^2 = 1\,. \qquad (13.4.20)$$

ϕ_k are angles with the period 2π in each plane $x^{2k-1} - x^{2k}$, and μ_k are direction cosines with respect to these planes having the range $0 \le \mu_k \le 1$. The functions Π and F which enter the formulas are

$$F = 1 - \sum_{k=1}^{d} \frac{a_k^2\, \mu_k^2}{r^2 + a_k^2}\,, \qquad \Pi = \prod_{k}^{d}(r^2 + a_k^2)\,. \qquad (13.4.21)$$

This metric admits the following Killing vectors

$$\xi_{(t)}^A\, \partial_A = \partial_t\,, \qquad \xi_{(k)}^A\, \partial_A = \partial_{\phi_k}\,. \qquad (13.4.22)$$

The outer boundary of the ergosphere is the surface of infinite redshift (the static limit) where $\xi_{(t)}^2 = 0$. The equation of this surface is

$$\Pi F - m\, r^{2-\epsilon} = 0\,. \qquad (13.4.23)$$

Denote by

$$\eta^A = \xi_{(t)}^A + \sum_{k=1}^{d} \Omega_k\, \xi_{(k)}^A\,. \qquad (13.4.24)$$

Outside the event horizon one can always choose the coefficients Ω_k in such a way that η is a timelike Killing vector. It means that a stationary reference frame is possible in this region. At the event horizon the linear combination (13.4.24) is a spacelike or null vector, η being null for $\Omega_k = \Omega_k^H$, where Ω_k^H are the angular velocities of the black hole. The value of Ω_k^H as well as the position of the horizon are defined by the following system of equations

$$\left[\frac{\partial \eta^2}{\partial \Omega_k}\right]_{H,\Omega_k=\Omega_k^H} = 0\,, \qquad \left[\eta^2\right]_{H,\Omega_k=\Omega_k^H} = 0\,. \tag{13.4.25}$$

By solving these equations, one gets

$$\left[\Pi - m\, r^{2-\epsilon}\right]_{r=r_H} = 0\,, \quad \Omega_k^H = \frac{a_k}{r_H^2 + a_k^2}\,. \tag{13.4.26}$$

Both the event horizon and the infinite red-shift surface have the topology $S^{D-2} \times R$. The external component of the infinite red-shift surface is outside the event horizon. They coincide when $\mu_k = 0$ for each $a_k \neq 0$.

For an arbitrary D the position of the horizon r_H cannot be found analytically. Myers and Perry (1986) showed that for a broad range of the parameters the general solution for a spinning black hole in higher dimensions contains horizons which surround the singularities. They also showed that the properties of rotating multi-dimensional black holes are similar to the properties of the Kerr solution in (3+1) dimensions. There are only two essential differences which occur in higher dimensions: For $D \geq 6$, black holes exist with arbitrary large angular momentum for a fixed mass, and for odd D, black holes are possible for the negative values of the mass parameter M.

The four laws of black hole thermodynamics (see Chapter 12) can be easily extended to higher dimensions [Myers and Perry (1986)].

13.4.2 D-dimensional black holes in string theory

There is a straightforward generalization of the Einstein D-dimensional black hole solutions to the solutions for the stringy action [Gibbons and Maeda (1988)]

$$W = \int d^D x\, \sqrt{-g}\, e^{-2\varphi}\left[R + 4(\nabla\varphi)^2 - F_{AB}\,F^{AB}\right]. \tag{13.4.27}$$

One simply starts with the D-dimensional Schwarzschild solution (13.4.4) and make the transformation similar to (13.2.3). The corresponding solution is

$$\begin{aligned} ds^2 = &-\left(1 - \frac{C\,m}{r^n}\right)\left(1 + \frac{C\,m}{r^n}\sinh^2\alpha\right)^{-2} dt^2 \\ &+ \left(1 - \frac{C\,m}{r^n}\right) dr^2 + r^2\, d\omega_{n+1}^2\,, \end{aligned} \tag{13.4.28}$$

$$A_B = \delta_B^t\, A_t\,, \quad A_t = -\frac{C\,m \sinh 2\alpha}{2\sqrt{2}\left[r^n + C\,m \sinh^2\alpha\right]}\,, \tag{13.4.29}$$

$$e^{-2\varphi} = 1 + \frac{C\,m}{r^n}\,\sinh^2\alpha\,. \tag{13.4.30}$$

Here $n = D - 3$, and C is a dimension-dependent constant. The mass and the charge are given by

$$M = m\left(1 + \frac{2n}{n+1}\sinh^2\alpha\right)\,, \qquad Q = \frac{C}{\sqrt{2}}\,m\,n\cosh\alpha\ \sinh\alpha\,. \tag{13.4.31}$$

In order to obtain the metric in the Einstein frame, one must multiply (13.4.28) by $\exp[-4\varphi/(n+1)]$.

13.4.3 Kaluza-Klein black holes

In the Kaluza-Klein theory one makes an assumption that the spacetime has the number of dimensions D greater than four. The physical spacetime appears as a result of compactification of the "superfluous" $D - 4$ dimensions. The metric of D-dimensional spacetime manifests itself as a set of physical fields in the physical spacetime. One of these fields is a four-dimensional metric. For example, the simplest version of such a theory (for $D = 5$) contains, in addition to the gravitational field, the electromagnetic and massless scalar fields. Black hole solutions in the Kaluza-Klein theory were studied by Dobiasch and Maison (1982), Chodos and Detweiler (1982), Gibbons (1982, 1984, 1986), Pollard (1983), Gibbons and Wiltshire (1986), and Gibbons and Maeda (1988).

As an example of the Kaluza-Klein black hole, we consider the original Kaluza-Klein theory. This is simply 5-dimensional general relativity

$$W = \frac{1}{16\pi G^{(5)}}\int d^5x\,\sqrt{-g^{(5)}}\,R^{(5)}\,, \tag{13.4.32}$$

with an additional requirement that the fifth dimension is a Killing one; that is, the fields are independent of the 5th coordinate, x^5. In accordance with this ansatz, we write the 5-dimensional metric in the following parametric form [Gibbons and Wiltshire (1986)]

$$ds^2_{(5)} = e^{4\varphi/\sqrt{3}}\,(dx^5 + A_\alpha\,dx^\alpha)^2 + e^{-2\varphi/\sqrt{3}}\,g_{\alpha\beta}\,dx^\alpha\,dx^\beta\,. \tag{13.4.33}$$

The 5-dimensional action (13.4.32) reduces to

$$W = \frac{1}{16\pi}\int d^4x\,\sqrt{-g}\left[R - 2(\nabla\varphi)^2 - e^{2\sqrt{3}\,\varphi}F_{\alpha\beta}\,F^{\alpha\beta}\right]. \tag{13.4.34}$$

This is the Einstein action modified by the presence of a scalar massless field φ and a Maxwell field $F_{\alpha\beta} = A_{\beta,\alpha} - A_{\alpha,\beta}$.

Let us consider a 5-metric of the form

$$ds^2_{(5)} = (dx^5)^2 + g_{\alpha\beta}\,dx^\alpha\,dx^\beta\,, \tag{13.4.35}$$

where $g_{\alpha\beta}$ is a solution of the vacuum four-dimensional Einstein equations. It is evident that the 5-metric is a solution to (13.4.32). In the particular case where the 4-metric is the metric of a black hole, the 5-dimensional solution is called a *black string*.[3]

Static solutions of the form (13.4.35) have two commuting Killing vectors ξ_5 and ξ_t. Given a static four-dimensional vacuum metric one can take its product with a real line. If one boosts this 5-dimensional solution in the 5th direction

$$t \to \cosh\alpha\ t + \sinh\alpha\ x^5\,, \qquad x^5 \to \sinh\alpha\ t + \cosh\alpha\ x^5\,, \tag{13.4.36}$$

it still obeys the 5-dimensional equations. However, when it is rewritten in the form (13.4.33), one gets non-trivial electric and scalar fields. In particular, starting from the four-dimensional Schwarzschild metric after boosting the black string (13.4.35), we have a solution [Dobiasch and Maison (1982), Chodos and Detweiler (1982), Gibbons (1982)]

$$ds^2_{(5)} = B^2(dx^5 + A_t\,dt)^2 + B^{-1}\left[-\frac{\Phi}{B}\,dt^2 + \frac{B}{\Phi}\,dr^2 + r^2 B\,d\omega^2\right], \tag{13.4.37}$$

where

$$\Phi = 1 - \frac{2m}{r}\,, \qquad B^2 = 1 + \frac{2m}{r}\sinh^2\alpha\,. \tag{13.4.38}$$

For this solution $\varphi = \left(\sqrt{3}\,/2\right)\ln B$. The mass M, electric charge Q, and scalar charge Σ are

$$M = m\left(1 + \frac{1}{2}\sinh^2\alpha\right), \quad Q = m\sinh\alpha\cosh\alpha\,, \quad \Sigma = \frac{\sqrt{3}}{2}\,m\,\sinh^2\alpha\,. \tag{13.4.39}$$

This solution can be easily generalized to the case of a rotating black hole [Frolov, Zel'nikov, and Bleyer (1987)].

[3]One can extend a black hole solution in D dimensions by taking its product with p-dimensional flat space R^p. The corresponding $(D+p)$-dimensional solution is called *black p-brane*. A black string in this terminology is simply a black 1-brane.

Chapter 14

The Interior of a Black Hole

14.1 Introduction

The collapse of a realistic star and formation of a black hole are initially highly dynamical processes. However, the external gravitational field settles down very quickly to the unique stationary state – a Kerr-Newman black hole. Dynamical deviations from this state produce the emission of gravitational radiation. This radiation interacts with the spacetime curvature. While part of it escapes to infinity, another part is backscattered and absorbed by the resulting black hole. In the end all deviations from the Kerr-Newman field decay to zero.

An analogous process occurs in the case of any subsequent perturbations of the Kerr-Newman black hole. The gravitational radiation which is backscattered by the curvature of spacetime and falls through the horizon decays very fast, typically according to an inverse power of advanced time.

What can one say about the interior of a black hole? Does the internal geometry also approach the Kerr-Newman form soon after the collapse?

This problem was the subject of very active investigation in recent decades. There has been great progress in these researches. We know some important properties of the realistic black hole interior, but some details and crucial problems are still the subject of much debate.

A very important point for understanding the problem of the black hole interior is the fact that the path into the gravitational abyss of the interior of a black hole is a progression in time. We recall that inside a spherical black hole the radial coordinate is timelike. This means that the problem of the black hole interior is an *evolutionary problem*. In this sense it is completely different from the problem of the internal structure of other celestial bodies; for example, stars .

In principle, if we know the conditions at the surface of a black hole, we can integrate the Einstein equations in time and learn the structure of progressively deeper layers inside the black hole. Conceptually it looks simple, but there are two difficulties of principle which prevent consistent realization of this idea.

The first difficulty is the following. As we shall see in the subsequent discussion, the internal structure of a general black hole even soon after its creation depends crucially on the conditions on the event horizon in the remote future of an external observer (formally in the infinite future). It is because a lightlike signal can propagate inside a black hole from the remote future to these regions if they are deep enough in the hole.

This means that the structure of these regions depends crucially on the fate of the black hole in the infinite future of an external observer. For example, it depends on the final state of the black hole evaporation, on possible collisions of the black hole with another black hole, and it depends on the fate of the Universe itself. It is clear that theorists feel uncomfortable under such circumstances.

The second serious problem is related to the existence of a singularity inside a black hole. Close to the singularity, where the curvature of spacetime approaches Plank scales, the classical theory of general relativity is not applicable. We do not yet have the final version of a quantum theory of gravity; thus, any extension of the discussion to physics in this region would be highly speculative. Fortunately, as we shall see, these singular regions are deep in the black hole interior and they are *in the future* with respect to overlying and *preceding* layers of the black hole where curvatures are not so high and which can be described by well-established theory.

The first attempts to extend the no hair theorem to the interior of a Schwarzschild black hole were made by Doroshkevich and Novikov (1978) and Poisson and Israel (1988). They demonstrated that the interior regions of a black hole located long after its formation are virtually free of perturbations because the gravitational radiation from aspherical initial perturbations becomes infinitely diluted as it reaches these regions (see Section 14.2). But this result breaks down in the general case where the angular momentum or the electric charge is not equal to zero. The reason for this is related to the fact that the causal structure of the interior of the Kerr-Newman solution differs drastically from Schwarzschild. The key point is that the interior of the Kerr-Newman black hole possesses a *Cauchy horizon*. This is a surface of infinite blueshift. Infalling gravitational radiation propagates inside the black hole along paths approaching the null generators of the Cauchy horizon, and the energy density of this radiation will suffer an infinite blueshift as it approaches the Cauchy horizon. As we will demonstrate in following sections, this infinitely blueshifted radiation together with the radiation scattered on the curvature of spacetime inside the black hole leads to the formation of a curvature singularity instead of the Cauchy horizon.

The discussion in this chapter is organized as follows. In Section 14.2 we discuss the *no hair theorem for the interior* of the Schwarzschild black hole. Section 14.3 is devoted to the analysis of the instability of the Cauchy horizon in a Reissner-Nordström black hole to small perturbations in the linear approximation. In Section 14.4 we discuss the problem of the internal structure of realistic black holes taking nonlinear effects into account. In Section 14.5 we analyze the effect of the process of electron-positron pair creation inside a charge black hole on its internal structure. Section 14.6

discusses the black hole interior allowing for quantum evaporation. And in the final Section 14.7 of this chapter we speculate what might happen beyond the singularity in the black hole abyss.

14.2 Physical Fields Inside a Schwarzschild Black Hole

The structure of spacetime inside a non-rotating spherically symmetric black hole has been discussed in Chapter 2. Here we will discuss the behavior of physical fields and the problem of the stability of the interior part of a Schwarzschild black hole with $r < r_g$, by analogy with what has been performed for the space outside a black hole in Chapter 4.

This problem was solved by Doroshkevich and Novikov (1978), and Poisson and Israel (1988). The properties of spacetime inside a black hole are essential for the problem of gravitational collapse and the nature of the singularity even though this region is inaccessible to an observer who stays outside the black hole. The general theorems about the properties of black holes, discussed in Chapter 5, do not give a specific expression for the structure of spacetime inside a black hole. It has sometimes been conjectured that all radiation fields and all perturbations grow beneath the black hole horizon and become nonlinear, and that the structure of the metric must be most peculiar. Furthermore, we have seen in Section 6.6 that the structure of the analytic continuation of the solution for the spacetime metric inside a charged and rotating black hole is very complicated, even fantastic. Is this structure indeed realized to some extent when a black hole is actually formed?

This question will be answered in this and the following sections. First we analyze the propagation of physical fields inside a Schwarzschild black hole and the stability of its internal structure.

Consider a perturbation created by a test object falling into the black hole; this object is the source of the various fields of interest (scalar, electromagnetic, gravitational, etc.). We wish to clarify the properties of wave fields long after the object has fallen into the black hole, that is, when the observation point tends to I^+ (see, e.g., Figure 5.4c) in the T_--region (also designated as region II).

This choice of asymptotic procedure signifies that if the coordinates (2.2.1) are employed, then $t \to \infty$. An analysis of "tails" of radiation (see Chapter 4) shows that in the external region R the radiative modes fall off, for a fixed r, according to a power law as $t \to \infty$. As for the horizon itself, $r = r_g$, this means that if one introduces an affine parameter V along all geodesics forming the horizon, the perturbations will also fall off according to a power law as $V \to \infty$. The evolution of the wave field in the T_--region (for $r < r_g$) is determined by equations similar to equation (4.2.4).

The required mathematical analysis was carried out in the already-cited paper of

Doroshkevich and Novikov (1978). The following results were obtained. For scalar perturbations,

$$\Phi \approx D_1 t^{-2(l+1)} + D_2 t^{-(2l+3)} \ln r \,, \tag{14.2.1}$$

where D_1 and D_2 are constants. For perturbations of massless fields of spin $s \neq 0$ (including metric perturbations), the main term of the r-dependent component has the following form for radiative multipoles $l \geq s$:

$$\Phi_1 \approx D_3 t^{-(2l+3)} r^{-n} \tag{14.2.2}$$

(D_3 and n are constants).

Hence, if r is fixed and $t \to \infty$, radiative modes of perturbations due to external sources are damped out and the spacetime tends to a "stationary" state described by the Schwarzschild solution. The situation is similar to that discussed in Chapter 4 for the exterior region of the black hole. However, important differences manifest themselves when $r < r_g$.

First, the coordinate r for $r < r_g$ plays the role of a time coordinate, and t acts as a spatial coordinate. It would be more correct, therefore, to refer to the tendency to a state that depends only on r, not to a tendency to a stationary state.

The second and more important difference is that perturbations grow infinitely as the point approaches the singularity at a fixed t. The general solution close to the singularity, ignoring quantum effects, was constructed by Belinsky, Lifshitz, and Khalatnikov (1970). The method of small perturbations does not work near the singularity. The boundary of the region where perturbations cease to be small is given by the expression (for $s \neq 0$):

$$r^n \approx D_3 t^{-(2l+3)} \,. \tag{14.2.3}$$

As t increases, this region contracts to the singularity. The Schwarzschild solution is stable with respect to small perturbations everywhere inside the black hole, with the exception of this region, which continuously contracts as $t \to \infty$, and all radiative modes are damped out by a power law as t increases.

Radiation from elementary perturbations produced in the region of $r < r_g$ propagates only to a finite (small) region inside the black hole because the emitted signals "ram" into the singularity. These perturbations do not affect in any way the properties of the black hole interior at $t \to \infty$. The most important conclusion that can be drawn from this discussion is that the Schwarzschild solution is as stable in the T_--region as it is in the R-region.

Several words are necessary on non-radiative perturbation multipoles connected with the particles falling into the black hole or with the collapsing body that produces the black hole. These multipoles are: (i) $l = 0$ for electromagnetic perturbations (the Coulomb field of a falling charge), and (ii) $l = 0$ (the field of accreted mass) and $l = 1$ (the field of accreted angular momentum) for gravitational perturbations.

These multipoles do not damp out as $t \to \infty$. They increase in the familiar manner as $r \to 0$, and drastically alter the metric near $r = 0$ because this situation corresponds to the transition to the Reissner-Nordström metric if electric charge is added and to the Kerr metric if angular momentum is added. We will later discuss these metrics in relation to the internal structure of black holes. Here, we only emphasize that if the corrections to the metric and the perturbation fields become significant sufficiently close to the singularity $r = 0$, they have no immediate physical meaning. The point is that quantum processes become important in the vicinity of the singularity where the spacetime curvature exceeds the Planck value (i.e., becomes greater than $1/l_{\rm Pl}^2$); for instance, for the quadratic invariant of curvature, we have

$$R_{\alpha\beta\mu\nu} R^{\alpha\beta\mu\nu} = \frac{12\, r_{\rm g}^2}{r^6} > \frac{1}{l_{\rm Pl}^4}\,. \tag{14.2.4}$$

The entire region defined by (14.2.4) may be regarded as a "singularity" from the standpoint of the classical theory.

The last point is the case of non-radiative multipoles of physical fields, connected with external sources. If the sources are stationary, that is, the characteristic time of field variation $t \gg r_{\rm g}/c$, then, as shown in Section 4.3, the fields of these sources with $k = 0$ penetrate unobstructedly into the black hole through the potential barrier. We assume these fields to be weak at the event horizon $r_{\rm g}$. The fields are independent of the coordinate t both inside and outside the black hole.

Typical examples to be discussed here are the effect of an external stationary quadrupole gravitational field and that of a stationary magnetic field, homogeneous far from the black hole, on the internal structure of this black hole. The exact solution describing a black hole in an external quadrupole field [Doroshkevich *et al.* (1965)] was given in Section 7.5.4. Expressions (7.5.25)–(7.5.27) show that if the parameter q of the quadrupole moment is sufficiently small so that the corrections to the metric at $r = r_{\rm g}$ are small, then the corrections remain small everywhere for $r < r_{\rm g}$ up to the singularity at $r = 0$. No instability is produced thereby inside the black hole.

Now we will consider a magnetic field that is homogeneous at spatial infinity. Let us assume this field to be weak. Then we find that the nonzero components of the electromagnetic field tensor are [Ginzburg and Ozernoi (1964)]

$$F^{\theta\phi} = \frac{B_0 \cot\theta}{r^2}\,, \qquad F^{\phi r} = -\frac{B_0}{r}\left(1 - \frac{r_{\rm g}}{r}\right), \tag{14.2.5}$$

where B_0 is the magnetic field strength at infinity. The component $F^{\phi r}$ describes the electric field inside the black hole. As $r \to 0$, the components of the electromagnetic field tensor grow infinitely. However, this does not restructure the metric because the components of the energy-momentum tensor constructed of components (14.2.5) grow at a slower rate than the corresponding expressions describing spacetime curvature. The electromagnetic field of a charge at rest outside the black hole has a similar behavior in the black hole interior [see equation (7.2.2)]. A massive vector field in the black hole interior generated by a static external charge has new features: it is oscillating, and generically the corresponding stress-energy tensor grows near $r = 0$ as r^{-4} [Frolov (1978a, 1986)].[1]

Let us again point out that a physical "singularity" for which consideration of quantum gravity processes is essential is created in the vicinity of the true mathematical singularity $r = 0$. Moreover, quantum phenomena in the presence of the fields other than the gravitational (e.g., in a charged black hole) can also be important far away from the singularity (where the curvature is not large). In some cases, these phenomena greatly influence the metric. They are to be considered in the subsequent sections. Finally, the Hawking evaporation of black holes affects the metric on very long time scales. Since inside the black hole the "time" coordinate t is spacelike, the effects connected with the Hawking evaporation are important at large distances from the collapsing body (but inside the event horizon!).

14.3 Instability of Cauchy Horizons Inside a Black Hole

14.3.1 Interior of a charged spherical black hole

Consider the behavior of small perturbations of the gravitational and electromagnetic fields inside a charged spherically symmetric black hole. A qualitatively new factor appears here in comparison with the case of the Schwarzschild black hole: the presence of the Cauchy horizon (see Section 6.6). Figure 14.1 shows a fragment of the Penrose diagram with the internal part (region II) of the charged black hole and the external space. If the charged black hole is a result of the collapse of a charged body from space I, then the other external space (I′ in Figure 6.4) is absent, as it is in the case of the collapse of an uncharged spherical body which becomes a Schwarzschild black hole (see Section 2.7). This is why region I′ is not shown in Figure 14.1.

There are serious reasons for expecting that small perturbations can grow infinitely in the neighborhood of $r_{-,1}$, [Penrose (1968)]. Indeed, let us consider a small perturbation of the gravitational and (or) electromagnetic field outside the black hole in region I. We have already shown in Section 4.5 and 4.8 that the "tails" of radiation induced by perturbations damp out following a power-law dependence v^{-p} at $r =$ const, as a result of scattering on the spacetime curvature (v is the advanced time). This decreasing radiation flux crosses the event horizon r_+ and concentrates along the horizon $r_{-,1}$ (see Figure 14.1). An observer traveling along a timelike world line and crossing the horizon $r_{-,1}$ will receive this radiation close to $r_{-,1}$ in a finite proper time (this radiation sinks into the black hole after an infinite time of the external observer). When the observer approaches $r_{-,1}$, the radiation he receives has an

[1] A similar oscillatory behavior was recently found in the exact solutions of the Einstein-Yang-Mills equations for the black hole interior [Gal'tsov, Donets, and Zotov (1997)].

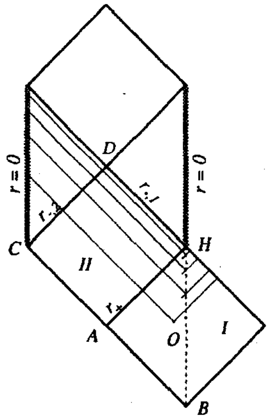


Figure 14.1: Part of the conformal Penrose-Carter diagram for a charged black hole. The propagation of radial rays is shown immediately after the flash (O) and after the scattering on the spacetime curvature. The scattering perturbations occurs in the black hole exterior along the line BH. The Cauchy horizon defined by the equation $r = r_-$ consists of two parts, which we denote by $r_{-,1}$ and $r_{-,2}$, respectively.

infinite blueshift. The blueshift factor grows as $e^{\kappa_- v}$, where κ_- is the surface gravity of the inner (Cauchy) horizon. It is natural to expect that such a concentration of energy restructures the spacetime and produces a true spacetime singularity in place of $r_{-,1}$. At the same time, no concentration of energy appears along the horizon $r_{-,2}$ (except at the point D), so that we cannot expect that perturbations appearing in region I "create" a singularity at $r_{-,2}$.

A mathematical analysis of the evolution of small perturbations confirms these intuitive arguments: perturbations are unstable near $r_{-,1}$. It should be emphasized that the method of small perturbations does not take into account the back action of the perturbations onto the gravitational field. For this reason, it cannot answer the question what kind of singularity is formed as a result of instability of the Cauchy horizon. We return to this problem later.

14.3.2 Linear instability of a Cauchy horizon of a charged spherical black hole

The evolution of small perturbations was analyzed by McNamara (1978a,b), Gürsel *et al.* (1979a,b), Matzner *et al.* (1979), and Chandrasekhar and Hartle (1982). Here, we follow the last of these papers.

The metric of a charged black hole is

$$ds^2 = \frac{-(r-r_+)(r-r_-)}{r^2}\,dt^2 + \frac{r^2}{(r-r_+)(r-r_-)}\,dr^2 + r^2\,d\omega^2 . \qquad (14.3.1)$$

We are interested in the spacetime region $r_- < r < r_+$. The system of equations describing the perturbations of the gravitational and electromagnetic fields of a charged Reissner-Nordström black hole in this region allows separation of variables and can be decoupled in the same way as it was done for the black hole exterior (see Section 4.4.3). Thus, the problem is reduced to solving radial equations similar to (4.2.7). Perturbations are subsumed under two classes: axial [superscript (+)] and polar [superscript (−)]. A perturbations of each class can be expressed in terms of a pair of gauge-invariant scalar functions $Z_i^{\pm}(r)$ $(i = 1, 2)$ that satisfy the equation [Chandrasekhar (1979b,1983), Chandrasekhar and Xanthopoulos (1979)]

$$\frac{d^2 Z_i^{(\pm)}}{dr_*^2} + \omega^2 Z_i^{(\pm)} = V_i^{(\pm)} Z_i^{(\pm)} . \tag{14.3.2}$$

Here

$$r_* = r + \frac{1}{2\kappa_+} \ln|r_+ - r| - \frac{1}{2\kappa_-} \ln|r_- - r| \tag{14.3.3}$$

is a "tortoise" radial coordinate, similar to (7.2.6), and

$$\kappa_+ = \frac{r_+ - r_-}{2r_+^2}, \qquad \kappa_- = \frac{r_+ - r_-}{2r_-^2} \tag{14.3.4}$$

are the surface gravity factors for the outer (+) and inner (−) horizons. The effective potentials $V_i^{(\pm)}$ are defined as

$$V_i^{(-)} = \frac{\Delta}{r^5}\left[(\mu^2 + 2)r - q_j + \frac{4Q^2}{r}\right], \tag{14.3.5}$$

$$V_i^{(+)} = V_i^{(-)} + 2q_j \frac{d}{dr_*} \frac{\Delta}{r^2(\mu^2 r + q_j)}, \tag{14.3.6}$$

$$\Delta = (r - r_+)(r - r_-) = r^2 - 2Mr + Q^2, \qquad \mu^2 = (l-1)(l+2), \tag{14.3.7}$$

$$q_1 = 3M + \sqrt{9M^2 + 4Q^2\mu^2}, \qquad q_2 = 3M - \sqrt{9M^2 + 4Q^2\mu^2}. \tag{14.3.8}$$

In formulas (14.3.5), (14.3.6), $i, j = 1, 2$; $i \neq j$. It is found that $Z_i^{(+)}$ is related to $Z_i^{(-)}$ by a simple transformation.

We have already encountered equations of type (14.3.2) in the problem of the behavior of physical fields outside a black hole (see Chapter 4). We are interested in that solution of these equations which describes the transmission and reflection of the wave incident on $V_i^{(\pm)}$. In the present case, $V_i^{(\pm)}$ represents a potential well, not a barrier as was the case for the exterior space of the black hole, but the situation remains qualitatively unchanged. It is found that the reflection and transmission coefficients for $Z_i^{(+)}$ can be expressed simply in terms of the corresponding coefficients for $Z_i^{(-)}$.

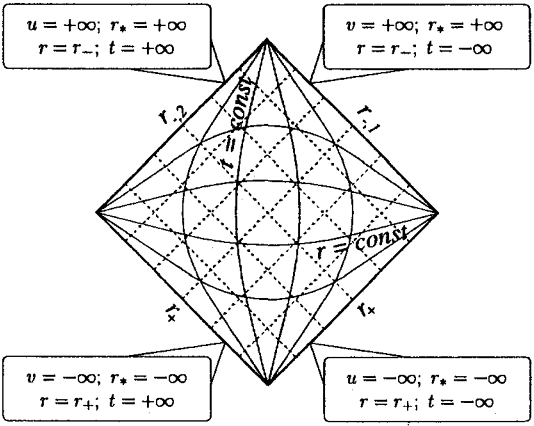


Figure 14.2: The definition of the coordinates u and v used in the domain between the two horizons of the Reissner-Nordström black hole. Note that a coordinate t is chosen so that t increases from $-\infty$ to $+\infty$ as one moves from right to left along lines r = const.

Let us analyze the behavior of wave perturbations that arrive at region II across the horizon r_+ from region I (see Figure 14.1). To do this, we consider the dispersion of a wave which has unit amplitude on $r = r_+$ $(r_* \to -\infty)$. The solution of equations (14.3.2) must satisfy the following boundary conditions (the upper and lower indices are dropped because the analysis holds for all their values):

$$Z(r_*, \omega) \to \begin{cases} A(\omega)\, e^{-i\omega r_*} + B(\omega)\, e^{+i\omega r_*}, & (r_* \to +\infty\,,\ r \to r_-)\,, \\ e^{-i\omega r_*}, & (r_* \to -\infty\,,\ r \to r_+)\,. \end{cases} \tag{14.3.9}$$

The coefficients $A(\omega)$ and $B(\omega)$ describing the transmission and the reflection of the wave at $V(r_*)$, can be found, in principle, by standard techniques provided the form of $V(r_*)$ is known [see (14.3.5) and (14.3.6)].

The analysis is facilitated by introducing the null coordinates [see Figure 14.2]

$$u = r_* + t\,, \qquad v = r_* - t\,. \tag{14.3.10}$$

The lines of constant u and v are represented by dashed straight lines. The event horizon corresponds to $u = -\infty$; the horizon $r_{-,1}$, to $v = \infty$, and the horizon $r_{-,2}$, to $u = +\infty$. The boundary conditions (14.3.9) are rewritten in the form

$$Z(r_*, t) \to \begin{cases} A(\omega)\, e^{-i\omega v} + B(\omega)\, e^{+i\omega u}, & (r_* \to \infty\,,\ r \to r_-)\,, \\ e^{-i\omega v}, & (r_* \to -\infty\,,\ r \to r_+)\,. \end{cases} \tag{14.3.11}$$

Consider a perturbation $Z_{\rm pert}(v)$ which intersects the horizon r_+, that is, which is given for $u \to -\infty$. Its Fourier transform is

$$\widetilde{Z}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Z_{\rm pert}(v)\, e^{i\omega v}\, dv\,. \tag{14.3.12}$$

Having undergone dispersion in region II, the perturbation reaches the horizon r_-. Its amplitude there can be written in the form[2]

$$Z_{\rm scat}(r_*, t) \to X(v) + Y(u) \qquad (r_* \to \infty), \tag{14.3.13}$$

where

$$X(v) = \int_{-\infty}^{\infty} \widetilde{Z}(\omega)\,[A(\omega) - 1]\, e^{-i\omega v}\, d\omega\,, \tag{14.3.14}$$

$$Y(u) = \int_{-\infty}^{\infty} \widetilde{Z}(\omega)\, B(\omega)\, e^{+i\omega u}\, d\omega\,. \tag{14.3.15}$$

We are interested in finding the radiative flux received by an observer crossing the horizon r_-. This flux is proportional to the square of the amplitude

$$F = u^\alpha\, Z_{,\alpha}\,|_{r \to r_-}\,, \tag{14.3.16}$$

where u^α is the four-dimensional velocity of the observer. The horizon r_- is either stable or unstable (in the linear approximation) with respect to small perturbations depending on whether this flux is finite or infinite.

Matzner *et al.* (1979) and Chandrasekhar (1983) have shown that the quantity F at the Cauchy horizon is of the form:

$$F_{r_{-,1}} = \frac{2r_-^2}{r_+ - r_-}\,|\,E\,|\, \lim_{v \to \infty} e^{\kappa_- v}\, X_{,-v} \quad \text{on} \quad r_{-,1}\,, \tag{14.3.17}$$

$$F_{r_{-,2}} = -\frac{2r_-^2}{r_+ - r_-}\, E\, \lim_{u \to \infty} e^{\kappa_- u}\, Y_{,u} \quad \text{on} \quad r_{-,2}\,, \tag{14.3.18}$$

where E is the constant time-component of the covariant four-velocity u_α of the observer. Expressions (14.3.17) and (14.3.18) are divergent or finite, depending on the behavior of the functions

$$e^{\kappa_- u}\, X_{,-v}\,|_{v \to \infty} \quad \text{and} \quad e^{\kappa_- u}\, Y_{,u}\,|_{u \to \infty}\,. \tag{14.3.19}$$

We assume that the form taken by the perturbing radiation $Z_{\rm pert}(v)$ crossing the horizon r_+ satisfies the following conditions:

$$Z_{\rm pert}(v) = 0 \quad \text{for} \quad v < v_0\,, \tag{14.3.20}$$

[2] In the absence of scattering the amplitude on $r_{-,2}$ is equal to one. The scattering amplitude here is defined as the difference $A(\omega) - 1$.

$$Z_{\text{pert}}(v) \text{ for } v \to \infty \text{ tends to zero at least as } v^{-1}. \tag{14.3.21}$$

It is these conditions that any real radiation must satisfy, for example, the radiation from an object falling into the black hole or from an elementary perturbation that has occurred in region I. Indeed, the second condition must hold for $v \to \infty$ at the horizon of an uncharged black hole, in accordance with the asymptotic behavior of perturbing radiation analyzed in the preceding section. A charge $Q < M$ does not change anything in this respect [see, e.g., Bičák (1972)]. The power-law decay of radiation "tails" from perturbations is typical for practically any perturbation.

The first condition is definitely met if v_0 is interpreted as an affine parameter corresponding to the moment when the horizon is intersected by the first ray reaching it from the perturbation.[3]

Chandrasekhar and Hartle (1982) proved that the quantity $F_{r_{-,2}}$ of (14.3.18) remains finite for both (axial and polar) perturbations; that is, the horizon $r_{-,2}$ is stable with respect to small perturbations in region I. In the contrast, the quantity $F_{r_{-,1}}$ diverges as $v \to \infty$ at least as $\exp(\kappa_- - \kappa_+)v$ or even faster (the rate of divergence depends on the type of perturbation). This means that as $r_{-,1}$ is approached, the observer receives a radiation flux of infinite density.

The mathematical analysis presented above has totally confirmed Penrose's intuitive arguments given at the beginning of the section. Perturbation analysis demonstrates an infinite concentration of energy density close to $r_{-,1}$. The method of small perturbations is thus invalid close to $r_{-,1}$. In the next section we shall discuss possible results of this instability.

Another remark needs to be made here. Let the sources of a constant external (electromagnetic, gravitational, or some other) field be present outside a black hole. If the black hole is charged, these fields penetrate through r_+ into the interior region, as happens in the case of an uncharged Schwarzschild black hole (see Section 14.2). If such stationary fields outside the black hole are weak and hardly affect the metric, they remain weak inside the black hole as well. Thus, they are weak on r_- and do not produce any instability [for a justification of this statement, see Gürsel *et al.* (1979b)].

14.3.3 Instability of the Cauchy horizon of a rotating black hole

In the previous section we demonstrated the linear instability of the Cauchy horizon of a spherical charged black hole. The key factor producing the instability is the infinite concentration of energy density close to $r_{-,1}$ as seen by a free falling observer. This infinite energy density is produced by the ingoing radiative "tail". The analytic continuation of the metric of a rotating black hole inside the event horizon has a

[3]Sufficiently small v are of no interest to us when a "non-eternal" black hole is considered, that is, a black hole created by collapse.

similar structure. In particular, there also exists a Cauchy horizon, which is similar to the one in the Reissner-Nordström geometry. The properties of the "tails" of the ingoing radiation on the event horizon of the rotating black hole are the same as in the spherical charged black hole (see Section 4.8). These facts give serious reasons to expect that the Cauchy horizon of a rotating black hole is linearly unstable in the same way as the Cauchy horizon of a spherical charged black hole [Novikov and Starobinsky (1980a,b), see also Ori(1997)]. Possible results of this instability will be discussed in Section 14.4.5.

14.4 Structure of a Classical Black Hole Interior

14.4.1 Formulation of the problem and overview

The preceding section discussed the instability of the Cauchy horizon under small external perturbations. However, the method of small perturbations employed cannot answer the question about how the metric is changed by the growth of initially small perturbations, and whether a true spacetime singularity will thereby be produced.

This section discusses the nonlinear effects which trigger the formation of a singularity at the Cauchy horizon inside a black hole. In the Introduction 14.1 we emphasized that the problem of the black hole interior is an evolutionary problem, and it depends on the initial conditions at the surface of the black hole for all moments of time up to infinity. To specify the problem, we will consider an isolated black hole (in an asymptotically flat spacetime) which was created as a result of a realistic collapse of a star without assumptions about special symmetries. At the beginning we will neglect the influence of the quantum evaporation on the internal structure of a black hole and leave this topic for discussion at the end of the chapter.

The initial data at the event horizon of an isolated black hole, which determine the internal evolution at fairly late periods of time, are known with precision because of the no hair property. Near the event horizon we have a Kerr-Newman geometry perturbed by a dying tail of gravitational waves (see Chapter 4). The fallout from this tail produces an inward energy flux decaying as an inverse power v^{-p} of advanced time v, where $p = 4l + 4$ for a multipole of order l [Price (1972b), Gundlach, Price, and Pullin (1994a,b), Ori (1997)].

Now we should integrate the Einstein equations with the known boundary conditions to obtain the internal structure of the black hole. In general, the evolution with time into the black hole depths looks as follows. The gravitational radiation penetrating the black hole and partly backscattered by the spacetime curvature can be considered, roughly speaking, as two intersecting radial streams of infalling and outgoing gravitational radiation fluxes, the nonlinear interaction of which leads to the formation a non-trivial structure of the black hole interior. However, in such a formulation it is a very difficult and still unsolved mathematical problem, and it is necessary to make a few simplifying assumptions. In subsequent subsections we will investigate the problem step by step starting from very simple models and little by little will reveal the main properties of this evolutionary process of the formation of the internal structure of a black hole.

A detail discussion of different steps in the analysis of the problem can be found in the following works: Poisson and Israel (1990), Ori (1992, 1997), Bonanno et al. (1995), Israel (1997) and references therein.

14.4.2 Spherical symmetric charged black hole. Vaidya solution

Our first simplification is that we will consider a spherical charged black hole initially described by the Reissner-Nordström metric (see Section 6.6). A motive for this simplification is that the global structures of the rotating and the charged black holes are similar, and as we shall see, the essential physics is the same in both models. Of course, the assumption of spherical symmetry greatly simplifies the mathematical analysis.

The second simplification is that we imitate the ingoing and outgoing gravitational radiation (see section 14.4.1) by two intersecting radial streams of ingoing and outgoing lightlike particles following radial null geodesics. We assume that the streams do not interact with each other and do not scatter on the background curvature. We will see that this approximation also works very well. The main reason is that the ingoing flux is subject to infinite blueshift near the Cauchy horizon (which is a crucial aspect of the whole analysis), and any further scattering should not be important. On the other hand, it turns out that the nature of the outgoing flux is not important at all for the formation of the essential properties of the black hole interior. Under these simplifications, the problem was analyzed by Poisson and Israel (1990).

We start by considering an even simpler model. Namely, we will not consider the outgoing flux inside the black hole, and will consider the structure of the interior of a black hole model where there is only an ingoing lightlike flux inside the event horizon. This allow us to understand the manifestation of the blueshifting of the ingoing flux near r_-,[1] in its purest form. After that we will come to more complicated models.

The model for ingoing lightlike radial flux in a charged black hole can be written in the form of the *charged Vaidya metric* [Bonnor and Vaidya (1970), Sullivan and Israel (1980)]:

$$ds^2 = dv(2\,dr - f_{\text{in}}\,dv) + r^2\,d\omega^2\,, \quad f_{\text{in}} = 1 - \frac{2m_{\text{in}}(v)}{r} + \frac{e^2}{r^2}\,, \tag{14.4.1}$$

where v is a radial null coordinate (advanced time), running forward inside the black hole event horizon; $d\omega^2$ is the line element on the unit two-sphere; $m_{\text{in}}(v)$ is an arbitrary function of v and must be specified by the boundary conditions; e is the electric charge, assumed fixed. Metric (14.4.1) is a solution of the Einstein equations with the stress-energy tensor (omitting the electrostatic contribution)

$$T_{\alpha\beta} = \rho_{\text{in}}\, l_\alpha\, l_\beta\,, \tag{14.4.2}$$

where $l_\alpha = -v_{,\alpha}$ is a radial null vector pointing inwards, and $4\pi r^2 \rho_{\text{in}} = dm/dv$. If $\rho_{\text{in}} \equiv 0$, we have the Reissner-Nordström solution.

We should specify the mass function in such a way that it describes the ingoing "tail" of radiation at late advanced time (see Section 14.4.1). We can write down it in the form

$$m_{\text{in}}(v) = m_o - \mu(v), \tag{14.4.3}$$

where $\mu(v) \sim v^{-(p-1)}$, and m_0 is a constant. We want to find an asymptotic behavior of the energy density ρ_{obs} measured by a free falling observer with the four-velocity U^α when he comes to the Cauchy horizon $v = \infty$. Simple calculations [see Appendix B of the paper by Poisson and Israel (1990)] give

$$\rho_{\text{obs}} = T_{\alpha\beta}\, u^\alpha\, u^\beta = \rho_{\text{in}} (l_\alpha\, u^\alpha)^2 \sim v^{-p} \exp(2\kappa_0\, v), \tag{14.4.4}$$

where $\kappa_0 \equiv (m_0^2 - e^2)^{1/2}/r_{CH}^2$ is the surface gravity of the Cauchy horizon $r = r_{CH} = m_0 - \sqrt{m_0^2 - e^2}$. Expression (14.4.4) is infinite at the Cauchy horizon. Note that near the Cauchy horizon $m(v)$ tends to m_0 when $v \to \infty$, and the function f is regular. From this it is almost obvious that the metric is regular at the Cauchy horizon. This is not clear from (14.4.1) since the coordinates (v, r) themselves become singular. Poisson and Israel demonstrated explicitly using appropriate regular coordinates that the metric is perfectly regular at and near the Cauchy horizon. Some of the second derivatives of the metric tensor are divergent, however, the curvature invariant

$$R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta} = \frac{48m^2}{r^6} - \frac{96me^2}{r^7} + \frac{56e^4}{r^8} \tag{14.4.5}$$

is finite. In other words, though a free falling observer will register an infinite increase of the energy density, tidal forces in his reference frame remain finite as he crosses the Cauchy horizon. The reason is that infinite ρ_{obs} is situated along a null surface which is an apparent horizon. This is a so-called weak nonscalar, whimper singularity [see Ellis and King (1974), King (1975)].

14.4.3 Mass inflation

Our next step will be consideration of a model of a charged black hole with two intersecting radial streams of ingoing and outgoing lightlike particles, as mentioned in section 14.4.2. We will see that this simple model incorporates all the essential physics of the formation of a black hole interior for the general case of collapse of an asymmetric rotating body. The crucial point here is a tremendous growth of the black hole internal mass parameter, which was dubbed *mass inflation* [Poisson and Israel (1990)].

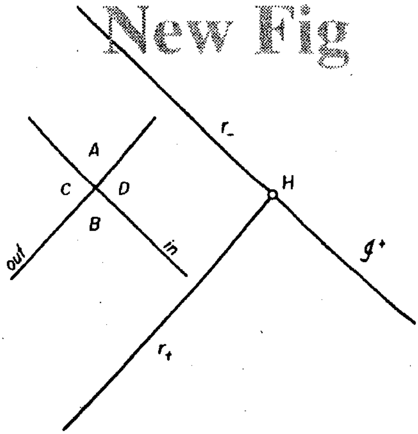


Figure 14.3: The world lines of two crossing ("out" and "in") shells moving with the speed of light between the horizons r_+ and r_-. A, B, C, and D denote four different spacetime regions bounded by shells.

We start by explaining the mechanism responsible for the mass inflation [see Dray and t'Hooft (1985), Barrabès *et al.* (1990), Bonanno *et al.* (1995)]. Consider a concentric pair of thin spherical shells in an empty spacetime without a black hole [Droz, Israel, and Morsink (1996)]. One shell of mass m_{con} contracts, while the other one of mass m_{exp} expands. We assume that both shells are moving with the speed of light (for example, "are made of photons"). The contracting shell, which initially has a radius greater than the expanding one, does not create any gravitational effects inside it, so that the expanding shell does not feel the existence of the external shell. On the other hand, the contracting shell moves in the gravitational field of the expanding one. The mutual potential of the gravitational energy of the shells acts as a debit (binding energy) on the gravitational mass energy of the external contracting shell. Before the crossing of the shells, the total mass of both of them, measured by an observer outside both shells, is equal to $m_{\text{con}} + m_{\text{exp}}$ and is constant because the debit of the numerical increase of the negative potential energy is exactly balanced by the increase of the positive energies of photons blueshifted in the gravitational field of the internal sphere.

When the shells cross one another, at radius r_0, the debit is transferred from the contracting shell to the expanding one, but the blueshift of the photons in the contracting shell survives. As a result, the masses of both spheres change. The increase of m_{con} is called mass inflation. The exact calculation shows that the new masses m'_{con} and m'_{exp} are

$$m'_{\text{con}} = m_{\text{con}} + \frac{2m_{\text{con}}\, m_{\text{exp}}}{\epsilon}, \quad m'_{\text{exp}} = m_{\text{exp}} - \frac{2m_{\text{con}}\, m_{\text{exp}}}{\epsilon}, \tag{14.4.6}$$

where $\epsilon \equiv (r_0 - 2m_{\rm exp})$. The total mass-energy is, of course, conserved $m'_{\rm con} + m'_{\rm exp} = m_{\rm con} + m_{\rm exp}$. If ϵ is small (the encounter is just outside the horizon of $m_{\rm exp}$), the inflation of mass of $m_{\rm con}$ can become arbitrarily large. In this case $m'_{\rm exp}$ becomes negative. This means that the corresponding shell is trapped inside a black hole with a mass equal to $m_{\rm con} + m_{\rm exp}$.

It is not difficult to extend this result to the shells crossing inside a charged black hole. Imagine the spacetime inside a Reissner-Nordström black hole which is split into four regions (A, B, C, D) by the two crossing shells (see Fig.14.3). Such regions are described by the solutions with metric functions

$$f_i = 1 - \frac{2m_i}{r} + \frac{e^2}{r^2}. \tag{14.4.7}$$

The index i $(i = A, B, C, D)$ enumerates the four different regions, and m_i is the value of mass in these regions.

The exact solution shows that the mass of the A region (the region after crossing of the shells, which is of interest to us) can be written in the form[4]

$$m_A = m_B + m_{\rm in} - m_{\rm out} + \frac{2m_{\rm out}\, m_{\rm in}}{r_0 f_B}, \tag{14.4.8}$$

where we have denoted

$$m_{\rm out} = m_B - m_C\,, \qquad m_{\rm in} = m_D - m_B, \tag{14.4.9}$$

$m_{\rm out}$ and $m_{\rm in}$ are the masses of the outgoing and ingoing shells, correspondingly. Note that the outgoing shell moves inside the black hole from a collapsing star to the Cauchy horizon and is trapped inside the black hole. That is why the value $m_{\rm out}$ should be negative, $m_{\rm out} < 0$. When the ingoing shell is very close to the Cauchy

[4]For spherically symmetric intersecting null shells in any spherical metric with $g^{rr} = f(r)$, one has:

$$f_A(r_0)\, f_B(r_0) = f_C(r_0)\, f_D(r_0)\,.$$

For the Reissner-Nordström metric (14.4.7) these relations can be written in a number of equivalent ways, including the following:

$$\frac{m_A - m_D}{m_C - m_B} = \frac{f_D}{f_B}, \qquad \frac{m_C - m_A}{m_B - m_D} = \frac{f_C}{f_B};$$

that is, $m'_{\rm in} = (f_D/f_B) m_{\rm in}$, and $m'_{\rm out} = (f_C/f_B) m_{\rm out}$. To obtain relation (14.4.8), it is sufficient to multiply the first of these equations by $m_{\rm out}$, the second, by $m_{\rm in}$, and to subtract the obtained expressions from one another.

The above relations imply also conservation of total energy

$$m'_{\rm in} + m'_{\rm out} = m_{\rm in} + m_{\rm out}\,.$$

If the intersection occurs at a radius r_0 close to an inner horizon of the outgoing shell: $f_B = -\epsilon$, $F_D < 0$, so that $m'_{\rm in} \sim m_{\rm in}/\epsilon$.

horizon, f_B is numerically very small (and negative), and the last term in (14.4.8) is positive and very large. It describes the mass inflation.

Now we can make the next step and improve the model, schematically describing the evolutionary formation of the black hole interior. Ori (1991) considered a continuous influx (imitating the "tail" of ingoing gravitational radiation) and the outflux as a thin shell (a very rough imitation of the outgoing gravitational radiation scattered by the spacetime curvature inside a black hole). He specified the mass $m_{\text{in}}(v)$ to imitate the Price power-law tail (see section 14.4.2) and found that the mass function diverges exponentially near the Cauchy horizon as a result of the crossing of the ingoing flux with the outgoing shell:

$$m_A \sim e^{\kappa_0 v_-}(\kappa_0 v_-)^{-p}, \quad v_- \to \infty, \tag{14.4.10}$$

where v_- is the advanced time in the region lying to the past of the shell. Expression (14.4.10) describes *mass inflation*. In this model, we have a scalar curvature singularity since the Weyl curvature invariant Ψ_2 (Coulomb component, see Appendix E.2) diverges at the Cauchy horizon. Ori (1991) emphasizes that in spite of this singularity, there are coordinates in which the metric is finite at the Cauchy horizon. He also demonstrated that though the tidal force in the reference frame of a freely falling observer grows infinitely, its integral along the world line of the observer remains finite. This singularity, though much stronger than the whimper singularity in the case of the Vaidya model (see section 14.4.2), is still quite weak (*mild singularity*).

There is one more effect caused by the outgoing flux. This is the contraction of the Cauchy horizon (which is singular now) with retarded time due to the focusing effect of the outgoing shell-like flux. This contraction continuous until the Cauchy horizon shrinks to $r = 0$, and a stronger singularity occurs. Ori (1991) has estimated the rate of approach to this strong singularity $r = 0$.

14.4.4 More realistic models of the classical black hole interior

The next step in our discussion is the following. We replace the outgoing thin shell in the previous model by a continuous arbitrary outflux which starts a finite time after the event horizon (see Figure 14.4). In this model, the continuous outflux represents the gravitational radiation scattered by the curvature of spacetime inside a black hole. This model incorporate many properties of the realistic situation. In particular, the switch on of the outgoing flux a finite time after the event horizon partly reflects the fact that the potential barrier at which radiation is scattered is peaked around the radius $r_b = e^2/m$. This region where the scattering mainly occurs is well inside the event horizon (see section 14.3). Even this model still does not describe some of the properties of the radiation scattered by the internal potential barrier. We will come back to this in the subsequent discussion in this section.

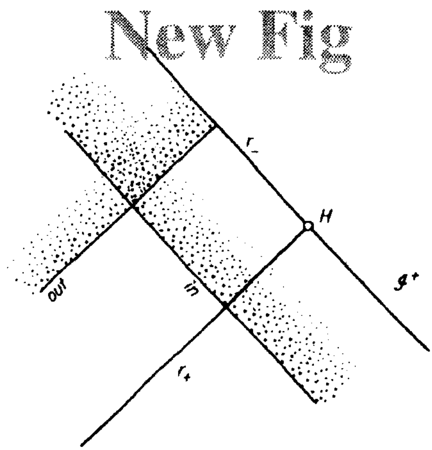


Figure 14.4: A spacetime similar to the one presented in Fig. 14.3 with two shell replaced by continuous influx and outflux which starts at some moment of v and u time, respectively.

The corresponding model with two continuous fluxes was analyzed in the pioneer paper by Poisson and Israel (1990)[see also Bonanno *et al.* (1995)]. It turns out that the main properties of mass inflation do not depend on the concrete nature of the outgoing flux, only the fact of its existence is essential. Because of this all the main properties of the Ori's model survive in this more general one. In particular, mass inflation leads to a divergence of the mass as

$$m \sim e^{\kappa_0 v}, \qquad v \to \infty. \tag{14.4.11}$$

The singularity which forms along the Cauchy Horizon is the same type of mild singularity as in Ori's model.

All previous mass inflation analysis suffers some limitations. In the picture presented it is always assumed that the outflux is turned on abruptly after some limit time behind the event horizon (see Figure 14.4). It is clear that in this model the null portion of the Cauchy horizon before meeting with the beginning of the outflux, is the Vaidya type null-like whimper singularity (see Section 14.4.2). This fact – the existence of the initial null portion of the Cauchy horizon – was important for the analysis. But what happens if we dispense with the assumption of an abrupt turn-on the outflux and consider a realistic outflux in the interior of a black hole? What is its influence on the early parts of the Cauchy horizon (in the vicinity of point H in Figure 14.4)?

This problem was analyzed by Bonanno *et al.* (1995). A star collapsing through its event horizon provides two types of outflux inside the black hole. First, there exists a direct outflux from the star after the event horizon is passed. For gravitational radiation its source could be deviations from sphericity in the distribution and

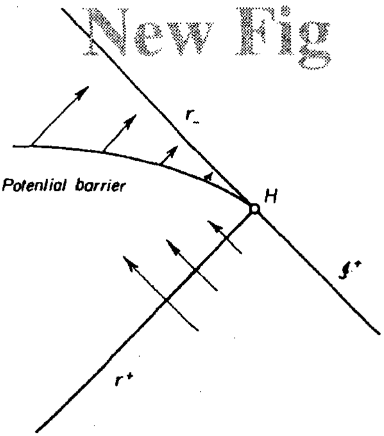


Figure 14.5: The outflux inside the black hole arising as the result of scattering of the influx off the internal potential barrier.

motion of the stellar matter, and perturbations in the gravitational field near the star. Second, there is the outgoing flux generated by scattering of the ingoing flux by the spacetime curvature inside a black hole. We recall that this ingoing flux is itself a part of radiation from outside which has been scattered by the black-hole external potential barrier and has fallen into the black hole.

The first type of outflux has a negligible effect on the vicinity of the point H because of the great dilution of the radiation. Practically all the radiation of this type falls down into the singularity not far from the collapsing star and does not reach the vicinity of the point H.

The second type of the outflux was analyzed by Bonanno *et al.* (1995). This outflux arises as a result of the scattering of the influx on an internal potential barrier which is peaked around the radius $r_b = e^2/m$ (see Figure 14.5). The outflux irradiates the Cauchy horizon during entire period of its existence. If the outflux in the vicinity of H is too strong, a spacelike singularity forms, and the Cauchy horizon does not survive. General stability arguments [Yurtsever (1993)] and early numerical studies [Gnedin and Gnedin (1993)] suggested that the Cauchy horizon is preceded by a spacelike $r = 0$ singularity; thus, the Cauchy horizon does not survive. Subsequent investigations both analytical [Bonanno *et al.* (1995), Flanagan and Ori (1996)], and numerical [Brady and Smith (1995)] did not confirm this conclusion. So, probably, the mass-inflation scenario is valid, and a part of the Cauchy horizon does exist.[5]

[5]As we mentioned above in the description of Ori's model, and as we will discuss later (see this and next sections), the generators of the Cauchy horizon are focused by the gravity of the outflux and eventually focus to a zero-radius spacelike singularity. But the mass inflation singularity precedes the $r = 0$ singularity deep inside a black hole.

From this discussion it is clear how important is the analysis of the spacetime near the point H.

In the region between the event horizon r_+ and internal barrier r_b there exists mainly an ingoing flux which produces only small perturbations (see Section 14.3). That is why this part of the black hole interior can be approximated as a Reissner-Nordström solution (with small perturbations). The task is to investigate deeper layers of the black hole. The method of Bonanno *et al.* (1995) is to solve the Einstein equations in the interior region after the scattering potential with the initial conditions imposed just after the potential barrier at r_b. These initial conditions must describe both the influx from outside and outflux just produced by scattering off the potential barrier around $r_b = e^2/m$.

Both these fluxes are weak in the vicinity of r_b and can be treated as perturbations (see Section 14.3). Bonanno *et al.* (1995) found an approximate solution which satisfies the initial conditions given by scattering from the potential barrier, and gives an accurate approximation close to the Cauchy horizon (where perturbations are not small). In Section 14.1 we formulated the problem of the black hole interior as an evolutionary problem with the boundary conditions imposed on the event horizon. Now we made a step forward and put the initial conditions inside the black hole at the potential barrier $r_b = r^2/m$, assuming that they are known from the previous analysis (as they definitely are). The great advantage of it is that on r_b we can specify not only the ingoing flux (as it originated from r_+) but also the outgoing flux. It is this outgoing flux that determines the structure of the Cauchy horizon in the vicinity of H. Bonanno *et al.* (1995) found the corresponding solutions which describe the interior of a black hole near the Cauchy horizon and satisfy the realistic initial conditions on the potential barrier r_b. Their conclusion is that the Cauchy horizon with the mild singularity (described above) really survives.

14.4.5 General structure of a classical black hole interior

We have demonstrated that a scalar curvature singularity forms along the Cauchy horizon of a spherical charged black hole in simplified models. This singularity is a result of mass inflation – the exponential divergence of the mass function with advanced time. The key ingredients producing the inflation are the blueshifted radiative flux along the Cauchy horizon and some transverse energy flux. In the case of a generic collapse of a rotating body without symmetry, the external gravitational field relaxes to a Kerr-Newman field. The analytic continuation of the Kerr-Newman solution inside the black hole also has a Cauchy horizon and the "tails" of the ingoing radiation on the event horizon are the same as in the case of a spherical charged black hole. In the pioneer paper by Poisson and Israel (1990) it was argued that the key physics underlying the analysis of a spherical charged black hole was sufficiently general and that a similar result should hold for generic collapse of a rotating body without spherical symmetry. Further analyses have supported this conjecture [Ori

(1992, 1997), Bonanno *et al.* (1994), Brady and Chambers (1995), Israel (1997)].

In the case of relaxation of the assumption of spherical symmetry, the essential new factor is shear. What are the effects of shear? Detailed analysis can be found in the papers referenced above.[6] Here we will illustrate the main effects [see Bonanno *et al.* (1995)] without getting entangled in all details of the full analysis.

For a region very close to the Cauchy horizon, the unperturbed Cauchy horizon is indistinguishable from a plane. It therefore suffices to consider the effect of shear on a crossflow of two lightlike streams with plane symmetry. For this problem the metric can be written in the following form:

$$ds^2 = -2r^{-1}e^{-2\Lambda}\,du\,dv + r^2(e^{2\beta}\,dx^2 + e^{-2\beta}\,dy^2)\,, \tag{14.4.12}$$

where r, Λ and β depend on u and v only so the shear β is homogeneous. Note that any (x, y)-dependence would get damped away exponentially during inflation.

The analysis of the solution of the corresponding field equations in a neighborhood of the Cauchy horizon shows that β remains bounded for $r \neq 0$ even when the influx diverges at the Cauchy horizon. Comparison with the spherical case shows that Λ inflates like $(1/2)\ln[m(u, v)]$ near the Cauchy horizon. Introducing the complex null tetrad $(l, n, m, \bar{m})$ defined by

$$l_\alpha\,dx^\alpha = -\,r^{-1}\,e^{-2\Lambda}\,dv\,, \quad n_\alpha\,dx^\alpha = -\,du\,, \tag{14.4.13}$$

$$m_\alpha\,dx^\alpha = 2^{-1/2}\,r(e^\beta\,dx + i\,e^{-\beta}\,dy)\,, \tag{14.4.14}$$

we obtain for the tetrad components of Weyl curvature [see (E.10)]

$$\begin{aligned}
\Psi_0 &= r^2\,e^{4\Lambda}\left[2\,\beta_{,v}\,\Lambda_{,v} - r^{-3}(r^3\,\beta_{,v})_{,v}\right], \\
\Psi_2 &= \frac{1}{3}\,r\,e^{4\Lambda}\left[\left(\Lambda + \frac{3}{2}\ln r\right)_{,uv} + 2\,\beta_{,u}\,\beta_{,v}\right], \\
\Psi_4 &= 2\,\beta_{,u}\,\Lambda_{,u} - r^{-3}(r^3\,\beta_{,u})_{,u}\,, \quad \Psi_1 = \Psi_3 = 0\,.
\end{aligned} \tag{14.4.15}$$

If the outflux is sheared ($\beta_{,v} \neq 0$), a divergence of Ψ_0 is superposed on the mass-inflationary divergence of Ψ_2 as anticipated earlier. Thus, the singularity remains lightlike and locally mild. Now, however, the mass inflation is overlaid by a gravitational shock wave imploding along the Cauchy horizon.

Let us summarize and give the general picture of the interior of the classical (without quantum effects) black hole (see Figure 14.6). The Cauchy horizon is a slowly contracting (with retarded time) lightlike three-cylinder which is mildly singular. It shrinks down and forms a strong spacelike singularity at late retarded time inside the black hole.

This is the picture according to classical general relativity. We shall see in subsequent sections that quantum effects probably add very essential traits to this picture.

[6] We should emphasize that the work on this question is still in progress (Autumn of 1997).

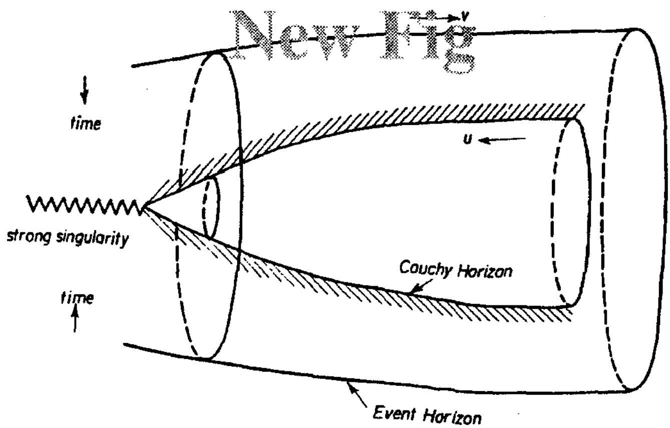


Figure 14.6: The diagram illustrating a structure of the interior of a generic classical black hole. Possible quantum effects are not taken into account.

14.5 Quantum-Electrodynamical Instability of Cauchy Horizons

The present section is devoted to quantum electrodynamic processes generated inside a charged non-rotating black hole; they were ignored in our previous analysis of the internal structure. It will be shown that quantum creation of electron-positron pairs lead to the instability of the Cauchy horizon and restructure the spacetime. A self-consistent solution is successfully constructed thereby, taking into account the backreaction of the created particles on the electromagnetic field and the metric. Besides, it is shown in the framework of this solution how the metric is modified, and it is proved that a true spacetime singularity is indeed created instead of the Cauchy horizon. This problem was solved by Novikov and Starobinsky (1980c) whose paper we follow here [see also Berezin (1980)].

Consider the restrictions imposed on the physical conditions inside black holes with different values of mass M and electric charge Q (Figure 14.7). First, a black hole is formed only if $Q \leq \sqrt{G}\,M$ [or $Q/e \leq 5\times 10^5\,M$ (grams) where e is the electron charge], that is, if the parameters lie below line 1 in Figure 14.7. If the charge of the black hole is sufficiently large, electron-positron pairs are created near it [Markov and Frolov (1970), Gibbons (1975), Damour and Ruffini (1975)]. One particle escapes to infinity, the other (its charge sign is apposite to that of the black hole) is absorbed by the black hole and diminishes its charge. It can be shown that in a very short

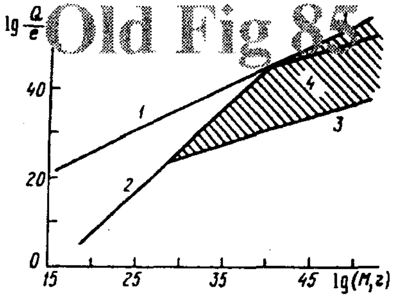


Figure 14.7: Various ranges of values of the black hole charge Q and mass M (on the boundaries of these ranges, see text).

time of order r_+/c the initial charge of the black hole is reduced to the value[7]

$$Q_2 = \frac{\pi\, m^2\, r_+^2\, c^3}{e\hbar}\,(\ln A - \ln\ln A)^{-1}, \tag{14.5.1}$$

where $A = e^2\,(2\pi\,\hbar\, c)^{-1}(r_+/\lambda)^2$; e is the electron charge, and m is its mass. Subsequently, the black hole charge remains practically constant. Numeral 2 in Figure 14.7 marks the line corresponding to equation (14.5.1). The region of allowed values of black hole parameters lies below lines 1 and 2.

Note that if the black hole charge is sufficiently small, the Cauchy horizon lies so close to the true singularity that the spacetime curvature here exceeds the critical value at which quantum gravity effects are appreciable. This entire region must be regarded as singular from a physical standpoint. We will call it the "Planck region of the Reissner-Nordström metric", or simply the "Planck region". A Cauchy horizon in the classical or semiclassical theory can be considered to have a real physical existence only if it lies outside the "Planck region". The curvature invariant $R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta}$ is of dimension (length)$^{-4}$. The boundary of the "Planck region" is found from the condition $R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta} = 1/l_{\rm Pl}^4$. For the Reissner-Nordström metric, the condition of r_- belonging to the boundary of the "Planck region" is the expression (for $r_- \ll r_+$)

$$R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta}\Big|_{r=r_-} \approx \frac{12\, r_+^2}{r_-^6} = \frac{1}{l_{\rm Pl}^4} \tag{14.5.2}$$

or

$$\frac{Q_3}{c} \approx G^{1/3}\,(\hbar\, c)^{1/6}\,\frac{M^{2/3}}{c} \approx 3\times 10^4\,\left(\frac{M}{1\rm g}\right)^{2/3} \tag{14.5.3}$$

[7] Here we consider black holes with $r_+ > \lambda$ where $\lambda = \hbar/mc$ is the Compton length of the electron [for the analysis of the opposite case, see Page (1977)]. The pairs are effectively created if the electric field Q/r_+^2 near the black hole is of the order of the critical field $E_{\rm cr}$. The latter is defined by the condition $eE_{\rm cr}\lambda \sim mc^2$ and is equal to $E_{\rm cr} = \pi m^2\, c^3/e\hbar$ (see Section 10.1).

(line 3 in Figure 14.7). On this boundary, $r_- \gg l_{\rm Pl}$. If the parameters of the black hole lie to the right of and below line 3, the Cauchy horizon lies inside the "Planck region". In actual astrophysical conditions, the relations given at the beginning of Section 3.6 are satisfied; hence, black holes with mass $M < 10^{60}\,$g definitely have $Q < Q_3$, and thus the Cauchy horizon is always inside the "Planck region".

The last step is an analysis of the region hatched in Figure 14.7. Quantum-electrodynamic processes inside a black hole are important if the parameters fall into this region. Let us introduce the following reference frame, similar to (2.4.8), in region II of the Reissner-Nordström metric:

$$ds^2 = -d\tau^2 + a^2\,dx^2 + r^2\,d\omega^2\,, \tag{14.5.4}$$

where $a = a(r)$ and $r = r(\tau)$ are defined by relations

$$\tau = -\int_{r_-}^{r} dr\,\frac{r}{[(r_+ - r)(r - r_-)]^{1/2}}\,, \tag{14.5.5}$$

$$a(r) = \frac{[(r_+ - r)(r - r_-)]^{1/2}}{r}\,. \tag{14.5.6}$$

The former coordinate t (for $r > r_+$) is now spacelike. We denote it by x ($x \equiv t$). The functions $a = a(\tau)$ and $r = r(\tau)$ are defined by formulas (14.5.5) and (14.5.6). The reference frame associated with metric (14.5.4) has a homogeneous (but anisotropic) three-dimensional space, which makes it especially convenient for calculations. The topology of this space is $R^1 \times S^2$. For an eternal black hole the three-dimensional space is infinite in the coordinate x ($-\infty < x < +\infty$). This reference frame lives for a finite proper time. Its evolution begins at the moment corresponding to the value $r = r_+$ when the expansion, described by $a(\tau)$, begins in the direction x. At the start of the expansion, $a = 0$. Here (at $r = r_+$), we find a false (coordinate) singularity. Cross-sections of our reference frame in transverse directions are spheres of radius $r(\tau)$. As τ increases, the spheres monotonically contract from the initial values $r = r_+$. At later τ, the expansion in the x coordinate is replaced by contraction, so that at $r = r_-$ the quantity a again vanishes; that is, we again come to a coordinate singularity. The radius of the spheres at this moment is $r = r_-$.

The electromagnetic field in the outlined reference frame is purely electric (the nonzero component of the electromagnetic tensor is $F_{\tau x}$); it is directed along x, and is independent of x. As τ increases, the strength of this field increases in inverse proportion to r^2. If this field is sufficiently strong, it creates electron-positron pairs. Particles created in region II cannot escape from the black hole because its boundary r_+ lies in the absolute past. As a result, they cannot affect the properties of the external space I but may greatly modify the situation inside the black hole. We will show that the gravitation of the created particles changes the metric in region II, whereby producing a true singularity instead of the Cauchy horizon.

Let us look into the details of this process. In Figure 14.7, we single out the values of the parameters so that the electric field in region II ($r_- < r < r_+$) reaches the value $E_{cr} = \pi m^2 c^3/e\hbar$ at which electron-positron pairs are intensively produced. The condition under which the electric field $E = Q/r^2$ assumes the critical value at the inner horizon r_- is $Q/r_-^2 = E_{cr}$. This relation can be rewritten in the form (assuming $Q < \sqrt{G}\,M,\ c = 1$)

$$\frac{Q}{e} = \left(\frac{4}{\pi}\frac{\hbar M^2}{m^2 e^2}\right)^{1/3} \approx 6 \times 10^{18} \left(\frac{M}{1\mathrm{g}}\right)^{2/3} \tag{14.5.7}$$

(line 4 in Figure 14.7). The condition $E > E_{cr}$ in region II is satisfied for parameters lying between lines 2, 3, and 4. In the region that we consider now, $Q \ll \sqrt{G}\,M$, that is, $r_- \ll r_+$. In this case ($r_- \ll r \ll r_+$), the classical evolution of system (14.5.4) obeys the law

$$r(\tau) \sim \left(\frac{3}{2}\right)^{2/3} r_+^{1/3}\,|\,\tau\,|^{2/3}, \qquad a(\tau) \sim \left(\frac{r_+}{r}\right)^{1/2} = \left(\frac{2}{3}\right)^{1/3} \left|\frac{r_+}{\tau}\right|^{1/3}. \tag{14.5.8}$$

At this stage, the electric field cannot yet influence the evolution of the metric (it is the same as for $Q = 0,\ r_- = 0$). If solution (14.5.8) were continued to $\tau = 0$ (as in the case of $Q \equiv 0,\ E \equiv 0$), it would result in a true singularity $\tau = 0$. In the course of contraction of the system, the electric field varies as $E = Q/r^2$ and, in our case, reaches the level E_{cr} at the stage (14.5.8), after which electron-positron pairs are intensively created and accelerated by the electric field, so that an electric current is produced. This current induces an important backreaction on the electric field. Without this reaction the field would become higher than E_{cr} and would finally change the form of (14.5.8) at r comparable with r_- and produce a coordinate singularity at the Cauchy horizon.

Novikov and Starobinsky (1980c) showed that the field E in region II cannot be appreciably greater than E_{cr}; otherwise, the current due to the pair production would bring the field down to E_{cr} in a short time $\tau_0 \approx 10^{-18}$–10^{-20} s. It is of interest to note that the field is reduced, not in a monotone manner, but through oscillations in which the sign of the field and the current direction are reversed. The relaxation of oscillations proceeds via radiative loss by particles and plasma instabilities.

The electric field thus cannot be substantially greater than E_{cr} as long as the characteristic time of field enhancement during the contraction of the system (we ignore the effect of particles being created) is greater than the relaxation time ($|\,\tau\,| > \tau_0$). In view of this, the metric is not affected by the electric field. Neither is it influenced by the gravitation of the particles created. Indeed, the energy density of the new particles transported by the field along the x axis increases in proportion to $a^{-2} r^2 \sim |\,\tau\,|^{-2/3}$, and after relaxation, in proportion to $(a\,r^2)^{-4/3} \sim |\,\tau\,|^{-4/3}$. At the same time, the spacetime curvature increases faster ($\sim 1/|\,\tau\,|^2$) than the energy density.

When the stage $|\tau| < \tau_0$ is reached in the course of evolution, the creation of particles and their motion cease to significantly affect the electric field. As a result, the electric field at $|\tau| < \tau_0$ grows in proportion to r^{-2}:

$$E \approx E_{cr}(\tau_0/|\tau|)^{4/3}. \tag{14.5.9}$$

The field begins influencing the metric when the quantity $8\pi G T_0^0$ (T_0^0 being a component of the energy-momentum tensor of the electric field) becomes equal, in order of magnitude, to the leading terms in Einstein's equation describing the spacetime curvature. These are terms of order $1/\tau^2$. Having equated

$$8\pi G T_0^0 \approx G E^2 \approx G E_{cr}^2(\tau_0/|\tau|)^{8/3} \quad \text{to} \quad 1/\tau^2,$$

we find

$$|\tau| \approx G^{3/2} E_{cr}^3 \tau_0^4 . \tag{14.5.10}$$

Using the numerical value E_{cr} and $\tau_0 \approx 10^{-18}$–10^{-20} s, we obtain $|\tau| \approx 10^{-46}$–10^{-54} s, i.e., $|\tau| \ll \tau_{Pl}$. Recall that a true physical singularity occurs at $|\tau| \approx \tau_{Pl}$. Hence, the quantum effects change the electric field and, consequently, the metric, restructuring it in such a way that it is now described by expressions (14.5.4) and (14.5.8) up to $|\tau| \approx \tau_{Pl}$. Furthermore, a true singularity appears instead of the Cauchy horizon.

The point to be emphasized is that we have constructed a self-consistent solution without resorting to the method of small perturbations as we did in Sections 14.2 and 14.3. The solution obtained gives an exact description (of course, as long as the theory is valid) of the formation of the true singularity; what is left is the region of parameters between lines 4 and 1 in Figure 14.7. Here, the electric field in region II is less than E_{cr} so that E is not appreciably affected by pair creation. Nevertheless, pairs get produced, and even a small number of them result in the formation of a true singularity.

Qualitatively, this is easy to understand. The electric field accelerates the particles created (carrying e^+ and e^- in opposite directions), and thus produces an electric current. The net three-dimensional momentum of the beams is zero. Macroscopically, one can assume that the plasma as a whole is at rest in reference frame (14.5.4), although its pressure is extremely anisotropic (unless the relaxation of the fluxes has taken place). The world lines of elementary volumes of plasma coincide with the world lines of reference frame (14.5.4) until we take into account the backreaction of the new particles on the metric. We find that, in this case, the world lines concentrate along the Cauchy horizons like the radiation due to perturbations that concentrated in region I in Figure 14.1.

Novikov and Starobinsky (1980c) constructed a self-consistent solution describing this situation. The gravitation due to the newborn plasma starts affecting the solution when r approaches r_-. Solution (14.5.8) ceases to hold when $r \approx r_-$. The reconstructed solution is of the form

$$a \sim |\tau|, \qquad r \sim \sqrt{\ln|\tau|}, \qquad \varepsilon_e \sim \frac{1}{\tau^2 \ln|\tau|} \gg \varepsilon_{(em)}, \tag{14.5.11}$$

where ε_e is the energy density of pairs created, and $\varepsilon_{(\text{em})}$ is the energy density of the electric field. Solution (14.5.11) continues until the true singularity is produced. No Cauchy horizon is thus formed in this case. What is formed is a true singularity. This is the property of self-consistent solution describing the process. This result cannot be obtained by the method of small perturbations.

In conclusion, we compare the instabilities of Cauchy horizon due to quantum effects with the classical instabilities due to external perturbations that we treated in the preceding section. Which instabilities are stronger? Obviously, a quantum instability is stronger when pairs are created intensively (for $E \approx E_{\text{cr}}$) far from r_- because it restructures the solution also far from r_-. If $E \ll E_{\text{cr}}$, both instabilities manifest themselves only very close to r_-, and the classical instability may prove stronger.

Novikov and Starobinsky (1980a,b) argued that quantum effects might be also important near the Cauchy horizon of a rotating black hole. These effects may replace the Cauchy horizon by a true singularity.

14.6 Complete picture?

In previous sections we discussed many aspects of the internal structure of black holes. We emphasized a few times that this is a problem of the evolution in time starting from boundary conditions on the event horizon for all moments of time up to the infinite future of the external observer.

It is very essential to know the boundary conditions up to infinity because we observed that the essential events (mass inflation and singularity formation) happened along the Cauchy horizon which brought information from the infinite future of the external spacetime. However, even an isolated black hole in the asymptotically flat spacetime cannot exist forever. It will evaporate by Hawking quantum radiation. So far we discussed the problem without taking into account this ultimate fate of black holes. Even without going into details, it is clear that quantum evaporation of the black holes is crucial for the whole problem. In Chapter 15 we will discuss some very idealized models of the internal structure of the black holes which account for quantum evaporation. These models have been constructed specially to discuss the possible issue of quantum evaporation or for the possible extension of the discussion of the evolution beyond a Planck singularity (see next section).

But what can one say about general picture of the black hole interior accounting quantum evaporation? We have to restrict ourselves to general remarks only because work on the problem is still in progress. To account for quantum evaporation, we have to change the boundary conditions on the event horizon as compared to the boundary conditions discussed in Section 14.4.4. Now they should include the flux of negative energy across the horizon (see Section 15.3 and references therein). If the black hole is charged or/and rotates, then the process of decrease of the charge and angular momentum is faster than decay of the mass (see Section 10.5.3). The last stage of the quantum evaporation when the mass of the black hole becomes comparable to the Planck mass is unknown. At this stage the spacetime curvature near the horizon reaches $l_{\rm Pl}^{-2}$. This means that from the point of view of semi-classical physics a singularity arises here. Probably, at this stage the black hole has the characteristics of an extremal black hole (see Section 15.3.2), and its external r_+ and internal r_- horizons coincide.

This means that the structure of the singularity inside an isolated black hole, which we discussed in Section 14.4.4 probably is more complicated. It depends on the initial parameters of the black hole as well as on properties of the process of the quantum evaporation, the final stage of which is not known. We should emphasize that in any case the solution cannot be continued beyond the true singularity. As a result, anything similar to the analytic continuation of the solution, shown in Figure 6.4, would be impossible. As for the processes inside a true singularity, they can be treated only in the framework of a unified quantum theory incorporating gravitation. Some speculation about this will be entered into the next section and in Chapter 15.

14.7 New Worlds Inside a Black Hole?

How the effects of quantum gravity could modify the structure of the spacetime singularity inside the black hole. To analyze this, let us consider a black hole which arises as a result of a spherically symmetric gravitational collapse. We know that the spacetime inside the black hole outside the collapsing matter can be described as an evolution of anisotropic homogeneous three-dimensional space. The Schwarzschild metric in the region $r \ll 2M$

$$ds^2 = -\frac{dr^2}{\frac{2M}{r}-1} + \left(\frac{2M}{r}-1\right) dt^2 + r^2\, d\omega^2 \tag{14.7.1}$$

can be approximated as

$$ds^2 \approx -d\tau^2 + \left(-\frac{3\tau}{4M}\right)^{-2/3} dt^2 + \left(-\frac{3\tau}{4M}\right)^{4/3} (2M)^2\, d\omega^2 . \tag{14.7.2}$$

As proper time $\tau(<0)$ grows, the radius $r \approx (-\tau)^{-1/3}$ decreases. This metric has the Kasner-type asymptotic behavior near the singularity: the contraction of space in two directions is accompanied by expansion in the third direction. The curvature invariant $\mathcal{R}^2 = R_{\alpha\beta\gamma\delta}\, R^{\alpha\beta\gamma\delta}$ grows as $\mathcal{R}^2 = 48M^2/r^6$.

Such behavior is a consequence of the classical equations which are valid until the spacetime curvature becomes comparable to the Planck one. Particle creation and vacuum polarization may change this regime. Quite general arguments allow one to suggest that the quantum effects may result in a decrease of the spacetime anisotropy

[Zel'dovich and Novikov (1983), Hu and Parker (1978), Hartle and Hu (1978, 1980)]. Poisson and Israel (1988) have shown that the anisotropy may be damped during this phase of contraction, while at the same time the curvature tensor remains of the order of the Planck value. Unfortunately, one cannot prove this result rigorously without knowing the physics at Planck scales.

Under these circumstances it is natural to use the following approach. One might assume that the notion of quantum average of a metric $g = \langle \hat{g} \rangle$ is still valid in the regions under consideration, and the average metric g obeys some effective equations. We do not know these equations at the moment, but we might assume that these equations and their solutions obey some general properties and restrictions. For example, it is natural to require that the effective equations for g in the low curvature limit reduce to the Einstein equations with possible higher-curvature corrections. It is also possible to assume that the future theory of quantum gravity would solve the problem of singularities of classical general relativity. One of the possibilities is that the equations of the complete theory would simply not allow dynamically infinite growth of the curvature, so that the effective curvature $\mathcal{R}$ of g is bounded by the value of order $1/l^2$.[8] This *limiting curvature principle* was proposed by Markov (1982, 1984). This principle excludes curvature singularity formation, so that the global properties of the solutions must change.

A special form of the gravitational action for cosmological models providing the limiting curvature principle was considered by Markov and Mukhanov (1985). It was shown that a collapsing *homogeneous isotropic universe* must stop its contraction and begin expansion, while during the transition phase its evolution is described by a metric close to the de Sitter one. Mukhanov and Brandenberger (1992) proposed a general nonlinear gravitational action which allows only regular homogeneous isotropic solutions. Polchinski (1989) proposed a simple realization of the limiting curvature principle by modifying the action and inserting inequality constraints into it, restricting the growth of curvature. In the case of the collapse of an inhomogeneous universe, formation of a few *baby universes* can be expected [Markov (1984)].

In the application to the problem of black hole interiors the limiting curvature principle means that the singularity which, according to the classical theory exists inside a black hole, must be removed in the complete quantum theory. We cannot hope to derive this result without knowledge of the theory, but we may at least discuss and classify possibilities. Such approach is a natural first step, and it was used in a number of publications. We discuss here[9] a singularity-free model of a black hole interior proposed by Frolov, Markov, and Mukhanov (1989, 1990). According to this FMM-model, inside a black hole there exists a closed Universe instead of a

[8] The limiting curvature does not necessarily coincide with the Planck curvature and might differ from it by some dimensional factor. For example it can be several orders of magnitude smaller. The difference between limiting curvature and Planck scales is not important here because the discussion in this section is mainly qualitative.

[9] Other models will be discussed in Chapter 15.

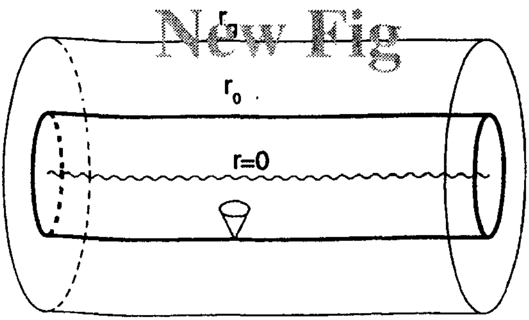


Figure 14.8: Spacetime inside a spherically symmetric eternal black hole. r_g is the surface of the gravitational radius (event horizon). Inside the horizon r is a timelike coordinate. The direction of time coincides with the direction of decrease of r. A spacelike surface Σ of constant r has the topology $S^2 \times R$. The curvature invariant at this surface is $\mathcal{R}^2 = R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} = 48M^2/r^6$. $\mathcal{R}^2$ reaches its limiting value l^{-4} at the surface $r = r_0$. The limiting curvature principle implies that in the region lying to the future of this surface (that is inside r_0 in the figure) the solution is modified.

singularity.

First consider an eternal black hole. Its interior in the collapsing T_--region is schematically presented in Figure 14.8. The curvature reaches its maximal limiting value on the surface Σ_0 where $r = r_0 = 48^{1/6}(M/l)^{1/3}\,l$. For macroscopic black holes r_0 is much larger than l. For example, for a black hole of stellar mass $r_0/l \sim 10^{13}$. The surface $r = r_0$ is spacelike, and its topology is $S^2 \times R$; that is, it is an infinite (in direction t) "tube" of a radius r_0. In the FMM-model it is assumed that inside this surface there exists a transition region after which the metric takes the de Sitter-like form

$$ds^2 = -\frac{dr^2}{\frac{r^2}{l^2} - 1} + \left(\frac{r^2}{l^2} - 1\right) dt^2 + r^2\, d\omega^2 . \tag{14.7.3}$$

It is assumed that the transition between Schwarzschild and de Sitter regimes is so fast that the transition region can be approximated by a thin spacelike shell. The global structure of spacetime in the FMM-model is shown in Figure 14.9. The spacetime passes through a *deflation* stage and instead of the singularity a new inflating universe is created. The parameter l determines the size of the spatial section at maximum contraction.

The equations for the gravitational evolution during the transition from the Schwarzschild regime to the de Sitter one can be written in the Einstein-type form

$$R_\mu^\nu - \frac{1}{2}\,\delta_\mu^\nu\, R = 8\pi\, T_\mu^\nu , \tag{14.7.4}$$

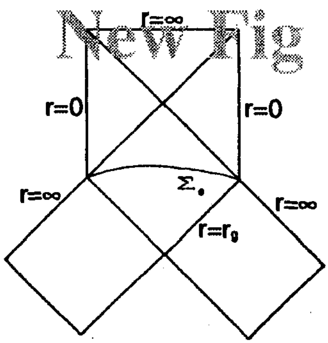


Figure 14.9: Penrose-Carter conformal diagram for an eternal black hole with a de Sitter-type world in its interior. Spacelike surface Σ_0 represents a rapid transition from the Schwarzschild regime (below Σ_0) to the de Sitter one (above Σ_0).

where T_μ^ν is a specially chosen effective stress-energy tensor. The rapid transition regime is characterized by layer parameters

$$S_m^n = \int_{\substack{\text{transition}\\ \text{layer}}} d\tau\, T_m^n\,, \quad m, n, \ldots = 1, 2, 3\,. \tag{14.7.5}$$

For $r \ll 2M$ one has

$$S_\mu^\nu = \frac{1}{4\pi l}\,\text{diag}\left(s_t^t, s_\theta^\theta, s_\phi^\phi\right), \tag{14.7.6}$$

$$s_t^t = 1 - (12)^{-1/4}, \quad s_\theta^\theta = s_\phi^\phi = 1 - \frac{1}{4}(12)^{-1/4}\,. \tag{14.7.7}$$

The concrete form of the layer parameters is not important; what is really important is that for $r \ll 2M$ they do not contain the "large parameter" $2M/l$. In other words, if T_m^n reaches the limiting value, it is sufficient that the transition continues for the time l/c. This condition makes the model self-consistent.

In the presence of the collapsing matter which produces the black hole the global structure of spacetime is modified. Figure 14.10 illustrates this change. The main difference of this model from the model for an eternal black hole shown in Figure 14.9 is that the surface Σ_0 of limiting curvature is restricted from the left. It is also restricted from the right as a result of the black hole evaporation. The proper length of this surface in the t direction is $L \sim (M/l)^{10/3}\,l$. For a stellar mass $L \sim 10^{94}$cm.

The presented model can be considered as "the creation of a universe in the laboratory". The effective stress-energy tensor (14.7.5) violates the weak energy

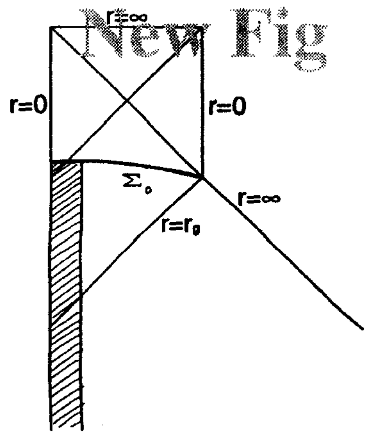


Figure 14.10: Penrose-Carter conformal diagram for the spacetime of a black hole with a de Sitter-type world in its interior. The black hole is formed by a collapsing spherically symmetric dust cloud. The spacelike surface Σ_0 represents a rapid transition from the Schwarzschild regime (below Σ_0) to the de Sitter one (above Σ_0).

condition.[10] It is also evident that the FMM-model possesses a Cauchy horizon.[11] That is why the spacetime has no singularity even though the other conditions of the Penrose theorem on singularities are satisfied. The conclusion by Farhi and Guth (1987) that the creation of a new baby universe in asymptotically flat spacetime is possible only if there exists an initial singularity is not applicable to the FMM-model because it violates the energy dominance condition.

Other (e.g., smooth) regimes of transition to the de Sitter-like phase were also considered in [Frolov, Markov, and Mukhanov (1989, 1990)]. Balbinot and Poisson (1990) proved the stability of the FMM-model with respect to small perturbations of the transition layer. Later Morgan (1991) showed that a similar result (formation of a contracting closed de Sitter-like universe with its further inflation), can be obtained in the framework of the Polchinski (1989) approach to the limiting curvature principle.

One of the assumptions of FMM and other similar models is that a "phase transition" to the de Sitter-like phase takes place at a homogeneous spacelike surface

[10]The violation of energy conditions is a characteristic property of a theory obeying the limiting curvature conjecture. One can relate this violation to quantum effects that are important in a strong gravitational field. It should be emphasized that we are not discussing here another quantum option, the quantum tunneling effect. Possibility of the creation of a universe in the laboratory by quantum tunneling was considered by Berezin, Kuzmin, and Tkachev (1988) and by Farhi, Guth, and Guven (1990).

[11]The results on the instability of Cauchy horizons discussed earlier in this chapter are not directly applicable in a theory obeying the limiting curvature conjecture. In the FMM-model the curvature near the Cauchy horizon already has the limiting value, so that it cannot be increased by focusing of blueshifted perturbations.

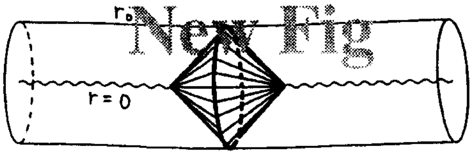


Figure 14.11: Formation of a single de Sitter bubble.

$r = r_0$. The presence of perturbations and quantum fluctuations, growing as $r \to 0$, could spoil the homogeneity. The bubbles of the new de Sitter-like phase could be formed independently at points separated by spacelike distances. For these reasons, one could expect that different parts of a black hole interior, can create spatially disconnected worlds.

A model of multiverse creation inside a black hole was proposed by Barrabès and Frolov (1996). In this model, it was suggested that spherical bubbles of the new de Sitter-like phase which are formed independently are separated from the old (Schwarzschild) phase by relativistically moving boundaries. Under this assumption the problem is reduced to the study of the evolution of lightlike shells representing the boundaries, and their intersection. The general theory of lightlike shells was developed by Barrabès and Israel (1991). This approach is purely kinematic in the following sense. It allows one to take into account the conservation of energy and momenta during the process of nucleation and the further evolution of the boundaries, including possible intersection of the boundaries of two different bubbles. But it does not answer questions concerning the probability of bubble formation or the structure of the transition regions between phases. If bubbles of the new de Sitter phase do not intersect one another, one can expect creation of $N_0 \sim L/l \sim (M/l)^{10/3}$ new baby universes. The analysis shows that disconnected de Sitter baby universes can be effectively generated even if their boundaries intersect one another, so that the total number of new-born universes can reach $N_1 \sim (M/l)^{2/3} N_0 \sim (M/l)^4$. Multiple baby universe creation in the black hole interior might have an interesting application to the information-loss puzzle (Section 15.4).

Chapter 15

Ultimate Fate of Black and White Holes

15.1 Role of Planck Scales in Black Hole Physics

On several occasions when discussing black hole properties, we mentioned Planck scales. The theory of general relativity we have used in the most of the book certainly is not a complete theory and requires modifications because of quantum effects. Quantum effects become dominant at Planck scales, and quantum gravity is required there. In this section we comment on some of the problems of black hole physics that require quantum gravity for their solutions.

First of all, it should be mentioned that besides these problems there are also situations when one formally appeals to the theory at super-Planckian energies but the final results basically are independent of the details of the physics in this region and are completely determined by low energy physics. The derivation of the Hawking radiation (Chapter 10) gives us an important example. The Hawking effect can be described as an effect of parametric excitation of zero-point fluctuations. In this process a black hole works like a boost machine. It processes high frequency input and transforms it into low frequency output. The redshift boost grows exponentially as $\exp(t/4M)$, where t is the time elapsed after the black hole formation. Particles of energy $\omega \sim M^{-1}$ emitted by the black hole at time t originate from initial zero-point fluctuations of frequency $\Omega \sim M^{-1}\exp(t/4M)$. Very soon after the black hole formation Ω becomes much larger than the Planck energy. Sometimes black holes are compared with a microscope that allows one to resolve tiny details and see something of the nature of these short distance fluctuations.[1] One might expect that details of the Hawking radiation might depend on the properties of the theory at super-Planck energies. In particular, if we assume that the spectrum of zero-point fluctuations is bounded by the Planck energy, there will be no Hawking flux of radiation after time

[1] There is a close similarity of this effect with the effect of the amplification of zero-point fluctuations in the inflating universe.

$t \sim 4M \ln(M/m_{\rm Pl})$. This conclusion might be too hasty. One cannot simply truncate the frequency of the zero-point spectrum without breaking the Lorentz invariance of the theory. Nevertheless, the problem is still with us: Does Hawking radiation require for its derivation and explanation the knowledge of Planck-scale physics? This question was analyzed recently by Jacobson (1993, 1995, 1996), Unruh (1995), Brout *et al.* (1995), Hambli and Burgess (1996), Corley and Jacobson (1996).

In 1981 Unruh showed that thermal emission is not only characteristic of black holes, but it is also characteristic of *dumb holes*, the sonic analogue of black holes [Unruh (1981), Jacobson (1991)]. A dumb hole forms when the velocity of the fluid exceeds the sound velocity. The surface where it occurs, forms a *sonic horizon*, which is similar to the event horizon. The propagation of sound in such a hypersonic fluid is similar to scalar field propagation in the spacetime of a black hole. A dumb hole emits phonons with thermal spectrum. To derive this spectrum, one makes basically the same calculations as in the black hole case. So the question of the origin of the thermal phonon radiation and of the role of ultra-high frequencies can be also referred to this model. Since one is dealing with conventional hydrodynamics, one can expect that all the divergences characteristic of quantum gravity simply do not arise. The cutoff (even if it is required) is given by the finite distance between atoms that form the liquid.

For calculations a simplified two-dimensional model was used. The dumb hole line element is

$$ds^2 = -\,dt^2 + (dx - v(x)\,dt)^2\,, \tag{15.1.1}$$

where $v(x)$ is the fluid velocity. We have assumed a constant velocity of sound ($c = 1$), and constant background density of matter. The model equation of motion of the scalar field φ describing the propagation of sound follows from the action

$$W = \frac{1}{2}\int dt\,dx\,\left[[(\partial_t + v\,\partial_x)\varphi]^2 - [iF(-i\partial_x)\varphi]^2\right]. \tag{15.1.2}$$

The function F is an odd function of ∂_x in order to make it explicit that the operator $iF(-i\,\partial_x)$ is real. This function determines the dispersion relation in a frame co-moving with the liquid:

$$\omega^2 = F^2(k)\,. \tag{15.1.3}$$

Because of the higher spatial derivatives in the action, the dispersion relation $\omega = \omega(k)$ differs at high wave numbers from the ordinary wave equation. In particular, one can exclude very high frequencies by assuming that F reaches a constant value for $k \to \infty$, or even decreases there. Unruh (1995) used the special form $F(k) = k_0 \tanh^{1/n}[(k/k_0)^n]$. By numerical integration of the wave equation he studied the propagation of wavepackets and established that, to within the numerical accuracy of his computations, Hawking radiation still occurs and is unaffected by the change of the dispersion relation. A wavepacket sent backwards in time towards the horizon reaches a minimal distance of approach, then reverses direction and propagates away from the horizon. The blueshift at the closest approach is independent of the time when the wavepacket was sent. Corley and Jacobson (1996) repeated these numerical calculations with higher accuracy for the dispersion function $F^2(k) = k^2 - k^4/k_0^2$ and demonstrated that the thermal Hawking flux, generated by "mode conversion" at the sonic horizon, is extremely close to being perfectly thermal. The difference between the two is of order $(T_H/k_0)^3$, where T_H is the temperature of the dumb hole. Brout *et al.* (1995) made an analytical study of the Unruh model, and arrived at similar conclusions.

These results indicate that Hawking radiation is a robust feature of the low-energy theory, which does not depend on the details of the theory at ultra-high energies. This point of view is supported by the following observation. Fredenhagen and Haag (1990) calculated the outgoing Hawking flux starting with the short-distance form of Green's function $G(x, x')$, when x and x' approach one another and the event horizon. Hambli and Burgess (1996) demonstrated that, by using a *Pauli-Villars regularization*, one can change the asymptotic short-distance form of the Green's function, and make it singularity-free. On the other hand, if the masses of the Pauli-Villars regulators are high, the Hawking radiation calculated for the regularized Green's function remains practically the same. A simple explanation of this fact is the following. In the Pauli-Villars regularization scheme $G(x, x')$ is changed to

$$G^{\mathrm{reg}}(x, x') = G(x, x') + \sum_i \epsilon_i\, G_i(x, x'), \tag{15.1.4}$$

where G_i are Green's functions of the regulators, and $\epsilon_i = \pm$ keep track of the sign of the corresponding fields' kinetic energy. This representation for the regularized Green's function implies that the Hawking radiation is a sum (with extra signs ϵ_i) of fluxes for each of the fields. Since the regulators are heavy, their thermal radiation is suppressed by the factor $\exp(-m_i/T_H)$. That is why the regulators practically do not affect the Hawking flux.

In the discussion of black hole entropy we also appealed to Planck scale physics. The statistical-mechanical explanation of the Bekenstein-Hawking entropy required assumptions on the material content of the underlying theory, so that the entropy is obtained by counting the states of constituents (e.g., strings). On the other hand, concrete calculations in string theory and in Sakharov's induced gravity indicate that there exists a universality; namely, the same answer for the entropy is given by different models provided they induce the same low-energy effective theory of gravity. It should be emphasized that only these universal thermodynamical characteristics can be measured in (gedanken) experiments.

One can expect that this is a generic situation. Namely, consider a massive black hole with mass M much larger than the Planck mass m_{Pl}. In order to describe the results of (gedanken) experiments on such a black hole made by an external observer, it is sufficient to know the low-energy theory. Effects that depend on the underlying theory are controlled by the parameter m_{Pl}/M and are usually negligibly small. One can call this a *low-energy censorship conjecture.*

Certainly, there are problems of black hole physics that for their solution require knowledge of the theory at Planck scales. One of them is the problem of the final stage of evaporation when the black hole mass shrinks to m_{Pl}. Planck scale physics is also important inside the black hole in the region where, according to the classical theory, there exists a singularity. Another example is a white hole, where a system starts its classical evolution from the curvature singularity. In all these cases quantum gravity is really important. Since there is no complete theory of quantum gravity,[2] the discussion of these problems is often qualitative, and is based on simplified models. In this chapter we collect some of the fragmentary results and models. The variety of possibilities greatly increases in modern unified theories (e.g., in the string theory), which open a door to extra spacetime dimensions, a wider class of fundamental fields, and new symmetries.

15.2 White Hole Instability

15.2.1 Classical instability

We start by discussing the instabilities of white holes. Formally, the solutions of Einstein's equations that describe black and white holes have a number of similar properties. Thus, the relationship between the solutions describing the formation of a black hole and the explosion of the white hole can be found by using the symmetry of Einstein's equations under time reversal. In fact, the physical properties of black and white holes, including the observational manifestations and the type of interaction with the surrounding matter, are very different. This is not surprising because the identical behavior of black and white holes would merely signify that the behavior of the surrounding matter and the characteristics of the external observer remain unchanged under the time reversal that transforms white and black holes into each other. However, this is wrong. An observer invariably moves *forward* in time and receives information on the processes in the field of a hole using *retarded* signals.

A spectacular manifestation of the asymmetry of properties inherent in white and black holes is the instability of the former. The instability of white holes may result both from classical processes due to the interaction with the surrounding matter [Eardley (1974), Frolov (1974), Eardley and Press (1975), Redmount (1984)] and from the processes due to the quantum creation of particles in the gravitational field of the holes [Zel'dovich *et al.* (1974)]. This section will be devoted to describing possible mechanisms of the instability of white holes.

Let us begin with the instability of white holes with respect to ordinary matter falling into them. What is this instability? It is the non-explosion of the white hole,

[2]On the status of quantum gravity see e.g., Isham (1995).

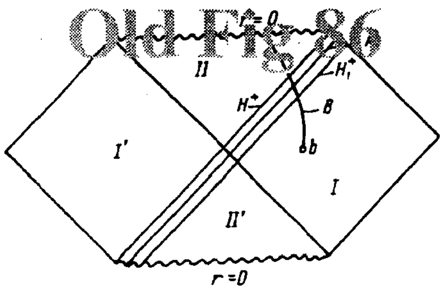


Figure 15.1: A scheme clarifying the reason for the instability of the white hole with respect to the accretion of matter from the surrounding space (see text).

however strange it may seem (recall the definition of the white hole). The spacetime of an exploding white hole is shown in Figures 2.9 and 2.10. Consider an external (i.e., at $r > r_g$) observer long before the white hole explodes. We wish to show that if a small mass of matter, δM, starts to fall into the white hole at a moment of time t_0 (for the sake of simplification, we consider the infall of a thin spherical shell), an explosion of the white hole very soon becomes impossible (by the clock of the external observer). Without accretion, the matter of the white hole would expand from the singularity and after some time would emerge from under the gravitational radius (as shown in Figure 2.9). Now accretion compels the matter to stay inside (the white hole does not explode).

The reason for the instability is as follows. Let us trace the motion of the boundary A of the exploding white hole on a Penrose diagram (Figure 15.1). For the sake of simplification, we assume that the boundary expands at an ultra-relativistic velocity; that is, it is represented by a null geodesic (this assumption does not affect the result). The longer the delay of the explosion, the closer the boundary A lies to the horizon H^+.

Let the mass δM start falling into the hole at a point b (the world line of this mass is shown by the curve B). Let us take into account the backreaction of the mass δM on the metric. Now the gravitational radius r'_g is

$$r'_g = r_g + 2\,\delta M, \tag{15.2.1}$$

where $r_g = 2M$ is the former gravitational radius. In view of the change in the metric, the world line of the horizon is H_1^+. (The shifting of I^+ and other lines due to the change in the metric are not shown in order to avoid crowding the figure.) Now it is quite clear if the world line A is to the left of H_1^+, the matter of the white hole can never emerge from under the horizon into region I where the observer is. In other words, the white hole will never explode.

Let us make several order-of-magnitude estimates. If the mass δM, moving through an unperturbed spacetime (i.e., spacetime described by unperturbed r and

t), gets closer to r_g than the perturbed horizon r'_g, the explosion of the white hole becomes impossible. Formula (2.3.4) implies that if the mass δM falls from a distance r equal to several r'_g, the fall lasts for

$$\Delta t = t - t_0 \approx \frac{r_g}{c} \ln \frac{r_g}{r - r_g}. \tag{15.2.2}$$

Replacing r with r'_g from formula (15.2.1), we obtain an estimate for the interval Δt after which the explosion of a white hole becomes impossible:

$$\Delta t \approx \frac{r_g}{c} \ln \frac{M}{\delta M}. \tag{15.2.3}$$

Obviously, a white hole retains the possibility of exploding for a short time even if δM is extremely small.

Accretion of matter onto white holes causes their instability and ultimately transforms them into a sort of black hole. The fate of such a white hole must, therefore, be discussed together with that of black holes. We will return to this situation in the next section. Frolov (1974) analyzed the changes in the motion of the expanding matter of a white hole when it collides with matter in the T_+-region (region II$'$ in Figure 15.1).

15.2.2 Quantum instability

We will now consider the quantum instability of white holes [Zel'dovich *et al.* (1974)]. This instability arises because the particles that are created with high intensity near the Schwarzschild singularity of the white hole move outward in the T_+-region and, therefore, can strongly influence the metric far from the singularity. They can also emerge from under the gravitational radius, and thus reduce the mass of the white hole. It is found that any change due to particles created in the white hole obstructs the explosion of the retarded core. Finally, another aspect of the problem arises because a white hole must have existed not in empty space but in the earliest expanding Universe. This means that the surrounding matter interacted strongly with the white hole and the particles created therein at the earliest stages of the cosmological expansion.

Let us begin by analyzing the creation of particles in the neighborhood of the Schwarzschild singularity in the T_+-region. Consider an "eternal" white hole (see Section 2.7). The singularity in it is spacelike and homogeneous. Hence, the centroids of each volume element of the created particles must be at rest in a reference frame with homogeneous space. The general form of such a reference frame (taking into account the effect of the created particles on the metric) in the spherically symmetric case is [Novikov (1964b)]

$$ds^2 = -dt^2 + e^\lambda dR^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{15.2.4}$$

where λ and r are functions only of t.

We choose t so that the singularity is at $t = 0$. As $t \to 0$,

$$r \propto t^{2/3}, \qquad e^{\lambda} \propto t^{-2/3}. \tag{15.2.5}$$

In such a metric, particles are created close to the singularity, perhaps at $t \approx t_{\rm Pl}$ [Zel'dovich and Starobinsky (1971)]. The energy density of these particles is

$$\varepsilon_{\rm Pl} \approx \frac{1}{G\, t_{\rm Pl}^2}\,. \tag{15.2.6}$$

At $t > t_{\rm Pl}$, the rate of particle creation drops sharply and can be neglected. Later on, the density decreases because of volume expansion. If the equation of the state of the new-born matter is known, then the evolution of the system can be calculated. Zel'dovich *et al.* (1974) constructed models for a number of equations of state. Not all of these models are realistic, but they possess a number of common properties that reflect the specifics of the respective problems and sometimes make it possible to completely solve the problem.

The simplest (unrealistic) assumption is that the pressure of the created particles is zero ($p = 0$). The solution is written in parametric form

$$r = \frac{1}{2}\, r_{\rm g}(1 - \cos\xi\,)\,, \quad t = \frac{1}{2}\, r_{\rm g}\,(\xi - \sin\xi\,)\,,$$

$$e^{\lambda/2} = \cot\frac{\xi}{2} + \alpha\left(1 - \frac{1}{2}\cot\frac{\xi}{2}\right), \quad 8\pi G\,\varepsilon = \alpha\, r^{-2}\, e^{-\lambda/2}\,, \tag{15.2.7}$$

where

$$\alpha \approx \frac{r_{\rm g}}{r_{\rm Pl}} \gg 1\,, \qquad 0 \le \xi \le 2\pi\,, \qquad -\infty < R < \infty\,. \tag{15.2.8}$$

This solution describes the uniform expansion of the mass of new-born particles from the moment $t \approx t_{\rm Pl}$ to the moment t_1 at which $r = r_{\rm g}$ and energy density is $\varepsilon = (8\pi G\, r_{\rm g}^2)^{-1}$, followed by a subsequent compression of the matter to the singularity.

In order to clarify the physical meaning of the solution, we demand that the following condition be satisfied: Particles are created near $t \approx t_{\rm Pl}$, on the interval of R from $-\infty$ to some R_1, with no particles created at $R > R_1$. (Later we will show how to make this assumption realistic.) The structure of the spacetime then has the form shown in Figure 15.2. The entire mass of the created particles lies under the gravitational radius and does not leave the white hole.

Now we analyze not an "eternal" white hole but a hole with a retarded expansion of the core. It is not difficult to show that the particles created close to the Schwarzschild singularity do not allow this core to escape from under the gravitational radius. Indeed, a prolonged retardation of the expansion indicates that the boundary of the core must lie at $R = R_2$ at the moment $t \approx t_{\rm Pl}$ (close to $r = 0$),

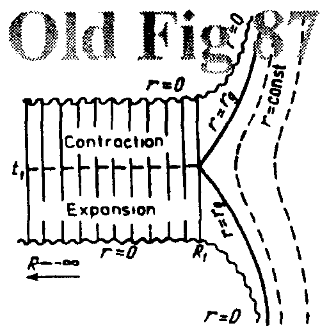


Figure 15.2: Schematic presentation of the expansion and compression of matter at $p = 0$ inside a black hole. The matter is created via quantum processes close to $r = 0$ to the left of R_1.

far to the left of the point R_1 ($R_2 \ll R_1$). A signal leaving R_2 at a moment $t_{\rm Pl}$ and traveling to the right from R_1 covers only a finite distance ΔR during the entire time of expansion t_1. Estimates show that

$$\Delta R \approx 7\, r_{\rm Pl} \ll r_{\rm g}\,. \tag{15.2.9}$$

If $R_1 - R_2 \gg \Delta R$, the signal does not reach R_1 by the time t_1 when expansion is terminated. Not only will the matter of the retarded core be unable to escape to a distant observer, but no signals emitted by the explosion will escape from the white hole and reach R_1. The retarded core will be buried under the mass of particles created.

Let us discuss the assumption of no particle creation near $r = 0$ to the right of the coordinate R_1. One has to remember that the white hole evolves not in an empty space but in the expanding Universe [Novikov (1964b), Ne'eman (1965)]. If the matter of homogeneous cosmological model surrounding the white hole is located to the right of R_1 close to $r = 0$, there is almost no particles production in this region (under standard assumptions; see Zel'dovich and Starobinsky (1971)]. If we also assume that the pressure in the surrounding matter is zero (which is not realistic), the matter produces no effect on the region to the left of R_1. The structure of spacetime in this model is shown in Figure 15.3.

Consider more realistic models. Assume that the new-born particles do not interact with one another, and thus form two colliding beams that travel along the radial coordinate with the speed of light. In this case, $-T_0^0 = T_1^1 = \varepsilon$, the remaining $T_\nu^\mu = 0$. Another assumption is that the interaction of the created particles produces Pascal pressure $p = \varepsilon/3$. The solutions of these cases are similar to (15.2.7) for $p = 0$; see Zel'dovich *et al.* (1974). They also describe the expansion of the system to some $r_{\rm max}$ and the subsequent contraction to a singularity. Here again the signal traveling at a speed of light covers only a finite small distance along R over the system's entire

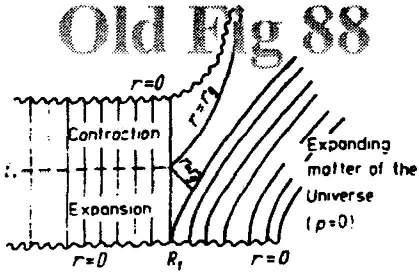


Figure 15.3: A white hole with created matter ($p = 0$) in the cold Universe model (with matter also at $p = 0$).

expansion time. This is why the created particles do not allow a retarded-expansion core (as in the case of $p = 0$) to explode and expand until it reaches the external observer. An essential difference in comparison with the case of $p = 0$ is that a matter flux to the right across the boundary R_1 arises if $T_1^1 \neq 0$. This flux may escape from under r_g and thereby reduce the mass of the white hole.

If the white hole is in a cold Universe with matter at zero pressure ($p = 0$), then the reduction in the white hole mass due to the spontaneous leaking of the new-born matter from the hole can be quite high, as was shown in the already-cited paper by Zel'dovich *et al.* (1974).

However, if the white hole is considered in a real hot Universe with matter, and its equation of state is $p = \varepsilon/3$, the situation is changed. The pressure of the surrounding hot matter constrains the leaking of the new matter from the white hole. In this case it is very likely that the loss of mass due to leaking is considerably lowered. We will not go into the details of this phenomenon since it is more a problem of cosmology (on the accretion of matter onto compact cores in the hot Universe, see Section 9.8).

15.3 What is Left after the Quantum Decay of a Black Hole?

15.3.1 Possible outcomes of black hole evaporation

Unfortunately, it is not possible to date to give an unambiguous answer to this question. The point is that any attempt to solve this problem brings one face to face with other problems that belong in the realm of quantum gravity. Theoretically, quantum gravity appears to be quite far from completion, while the difficulties inherent in it (divergences, non-renormalizability, ambiguity of going off the mass shell, the possibility of changes of spacetime topology) are fundamental. As a result, there is no complete self-consistent quantum theory of evaporating black holes. A natural approach to this situation is to analyze models that reflect specific facets of the complete problem.

In principle, there are several possible outcomes of the black hole evaporation. The semi-classical approximation should be good until the black hole has shrunk nearly to the Planck mass. After this:

1. The black hole can completely disappear, and the information captured in the black hole interior would disappear from our world.
2. The black hole can disappear, but the information is released back.
3. There remains a stable black hole remnant.

Spacetime structure for an evaporating black hole

In order to illustrate the possible spacetime structure for each of these outcomes, let us consider the following model. We limit the analysis to the spherically symmetric case.[3] The corresponding average metric $g_{\mu\nu} = \langle \hat{g}_{\mu\nu} \rangle$ is conveniently written in the form [Bardeen (1981)]

$$ds^2 = -\,e^{2\psi}\,F\,dv^2 + 2\,e^{\psi}\,dr\,dv + r^2\,d\omega^2. \tag{15.3.1}$$

Here v is the null coordinate of the advanced time, and ψ and F are functions of r and v with the following invariant meaning:

$$F(r,v) \equiv g^{\mu\nu}\,r_{,\mu}\,r_{,\nu}\,, \qquad e^{-\psi(r,v)} = g^{\mu\nu}\,r_{,\mu}\,v_{,\nu}\,. \tag{15.3.2}$$

Assume that the spacetime is asymptotically flat and demand that

$$\lim_{r\to\infty} F(r,v) = 1\,, \qquad \lim_{r\to\infty} \psi(r,v) = 0\,. \tag{15.3.3}$$

Of course, the range of applicability of geometrical description in terms of the averaged metric $g_{\mu\nu} = \langle \hat{g}_{\mu\nu} \rangle$ is limited. Thus, it is not valid on scales less than $l_{\rm Pl}$ owing to the strong quantum fluctuations of the gravitational field.

Essential information on the properties of spacetime can be obtained by analyzing the behavior of the surfaces $F = $ const of the function F. Thus, the exterior part of the surface $F = 0$ coincides with the apparent horizon. If the created black hole were static, the apparent horizon would coincide with the event horizon, and the surface $F = 0$ would be described by the equation $r = 2M$, where M is the mass of the created black hole. The quantum evaporation of the black hole makes the horizon

[3]The theorem on the "falling of hair" inside the black hole (see Section 14.1) states that the farther we are from a collapsing non-rotating body, the less the deviation of the spacetime in the T_--region is from spherical symmetry. Hence, this theorem gives grounds for assuming that the conclusions drawn for spherically symmetric black holes may prove to be valid for more general situations.

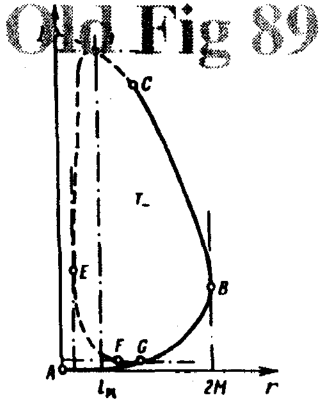


Figure 15.4: Versions of the possible behavior of the apparent horizon in the quantum evaporation of a black hole.

nonstatic, with the size decreasing with time (curve BC in Figure 15.4). If $r = \rho(v)$ is the equation of the outgoing radial null rays, we find that on the $F = 0$ surface,

$$\frac{d\rho}{dv} = \mathrm{e}^{\psi} F = 0, \qquad \frac{d^2\rho}{dv^2} = (\mathrm{e}^{\psi} F)_{,v} = \mathrm{e}^{\psi} F_{,v}. \tag{15.3.4}$$

Since $F < 0$ in T_--region and F changes its sign on the line BC, one has $d^2\rho/dv^2 > 0$ on the segment BC.

Using expression (15.3.1) for the metric, we can calculate the corresponding Ricci tensor and verify that in this general case the metric satisfies Einstein's equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi T_{\mu\nu} \tag{15.3.5}$$

with a nonzero right-hand side. In particular,

$$T_{vv} = \frac{1}{8\pi r}\left\{\mathrm{e}^{\psi} F \left[\frac{\mathrm{e}^{\psi}}{r}\,\partial_r(r(1-F)) + \psi_{,v}\right] - (\mathrm{e}^{\psi} F)_{,v}\right\}. \tag{15.3.6}$$

This relation simplifies for the surface $F = 0$ (the apparent horizon):

$$T_{vv}|_{F=0} = -\frac{1}{8\pi r}\left[(\mathrm{e}^{\psi} F)_{,v}\right]\Big|_{F=0} = -\frac{1}{8\pi r}\frac{d^2\rho}{dv^2}\Big|_{F=0}. \tag{15.3.7}$$

It shows that there is a flux of negative energy density across the segment BC of the apparent horizon, in complete agreement with the results presented in Chapter 11.

For describing the processes over the entire time interval v during which the black hole mass $m(v)$ [e.g., we can choose $m(v) = r/2|_{F=0}$] is much greater than the Planck mass m_{Pl}, and hence the rate of change of the apparent horizon size, $d(r|_{F=0})/dv$, is small in comparison with the speed of light, one can use the *quasi-static approximation* [Hajicek and Israel (1980), Bardeen (1981), Frolov (1981), Nityananda and Narayan (1981)].[4]

The last stage of evaporation at which the mass of the black hole becomes comparable to the Planck mass is the most difficult one to describe. The spacetime curvature near the apparent horizon may reach $1/l_{\rm Pl}^2$ at this stage so that in the general case the mean metric can be found only if the effective action, including all quantum corrections, is known. We can conclude only that if the surface $F = 0$ intersects $r = 0$, according to our semi-classical picture, a singularity arises in which curvature invariants tend to infinity [Frolov and Vilkovisky (1979, 1981), Kodama (1979, 1980)]. It is more reasonable to expect that the semi-classical picture is not valid in this region and quantum fluctuations become important. It is more probable that instead of creating a singularity the system makes a quantum jump with a change of spatial topology. As a result of this quantum-gravity process, the black hole disappears completely. This point of view was advocated by Hawking (1976b).

Frolov-Vilkovisky model

In order to illustrate the second possible outcome of black hole evaporation listed on page 585, let us consider the following model proposed by Frolov and Vilkovisky (1979, 1981, 1982) [see also Tomboulis (1980), and Hasslacher and Mottola (1981)]. In this model the $F = 0$ surface is closed and never intersects the line $r = 0$ (line BCDEFG in Figure 15.4). In this case, the singularity inside the black hole disappears.[5]

The spacetime near $r = 0$ is locally flat for this solution, and we may expect that its curvature at $r \leq l_{\rm Pl}$ is of order $l_{\rm Pl}^{-2}$, with the internal part of the line $F = 0$ (FED) being separated from $r = 0$ by some distance. It can be shown, using the general relation (15.3.7), that $T_{vv} < 0$ on the segment EDB and $T_{vv} \geq 0$ on the segment $EFGB$, where E and B are the points at which $F = 0$ is tangent to the lines $r =$ const [Roman and Bergman (1983)].

A spacetime with closed horizon $F = 0$ does not have an event horizon. Rigorously speaking, a black hole does not exist in this situation. However, a region that does not let any signals out exists throughout the time of quantum evaporation. If the initial mass of such an object is much greater than the Planck mass, its manifestations are indistinguishable from those of a black hole for a very long time (time of its complete evaporation).

A number of fundamental questions arise when this model of a singularity-free

[4]For detailed study of the geometry of evaporating black holes, see also Volovich *et al.* (1976), Hiscock (1981), Balbinot and Bergamini (1982), Balbinot *et al.* (1982), Balbinot (1984), Kuroda (1984a,b).

[5]Recall that owing to quantum effects, the total effective energy-momentum tensor in Einstein's equations does not satisfy, in the general case, the conditions of positive energy density and positive pressure. As a result, quantum effects may violate the conditions of existence of a singularity inside a black hole (see Section 5.6), so that the singularity may be absent.

"black hole" is analyzed. One of them concerns the conservation of the baryon charge in this system. Assume that a collapsing system has a considerable baryon charge. This charge cannot change substantially in the process of quantum evaporation, owing to the symmetry of baryon and antibaryon creation.[6] Then again if this "black hole" completely evaporates, the original baryon charge disappears. We thus come to an obvious violation of baryon charge conservation.

This situation could be regarded as a difficulty inherent in the model if processes that do not conserve baryon charge were impossible. In fact, such non-conserving processes have been widely discussed in connection with Grand Unification Theories; they involve supermassive (with masses of 10^{14}–10^{15} GeV) vector X and Y bosons. When the collapsing matter is compressed to a density $\rho \sim 10^{74}$–10^{78} g/cm^3, corresponding to the mass of these particles, the system almost immediately becomes neutral with respect to baryon charge, regardless of its initial value.[7] As a result, the matter can get rid off its original baryon charge, even before the Planck density $\rho_{\rm Pl} \sim 10^{94}\,{\rm g/cm^3}$ is reached [Bolashenko and Frolov (1984, 1986)]. There exist other mechanisms. For example, Hawking (1984) suggested purely quantum-gravity mechanism of baryon charge non-conservation that may prove to be important at Planck densities. Coleman and Hugh (1993) showed that, in general, all global charges are extinguished before the infalling matter crosses the singularity by the wormhole-induced global-charge violation mechanism.

The motion of particles and light beams in a spacetime with a closed horizon $F = 0$ is characterized by a number of unusual features. Particles falling along radii, cross the T_--region in a short proper time of order r_g/c, reach the line $r = 0$, and start moving away from the center. However, they cannot cross the line ED again and enter the T_--region. Therefore, all such particles accumulate close to ED (in the classical description) and escape to the external space after the "black hole" has evaporated (in a proper time of order r_g/c). The particles have a "blueshift" $\propto \exp(\kappa_- V_{BH})$ where

$$\kappa_- = \frac{1}{2}\left|\frac{\partial(e^{\psi}F)}{\partial r}\right|_{F=0} \tag{15.3.8}$$

is an analogue of the surface gravity for the interior horizon (on the line ED), and

[6] Zel'dovich (1976) pointed out that if the Hawking radiation creates heavy particles whose decay violates the CP parity, an excess of the baryon or antibaryon charge may arise. These processes were analyzed in detail by Dolgov (1980a,b, 1981). These processes are essential only at a relatively late stage of evaporation, when the black hole temperature reaches a value $\theta = 1/8\pi M \sim 10^{14}$–$10^{15}$ GeV; hence, the baryon asymmetry of the decay cannot substantially change the baryon charge of the black hole of mass $M \gg 1$ g, consisting of baryons [see the review of Dolgov and Zel'dovich (1980, 1981)].

[7] These processes were analyzed in detail in connection with the origin of the baryon asymmetry of the Universe [see, e.g., Dolgov and Zel'dovich (1980, 1981), Barrow (1982), Kolb and Turner (1983)]. Estimates of the rate of neutralization of the baryon charge in the superdense matter in Grand Unification Theories can be found in Fry, Olive, Turner (1980a,b,c) and Kolb and Turner (1983).

V_{BH} is the lifetime of the "black hole". A similar "blueshift" effect must take place for waves trapped in such a "black hole". In quantum analysis, this effect results in the extremely intensive creation of particles in the decay of the "black hole". As this energy release cannot be higher than a quantity of the order of the Planck mass (otherwise, the energy conservation law in the external space would be violated), one can conclude that the surface gravity κ_- must be less than, or of order, V_{BH}^{-1} provided the estimate of radiation based on using quantum theory in a given averaged metric is correct [Bolashenko and Frolov (1984, 1986)].[8]

The absence of an event horizon in a model with a closed horizon could lead to another extremely interesting corollary. The creation of an outgoing particle in a black hole is accompanied by a particle emerging inside it. A distant observer records only some of the particles; hence, the radiation of the black hole possesses entropy and is described by a density matrix (Section 10.3). There is no event horizon in the model with a closed horizon, so that the particles created inside the "black hole" can leave it after evaporation is completed. As a result, the quantum state could again be pure from the standpoint of the distant observer. In other words, the growth of entropy in the external space due to the thermal radiation of the black hole at a stage when its mass is still much greater than the Planck mass is predicted as being replaced by a sharp drop to zero at the last stage of decay. (We shall discuss this possibility in section 15.3.4.)

The analysis above operated in terms of the approximation in which the created particles are assumed non-interacting and the fluctuations of the gravitational field are neglected. Both these assumptions seems to fail for the description of the propagation of particles in the region close to the internal horizon ED. The interactions between particles inside a black hole and the scattering of particles on gravitational field fluctuations are such that particles may "forget" their phases[9] and no entropy decrease occurs when the black hole decays.

15.3.2 Elementary black holes: maximon, friedmon, and so on

In addition to the scenarios discussed above (the formation of a naked singularity and a model with a closed horizon in which a black hole burns out completely), another version is possible in which a residue is left after the black hole has evaporated. An *elementary black hole* of a mass of the order of one Planck mass is a possible form of this residue.[10] In Figure 15.4, this case would correspond to the following behavior

[8]We emphasize that this conclusion is obtained without taking into account the gravitational field fluctuations. On the possible relationship between the fluctuations of the apparent horizon and quantum radiation of black holes, see Kodama (1980).

[9]On the mechanism of coherence loss in the scattering by quantum-gravity fluctuations, see Hawking (1984, 1988a,b), Lavrelashvili *et al.* (1988).

[10]Note that the existence of heavy magnetic monopoles in nature, as predicted by Grand Unification Theories, could have a curious consequence for small black holes [Gibbons (1977), Hiscock

of the $F = 0$ line: Its outer part being infinitely extended into future never ends and never crosses the $r = 0$ line. (In the model without singularity it occurs if the inner and outer parts of the line $F = 0$ continue infinitely along the coordinate v, getting very close to each other or even merging.) An analysis of the spherically symmetric collapse of a system of less than the Planck mass shows that quantum effects make the "averaged metric" $g_{\mu\nu} = \langle \hat{g}_{\mu\nu} \rangle$ everywhere regular. There is no apparent horizon, and hence no event horizon [Frolov and Vilkovisky (1979, 1981, 1982)]. This result indicates that black holes of less than the Planck mass cannot exist; that is, if elementary black holes exist at all, their masses must be of the order of the Planck mass.

Stable elementary black holes of Planck mass [Hawking (1971)] may play the role of the heaviest possible fundamental particles [Markov (1966, 1976, 1981c)]. If we relate the size of a particle with a Compton wavelength $\lambda = \hbar/mc$, this size is found to be smaller than the gravitational radius for particles of masses $m > m_{\rm Pl}$.[11] For this reason, Markov (1965, 1966) proposed the name *maximon* for such an elementary black hole.

The problem of stability with respect to quantum decay is one of the basic problems for the hypothesis of elementary black hole existence. A black hole does not emit Hawking radiation when its temperature vanishes. This occurs if the black hole is extremal. In Einstein gravity this happens when its parameters—the electric (Q) and magnetic (P) charges and the angular momentum (J)—are related to the mass M of the black hole by the formula

$$M^2 = \frac{Q^2 + P^2}{G} + \frac{J^2 c^2}{G^2 M^2}. \tag{15.3.9}$$

Owing to the quantum effect of vacuum polarization, the modification of the Einstein-Maxwell equations may change conditions (15.3.9) for the vanishing of the black hole temperature. If an extremal black hole is charged or rotating, it still can emit particles because of the superradiance effect.

Extremally charged elementary black holes[12] were considered by Markov and Frolov (1970, 1972), who showed that because of creation of charged particles they

(1983)]. The extremal (with a magnetic charge) black holes of mass $M > 150 \times (10^{17}\,\mathrm{GeV}/\mu)^2$ g (μ is the monopole mass in GeV) would have a lifetime greater than that of the Universe because the Hawking temperature of such a hole is zero and the process of monopole creation is suppressed owing to their high mass.

[11]Note that in this situation the Planck length acts as a sort of fundamental length. It can be shown [Ginzburg and Frolov (1976)] that, in the most general case, the introduction of a fundamental length yields a restriction from below on the spectrum of possible black holes.

[12]Note that extremally charged black holes are of great interest for analyzing the problem of self-energy of charged particles. In the framework of classical theory, the gravitational mass defect creates a difference between the mass M observed at infinity and the internal mass M_0 of the system. If the system is neutral, configurations are possible at a fixed value of M_0 at which M is arbitrarily small [Zel'dovich (1962a)] or identically vanishes (e.g., the case of M_0 forming a closed world). In charged systems (with charge Q), the value of M is bounded from below by the value $Q/\sqrt{G}$ (for the magnetic charge, by $P/\sqrt{G}$) [Arnowitt *et al.* (1963), Markov and Frolov (1970, 1972), Gibbons and

are unstable unless their charge is less than $137e$, where e is the charge of the electron. The smallest possible elementary black hole with the charge e was called a *friedmon* [Markov and Frolov (1970, 1972)]. The mass $e/\sqrt{G}$ is 1.86×10^{-6} g, less by an order of magnitude [by a factor $(\hbar c/e^2)^{1/2} \approx 11.7$] than the Planck mass.

More recently interest in the possible remnants of black holes was stimulated by discussion of the *information loss* puzzle (see Section 15.3.5). Modifying the low-energy gravity action by introducing dilaton and other fields opens new dimensions of the problem. For example, the extremal version of the magnetically charged dilation black hole (13.2.10) (called *cornucopin* or *horned particles*) was proposed as a possible candidate for a black hole remnant [see, e.g., Banks (1995) and references therein].

The elementary black holes (maximons, friedmons, and so on) would have the smallest admissible mass. A loss of arbitrarily small mass results in their complete decay. It is logical that this process should yield quanta with a characteristic energy $\varepsilon \sim m_{\rm Pl}c^2$, whose wavelength $\lambda \sim \hbar c/\varepsilon$ is comparable to their gravitational radius. Presumably, the approximation based on assuming that new-born particles exert a negligible effect on the metric is unacceptable under these conditions. When we are speaking about elementary black holes as possible final state of an evaporating black hole, we assume that they are stable. Decaying elementary black holes are just another version of the case 1 (complete evaporation) where the last phase takes a long enough time.

If nature admits elementary black holes, they possess a number of fascinating properties [Markov (1966)]. They are characterized by an extremely small interaction cross-section of order $10^{-66}\,\text{cm}^2$. When a maximon falls in the gravitational field of the Earth, it gains energy of order $10^{20}\,\text{eV}$. However, it seems to be impossible to observe maximons, using their ionizing ability, even if they are charged and the interaction with matter is sufficiently strong because of their low velocities. Maximons are difficult to retain and accumulate in ordinary matter on the surface of the Earth because the terrestrial gravitational field imparts energy of order $10^3\,\text{eV}$ on the length scale of the intermolecular spacing, which is much more than the energy of the intermolecular interaction.

As a result of the weakness of interaction between matter and neutral maximons, a considerable (and even predominant) part of matter in the Universe could consist of maximons. Thus, maximons could act as the the dark matter which is now recognized as a reality in cosmology [Markov (1981b)].[13]

It appears that the most promising method of searching for maximons is based

Hull (1982), Ludvigsen and Vickers (1983)]. A system with large internal mass M_0 and small charge $Q \ll \sqrt{G}\,M_0$ in its lowest energy state forms a *semi-closed world*. For example, such a solution can be obtained by "inserting" the charge into the initially closed Friedmann-Robertson-Walker world [Markov and Frolov (1970, 1972)]. In this picture the entire internal universe looks like a black hole remnant to an external observer.

[13]Strong restrictions on the admissible mean density of maximons in the Universe can be obtained using arguments similar to those employed in the derivation of the restrictions on the number of monopoles and other massive particles [see, e.g., Polnarev and Khlopov (1985)].

on detecting their decay products. If one assumes the existence of a bound state of many maximons [Markov and Frolov (1979)] or of a small number of them, for instance, of a maximon pair [Aman (1983)], a merger of a pair of maximons into one can be expected in the evolution of such systems, with energy release of order 10^{28} eV. Processes of this type could presumably by recorded in DUMAND-type experiments [for details, see Markov (1981a) and Markov and Zheleznykh (1981)].

15.3.3 Virtual black holes

Elementary black holes, even if they are unstable, are important for particle theory for at least one more reason. Indeed, when calculations are carried out in modern quantum theory (for example, calculations of the self-energy of particles), it is usually necessary to take into account the contribution of intermediate states with an arbitrarily high energy; the result is the familiar divergences. If the gravitational interaction of the appropriate virtual particles is included in the calculations, and the possibility of the appearance of *virtual* (short-lived) black holes in the intermediate states is taken into account, these divergences may be removed [Markov (1971)].

Virtual black holes may also appear in the vacuum as a result of quantum fluctuations. Quantum fluctuations of the gravitational field are the greater, the shorter the length scale. The fluctuations of the metric are comparable to the metric itself over distances of the order of the Planck length. Such fluctuations signal the possibility of strong deviations from the flat geometry and Euclidean topology. In other words, continuous creation and annihilation of virtual black holes makes the spacetime on small scales resemble a soapy foam.

The concept of a foam-like (cell) structure of spacetime, formulated by Wheeler in the 1950's, was developed by Hawking and his coworkers [Hawking (1978, 1984, 1996), Hawking *et al.* (1979, 1980), Warner (1982)].

Some interesting applications of these ideas deserve special mention: (1) possible violation of quantum coherence and the transformation of a pure state into a mixed one as a result of the interaction of the quantum field with the fluctuations of the gravitational field [Hawking (1984, 1988a,b, 1996) and Lavrelashvili *et al.* (1988)], and (2) non-conservation of the baryon and lepton numbers in the process of interaction between elementary particles and virtual black holes (the space "foam") [Hawking (1984)]. Even though the proton lifetime with respect to this process is many orders of magnitude greater than that predicted in Grand Unification Theories, the very possibility of such processes may be of fundamental importance, especially in the discussion of the origin of the Universe.

15.4 Information loss puzzle

The Hawking radiation from a black hole is generated by uncorrelated thermal emission in each mode (Chapter 10). This result is obtained in the semi-classical approximation, which is supposed to be good until the black hole has shrunk to nearly the Planck mass. If the black hole disappears completely in this process, what is left is just thermal radiation. This occurs even in the case where the initial state that underwent gravitational collapse into a black hole was a pure quantum state. This process would correspond to a loss of information. Namely, even if the initial quantum state were precisely known, we cannot predict with certainty what the final quantum state will be. The best we can do is to assign probabilities to various alternatives.

In order to describe this type of process, Hawking (1976b, 1982) proposed to generalize quantum mechanics to allow for information loss. He proposed to replace the unitary S matrix, describing a unitary evolution of the quantum system and mapping an initial pure state to a pure final state, by a ***superscattering operator*** $\hat{\$}$ which acts on density matrices rather than state vectors. If the initial state of the system is given by the density matrix $\hat{\rho}^{\text{init}}$ and the final density matrix is $\hat{\rho}^{\text{final}}$, then

$$\rho_{ab}^{\text{final}} = \$_{ab}^{cd}\, \rho_{cd}^{\text{init}}\,. \tag{15.4.1}$$

Here indices a, b, c and d refer to the states in the Hilbert space of the quantum system. In the standard quantum mechanics, one has $\$_{ab}^{cd} = S_a^c\, \bar{S}_b^d$, where $\hat{S}$ is a S-matrix. In the case when information is lost, the superscattering operator $\hat{\$}$ does not allow such factorization. The *information loss* proposed by Hawking would be a new feature of quantum gravity, not seen in other quantum field theories.

If Hawking's conclusion is correct, then we cannot simply combine quantum mechanics and general relativity. Since this is quite a radical step requiring a change of the fundamental principles of modern theoretical physics, this issue has been much debated, but it has not been definitely resolved. Here we briefly discuss only some of the problems and proposals connected with the information loss puzzle. For more detailed discussion, see the reviews by Preskill (1992), Page (1994), Banks (1995), Giddings and Thorlacius (1995), and Strominger (1995).

First of all, it must be stressed that Hawking's derivation of the thermal radiation uses the semi-classical approximation and can be directly applied only to a macroscopic black hole. This prediction is derived by applying the ordinary dynamical evolution laws to the quantum field, so that no violation of any principles of local quantum field theory occurs. The loss of quantum coherence is directly related to the failure of the final time surface in the external space (after the black hole evaporates) to be a Cauchy surface.

Zel'dovich (1977b) argued that the indeterminacy connected with the loss of the information can arise as a result of the semi-classical treatment of the problem. Page (1980) indicates that the semi-classical approximation would be expected to break down from fluctuations of the black hole momentum, long before the black hole has shrunk to the Planck size. Although this particular fluctuation effect would not help restore the information, it illustrates the importance of going beyond the semi-classical approximation. Nevertheless, for a massive black hole the semi-classical approximation does work well for a time long enough to allow the black hole to emit a considerable part of its entropy. In order to escape the Hawking conclusion, one must indicate the mechanism by which the lost information is restored.

Since the Hawking suggestion is quite radical, there were a lot of attempts to escape it by finding a mechanism which preserves the validity of quantum mechanics. Let us examine some of the alternatives [for more detailed discussion, see Preskill (1992), Page (1994), and Giddings and Thorlacius (1995)].

1. Information comes out with the Hawking radiation. At first sight, this possibility looks the most natural. Indeed, we know from our ordinary experience that burning a book would destroy its information. But this information is lost in practice because we are not able to keep track of all the correlations for the radiation emitted in this process. Unfortunately, the case of a black hole is even worse. If a book is thrown into the black hole and the latter burns completely, the information is lost in principle, unless the black hole manages to record the information concerning the book in the Hawking radiation. The main problem here is that the collapse resulting in the black hole formation precedes in time the beginning of the Hawking radiation. If in the process of collapse bleaching of the information at the horizon does not occur, then only macroscopic violation of causality can help to transport the information from the collapsing body to the outgoing quantum radiation.

2. Information comes out "at the end". The Frolov-Vilkovisky model discussed above can be used to illustrate this possibility. The collapsing matter producing the "black hole" as well as the particles created inside the "black hole" in the Hawking process are accumulated inside the inner part of the surface $F = 0$ (in the region restricted by the line ED at Figure 15.4). After the apparent horizon disappears, all the matter returns to our space and quantum coherence would be restored. The stage of the final decay of the "black hole" can require a very long time [Aharonov, Casher, and Nussinov (1987)]. Carlitz and Willey (1987), and later Preskill (1992) by a more general argument, showed that the lower bound for the lifetime of a Planck mass remnant, which contained all the information originally in a black hole of mass M, would be of order

$$T(M) \sim t_{\rm Pl}\,(M/m_{\rm Pl})^4 . \tag{15.4.2}$$

This time can be estimated as follows [Preskill (1992)]. The remnant must emit about $N \sim S \sim (M/m_{\rm Pl})^2$ quanta to reinstate the information. Since the mass of the remnant is $m_{\rm Pl}$, a typical quantum has energy $m_{\rm Pl}^3/M^2$. To carry the required information, the quanta cannot be strongly correlated and must be emitted one after another without overlapping. Since the time of emission of a single quantum is $t_{\rm Pl}(M/m_{\rm Pl})^2$, and the number of emitted quanta $N \sim M^2$, the total time is given by equation (15.4.2). For this scenario, the black hole evaporates to the Planck size in a time $t_{\rm Pl}(M/m_{\rm Pl})^3$, but the black hole remnant disappears only after a much longer time $t_{\rm Pl}(M/m_{\rm Pl})^4$. For $M > 10^{10}$g the latter time is greater than the lifetime of the Universe. In other words, the remnants of black holes with initial mass $M > 10^{10}$g are practically stable.

3. The information is retained by stable black hole remnants. In this scenario, the final state of an evaporating black hole might be some version of the elementary black hole. Since the initial black hole could have been arbitrarily massive, the remnant must be able to carry an arbitrary large amount of information. One can consider it as a particle with an extremely large ($\sim (M/m_{\rm Pl})^2$) number of internal states. If there exist an interaction of these degrees of freedom with an external world, one would have the following problem. Consider a thermal ensemble at temperature T. The probability to have in this ensemble a specific black hole remnant is suppressed by the factor $\exp(-m_{\rm Pl}/T)$. But if there are an infinite number of species of remnants, the ensemble would be dominated by these remnants. This objection might not be so flagrant if the interaction with the internal degrees of freedom is so weak that the relaxation time in the system exceeds the lifetime of the Universe.

4. The information escapes to "baby universes". In this scenario offered by Dyson (1976), Zel'dovich (1977a), and Hawking (1988a,b, 1990a,b), the collapsing body does not produce a true singularity inside the black hole, but rather induces the nucleation of a closed *baby universe*. A concrete model for a single [Frolov, Markov, and Mukhanov (1988)] and multiple [Barrabès and Frolov (1996)] baby universe creation inside the black hole was discussed in Section 14.7. The baby universe is causally disconnected from our own universe. Yet the information is not lost in the enlarged system (our universe plus baby universes), and the standard quantum mechanics can be used in such a "multiverse" picture. The density matrix in our universe arises for the same reason as in the usual quantum mechanics since we dealing only with a subsystem. Nevertheless, there is still a question how to describe physics in our own universe. One of the possibilities is when a pure state collapses to form a black hole, and then evaporates, it evolves to a pure state [Hawking (1990a,b)]. If we repeat this experiment for the same initial state, we obtain the same final result, so that in this sense the result is predictable. But we are not able to predict the result from first principles until we also fix some additional parameters by observations. The number of these parameters conceivably is of order of $\exp(S(M)$, where $S(M)$ is the entropy of the black hole of mass M. But we still have no mechanism that allows a black hole to remember the initial state.

5. Quantum hair. Black hole hair provides partial information about the matter that has formed a black hole. One may say that black hole hair encodes this information. An interesting example is the quantum hair that we discussed in Section 13.6. Its energy vanishes so that quantum hair does not affect the black hole geometry. For this reason, in those theories that have many varieties of quantum hair, a black hole can store an accurate record of a number of properties of the collapsing matter. Still it is hard to believe that this solves the information puzzle. The theory must possess an infinite number of new conservation laws that are still present in the low-energy limit. The problem is to find such a theory.

Besides these there are a number of other different proposals [see e.g., the review by Page (1994)], but it seems that at the moment there is no resolution of the information loss puzzle that satisfies most of the specialists. If the information is really lost in black holes as Hawking suggested what might be the consequences of this? The most extreme point of view is: "The information loss paradox may be a genuine failing of 20th century physics, and a signal that we must recast the foundations of our discipline". [Preskill (1992)]. Because we have no idea now how to do this, it is instructive to begin by analyzing possible consequences of the information loss assumption.

Beyond the semi-classical approximation one can expect that for almost any initial quantum state there is a non-vanishing amplitude for black hole creation and evaporation to occur. These processes give rise to a non-vanishing probability for the evolution of a pure state into a mixed one [Ellis *et al.* (1984), Ellis, Mavromatos, and Nanopoulos (1992), Huet and Peskin (1995)]. One would expect deviations from the ordinary evolution laws in experiments where the energies involved are high enough for there to be a substantial probability of producing black holes of the Planck mass. For ordinary systems the interactions producing the decoherence normally provide an exchange of both entropy and energy with the environment. For this reason, one can expect that processes involving virtual black holes would violate energy-momentum conservation. Loss of coherence can be modeled by coupling a quantum system to a source of random noise. Banks, Susskind, and Peskin (1984) argued that this interaction results in a failure of energy conservation that would be unacceptably large at the scales of laboratory physics. These arguments were reanalyzed by Unruh and Wald (1995) who found a gap in their proof and showed that the rapid evolution from pure states to mixed states can occur with energy-momentum non-conservation that is unobservably small for all states which one could expect to produce in a laboratory.

To summarize, at the moment we have no proof that information loss if it occurs necessarily implies a violation of the physical laws that we should be able to verify in the laboratory. Nevertheless, one could expect that the attempts to solve this puzzle would help us to understand better the most fundamental problems of theoretical physics.

Chapter 16

Black Holes, Wormholes, and Time Machines

16.1 Topological and Causal Structure of Spacetime

One of the most intriguing features of the theory of general relativity is the possible existence of spacetimes with non-trivial topological structure. The wormhole solutions of the Einstein equations described by Wheeler [Misner and Wheeler (1957), Wheeler (1962)] are well-known examples of spaces with non-trivial topology. The simplest example of such a spacetime is the Einstein-Rosen bridge described in Chapter 2. The spacetime of an eternal black hole can be considered as the evolution of the Einstein-Rosen bridge. To the future of the moment of time symmetry the throat shrinks to zero size and the singularity arises. No causal signal can propagate through the throat from one asymptotically flat region (R_+) to the other one (R_+). This property is directly connected with the non-trivial causal structure of spacetime in the presence of a black hole.

Three-dimensional geometries containing a wormhole which connect different asymptotically flat regions were considered in Section 7.6.3 where the initial value problem for vacuum solutions of the Einstein equations was discussed. By identifying the asymptotic infinities of different regions, one gets a wormhole connecting separate regions in the same spacetime. Such a wormhole is shown in Figure 7.5. In the same manner as in the case of the eternal black hole, in the absence of matter the throat of the wormhole pinches off so quickly that it cannot be traversed even by light. This is a generic property of spacetimes with a nonsimply connected Cauchy surface. According to the **theorem** proved by Gannon (1975) [see also Lee (1976) and Galloway (1983)]:

Any asymptotically flat spacetime with a nonsimply connected Cauchy surface has singular time evolution if it satisfies the weak energy condition.

Moreover, such singularities arise so quickly that no information carrying signal can propagate through a wormhole to the asymptotic region before creation of the singularity. One can formulate ***topological censorship*** as the statement that no observer remaining outside a region with strong gravitational field has time to probe the topology of spacetime [Friedman, Schleich, and Witt (1993), see also Section 5.2.3].

In order to prevent shrinking of a wormhole and to make it traversable, one needs to fill its throat with matter violating the averaged null energy condition. This follows from the above theorem. On the possibility of the existence of matter violating the weak energy condition, see e.g., Thorne (1993), Visser (1995), and Flanagan and Wald (1996).

A spacetime with a traversable wormhole (if it exists) can be transformed into a spacetime with closed timelike curves. Namely, it was shown by Morris, Thorne, and Yurtsever (1988) and Novikov (1989) that closed timelike curves may arise as a result of the relative motion of the wormhole's mouths. This possibility in principle to create closed timelike curves in a spacetime where they were absent initially attracted new interest to the wormhole-like solutions. From a more general point of view, this result hints at deep relations between the topological and causal structure of spacetime.

The main features of such spacetimes can be illustrated by considering a simple model proposed by Morris, Thorne, and Yurtsever (1988). The simplest wormhole is obtained by removing two balls of equal radius from Euclidean space and identifying their surfaces. The surfaces then become the wormhole's mouths. In this process of identification of surfaces their extrinsic curvature jumps. It implies that if such a spacetime is a solution of the Einstein equations, there must be a delta-like distribution of stress-energy at the junction between the two mouths, which violates the null energy condition. In this particular model, this is obvious because any bundle of radially traveling null geodesics that passes through the wormhole is converging as it enters and diverging as it leaves, and therefore gets defocused by the wormhole. One can make the matter distribution smooth by connecting the mouths by a handle of finite length instead of gluing them together. In what follows we always assume that the length of the handle is small enough.

One can construct wormhole spacetimes such that the wormhole's mouths are moving along arbitrarily chosen world lines by removing world tubes enclosing these lines and identifying their surfaces with each other. The junction conditions requires that the intrinsic geometries of the tube surfaces must be the same. This may require a distortion of the spacetime geometry near the mouths if they are accelerated. But the distortion can be made arbitrarily small by taking the value of (acceleration)$\times$(mouth radius) to be small enough [Morris, Thorne, and Yurtsever (1988), Friedman *et al.* (1990)]. Since the mouths' intrinsic geometries are the same, the proper time interval between two identified events on the mouths must be the same as seen through either mouth. Thus, if we specify the identification of events on two mouths at some initial moment of their proper time, the identification for all subsequent moments would be fixed.

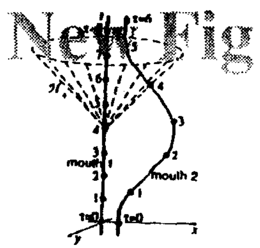


Figure 16.1: The chronological structure of spacetime with a moving wormhole. The wormhole is formed by removing two balls from Euclidean space and identifying their surfaces. One of the wormhole's mouths is at rest, while the other makes a "twins paradox trip".

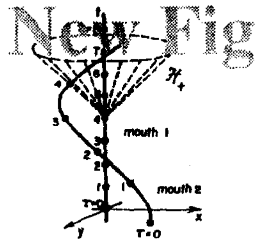


Figure 16.2: The chronological structure of spacetime with a moving wormhole. One of the wormhole's mouths is revolving around the other.

In order to accelerate a mouth, one must apply a force to it. This can be done if the mouths are charged (i.e., if there exists a flux of electric field through the mouth). In this case, it is sufficient to apply an electric field to one of the mouths. The mouth is set in motion in the external space. We assume that during this motion the position of the other mouth in the external space remains unchanged.

Figure 16.1 shows the chronological structure of a wormhole spacetime for the case where one mouth remains at rest in a chosen inertial frame, while the other (which initially also was at rest) makes a "twin-paradox-type trip" and returns to its initial position [Morris, Thorne, and Yurtsever (1988)]. One can choose the clocks located near both mouths to be synchronized in the external space before the beginning of the motion. As a result of the motion, the synchronization in the external space is lost. The moving clock (measuring proper time near one mouth) shows less time than the clock which remains at rest ("twin paradox"). Denote by $\Delta\tau$ this time difference. If $\Delta\tau > L/c$, where L is the distance between the final positions of the mouths, a *closed timelike curve* becomes possible. Denote by C the region formed by events through which a closed timelike curve passes. The region formed by C and by events which can be connected with events in C by causal past directed lines is restricted from the past by the ***future chronology horizon.***

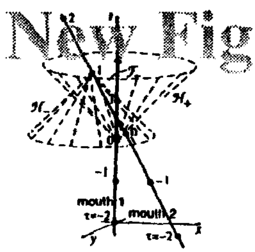


Figure 16.3: The chronological structure of spacetime with a moving wormhole. The wormhole's mouths are moving with constant velocities.

Figure 16.2 shows a spacetime with a wormhole in the case where one of the mouths is revolving around another in the external space [Novikov (1989)]. Closed timelike curves in this model arise for the same reason as in the above model with an accelerated mouth. As seen in the external space, there is a dilation of proper time on the moving mouth relative to the static one, but, as seen through the wormhole, there is no such time dilation.

Figure 16.3 shows an even simpler model where the mouths are moving with constant velocity with respect to each other and closed timelike curves also arise [Morris, Thorne, and Yurtsever (1988)]. In this case, closed timelike curves are confined to a bounded achronal region of spacetime: the region that begins at the future chronology horizon and ends at the *past chronology horizon.* (The latter is defined similarly to the future chronology horizon with the obvious change of the direction in time of the causal curves which enter the definition.)

The formation of achronal regions containing closed timelike curves is a generic property of spacetimes with wormholes. Generic relative motions of a wormhole's mouths will always produce closed timelike curves [Morris, Thorne, and Yurtsever (1988)], as will the gravitational redshifts when the wormhole's mouths are placed in a generic gravitational field [Frolov and Novikov (1990)].

16.2 Locally Static Multiply Connected Spacetimes

16.2.1 Non-potential gravitational fields

In this section we describe generic properties of wormholes in an external gravitational field and relate the conditions of closed timelike curves formation with topological invariants of the multiply connected spacetime [Frolov and Novikov (1990)].

We begin by showing at first that the gravitational field in a locally static multiply connected spacetime is generically non-potential. A spacetime is called *locally static* if in any simply connected region there exists a uniquely defined (up to normalization) timelike Killing vector ξ obeying the relations

$$\xi_{(\mu;\nu)} = 0\,, \qquad \xi_{[\mu;\nu}\,\xi_{\lambda]} = 0\,, \tag{16.2.1}$$

The first equation is the definition of a Killing vector, while the second equation means that ξ is locally "surface orthogonal". In the general case in a multiply connected spacetime (e.g., when a wormhole is present), it is impossible to define such a Killing vector globally. The existence of non-contractible closed paths (i.e., paths which pass through a wormhole and cannot be contracted to a point by continuous transformations) may create an obstacle for the existence of a global Killing vector field. In the general case, after continuous propagation of a Killing vector field along a closed non-contractible path, it will return to the initial point having the same direction as the initial vector but with a different norm. In order to be able to deal with this kind of situation, it is convenient to introduce a unit timelike vector $u^\mu \equiv \xi^\mu/|\,\xi^2\,|^{1/2}$ instead of ξ. This four-velocity of a Killing observer as well as his four-acceleration $w^\mu \equiv u^\nu\, u^\mu{}_{;\nu}$ do not depend on the norm of ξ, and hence they are well defined globally. It is easy to verify that the following equations

$$u_\mu\, u^\mu = -1\,, \qquad u_{\mu;\nu} = -\,w_\mu\, u_\nu\,, \qquad w_{[\mu;\nu]} = 0 \tag{16.2.2}$$

are locally equivalent to (16.2.1). Hence, a locally static spacetime can be equivalently defined as a spacetime which admits two globally defined vector fields u^μ and w^μ obeying the relations (16.2.2).

It is convenient to consider a locally static spacetime M as the collection S of Killing trajectories [Geroch (1971)], i.e., the set of integral lines $x^\mu(\tau)$ of u^μ ($dx^\mu/d\tau = u^\mu$). We assume that a unique Killing trajectory passes through each point of the spacetime M and that $M = T \times S$, where T topologically is R^1. One may refer to S as to the three-dimensional space under consideration. In the presence of a wormhole this space has a non-trivial topology. A non-contractible loop passing through a wormhole n-times in the chosen direction can be characterized by its winding number n. Two loops with the same winding number n are connected by a continuous transformation (are homotopic). These loops are also homological, i.e., they are boundaries of some two-dimensional surface. In other words, the first fundamental group π_1 and the first homology group H_1 calculated for S (as well as for the spacetime M itself) coincide with the group of integer numbers $\mathbb{Z}$.

The last equation in (16.2.2) shows that the acceleration form $w \equiv w_\mu\, dx^\mu$ is closed ($dw = 0$). According to the Stokes' theorem, the integral of this form over any closed path C_n

$$I_n[w] = \oint_{C_n} w \tag{16.2.3}$$

depends only on the winding number n of the path. One has $I_n = nI$. We choose the direction of the path with $n = 1$ in such a way that $I[w] \equiv I_1[w] \leq 0$.

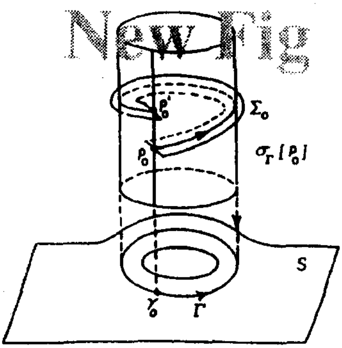


Figure 16.4: The clock synchronization in the multiply connected spacetime.

If the period $I[w]$ vanishes, then the form w is exact, i.e., there exists a globally defined scalar function φ (gravitational potential) such that $w = d\varphi$. For such a *potential gravitational field* the vector field $\xi^\mu = \mathrm{e}^\varphi u^\mu$ is a global Killing vector field.

In the general case (when $I[w] \neq 0$), the gravitational potential φ cannot be defined as a single-valued function. Such a gravitational field is called *non-potential.* The period $I[w]$ may be considered as the measure of non-potentiality of the field. The important physical property of a *non-potential gravitational field* is that the work done by such a field on a particle moving along a closed contour C_n passing through a wormhole does not vanish. This work is proportional to $e^{-I_n[w]}$.

16.2.2 Clock synchronization

Before considering the properties of a spacetime with a non-potential gravitational field, we discuss the quite general problem of *clock synchronization* in a multiply-connected spacetime. The presence of non-contractible paths makes this problem non-trivial already in a potential gravitational field. The reason is very simple. In a static gravitational field, it is possible to synchronize the clocks along any path lying in a simply connected region. But for the process of clock synchronization along the closed path passing through a wormhole, one cannot guarantee that the "synchronized clocks" show the same time as the initial ones. In the general case, there will be a non-vanishing time gap for the clocks' synchronization along a non-contractible path. The clock synchronization in a multiply-connected spacetime is illustrated in Figure 16.4

For a potential gravitational field , the value of the time gap for clock synchronization in a multiply connected spacetime is a topological invariant. In order to show this, we introduce the 1-form

$$\eta \equiv \mathrm{e}^{-\varphi}\, u_\mu\, dx^\mu\,. \tag{16.2.4}$$

One can verify that this globally defined 1-form is closed ($d\eta = 0$). In fact, using equations (16.2.2) and the definition of the acceleration, $w_\mu = \varphi_{,\mu}$, we have

$$\eta_{[\mu;\nu]} = e^{-\varphi}\left[u_{[\mu;\nu]} - \varphi_{[,\nu}u_{\mu]}\right] = e^{-\varphi}\left[w_{[\nu}u_{\mu]} - \varphi_{[,\nu}u_{\mu]}\right] = 0 \,.$$

The integral of the closed 1-form η over a closed path C_n

$$J_n[\eta] = \oint_{C_n} \eta \tag{16.2.5}$$

is a topological invariant depending only on the winding number. We now show that the value of $J_n[\eta]$ coincides with the time gap for clock synchronization along a closed path with a chosen winding number n.

For this purpose, we choose a point p_0 and denote by γ_0 a Killing trajectory passing through this point (see Figure 16.4). Consider now a curve $\sigma_\Gamma[p_0]$, which begins at p_0, passes once through the wormhole, and ends at the point p'_0 lying on the same Killing trajectory as p_0. We assume that $\sigma_\Gamma[p_0]$ is everywhere orthogonal to the four-velocity of a Killing observer. The latter condition means that the curve $\sigma_\Gamma[p_0]$ is formed by the events which are simultaneous (in the reference frame of Killing observers) with the initial event p_0. The projection Γ of this line into the space S is closed, but in the general case the line itself is not closed (see Figure 16.4). We add the part of a Killing trajectory lying between p_0 and p'_0 to this curve to make it closed. The evaluation of the integral (16.2.3) for this closed path with a winding number $n = 1$ is quite simple. Only that part of the path lying along the Killing trajectory gives a non-vanishing contribution. As a result, the value of the integral $J_1[\eta]$ calculated for the chosen path coincides with the time gap for the clock synchronization along this path. It is now evident that $J_1[\eta]$ does not depend on the particular choice of path with the given winding number $n = 1$ and that $J_n[\eta] = nJ_1[\eta]$.

In a ***non-potential gravitational field***, the time gap $\Delta_1(\tau_0) \equiv \tau_0 - \tau_1$ for clock synchronization is no longer time independent (see Figure 16.5). It can be shown [Frolov and Novikov (1990)] that it satisfies the equation

$$\frac{d\Delta_1(\tau_0)}{d\tau_0} = 1 - e^{I[w]} \,. \tag{16.2.6}$$

This relation shows that the time gap for clock synchronization along any closed path with a winding number $n = 1$ grows with time, and the value $I[w]$ may be considered as the measure of this growth. As soon as the gap becomes greater than the minimal time needed for a light signal to propagate along a closed path passing through the wormhole and return, a closed timelike curve can arise.

Frolov and Novikov (1990) demonstrated that the interaction of a wormhole with classical matter generically generates a non-potential component of the gravitational field. That is why a locally static wormhole is generically unstable with respect to the processes which transform it into a "time machine". Relative motion of the wormhole's mouths also generates a time gap for clock synchronization [Morris, Thorne,

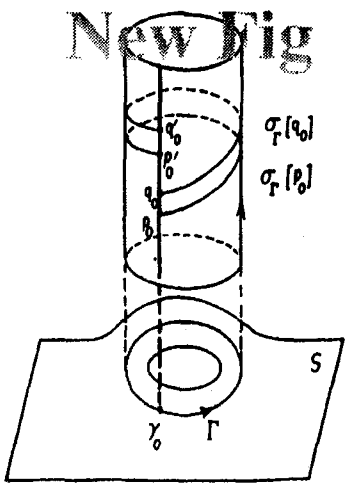


Figure 16.5: Clock synchronization in a multiply-connected spacetime with a non-potential gravitational field. Points p_0 and q_0 are located on the same Killing trajectory. The lines $\sigma_\Gamma[p_0]$ and $\sigma_\Gamma[q_0]$ represent sequences of events, along the closed spatial path Γ, which are simultaneous with the initial events p_0 and q_0. These lines intersect the initial Killing trajectory at points p'_0 and q'_0. If the gravitational field is non-potential, the difference of proper time between p'_0 and q'_0 is not equal to the difference of proper time between p_0 and q_0.

and Yurtsever (1988)]. One can interpret the above results in the following way. There exist inner relations between the topological and causal properties of a spacetime. The existence of closed timelike curve is a generic property of multiply connected locally static spacetimes [see also Visser (1995)].

16.3 Spacetimes with Closed Timelike Curves

16.3.1 Chronology horizon

In this section we discuss some general properties of spacetimes with closed timelike curves (for more details see [Friedman *et al.* (1990), Hawking (1992), Thorne (1993), and Visser (1995)].

Solutions of Einstein's equations which allow closed timelike curves have been known for a long time. The earliest example of such a spacetime is a solution obtained by Van Stockum (1937), which describes an infinitely long cylinder of rigidly and rapidly rotating dust. Another well-known example is Gödel's (1949) solution representing a stationary homogeneous universe with nonzero cosmological constant, filled with rotating dust. Closed timelike curves are also present in the interior of the eternal Kerr black hole in the vicinity of its ring singularity. Other examples of spacetimes with closed timelike curves were discussed by de Felice (1981).

In the general case, a spacetime can be divided into *chronal regions*, without closed timelike curves, and *achronal regions* that contain closed timelike curves. The boundaries between the chronal and achronal regions are formed by chronology horizons. Chronal regions end and achronal region begins at a *future chronology horizon*. Achronal regions end and chronal region begins at *past chronology horizon*. Thus, achronal regions are intersections of the regions bounded by both of these horizons. A future chronology horizon is a special type of future Cauchy horizon, and as such it is subject to all the properties of such horizons. In particular, it is generated by null geodesics that have no past endpoints but can leave the horizon when followed into the future. If the generators, monitored into the past, enter one or more compact regions of spacetime and never thereafter leave them, the future chronology horizon is said to be *compactly generated*. In a wormhole model with closed timelike curves, the future chronology horizon is compactly generated. The inner horizon of a Kerr-Newman solution is an example of a Cauchy horizon that is not compactly generated. A compactly generated chronology horizon cannot form in a spacetime developed from a spacelike non-compact surface without boundary if the null energy condition holds [Hawking (1992)].

The past-directed generators of the compactly generated future chronology horizon have no past end points. They will enter and remain in a compact region C. Hawking (1992) showed that there exists a nonempty set E of generators, each of which remains in the compact set C in the future direction, as well as in the past direction. The sets E generically contain at least one closed null geodesic. More exactly, Hawking (1992) **proved** that:

1. *If E contains such a closed null geodesic, small variations of the metric preserve this property.*

2. *If E does not contain a closed null geodesic, then in geometries obtained by small variation of the metric such curves do exist.*

The generators traced into the past either wander ergodically around C or they asymptote to one or more smoothly closed geodesics. In the latter case, followed forward in time they are seen to originate in fountains and spew out of them. That is why Thorne (1993) proposed to name such closed null geodesics *fountains*. Hawking's result indicates that in the generic case, C will contain such fountains, and it is likely that generically almost all the horizon generators will emerge from them [Thorne (1993)].

16.3.2 Possible obstacles to creation of a time machine

In order to create a time machine (region with closed timelike curves) by using a wormhole, one needs to assume that there exists the principle possibility to make them long living and traversable. This is impossible without violation of the averaged null energy condition. Moreover, the averaged null energy condition must be violated near the fountain of any compactly generated future chronology horizon [Hawking (1992)]. This means that it is impossible to create a time machine in a finite region of space time without violating the averaged null energy condition.

It is unclear whether it is possible to provide such a violation in physically acceptable conditions. It has been shown that the averaged null energy condition is satisfied for non-interacting quantized scalar and electromagnetic fields in flat spacetime [Klinkhammer (1992), Folacci (1992)], and in generic, curved (1+1)-dimensional spacetimes [Wald and Yurtsever (1991)]. On the other hand, in (3+1)-dimensional spacetime, both non-trivial topology [Klinkhammer (1992)] and curvature [Wald and Yurtsever (1991)] can induce violations of the averaged null energy condition. Moreover, in the latter paper it was shown that there are generic classes of spacetimes where quantum effects may violate the averaged null energy condition. Under these conditions, it is impossible at least for now to exclude the possibility of such a violation of the averaged null energy condition required for time machine creation. Later, Flanagan and Wald (1996) investigated whether the averaged null energy condition holds for self-consistent solutions of the semi-classical Einstein equations $R_{\alpha\beta} - \frac{1}{2} R_{\alpha\beta} = 8\pi\langle T_{\alpha\beta}\rangle$ where $\langle T_{\alpha\beta}\rangle$ is the expectation value of the renormalized stress-energy tensor of quantum fields. They studied a free, linear, massless scalar field with arbitrary curvature coupling in the context of perturbation theory about the vacuum solution. They argued that probably if traversable wormholes do exist as self-consistent solutions of the semi-classical equations, they cannot be macroscopic but must be Planck scale [see also Ford and Roman (1990, 1992, 1993, 1995, 1996)]. Still the question is open for the case of general conditions. Precise conditions under which quantum field theory permits the required violation of the averaged null energy condition are less than pellucid. This remains an area of active research [Visser (1995)].

If there are no eternal traversable wormholes in the Universe, then in order to create a time machine by the proposed mechanism, one needs first to create a wormhole. In the absence of wormholes, the space initially (on some spacelike surface S) was simply connected. If a wormhole is created, then a later spacelike surface S' has different spatial topology. If the processes connected with time machine formation are restricted to a spatially bounded region, then it is natural to assume that one can surround this region by a timelike cylinder T which intersect the spacelike surfaces S and S' in compact regions S_T and S'_T of different topology. In other words, the topology change occurs in a spacetime region M_T bounded by S, S', and T. In the absence of singularities and the impossibility for M_T to extend to infinity, M_T is compact. Hawking (1992) proved that the change of spatial topology inside M_T is impossible until it contains closed timelike curves. This is a generalization of a well-known theorem by Geroch (1967). This result means that even creation of a wormhole cannot be possible without creation of a time machine. Whether it is forbidden by some fundamental physical laws remains unclear.

There is another danger: quantum instability of the compactly generated future chronology horizon. We postpone the discussion of this problem to Section 16.5, and we now discuss the problem whether it is possible to have any reasonable physics in a spacetime with closed timelike curves.

16.4 Classical Physics in the Presence of Closed Timelike Curves

In this section we discuss whether and how the laws of physics can deal with closed timelike curves. Usually one worries that the laws of physics cannot deal reasonably with a time machine. The crucial problem here is the problem of causality. The existence of closed timelike curves allows one to travel into the past. At first sight, it inevitably leads to the possibility of changing the past, thereby producing causality violations. But it is not so.

In order to escape possible contradictions, it is necessary to impose the *principle of self-consistency.* This principle was proposed and discussed in the works by Novikov (1983), Zel'dovich and Novikov (1975), Novikov and Frolov (1989), Friedman *et al.* (1990), and Novikov (1992).[1] The meaning of this principle is the following.

In the case of an open timelike curve, any event x divides the other events on this curve into two sets: future events and past events with respect to x. All past events can influence x, but future events cannot. On a closed timelike curve the choice of the event x divides other events on the curve into future events and past ones only locally. In this case, events which locally are in the future with respect to x can influence the event x circularly around the closed timelike curve. There is no global division of events on the closed timelike curves into future and past. The future influences the present around the closed timelike line, with a locally future-directed half of a light cone at each event of the closed curve. Not only the future is a result of evolution of the past, but the past is a result of the future also. All events in a spacetime with closed timelike curves must be self-consistent. According to the principle of self-consistency, all events on closed timelike curves influence each other around the closed timelike line in a self-adjusted way.

A more precise formulation of this **principle** is: *the only solution to the laws of physics that can occur locally in the real Universe are those which are globally self-consistent.*

The principle of self-consistency by fiat forbids changing the past. All events happen only once, and cannot be changed.

In order to demonstrate how this principle works, we shall consider the so-called *billiard ball problem* [Friedman *et al.* (1990), Echeverria, Klinkhammer, and Thorne

[1]For discussion of other approaches to the problem of consistency, see Visser (1995).

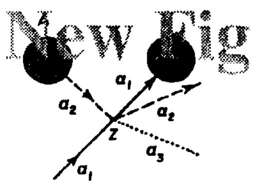


Figure 16.6: The self-inconsistent evolution in the billiard ball problem.

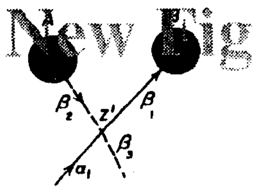


Figure 16.7: The self-consistent evolution in the same billiard ball problem as in Figure 16.6.

(1991), Mikheeva and Novikov (1993)], which is the following: A solid perfectly elastic ball moves relative to the mouths of the wormhole. Its speed is assumed to be small compared with the speed of light so it can be treated non-relativistically. The ball enters the wormhole through mouth B, emerges from A in the past and, continuing its motion, it can encounter and collide with itself.

At first glance, there is a "paradox" in this problem [Polchinski (unpublished)]. The initial position and velocity of the ball are chosen so that the ball moves along the trajectory α_1 (see Figure 16.6), enters mouth B, and exits from mouth A earlier in time than it entered into B. The ball continues its motion along the trajectory α_2.[2] The timing is just right for the ball to hit itself at the point Z, knocking its "younger" self along trajectory α_3 and thereby preventing itself from ever reaching mouth B. Such an evolution is not self-consistent, and hence it is impossible. It is not a solution of the evolution equations.

The mistake (the reason for the "paradox") is obvious: When at the beginning of our discussion we continued the trajectory α_1 after point Z, we did not take into account the influence of the impact and considered the motion of this ball after the passage of the mouths B and A along the trajectory α_2 without taking into account this impact. This means that we did not take into account the influence of the future on the past.

Echeverria, Klinkhammer, and Thorne (1991) demonstrated that for the same initial data which give a self-inconsistent "solutions" there exist also self-consistent solutions. The self-consistent solution is shown in Figure 16.7. The initial data (initial position and velocity of the ball) are the same as in Figure 16.6. The part of trajectory α_1 before the collision with the "older" self coming from the future is the same. This "older" ball moves along trajectory β_2 which differs only slightly from the path α_2 in Figure 16.6. The "older" ball on β_2 strikes itself on α_1 gently, deflecting itself into the slightly altered trajectory β_1. This altered trajectory β_1 takes the ball into the mouth B at a slightly altered point compared to the point in the Figure 16.6. The ball exits from the mouth A before it went into mouth B, and moves along the

[2]The trajectory α_2 is well defined if the trajectory α_1 is given (see Echeverria, Klinkhammer, and Thorne (1991)).

trajectory β_2 to the collision event. This solution is self-consistent.

The general method of construction of the self-consistent solutions is the following. Let us forget for a moment that the "older" ball is from the future and, during its motion after appearance from the mouth A we will treat it as a normal ball independent of the "younger" one. Let us suppose that we know the moment t_2 and place $\mathbf{x}_2$ of its emergence from the mouth A and its velocity $\mathbf{v}_2$ at t_2. We can treat them as "initial conditions" for the "older" ball. Now using the standard physical laws we can calculate the collision between the "younger" and "older" balls. This allows us to get the change of the trajectory of the "younger" one, as well as the place $\mathbf{x}_1$, velocity $\mathbf{v}_1$, and the moment t_1 of its arrival to the mouth B. Of course $\mathbf{x}_1$, $\mathbf{v}_1$, and t_1 are function of $\mathbf{x}_2$, $\mathbf{v}_2$, and t_2 (as well as of the initial data for the "younger" ball). Now we must take into account that the "older" ball is the "younger" one returned from the future and that $\mathbf{x}_2$, $\mathbf{v}_2$, and t_2 are not arbitrary but are determined by the motion of the "younger" ball. Thus, because we know the laws of passage through the time machine, we can express $\mathbf{x}_2$, $\mathbf{v}_2$, and t_2 in terms of $\mathbf{x}_1$, $\mathbf{v}_1$, and t_1. Thus, we have a system of equations allowing us to calculate all motions if the initial data are specified.

The natural questions are: 1) Does a solution to this system exist for any initial conditions? and 2) Is this solution unique?

A positive answer to the first question was obtained by Echeverria , Klinkhammer, and Thorne (1991) for an ideal elastic billiard ball and by Mikheeva and Novikov (1993) for an inelastic one. The answer to the second question is not trivial. Namely, Echeverria, Klinkhammer, and Thorne (1991) demonstrated that in the general case for the same initial conditions there could be an infinite number of self-consistent solutions. These solutions include a number of wormhole traversals. Figure 16.8 illustrates some possibles examples. On the other hand, there exist initial data for which there is only one solution. For example, a ball initially at rest far from the wormhole has only one solution to its equations of motion; namely, it remains forever at rest. One can argue that there is a single solution for any ball with an initial speed small enough and an initial path of motion that, if extended forever, remains far from the wormhole. Echeverria, Klinkhammer, and Thorne (1991) tried to find initial data which do not allow any self-consistent solutions at all, but none were found.

What does the multiplicity of self-consistent solutions mean? Does it have any physical meaning or does it simply indicate that the laws of physics cannot deal with time machines in a reasonable way?

There is no definite answer. But we know that physics is quantum mechanical (not classical) by nature. If one considers the classical problem as a limiting case of the quantum one, then the Cauchy problem turns out to be well posed in the formalism of quantum mechanics [Klinkhammer and Thorne (unpublished)].

So far we discussed plane motions of billiard balls. A three-dimensional billiard problem with a time machine was discussed by Mensky and Novikov (1996a). Slightly noncoplanar initial trajectories were discussed by Echeverria, Klinkhammer,

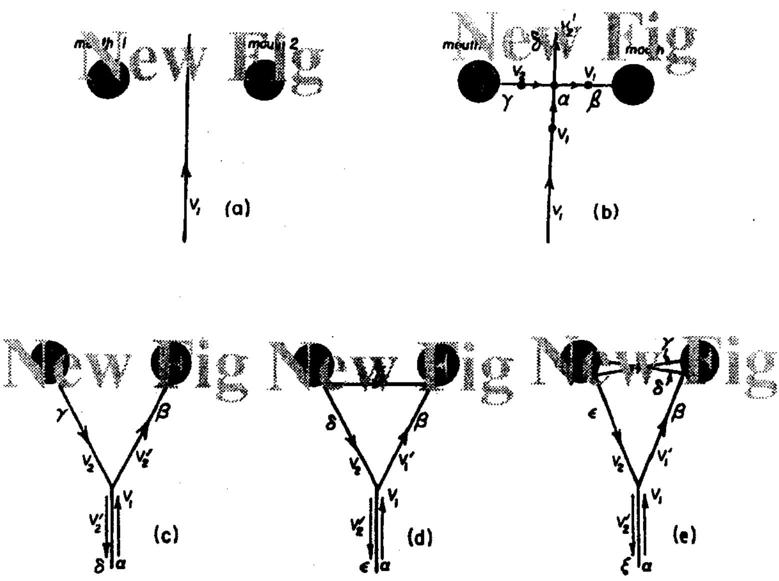


Figure 16.8: A specific example of an initial trajectory with infinite number of self-consistent solutions. (a) A trivial solution. (b) A solution with one collision and one wormhole traversal. (c) Another solution with one collision and one wormhole traversal. (d) A solution with one collision and two wormhole traversals. (e) A solution with one collision and three wormhole traversals. (Greek letters indicate the order of a ball motion in its proper time.) This set can be continued to infinity.

and Thorne (1991). In classical physics more complicated interacting systems than the billiard ball have been investigated [for example, see Novikov (1992), Lossev and Novikov (1992), Novikov (1993)]. Originally it was believed that at least for some initial data self-consistent evolution does not exist at all. But so far no clean example of such a thing have been exhibited. In the work by Carlini *et al.* (1995) it was demonstrated that for a billiard-ball problem the principle of self-consistency directly follows from the principle of least action, in which the initial and final positions of balls are fixed. This result motivates the authors to formulate the conjecture that the "principle of self-consistency" is a consequence of the "principle of least action" in the general case for all physical phenomena, not only for the simple mechanical problem considered there. Carlini and Novikov (1996) extended the analysis to the case of point-like "billiard balls" moving with relativistic velocity. It was shown that for the case under consideration the only possible trajectories for which the action is extremal are those which are globally self-consistent. This gives additional support for the conjecture.

16.5 Chronology Protection Conjecture

The chronology horizon seems to be stable against classical perturbations [Morris *et al.* (1988), Friedman *et al.* (1990)]. For quantum fields, because of the existence of zero-point fluctuations, which cannot be isolated or suppressed, the situation is quite different: The renormalized vacuum stress-energy tensor diverges near the future chronology horizon. This result follows from works of Kim and Thorne (1991), Frolov (1991), and Hawking (1992) [see also Thorne (1993), Lyutikov (1994), Visser (1995)]. Here, following the paper [Frolov (1991)], we demonstrate this result for a locally static spacetime with a wormhole, where a time machine is formed because of the non-potentiality of the gravitational field. The consideration in this case is somewhat simpler, but it allows us to demonstrate the main line of argument used in other cases.

Consider the amplitude $G(x, x')$ of a quantum particle propagating from a point x to a point x'. This amplitude $G(x, x')$ can be presented as a Feynman integral over all trajectories connecting these points. For nearby points in a multiply-connected spacetime, besides trivial paths, there exist homotopically non-equivalent classes of paths enumerated by an integer winding number n. In the geometrical optics approximation, the main contribution to the Feynman propagator is given by narrow tubes of paths located near geodesics connecting x and x'. In the limit $x' \to x$, the trivial direct path contribution diverges. This results in a divergence of the vacuum expectation value of the stress-energy tensor $\langle T_{\mu\nu} \rangle$. This divergence exists already in a topologically trivial flat spacetime. According to the standard renormalization scheme, this divergence is to be subtracted. Besides the regular part remaining after subtraction of the $n = 0$ contribution, the renormalized stress-energy tensor $\langle T_{\mu\nu} \rangle^{ren}$ in a multiply connected spacetime in the geometric optics approximation contains also contributions from closed geodesics with non-vanishing winding numbers. In the region lying to the past of the future chronology horizon these additional geodesics (with $n \neq 0$) are spacelike and their contribution generally remains finite. But slightly above the future chronology horizon these additional geodesics become null. Namely, consider a locally-static multiple-connected spacetime. For a chosen number n and for a chosen Killing trajectory, there exists a point p_n lying on it and a null geodesic which begins at p_n, passes through the wormhole exactly n times in a chosen direction, and returns to the initial point p_n. The directions of initial and final tangent vectors of these null geodesics are not necessarily the same. The points p_n for a given n form a so-called nth *polarized hypersurface* H_n. The hypersurfaces H_n lie above the chronology horizon, and for $n \to \infty$ the set of H_n is concentrated near it [Kim and Thorne (1991)].

We now demonstrate that the existence of closed null geodesics with $n \neq 0$ near the chronology horizon results in the divergence of $\langle T_{\mu\nu} \rangle^{ren}$ near the chronology horizon [Kim and Thorne (1991), Frolov (1991)].

For simplicity, we restrict ourselves to consideration of the quantum theory of a scalar massless field Φ obeying the equation

$$\Box\Phi - \xi R\Phi = 0, \tag{16.5.1}$$

where R is the scalar curvature, and the parameter ξ takes the value 1/6 for a conformal invariant field. There are two equivalent ways of quantizing the field Φ in a multiply-connected spacetime. The first is to work directly with solutions of equation (16.5.1) in the spacetime M. The second approach uses the universal covering space $\tilde{M}$ [Schulman (1968), Dowker (1972), Dowker and Banach (1978), Banach and Dowker (1979a,b)]. The discrete group of isometries Γ which is isometric to the fundamental group $\pi_1(M)$ acts on $\tilde{M}$ from the left, and the quotient $\Gamma\setminus\tilde{M}$ coincides with the original spacetime M. The universal covering space $\tilde{M}$ is simply connected by definition. That is why $\tilde{M}$ admits a *global* Killing vector field even though the Killing vector is not globally defined on the multiply-connected locally static spacetime M. The existence of a global Killing vector field in $\tilde{M}$ makes it possible to separate naturally the positive and negative frequencies and to develop the standard scheme of quantization. The physically interesting results in the spacetime M can be obtained by projecting the field theory in $\tilde{M}$ onto the field theory on M.

The basic idea of this projection is to identify M with a fundamental domain $\Gamma\setminus\tilde{M}$ and regard a field theory on M as a field theory on $\tilde{M}$ obeying certain conditions. Because the change of the fundamental domain is a symmetry transformation and the metric $d\tilde{s}^2$ on $\tilde{M}$ is invariant under these transformations $\gamma\in\Gamma$, invariance of the field action implies

$$\Phi(\gamma x) = a(\gamma)\,\Phi(x). \tag{16.5.2}$$

The group property of Γ gives $a(\gamma_1\gamma_2) = a(\gamma_1)\,a(\gamma_2)$. A field Φ on $\tilde{M}$ obeying condition (16.5.2) is known as an *automorphic field.*

The Hadamard function $G^1(x,x')$ in the physical spacetime M is connected with the Hadamard function $\tilde{G}^1(x,x')$ in the universal covering space $\tilde{M}$ by the relation [Banach and Dowker (1979a,b)]

$$G^1(x,x') = \sum_{\gamma\in\Gamma} a(\gamma^{-1})\,\tilde{G}^1(x,\gamma x'). \tag{16.5.3}$$

If one applies the operator

$$\begin{aligned} D_{\mu\nu} = {} & \frac{1}{2}\Big(\frac{1}{2}-\xi\Big)(\nabla_{\mu'}\nabla_\nu + \nabla_\mu\nabla_{\nu'}) + \Big(\xi-\frac{1}{4}\Big)\,g_{\mu\nu}\nabla_\rho\nabla^{\rho'} - \frac{1}{2}\,\xi(\nabla_\mu\nabla_\nu + \nabla_{\mu'}\nabla_{\nu'}) \\ & + \frac{1}{8}\,\xi\, g_{\mu\nu}\left(\nabla_\rho\nabla^\rho + \nabla_{\rho'}\nabla^{\rho'}\right) + \frac{1}{2}\,\xi\left[R_{\mu\nu} - \frac{1}{2}(1-3\xi)g_{\mu\nu}R\right] \end{aligned} \tag{16.5.4}$$

to $G^1(x,x')$, subtracts the divergences, and takes the coincidence limit $x'=x$, then one obtains the renormalized stress-energy tensor $\langle T_{\mu\nu}\rangle^{\rm ren}$ in a physical spacetime (see Christensen (1976, 1978), Birrell and Davies (1982) and Section 11.1). It should be emphasized that the divergences, which are to be subtracted are uniquely specified by the local geometry, so that they are the same for M and $\tilde{M}$. That is why the renormalized stress-energy tensor in the physical spacetime M can be written as

$$\langle T_{\mu\nu}\rangle_M^{ren} = \langle T_{\mu\nu}\rangle_{\tilde{M}}^{ren} + T_{\mu\nu}\,, \tag{16.5.5}$$

where

$$T_{\mu\nu}(x) = \sum_n{}' \, T^n_{\mu\nu}(x), \qquad T^n_{\mu\nu}(x) = \lim_{x'\to x} D_{\mu\nu}\, \tilde{G}^1(x, \gamma^n x')\,. \tag{16.5.6}$$

The prime in the sum indicates that the term with $n = 0$ is omitted from the summation. The universal covering space $\tilde{M}$ is a regular static spacetime, and generally $\langle T_{\mu\nu}\rangle_{\tilde{M}}^{ren}$ is regular and finite.

We now demonstrate that additional contributions $T^n_{\mu\nu}(x)$ to the stress-energy tensor connected with the non-trivial topology of M are divergent near the nth polarized hypersurface. To show this we choose a point x_0 on H_n and denote by C_n a null geodesic in $\tilde{M}$ connecting x_0 and $x_{0,n} \equiv \gamma^n x_0$. We assume that C_n lies in some causal domain $\Omega[C_n]$, so that the Hadamard function $\tilde{G}^1$ allows the following expansion [Hadamard (1923), DeWitt and Brehme (1960), Friedlander (1975)]

$$\tilde{G}^1(x, x') = \frac{\Delta^{1/2}}{4\pi^2}\left[\frac{1}{\sigma(x,x')} + v(x,x')\ln|\sigma(x,x')| + w(x,x')\right], \tag{16.5.7}$$

where $\sigma(x,x') = \frac{1}{2}\, s^2(x,x')$; $s(x,x')$ is a geodesic interval, and $\Delta = \Delta(x,x')$ is a Van Vleck-Morette determinant. The functions v and w are finite in the coincidence limit.

Consider now points $x^\mu = x_0^\mu - \delta t\, \xi^\mu(x_0)$ and $x_n^\mu \equiv \gamma^n x^\mu = x_{0,n}^\mu - \delta t\, A^n\, \xi^\mu(x_{0,n})$ located on the Killing trajectories passing through x_0 and $x_{0,n}$. For these points $\sigma(x, x_n) \simeq \delta t$. The leading (divergent as $\delta t \to 0$) part of $T^n_{\mu\nu}$ is of the form[3]

$$T^n_{\mu\nu} \simeq \frac{\Delta_n^{1/2}}{12\,\pi^2\,\sigma^3(x,x_n)}[\sigma_{\mu'}\,\sigma_\nu + \sigma_\mu\,\sigma_{\nu'} - \frac{1}{2}\, g_{\mu\nu}\,\sigma_\rho\,\sigma^{\rho'} - \frac{1}{2}\,(\sigma_\mu\,\sigma_\nu + \sigma_{\mu'}\sigma_{\nu'})]\,. \tag{16.5.8}$$

This relation shows that $T^n_{\mu\nu}$ is divergent near the nth polarized surface and this divergence is of order $(\delta t)^{-3}$. Near the chronology horizon (for large n) the renormalized stress-energy tensor describes null fluid propagation with a energy density diverging to infinity at the horizon. [For more detailed discussion of the divergence structure, see Frolov (1991).]

To summarize, we have shown that in a non-potential gravitational field of a locally static spacetime with a wormhole the renormalized stress-energy tensor $\langle T_{\mu\nu}\rangle^{ren}$ is divergent near each of the polarized hypersurfaces H_n. This result is in complete agreement with the result of calculations of Kim and Thorne (1991) who obtained the

[3]For simplicity, we write the result for a non-twisting ($a(\gamma) = 1$) conformal ($\xi = 1/6$) massless field.

renormalized stress-energy tensor near a chronology horizon for the model of moving wormhole's mouths in a flat spacetime.

The divergence of the renormalized stress-energy tensor near a chronology horizon means that the backreaction of the quantum effects on the background geometry is to be taken into account and one needs to consider a self-consistent problem which includes quantum effects. It is possible to make quite general remarks concerning the possible properties of such solutions. First of all, if we really believe in the one-loop result and assume that there exist only one massless scalar field, then we may conclude that time-machine formation is forbidden by quantum effects. In our concrete model, according to classical theory, a time machine is formed due to the non-potentiality of the gravitational field of a wormhole. The negative energy flux due to the vacuum polarization is directed in such a way that it automatically reduces the non-potentiality of the field. As long as no singularity arises, then this effect acts until the gravitational field becomes potential. In other words, the backreaction either prevents the formation of the chronology horizon or destroys it by creating a singularity.

In reality there exist many physical fields of different spins. The general conclusion about the divergence of the renormalized stress-energy tensor near the chronology horizon evidently remains valid also for massive fields because the leading (divergent at small geodesic distances) part of the massive propagator has the same singularity as for massless case. The sign of the vacuum energy density depends on the spin of the field and has the opposite sign for fermions. Nevertheless, the mutual cancellation of divergences for bosons and fermions is impossible [Klinkhammer (1992)]. If the contribution of fermions is dominant, the resulting flux of positive mass increases the non-potentiality of the gravitational field until a singularity is created.

One may expect a cutoff of the divergence due to quantum gravity effects, but no calculations have been done to prove it. Different estimations for this cutoff were proposed by Kim and Thorne (1991) and by Hawking (1992). For the more invariant cutoff proposed by Hawking, the curvature of the spacetime due to quantum effects must reach the Planckian value which effectively is to be considered as a singularity in the framework of the classical theory. Summarizing these results, Hawking proposed the chronology protection conjecture: *The laws of physics do not allow the appearance of closed timelike curves.*

This conjecture is unproved. There are examples in two and four dimensions, where for a *special choice of state* of a quantum field, $\langle T_{\mu\nu}\rangle^{\text{ren}}$ remains finite at the chronology horizon [Sushkov (1995), Krasnikov (1996)]. Recently, Visser (1997) demonstrated that even for a general choice of the state, the singularity of $\langle T_{\mu\nu}\rangle^{\text{ren}}$ can be made weaker if instead of one wormhole a number of wormholes are used. For special relative positions of the wormhole's mouths, and a large number N of wormholes, there is an additional factor $(b/L)^{2(N-1)}$ in the expression for the stress-energy tensor (b is the radius of the mouths, and L is the distance between them). This factor is connected with the defocusing effect, and formally arises from the

Conclusion

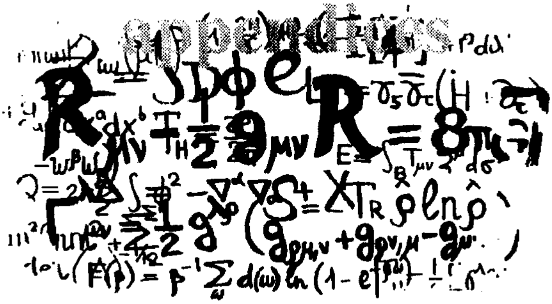
Why are the problems of black holes so important for modern physics and astrophysics? The answer is obvious: black holes are absolutely unusual objects. They are neither material bodies nor radiation fields. Black holes are clots of gravity. One may say that black holes open for us a door to a new very wide field of study of the physical world.

Some 30 years ago, very few scientists thought that black holes may really exist. Attention focussed on the black hole hypothesis after neutron stars had been discovered. It was rather surprising that astrophysicists immediately 'welcomed' them. Black holes found their place not only in the remnants of supernova explosions but also in the nuclei of galaxies and quasars.

The question about the cosmological role of small black holes became especially important after Hawking's discovery of the quantum evaporation of black holes. The hypothesis of elementary black holes (maximons) is interesting not only in its cosmological consequences but also in the aspects concerning particle physics. In all likelihood, virtual black holes will be an important element of the future quantum theory of gravitation. The investigation of the properties of black holes revealed profound relationships between gravitation, quantum theory, and thermodynamics. This progress (and especially the fact that the theory of black holes participating in physical processes requires qualitatively new ideas) has brought about the growth of an essentially novel branch of physics during the past 15 to 20 years: black-hole physics with its own object of study and its inherent problems. These problems are often of a very fundamental nature, while the object is so astonishing that this new field attracted quite a few researchers.

We wanted this book to explain the main phenomena in black hole physics. We fully understand that some aspects of this field deserve a more detailed presentation than we were able to provide. We feel slightly justified because incompleteness frequently reflects the current situation in the theory. Black-hole physics is a young and rapidly progressing science. We hope that this development will not only clarify the puzzling aspects but also make physicists happy with new, even less predictable results.

Appendices



Van Vleck-Morette determinant Δ_n in (16.5.8). Kay, Radzikowski, and Wald (1997) proved that even if the singularity of the averaged stress-energy tensor near the chronology horizon could be smoothed or be absent for a special choice of state of a quantum field, $\langle T_{\mu\nu}\rangle^{\mathrm{ren}}$ is always ill defined at least at some points of the horizon since the point-splitting prescription cannot exist, and in fact must diverge in some directions. This result can be viewed as supportive of the chronology protection conjecture (see also Cassidy and Hawking (1997)).

16.6 Quantum Theory and Time Machines

We have discussed the role of quantum effects (vacuum polarization) in the process of time machine formation. But there is another more general aspect of the problem. Namely, we may assume that a time machine exists. Is it possible to formulate quantum mechanics in such a world? Though this problem appears to be very complicated, the study in this direction has proved to be fruitful. This interesting subject goes far beyond the scope of the book, so that we confine ourselves to brief remarks and to giving the corresponding references.

First of all, in the presence of a time machine there is no well-defined global time. As a result, it is impossible to determine a quantum state at given moment of time and to write the Schrödinger equation describing evolution of a quantum state in time. Nevertheless, the Feynman's sum over histories approach can be used to obtain the amplitude of transitions (for example, from one spacetime point to another). It turns out that for non-relativistic particles [Politzer (1992)] and also for relativistic fields [Friedman *et al.* (1992a,b)], this approach enables one to compute unique probabilities for the outcomes of all measurements that one might reasonably try to make, even in the achronal regions of spacetime. However, if the particles or fields are self-interacting, quantum theory in the presence of time machine is peculiar. If a time machine was "operating" only during finite period, and there exist initial and final regions with the usual causal properties, nevertheless, the transition matrix from the initial region to the final region is non-unitary for an interacting field theory, order by order in perturbation theory [Politzer (1992), Boulware (1992), Friedman *et al.* (1992a,b)]. One might have standard Hamiltonian quantum mechanics in the chronal future domain, but only under the condition that the future influences probabilities in the chronal initial region [Friedman *et al.* (1992a,b)]. In this kind of situations the approach based on the *generalized quantum theories* developed by Hartle (1994, 1995) might be appropriate. For other aspects of quantum theory in the presence of a time machine, see also Hawking (1992), Deutsch (1992),Thorne (1993), Goldwirth *et al.* (1994), Mensky and Novikov (1996b).

These remarks might be important for the discussion of the role of quantum effects for time-machine formation. In the above discussion of quantum effects, it was implicitly assumed that before formation of the chronology horizon formation the standard local causal quantum field theory is valid. The situation would be completely different if it turns out that a non-contradictory quantum mechanics in a spacetime with closed timelike curves is possible only if there exist some global acausal restrictions on the choice of states and their evolution which modify the standard equations not only in regions with closed timelike curves, but also in the regions lying to the past of the chronology horizon. In this case, the above conclusions concerning the behavior of the renormalized stress-energy tensor near the chronology horizon might fail. In any case, a proof of the chronology protection conjecture (if it is valid) might require further development of the theory.

Appendix A

Mathematical Formulas

In this appendix, we have collected the most important formulas of Riemannian geometry and general relativity that are used throughout the book. We only list the main relations and give very brief comments because the derivations of these formulas and all necessary clarifications can be found in the available textbooks and monographs [see, e.g., Landau and Lifshitz (1975), Misner, Thorne, and Wheeler (1973), Hawking and Ellis (1973), Kramer *et al.* (1980), Vladimirov (1982), Wald (1984)].

A.1 Differential Manifold. Tensors

An n-dimensional, C^m, real manifold M is a set of points together with a collection of subsets $\{O_\alpha\}$ satisfying the following properties:

1. Each point $p \in M$ lies in at least one O_α , i.e., the $\{O_\alpha\}$ cover M.

2. For each α, there is a one-to-one, onto, map $\psi_\alpha : O_\alpha \to U_\alpha$, where U_α is an open subset of R^n (Euclidean n-dimensional space).

3. If any two sets O_α and O_β overlap, we can consider the map $\psi_\beta \circ \psi_\alpha^{-1}$ ($\circ$ denotes composition) which takes points in $\psi_\alpha[O_\alpha \cap O_\beta] \subset U_\alpha \subset R^n$ to points in $\psi_\beta[O_\alpha \cap O_\beta] \subset U_\beta \subset R^n$. We require this map to be C^m, i.e., m-times continuously differentiable.

Each map ψ_α is generally called a *chart* by mathematicians and a *coordinate system* by physicists. The collection of all charts is called an *atlas*. The map $\psi_\beta \circ \psi_\alpha^{-1}$ describes the transition from one ("old") coordinate system to another ("new") one. If x^μ are n numbers labeling points in U_α, and $x'^{\mu'}$ are n numbers labeling points in U_β, then the transition from one coordinate system (chart) to another one is given by functions $x'^{\mu'} = x'^{\mu'}(x^\mu)$. The requirement that this transition is given by smooth functions allows one to use differential equations to describe the physical laws. For simplicity, we assume that $m = \infty$.

A k-times covariant and l-times contravariant tensor $\mathbf{T}$ at point p of M in each coordinate system x^μ which covers p is described by n^{k+l} numbers $T^{\mu_1 \ldots \mu_l}_{\nu_1 \ldots \nu_k}$ (its components), and for any two such coordinate systems the components of $\mathbf{T}$ are related as

$$T^{\mu_1{}' \ldots \mu_l{}'}_{\nu_1{}' \ldots \nu_k{}'} = \frac{\partial x'^{\mu_1{}'}}{\partial x^{\mu_1}} \cdots \frac{\partial x'^{\mu_l{}'}}{\partial x^{\mu_l}} \frac{\partial x^{\nu_1}}{\partial x'^{\nu_1{}'}} \cdots \frac{\partial x^{\nu_k}}{\partial x'^{\nu_k{}'}} T^{\mu_1 \ldots \mu_l}_{\nu_1 \ldots \nu_k} . \tag{A.1}$$

Symmetrization $T_{(\mu_1 \ldots \mu_p)}$ and *antisymmetrization* $T_{[\mu_1 \ldots \mu_p]}$ of a tensor $T_{\mu_1 \ldots \mu_p}$:

$$T_{(\mu_1 \ldots \mu_p)} = \frac{1}{p!} \sum_{\substack{\text{over all} \\ \text{permutations} \\ (\mu_1, \ldots, \mu_p)}} T_{\mu_1 \ldots \mu_p} , \tag{A.2}$$

$$T_{[\mu_1 \ldots \mu_p]} = \frac{1}{p!} \sum_{\substack{\text{over all} \\ \text{permutations} \\ (\mu_1, \ldots, \mu_p)}} (-1)^J \, T_{\mu_1 \ldots \mu_p} , \tag{A.3}$$

where $J = 0$ if the permutation is even, and $J = 1$ otherwise.

A.2 Metric. Space and Time Intervals

A spacetime in general relativity is described as a four-dimensional differential manifold with a metric g on it. Points of this manifold correspond to different physical *events*. A spacetime metric $g_{\mu\nu}$ is a non-degenerate symmetric covariant tensor field which has signature $(- + + +)$. We use the following notations for *indices*:

- Greek indices $\alpha, \beta, \ldots$ run through the values $0, 1, 2, 3$.
- Roman indices $i, j, \ldots$ run through the values $1, 2, 3$.

A smooth curve $x^\mu(\lambda)$ is said to be spacelike, timelike, or lightlike at a point $\lambda = \lambda_0$ if the vector $u^\mu = dx^\mu / d\lambda$ tangent to it at this point satisfies the condition $u^\mu u_\mu > 0$, $u^\mu u_\mu < 0$, or $u^\mu u_\mu = 0$, respectively. The quantity $ds^2 = g_{\mu\nu} \, dx^\mu \, dx^\nu$ gives the interval between two close events x^μ and $x^\mu + dx^\mu$. For a line $x^\mu(\lambda)$ the proper interval between the points λ_1 and λ_2 is

$$S = \int_{\lambda_1}^{\lambda_2} \sqrt{| \, g_{\mu\nu} \, (dx^\mu / d\lambda) \, (dx^\nu / d\lambda) |} \, d\lambda . \tag{A.4}$$

For a timelike curve S is proper time, for a spacelike curve S is a proper distance.

Let x^μ be a coordinate system in the vicinity of an event p such that the coordinate lines $x^i =$ fixed are timelike lines. Then one can relate a reference frame to such a coordinate system by considering an ensemble of test particles moving along the above-specified x^0-coordinate lines. The coordinate x^0 plays the role of coordinate time, and it is related to proper time τ (registered by ideal clocks attached to the test particles) by

$$d\tau = \sqrt{|g_{00}|}\, dx^0 \,. \tag{A.5}$$

Let a test particle at a point p have the four-velocity u^μ. Then the set of events in its vicinity which are simultaneous with p in the reference frame of the particle is given by a local surface Σ passing through p and orthogonal to u^μ. The space interval between two neighboring test particles is defined as the proper distance between their trajectories along the surface Σ. Let u^μ be a tangent vector to the x^0-coordinate lines. (For a chosen reference frame and coordinate system $u^\mu \doteq \delta^\mu_0$; a symbol $\doteq$ indicates that equality takes holds only in a chosen coordinate system.) The tensor

$$p^\mu_\nu = \delta^\mu_\nu - \frac{u^\mu\, u_\nu}{u^2} \,, \tag{A.6}$$

is a projector onto the plane orthogonal to u^μ and it possesses the property

$$p^\mu_\alpha\, p^\alpha_\nu = p^\mu_\nu \,. \tag{A.7}$$

If dx^μ is the coordinate distance between two events on neighboring trajectories, then $p^\mu_\nu\, dx^\nu$ is the coordinate distance between the trajectories along Σ. The proper distance between the trajectories is

$$dl^2 = h_{\alpha\beta}\, dx^\alpha dx^\beta \equiv g_{\mu\nu}\, p^\mu_\alpha\, dx^\alpha\, p^\nu_\beta\, dx^\beta \,, \tag{A.8}$$

where

$$h_{\alpha\beta} = g_{\alpha\beta} - \frac{g_{\alpha 0}\, g_{\beta 0}}{}g_{00} \,. \tag{A.9}$$

It is easy to show that $h_{\alpha\beta} \doteq h_{ij}\, \delta^i_\alpha\, \delta^j_\beta$, and $h_{ij}\, g^{jk} = \delta^k_i$.

A.3 Causal Structure

A spacetime M is said to be *time-orientable* if there exists a globally defined and nowhere vanishing continuous timelike vector field.

The *chronological future* $I^+(Q)$ (*chronological past* $I^-(Q)$) of a set Q is the set of points for each of which there is a past-directed (future-directed) timelike curve that intersects Q. The boundary $\dot{P}$ of a set P is defined as $\dot{P} = \overline{P} \cap \overline{(M - P)}$, where M is the complete spacetime, and $\overline{P}$ is the closure of P. The boundary of the future $\dot{I}^+(Q)$ and the boundary of the past $\dot{I}^-(Q)$ of a set Q is lightlike except at Q itself. More precisely, if p is a point in $\dot{I}^+(Q)$ ($\dot{I}^-(Q)$) but not in the closure of Q, then there exists a past-directed (future-directed) null geodesic segment passing through p and lying in the boundary $\dot{I}^+(Q)$ ($\dot{I}^-(Q)$).

The curve $x^\mu(\lambda)$ is said to be causal (or *non-spacelike*) if its tangent vector $u^\mu = dx^\mu/d\lambda$ obeys the condition $u^\mu u_\mu \leq 0$ at each of its points. A non-spacelike (causal) curve between two points which is not a null geodesic curve can be deformed into a timelike curve connecting these points.

The *causal future* $J^+(Q)$ (*causal past* $J^-(Q)$) of a set Q is the set of points for each of which there is a past-directed (future-directed) causal curve that intersects Q.

The *future Cauchy domain* $D^+(Q)$ (*past Cauchy domain* $D^-(Q)$) of a set Q is the set of points such that any past-directed (future-directed) causal curve passing through it intersects Q.

A surface Σ is called spacelike, timelike, or null if its normal vector is timelike, spacelike, or null, respectively. A *global Cauchy surface* in a spacetime M is a non-timelike hypersurface that is intersected by each causal curve exactly once.

The *strong causality condition* is said to hold at p if every neighborhood of p contains a neighborhood of p which no non-spacelike curve intersects more than once. A set U is said to be a *globally hyperbolic region* if: (1) the strong causality holds on U, and (2) for every pair of points p and q in U the intersection $J^+(p) \cap J^-(q)$ is compact and contained in U.

Let p and q be two points in a globally hyperbolic region U that can be connected by a timelike or null curve. Then there exists a timelike or null geodesic between p and q which maximizes the length of timelike or null curves from p to q. The point q is *conjugate* to p along the (null) geodesic γ connecting these points if there exist an infinitesimally neighbouring (null) geodesic from p (different from γ) that intersects γ at q. For points on a null geodesic beyond the point q conjugate to p there is a variation of the geodesic γ that gives a timelike curve from p. Thus, γ cannot lie in the boundary of the future of p beyond the conjugate point q.

A.4 Covariant Derivative

The *Christoffel symbols* are defined as

$$\Gamma^\mu{}_{\alpha\beta} = g^{\mu\nu}\,\Gamma_{\nu\alpha\beta}\,, \quad \Gamma_{\nu\alpha\beta} = \frac{1}{2}\left(g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\alpha\beta,\nu}\right). \tag{A.10}$$

Covariant derivative:

$$\nabla_\alpha B^{\beta_1\ldots}{}_{\gamma_1\ldots} = \partial_\alpha B^{\beta_1\ldots}{}_{\gamma_1\ldots} + \Gamma^{\beta_1}{}_{\alpha\mu}\,B^{\mu\ldots}{}_{\gamma_1\ldots} + \ldots - \Gamma^\mu{}_{\alpha\gamma_1}\,B^{\beta_1\ldots}{}_{\mu\ldots} - \ldots\,. \tag{A.11}$$

Another notation for the covariant derivative: $\nabla_\alpha(\ ) \equiv (\ )_{;\alpha}$. Covariant derivatives with respect to the three-dimensional metric are denoted by $\nabla_i(\ ) \equiv (\ )_{:i}$. The notation for two-dimensional derivatives is $(\ )_{|a}$ $(a = 1, 2)$.

The covariant derivative possesses the following properties:

1. Linearity: For constants a and b $\nabla_\mu(aA^{\cdots}_{\cdots} + bB^{\cdots}_{\cdots}) = a\nabla_\mu A^{\cdots}_{\cdots} + b\nabla_\mu B^{\cdots}_{\cdots}$.

2. Leibnitz rule: $\nabla_\mu(A^{\cdots}_{\cdots}B^{\cdots}_{\cdots}) = \nabla_\mu(A^{\cdots}_{\cdots})B^{\cdots}_{\cdots} + A^{\cdots}_{\cdots}\nabla_\mu(B^{\cdots}_{\cdots})$.

3. Commutativity with contraction: $\nabla_\mu(A^{\cdots\beta\cdots}_{\cdots\beta\cdots}) = (\nabla_\mu A)^{\cdots\beta\cdots}_{\cdots\beta\cdots}$.

4. For a scalar field: $\nabla_\mu\varphi = \partial_\mu\varphi$.

5. Torsion free: $\nabla_\mu\nabla_\nu\varphi = \nabla_\nu\nabla_\mu\phi$.

6. $\nabla_\mu g_{\alpha\beta} = 0$.

A tensor field $A^{\alpha\dots}{}_{\beta\dots}$ undergoes *parallel transport* along a vector field ξ^μ if

$$\xi^\mu \nabla_\mu A^{\alpha\dots}{}_{\beta\dots} = 0.$$

A.5 Geodesic Lines

A *geodesic* $x^\alpha(\lambda)$ is defined as a solution of the equation

$$\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha{}_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = f(\lambda)\frac{dx^\alpha}{d\lambda}, \tag{A.12}$$

where $f(\lambda)$ is an arbitrary function. The function f can be set to zero by a reparametrization $\tilde\lambda = \tilde\lambda(\lambda)$. The parameter for which $f = 0$ is said to be affine. The affine parameter is defined up to a linear transformation. The affine parameter for timelike $(g_{\mu\nu}\dot x^\mu\dot x^\nu < 0)$ and spacelike $(g_{\mu\nu}\dot x^\mu\dot x^\nu > 0)$ geodesics is proportional to the proper length $\int |ds^2|^{1/2}$ along the curve.

A.6 Curvature

Riemann curvature tensor:

$$R^\mu{}_{\nu\alpha\beta} = \partial_\alpha\Gamma^\mu{}_{\nu\beta} - \partial_\beta\Gamma^\mu{}_{\nu\alpha} + \Gamma^\mu{}_{\sigma\alpha}\Gamma^\sigma{}_{\nu\beta} - \Gamma^\mu{}_{\sigma\beta}\Gamma^\sigma{}_{\nu\alpha}. \tag{A.13}$$

The Riemann curvature tensor possesses the symmetries:

$$R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu} = -R_{\nu\mu\alpha\beta} = -R_{\mu\nu\beta\alpha}, \quad R_{\mu[\nu\alpha\beta]} = 0, \tag{A.14}$$

and obeys the *Bianchi identities*

$$R_{\alpha\beta[\gamma\delta;\mu]} = 0\,. \tag{A.15}$$

The Riemann curvature tensor in four dimensions has 20 independent components. (In an N-dimensional space there are precisely $N^2(N^2-1)/12$ independent components of the Riemann tensor.) The following relation is valid for the *commutator of covariant derivatives*:

$$(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha) B_{\mu\ldots}{}^{\nu\ldots} = R_{\alpha\beta\mu}{}^{\sigma} B_{\sigma\ldots}{}^{\nu\ldots} + \ldots - R_{\alpha\beta\sigma}{}^{\nu} B_{\mu\ldots}{}^{\sigma\ldots} - \ldots\,. \tag{A.16}$$

Theorem. *A simply connected spacetime region V is flat (i.e., there exist "flat coordinates" in which $g_{\mu\nu} = \eta_{\mu\nu} = diag(-1,1,1,1)$ and $\Gamma^\alpha_{\mu\nu} = 0$) if and only if the Riemann tensor vanishes in V.*

Theorem. *Let γ be a curve connecting two points p and p' in a simply connected region V. The result of the parallel transport of any chosen vector u^α from p to p' does not depend on the particular choice of γ if and only if the Riemann tensor vanishes in V.*

Theorem. *Any tensor constructed from metric $g_{\mu\nu}$ and its partial derivatives up to the order N can be presented as a tensor function of the metric $g_{\mu\nu}$, the Riemann curvature tensor, and its covariant derivatives up to the order $N-2$.*

The *Ricci tensor* $R_{\mu\nu}$ and the *Einstein tensor* $G_{\mu\nu}$ connected with the Riemann curvature tensor are defined as

$$R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu}\,,$$

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}\, g_{\alpha\beta}\, R\,, \tag{A.17}$$

The Einstein tensor $G_{\mu\nu}$ obeys the relation

$$G^\mu{}_{\nu;\mu} = 0\,.$$

This relation directly follows from the Bianchi identities.

The *Weyl tensor* $C_{\mu\nu\alpha\beta}$ is defined as

$$C_{\mu\nu\alpha\beta} = R_{\mu\nu\alpha\beta} + g_{\nu[\alpha}\, S_{\beta]\mu} - g_{\mu[\alpha}\, S_{\beta]\nu}\,,$$

where $S_{\mu\nu} = R_{\mu\nu} - \frac{1}{6}\, g_{\mu\nu}\, R$. In four dimensions the Weyl tensor has 10 independent components. (In N-dimensional space the number of independent components of the Weyl tensor is $N^2(N^2-1)/12 - N(N+1)/2$ for $N \geq 3$, and 0 for $N \leq 3$.)

Theorem. *A simply connected spacetime region V is conformally flat (i.e., there exist "conformally flat coordinates" in which $g_{\mu\nu} = \Omega(x)\,\eta_{\mu\nu}$, where $\eta_{\mu\nu} = \mathrm{diag}\,(-1,1,1,1)$) if and only if the Weyl tensor vanishes in V.*

A.7 Lie- and Fermi-Transport

Diffeomorphism. Let M and N be differential manifolds. A map $\mathbf{f} : M \to N$ is said to be a diffeomorphism if it is one-to-one, onto, and it as well as its inverse is C^∞.

If ψ_α is a chart in the vicinity of a point $p \in M$, then $\psi_\alpha \circ \mathbf{f}^{-1}$ is a corresponding ("dragged along") chart in the vicinity of a point $p' = \mathbf{f}(p) \in N$. A tensor $\mathbf{T}^*$ at p' is said to be "dragged along" by the map $\mathbf{f}$ ($\mathbf{T}^* \equiv \mathbf{f}^* \mathbf{T}$) if its components in the corresponding "dragged along" chart are the same as the coordinates of the original tensor $\mathbf{T}$ at the initial point p in the original chart. If $\mathbf{f} : M \to M$ is a diffeomorphism, and $\mathbf{T}$ is a tensor field on M , one can compare T with $T^* \equiv \mathbf{f}^* T$. If $\mathbf{T} = \mathbf{T}^*$, then the tensor field is said to be invariant with respect to $\mathbf{f}$.

Lie derivative. Let M be a differential manifold, and let $\mathbf{f}_\lambda$ be a one-parameter group of diffeomorphisms. Denote $p_\lambda = \mathbf{f}_\lambda(p)$. For a chosen chart we define a function $x^\mu(\lambda; p) \equiv x^\mu(p_\lambda)$. Denote by

$$\xi^\mu(p) = \left. \frac{dx^\mu(\lambda; p)}{d\lambda} \right|_{\lambda=0} \tag{A.18}$$

a vector field generating a one-parameter group of diffeomorphisms. The Lie derivative $\mathcal{L}_\xi A^{\alpha\dots}{}_{\beta\dots}$ of a tensor field $A^{\alpha\dots}{}_{\beta\dots}$ along ξ^μ is defined by the relation

$$\mathcal{L}_\xi A^{\alpha\dots}{}_{\beta\dots} = \lim_{\lambda \to 0} \left[\frac{\mathbf{f}^*_{-\lambda} A^{\alpha\dots}{}_{\beta\dots} - A^{\alpha\dots}{}_{\beta\dots}}{\lambda} \right]. \tag{A.19}$$

The following explicit expressions are valid for the Lie derivative

$$\begin{aligned} \mathcal{L}_\xi A^{\alpha\dots}{}_{\beta\dots} &= \xi^\mu\, \partial_\mu A^{\alpha\dots}{}_{\beta\dots} - \partial_\mu \xi^\alpha\, A^{\mu\dots}{}_{\beta\dots} - \dots + \partial_\beta \xi^\mu\, A^{\alpha\dots}{}_{\mu\dots} + \dots \\ &= \xi^\mu\, \nabla_\mu A^{\alpha\dots}{}_{\beta\dots} - \nabla_\mu \xi^\alpha\, A^{\mu\dots}{}_{\beta\dots} - \dots + \nabla_\beta \xi^\mu\, A^{\alpha\dots}{}_{\mu\dots} + \dots . \end{aligned} \tag{A.20}$$

The Lie derivative obeys the following relations

$$\mathcal{L}_\xi \eta^\alpha \equiv [\boldsymbol{\xi}, \boldsymbol{\eta}]^\alpha = \xi^\mu\, \partial_\mu \eta^\alpha - \eta^\mu\, \partial_\mu \xi^\alpha, \tag{A.21}$$

$$\mathcal{L}_\xi\, \mathcal{L}_\eta - \mathcal{L}_\eta\, \mathcal{L}_\xi = \mathcal{L}_{[\xi,\eta]} . \tag{A.22}$$

The *Fermi derivative* $\mathcal{F}_\xi A^{\alpha\dots}{}_{\beta\dots}$ of a tensor field $A^{\alpha\dots}{}_{\beta\dots}$ along a vector field $\boldsymbol{\xi}$ ($\xi^\mu \xi_\mu \neq 0$) is defined as:

$$\mathcal{F}_\xi A^{\alpha\dots}{}_{\beta\dots} = \xi^\mu\, \nabla_\mu A^{\alpha\dots}{}_{\beta\dots} + \mathcal{F}^\alpha_{\ \sigma}\, A^{\sigma\dots}{}_{\beta\dots} + \dots - \mathcal{F}^\sigma_{\ \beta}\, A^{\alpha\dots}{}_{\sigma\dots} + \dots , \tag{A.23}$$

where

$$\mathcal{F}^\alpha_{\ \sigma} = (w^\alpha\, \xi_\sigma - \xi^\alpha\, w_\sigma)\, \varepsilon(\boldsymbol{\xi}), \tag{A.24}$$

$$w^\alpha = \xi^\beta \xi^\alpha{}_{;\beta}, \quad \varepsilon(\boldsymbol{\xi}) = \mathrm{sign}(-\xi^\mu \xi_\mu). \tag{A.25}$$

A tensor field $A^{\alpha\dots}{}_{\beta\dots}$ is said to be *Lie-transported* along $\boldsymbol{\xi}$ if

$$\mathcal{L}_\xi A^{\alpha\dots}{}_{\beta\dots} = 0, \tag{A.26}$$

and *Fermi-transported* if

$$\mathcal{F}_\xi A^{\alpha\dots}{}_{\beta\dots} = 0. \tag{A.27}$$

A.8 Symmetries and Conservation Laws

Let $g_{\alpha\beta}$ be a metric on a differential manifold M and $\mathbf{f}: M \to M$ be a diffeomorphism under which the metric is invariant: $g = g^* \equiv \mathbf{f}^* g$. Such a diffeomorphism is called an *isometry*. It is evident that the Riemann curvature $\mathcal{R}_{\dots}$ and all its covariant derivatives $\nabla \dots \nabla \mathcal{R}_{\dots}$ are also invariant under the isometry

$$\mathbf{f}^* \mathcal{R}_{\dots} = \mathcal{R}_{\dots}, \quad \mathbf{f}^* \nabla \dots \nabla \mathcal{R}_{\dots} = \nabla \dots \nabla \mathcal{R}_{\dots}. \tag{A.28}$$

Let $\mathbf{f}_\lambda$ be a one-parameter group of isometries. A vector field $\boldsymbol{\xi}$ generating the isometries is called a *Killing vector field.* A Killing vector field ξ^μ in the space with metric $g_{\mu\nu}$ obeys the equation

$$\mathcal{L}_\xi g_{\mu\nu} \equiv 2\,\xi_{(\mu;\nu)} = 0. \tag{A.29}$$

This equation implies $\mathcal{L}_\xi R_{\alpha\beta\gamma\delta} = 0$ and $\mathcal{L}_\xi \nabla_{\nu_1} \dots \nabla_{\nu_n} R_{\alpha\beta\gamma\delta} = 0$. The Killing vector field ξ^μ obeys the equation

$$\xi_{\mu;\alpha;\beta} = R_{\gamma\beta\alpha\mu}\, \xi^\gamma. \tag{A.30}$$

If ξ^μ and η^μ are two Killing vector fields, then $[\boldsymbol{\xi}, \boldsymbol{\eta}]^\mu \equiv \xi^\alpha \partial_\alpha \eta^\mu - \eta^\alpha \partial_\alpha \xi^\mu$ is also a Killing vector field.

A *Killing tensor field* is a symmetric tensor field $\xi_{\alpha_1\dots\alpha_m} = \xi_{(\alpha_1\dots\alpha_m)}$ satisfying the condition $\xi_{(\alpha_1\dots\alpha_m;\beta)} = 0$. A second order symmetric Killing tensor field $\xi_{\mu\nu}$ can be expressed in terms of an antisymmetric *Yano-Killing tensor* $f_{\mu\nu}$ obeying the relation $f_{\mu\nu;\lambda} + f_{\mu\lambda;\nu} = 0$ as follows: $\xi_{\mu\nu} = f_{\mu\lambda} f^\lambda{}_\nu$. For a freely moving particle in an external gravitational field which possesses a Killing tensor $\xi_{\alpha_1\dots\alpha_m}$ the following quantity is conserved (i.e., is constant along the trajectory):

$$P_\xi = \xi_{\alpha_1\dots\alpha_m} u^{\alpha_1} \dots u^{\alpha_m}. \tag{A.31}$$

In particular, for a Killing vector ξ^α a conserved quantity is $P_\xi = \xi_\alpha u^\alpha$.

A.9 Geometry of Congruence of Lines

A congruence of curves is a three-parameter family of curves $x^\mu(\lambda; y^i)$ (λ is a parameter along the curve, and y^i is a set of parameters that "labels" the curves) such that one and only one curve of the family passes through each point. If specific parameters λ and y^i are chosen on the congruence, we obtain a coordinate system. A congruence of timelike curves is sometimes known as a *reference frame.*

Let λ be a proper time parameter and $u^\alpha = dx^\alpha/d\lambda$ be a vector field ($u^\alpha u_\alpha = -1$) connected with a congruence $x^\alpha(\lambda; y^i)$. Then $u_{\alpha;\beta}$ admits an unambiguous representation of the form

$$u_{\alpha;\beta} = \omega_{\alpha\beta} + D_{\alpha\beta} - w_\alpha u_\beta\,, \tag{A.32}$$

where $w_\alpha = u^\beta u_{\alpha;\beta}$ is the acceleration; $\omega_{\alpha\beta}$ is the *vorticity tensor,*

$$\omega_{\alpha\beta} = \omega_{[\alpha\beta]} = \frac{1}{2}\left(u_{\alpha;\mu}\, p^\mu{}_\beta - u_{\beta;\mu}\, p^\mu{}_\alpha\right), \tag{A.33}$$

and the *rate of deformation tensor* $D_{\mu\nu}$ is

$$D_{\alpha\beta} = \frac{1}{2}\mathcal{L}_u p_{\alpha\beta} = \frac{1}{2}\left(u_{\alpha;\mu}\, p^\mu{}_\beta + u_{\beta;\mu}\, p^\mu{}_\alpha\right). \tag{A.34}$$

Here $p_{\alpha\beta} = g_{\alpha\beta} + u_\alpha u_\beta$ is the projection tensor (A.6) that projects vectors and tensors onto the three-dimensional space orthogonal to u^α. The *shear* $\sigma_{\mu\nu}$ is defined as the traceless part of the rate of deformation tensor $D_{\mu\nu}$, while the *expansion*

$$\theta = u^\alpha_{;\alpha} \tag{A.35}$$

is its trace, so that

$$D_{\mu\nu} = \sigma_{\mu\nu} + \frac{1}{3}\theta\, p_{\mu\nu}\,. \tag{A.36}$$

Raychaudhuri equation:

$$\frac{d\theta}{d\lambda} = w^\alpha{}_{;\alpha} + 2(\omega^2 - \sigma^2) - \frac{1}{3}\theta^2 - R_{\alpha\beta}\, u^\alpha u^\beta\,, \tag{A.37}$$

where

$$\omega^2 = \frac{1}{2}\omega_{\alpha\beta}\,\omega^{\alpha\beta}, \qquad \sigma^2 = \frac{1}{2}\sigma_{\alpha\beta}\,\sigma^{\alpha\beta}. \tag{A.38}$$

For the special choice of coordinates in which the lines of the congruence are described by the equations $x^i = \text{const}$ the differential invariants take simpler forms. In these coordinates

$$u^\mu = \frac{1}{\sqrt{-g_{00}}}\,\delta^\mu_0\,, \qquad u_\mu = \frac{1}{\sqrt{-g_{00}}}\,g_{0\mu}\,. \tag{A.39}$$

The projection tensor $p_{\alpha\beta}$ has only spatial non-vanishing components, which we denote by h_{ij}: $h_{ij} = p_{ij}$. The three-dimensional metric h_{ij} is of the form

$$h_{ik} = g_{ik} - \frac{g_{0i}\, g_{0k}}{g_{00}}. \tag{A.40}$$

The components of the three-dimensional tensor h^{ij} inverse to h_{ij} coincide with the spatial components of $g^{\mu\nu}$

$$h_{ki}\, g^{il} = \delta^l_k. \tag{A.41}$$

The acceleration vector w_μ is orthogonal to u^μ, and hence has only spatial components w_i. Define $w^i = h^{ij}\, w_j$. w^i coincides with the spatial components of the vector w^μ and can be written in the form

$$w^i = -\frac{\Gamma^i_{00}}{g_{00}}. \tag{A.42}$$

Sometimes it is convenient to use the vector a^i of free fall acceleration of a test particle with respect to the chosen frame. a^i is the field-vector of the gravitational-inertial forces which act in the chosen frame of reference. It is evident that $a^i = -w^i$. The scalar

$$a = \sqrt{a^i\, a^k\, h_{ik}} \equiv \sqrt{a_\mu\, a^\mu} \tag{A.43}$$

is the magnitude of the free fall acceleration of a body initially at rest in the chosen reference frame.

The deformation rate tensor $D_{\mu\nu}$ and the vorticity tensor $\omega_{\mu\nu}$ have only spatial components, which determine the angular velocity of rotation of the reference frame (ω^{ik}) and its rate of shear (D_{ik}). Equation (A.34) implies

$$D_{ik} = \frac{1}{2}\, \frac{1}{\sqrt{-g_{00}}}\, \frac{\partial h_{ik}}{\partial x^0}, \tag{A.44}$$

while for ω_{ij} one has

$$\omega^{ik} = \frac{1}{\sqrt{-g_{00}}}\, g^{\alpha[i}\, \Gamma^{k]}_{\alpha 0}. \tag{A.45}$$

The vector of the angular velocity of rotation of the reference frame, Ω_i, is calculated, using ω_{ij}:

$$\Omega_i = \frac{1}{2}\, \epsilon_{ikl}\, \omega^{kl}, \tag{A.46}$$

where ϵ_{ikl} is the totally antisymmetric object, $\epsilon_{123} = (g/g_{00})^{1/2}$. The scalar

$$\Omega = \sqrt{\Omega_i\, \Omega_k\, h^{ik}} \tag{A.47}$$

is the angular velocity of rotation per unit of proper time $d\tau = (-g_{00})^{1/2}\, dt$.

A geodesic congruence is a special type of a congruence in which the lines forming it are geodesics so that $w^\mu = 0$. Let $n^\mu(\lambda)$ be a vector that connects two nearby geodesics for identical values of the affine parameters λ along them. Then the following *geodesic deviation equation*

$$\frac{D^2 n^\alpha}{d\lambda^2} + R^\alpha{}_{\beta\gamma\delta}\, u^\beta\, n^\gamma\, u^\delta = 0 \tag{A.48}$$

is valid. Here

$$u^\beta = \frac{dx^\beta}{d\lambda}, \qquad \frac{D}{d\lambda} = u^\mu\, \nabla_\mu . \tag{A.49}$$

A.10 Stationary Congruences in a Stationary Spacetime

A.10.1 Killing congruence

Let ξ^μ be a timelike Killing vector and consider a congruence of timelike lines generated by a vector field u^μ $(u^2 = -1)$. The congruence is stationary if $\mathcal{L}_\xi u = 0$. An important example of a stationary congruence is the *Killing congruence* generated by the vector field

$$\tilde{u}^\mu = \xi^\mu/(-\xi^2)^{1/2} . \tag{A.50}$$

For this congruence

$$w^\mu = \tilde{u}^\nu\, \tilde{u}^\mu{}_{;\nu} = \frac{1}{2}\left[\ln(-\xi^2)\right]^{;\mu} , \tag{A.51}$$

$$\tilde{u}_{\alpha;\beta} = \frac{\xi_{\alpha;\beta}}{(-\xi^2)^{1/2}} - \tilde{u}_\alpha\, w_\beta . \tag{A.52}$$

By comparing the latter expression with (A.32) and (A.37), we get

$$D_{\alpha\beta} = 0 , \tag{A.53}$$

$$\omega_{\alpha\beta} = \frac{\xi_{\alpha;\beta}}{(-\xi^2)^{1/2}} - (\tilde{u}_\alpha\, w_\beta - \tilde{u}_\beta\, w_\alpha) . \tag{A.54}$$

The 3-metric h_{ij} defined by (A.40) takes the form

$$h_{ij} = g_{ij} - \frac{\xi_i\, \xi_j}{\xi^2} . \tag{A.55}$$

The parameter t is called a Killing time if $dx^\mu/dt = \xi^\mu$. The origin $t = 0$ of the Killing time on each line of the congruence is arbitrary, that is why there exists an arbitrariness in the choice of t:

$$t \to \tilde{t} = t + f(x^i) . \tag{A.56}$$

One can choose the Killing time t and the parameters x^i labeling the Killing trajectories as new coordinates (t, x^i). In these coordinates

$$ds^2 = \xi^2 (dt + g_i\, dx^i)^2 + h_{ij}\, dx^i\, dx^j , \tag{A.57}$$

where $g_i = g_{ti}/g_{tt}$ is the three-dimensional part of the vector $g_\mu = (1, g_i) = \xi_\mu/\xi^2$. The Killing equation implies that the metric coefficients in (A.57) are time-independent. Under the coordinate transformation (A.56), ξ^2 and h_{ij} are invariant, while g_i transforms as

$$g_i \to \tilde{g}_i = g_i - f_{,i} . \tag{A.58}$$

The vorticity tensor $\omega_{\alpha\beta}$ can be written as

$$\omega_{\alpha\beta} = -(-\xi^2)^{1/2}\, g_{[\alpha;\beta]} . \tag{A.59}$$

It is evident that ω_{ij} has only spatial non-vanishing components

$$\omega_{ij} = -(-\xi^2)^{1/2}\, g_{[i,j]} , \tag{A.60}$$

and they are invariant under the transformation (A.56).

A.10.2 Congruence of locally non-rotating observers

We discuss now the properties of another important stationary congruence in a spacetime (A.57). Namely, we consider a congruence of lines orthogonal to the sections $t =$ const. For this purpose, we rewrite the metric (A.57) in the form

$$ds^2 = -\alpha^2\, dt^2 + g_{ij}\, (dx^i + \alpha^i\, dt)\, (dx^j + \alpha^j\, dt) . \tag{A.61}$$

By comparing (A.61) with (A.57), we get

$$\alpha^2 = g_{ij}\, \alpha^i\, \alpha^j - \xi^2 , \tag{A.62}$$

$$\alpha^i = \xi^2\, \tilde{g}^{ij}\, g_j , \tag{A.63}$$

where $\tilde{g}^{ij}$ is the three-dimensional matrix inverse to g_{ij}: $\tilde{g}^{ij}\, g_{jk} = \delta^i_k$.

The matrix $g^{\mu\nu}$ has the following components

$$g^{tt} = -\frac{1}{\alpha^2} , \quad g^{ti} = \frac{\alpha^i}{\alpha^2} , \quad g^{ij} = \tilde{g}^{ij} - \frac{\alpha^i\, \alpha^j}{\alpha^2} . \tag{A.64}$$

A future-directed unit vector u^μ orthogonal to t = const is $u_\mu = -\alpha\, t_{,\mu}$. It has components

$$u_\mu = -\,\alpha\, \delta^t_\mu\,, \qquad u^\mu = \alpha^{-1}(1, -\alpha^i)\,. \tag{A.65}$$

For the stationary congruence generated by u^μ, the vorticity tensor $\omega_{\mu\nu}$ vanishes ($\omega_{\mu\nu} = 0$). That is why this congruence is sometimes called the congruence of locally non-rotating observers. The acceleration vector w^μ for this reference frame and the deformation rate tensor are

$$w_\mu = (\ln\alpha)_{;\mu} + \alpha^2 (t^{;\nu}\alpha_{;\nu})\, t_{;\mu}\,, \tag{A.66}$$

$$D_{\mu\nu} = -\left[\alpha\, t_{;\mu;\nu} + 2\,\alpha_{(;\mu}\, t_{;\nu)} + \alpha^2 (t^{;\lambda}\,\alpha_{;\lambda})\, t_{;\mu}\, t_{;\nu}\right]. \tag{A.67}$$

Their spatial components are

$$w_i = (\ln\alpha)_{,i}\,, \qquad D_{ij} = \alpha\,\Gamma^t_{ij}\,. \tag{A.68}$$

A.11 Local Reference Frames

Consider a world line γ of a (generally speaking, accelerated) observer and denote by τ proper time measured by the ideal observer's clock (i.e., the proper time distance along his trajectory). Let $u^\mu = dx^\mu/d\tau$ be the four-velocity of the observer, and denote by $w^\mu \equiv u^\nu u^\mu{}_{;\nu}$ his acceleration. Choose an orthonormal tetrad $\mathbf{e}_{\hat{\alpha}}$ at some initial point on γ and use the Fermi transport to define the tetrad along γ. In particular, one has $\mathbf{e}^\alpha_{\hat{0}} = u^\alpha$.

A *local reference frame* of the observer is constructed as follows. Choose an event $p(\tau)$ at the line γ at moment τ of proper time and consider spatial geodesics emitted from this point which are orthogonal to u^μ. A point on the geodesic is uniquely defined by the value s of the proper distance along the geodesic, while the geodesic itself can be specified by a unit vector $\mathbf{n}$ tangent to it at $s = 0$. Each event in the vicinity of the observer's world line is intersected by precisely one of the geodesics. That is why $x^{\hat{0}} = \tau$, and $x^{\hat{\imath}} = s\mathbf{n}\cdot\mathbf{e}_{\hat{\imath}}$ can be used as coordinates in the vicinity of γ. These coordinates are known as *Fermi coordinates.* In Fermi coordinates

$$ds^2 = -\left(1 + 2\, w_{\hat{\imath}}\, x^{\hat{\imath}}\right) dx^{\hat{0}^2} + \delta_{\hat{\imath}\hat{\jmath}}\, dx^{\hat{\imath}}\, dx^{\hat{\jmath}} + O\left(|\,x^{\hat{\jmath}}|^2\right) dx^{\hat{\alpha}}\, dx^{\hat{\beta}}\,. \tag{A.69}$$

The so-constructed local reference frame is non-rotating. In particular, this means that if an observer is at rest with respect to this frame and carries gyroscopes with him, he observes that they do not precess with respect to the frame.

More general local reference frames can be obtained if we adjoin to $\mathbf{e}^\alpha_{\hat{0}} = u^\alpha$ three other unit vectors $\tilde{\mathbf{e}}^\alpha_{\hat{m}}$ which together with $\mathbf{e}^\alpha_{\hat{0}}$ form an orthonormal frame. One has $\tilde{\mathbf{e}}^\alpha_{\hat{m}} = A_{\hat{m}}{}^{\hat{n}}\, \mathbf{e}^\alpha_{\hat{n}}$, where $A_{\hat{m}}{}^{\hat{n}}$ is an orthogonal matrix depending on time. If $\tilde{\mathbf{e}}_{\hat{\mu}}$ coincides with $\mathbf{e}_{\hat{\mu}}$ at an initial moment of time, then $\omega_{\hat{m}}{}^{\hat{n}} = dA_{\hat{m}}{}^{\hat{n}}/dx^{\hat{0}}$ is the vorticity tensor.

The vorticity tensor $\omega_{\hat{m}\hat{n}}$ is antisymmetric. $\omega^{\hat{l}} = (1/2)\,\epsilon^{\hat{l}\hat{m}\hat{n}}\,\omega_{\hat{m}\hat{n}}$ is the angular velocity vector. The line element in such a rotating local reference frame takes the form

$$ds^2 = -\left(1 + 2\,w_{\hat{i}}\,x^{\hat{i}}\right) dx^{\hat{0}^2} - 2\left(\epsilon_{\hat{l}\hat{m}\hat{n}}\,x^{\hat{m}}\,\omega^{\hat{n}}\right) dx^{\hat{0}}\,dx^{\hat{l}} + \delta_{\hat{i}\hat{j}}\,dx^{\hat{i}}\,dx^{\hat{j}} + \dots\,, \qquad (A.70)$$

where ... indicate omitted terms of order $O\left(|\,x^{\hat{j}}|^2\right)$.

In the absence of local rotation for the special case of zero acceleration (i.e., when an observer is moving along a geodesic line), one has $g_{\hat{\alpha}\hat{\beta}} = \eta_{\alpha\beta}$, and $\Gamma^{\hat{\alpha}}_{\hat{\beta}\hat{\gamma}} = 0$ along this line in Fermi coordinates. The following expression for the metric, accurate to second order in $|\,x^{\hat{j}}|$ is valid in the vicinity of the geodesic world line in Fermi coordinates:

$$\begin{aligned} ds^2 &= \left(-1 - R_{\hat{0}\hat{i}\hat{0}\hat{j}}\,x^{\hat{i}}\,x^{\hat{j}}\right) dx^{\hat{0}^2} - \frac{4}{3}\,R_{\hat{0}\hat{i}\hat{k}\hat{j}}\,x^{\hat{i}}\,x^{\hat{j}}\,dx^{\hat{0}}\,dx^{\hat{k}} \\ &+ \left(\delta_{\hat{i}\hat{j}} - \frac{1}{3}\,R_{\hat{i}\hat{j}\hat{k}\hat{l}}\,x^{\hat{j}}\,x^{\hat{l}}\right) dx^{\hat{i}}\,dx^{\hat{k}} + O\left(|\,x^{\hat{j}}|^3\right) dx^{\hat{\alpha}}\,dx^{\hat{\beta}}. \end{aligned} \qquad (A.71)$$

Consider now local reference frames associated with a stationary congruence in a stationary spacetime. Let $\tilde{u}^{\mu}$ be a unit timelike vector tangent to the Killing congruence (see (A.50)). Denote by $z^{m}_{\hat{m}}$ time-independent three-vectors normalized by $h_{mn}\,z^{m}_{\hat{m}}\,z^{n}_{\hat{n}} = \delta_{\hat{m}\hat{n}}$, then $\mathbf{e}^{\mu}_{\hat{m}} = (-\,g_i\,z^{i}_{\hat{m}}, z^{m}_{\hat{m}})$ are three unit vectors that together with $\mathbf{e}^{\mu}_{\hat{0}} = u^{\mu}$ form a local reference frame. The local frames along the Killing trajectory are connected by the Lie-transport along ξ^{μ}: $\mathcal{L}_{\xi}\mathbf{e}_{\hat{m}} = 0$. Thus, we have

$$\mathcal{L}_u\mathbf{e}^{\mu}_{\hat{m}} = (\mathbf{e}^{\nu}_{\hat{n}}\,w_{\nu})\,u^{\mu}\,, \qquad (A.72)$$

where $w_{\mu} = \frac{1}{2}\left(\ln(-\xi^2)\right)_{,\mu}$.

One can also introduce another local reference frame along the Killing trajectories that are propagated in the Fermi sense. The reference frame $\mathbf{e}_{\hat{\mu}}$ is rotating with respect to the *Fermi frame*. The vorticity tensor is

$$\omega_{\hat{m}\hat{n}} = \mathbf{e}^{\mu}_{\hat{m}}\,\mathbf{e}^{\nu}_{\hat{n}}\,u_{\mu;\nu}\,. \qquad (A.73)$$

For a Killing congruence the deformation rate tensor vanishes $D_{\mu\nu} = 0$ so that $\omega_{\hat{m}\hat{n}} = \mathbf{e}^{\mu}_{\hat{m}}\,\mathbf{e}^{\nu}_{\hat{n}}\,\omega_{\mu\nu}$.

A.12 Geometry of Subspaces

Consider a spacelike or timelike hypersurface Σ, and denote by n^{μ} a unit normal to it. Let $x^{\mu} = x^{\mu}(y^i)$ be the equation of Σ. Then

$$h_{ij} = \frac{\partial x^{\mu}}{\partial y^i}\,\frac{\partial x^{\nu}}{\partial y^j}\,g_{\mu\nu} \qquad (A.74)$$

is the *induced metric* on Σ, and

$$K_{ij} = -\,\frac{\partial x^{\mu}}{\partial y^i}\,\frac{\partial x^{\nu}}{\partial y^j}\,\nabla_{\nu}n_{\mu} \qquad (A.75)$$

is the *extrinsic curvature* of Σ. Both h_{ij} and K_{ij} are invariant under transformations of the external coordinates x^μ, and are symmetric tensors with respect to transformations of the internal coordinates y^i.

Consider a set of geodesics orthogonal to Σ. Each of these geodesics can be labeled by the coordinates y^i of the point where it intersects Σ. Denote by n the proper time or proper distance parameter along the geodesics. We assume that $n = 0$ on Σ. Four parameters (n, y^1, y^2, y^3) can be used as coordinates in the vicinity of Σ. These coordinates are known as *Gaussian normal coordinates.* The metric in Gaussian normal coordinates takes the form $(\epsilon(\boldsymbol{n}) = n_\mu\, n^\mu)$

$$ds^2 = \epsilon(\boldsymbol{n})\, dn^2 + h_{ij}\, dy^i\, dy^j \,,$$

and the extrinsic curvature K_{ij} is

$$K_{ij} = -\frac{1}{2}\, h_{ij,n} \,.$$

Gauss-Codazzi equations:

$$R^m{}_{ijk} = {}^{(3)}R^m{}_{ijk} + \epsilon(\boldsymbol{n})(K_{ij}\, K^m_{\ k} - K_{ik}\, K_j{}^m)\,, \tag{A.76}$$

$$R^n{}_{ijk} = -\,\epsilon(\boldsymbol{n})(K_{ij:k} - K_{ik:j})\,, \tag{A.77}$$

where $\epsilon(\boldsymbol{n}) = n_\mu\, n^\mu = \pm 1$; n^μ is a unit vector normal to the hypersurface Σ; $(\ )_{:i}$ is the covariant derivative in the induced metric h_{ij}, and ${}^{(3)}R^m{}_{ijk}$ is the curvature tensor of the three-dimensional space with this metric. A component of the curvature tensor constructed for basis vector $\boldsymbol{n}$ has index n.

(3+1)-split of the Einstein tensor:

$$G^n{}_n = -\frac{1}{2}\, {}^{(3)}R + \frac{1}{2}\,\epsilon(\boldsymbol{n})\left[K^2 - K_{ij}\, K^{ij}\right], \tag{A.78}$$

$$G^n{}_i = -\,\epsilon(\boldsymbol{n})\,[K_i{}^m{}_{:m} - K_{:i}]\,, \tag{A.79}$$

where

$${}^{(3)}R = h^{ij}\, {}^{(3)}R^m{}_{imj}\,, \qquad K = h^{ij}\, K_{ij}\,. \tag{A.80}$$

A.13 Integration in Curved Space

Element of volume: $d^4v = \sqrt{-g}\, d^4x$.

Element $d\sigma_\alpha$ *of a hypersurface* Σ defined by equations $x^\mu = x^\mu(y^i)$:

$$d\sigma_\alpha = \frac{1}{3!}\, e_{\alpha\beta_1\beta_2\beta_3} \det\left(\frac{\partial x^{\beta_i}}{\partial y^j}\right) dy^1\, dy^2\, dy^3, \tag{A.81}$$

where $e_{\alpha\beta_1\beta_2\beta_3}$ is the antisymmetric tensor:

$$e_{\alpha\beta_1\beta_2\beta_3} = \sqrt{-g}\, \epsilon_{\alpha\beta_1\beta_2\beta_3}\,.$$

[$\epsilon_{\alpha\beta_1\beta_2\beta_3}$ is the completely antisymmetric symbol ($\epsilon_{0123} = 1$)].

Element $d\sigma_{\alpha\beta}$ *of a two-dimensional surface* S defined by equations $x^\mu = x^\mu(z^a)$ $(a = 1, 2)$:

$$d\sigma_{\alpha\beta} = \frac{1}{4}\, e_{\alpha\beta\beta_1\beta_2} \det\left(\frac{\partial x^{\beta_a}}{\partial z^b}\right) dz^1\, dz^2. \tag{A.82}$$

Integration in the Riemannian space. Let φ be a scalar, φ^α be a vector, and $\varphi^{\alpha\beta}$ be an antisymmetric tensor field. The following integrals are then defined:

$$T_V[\varphi] = \int_V \varphi\, d^4v\,, \tag{A.83}$$

$$T_\Sigma[\varphi^\alpha] = \int_\Sigma \varphi^\alpha\, d\sigma_\alpha\,, \tag{A.84}$$

$$T_S[\varphi^{\alpha\beta}] = \int_S \varphi^{\alpha\beta}\, d\sigma_{\alpha\beta}\,. \tag{A.85}$$

The *generalized Stokes' theorem*:

$$\int_V d\,\omega = \int_{\partial V} \omega\,,$$

where ω is p-form; $d\omega$ is its exterior derivative, and V is a domain of a p-dimensional manifold. The following two relations are the consequences of the generalized Stokes' theorem

$$\int_V \varphi^\alpha{}_{;\alpha}\, d^4v = \int_{\partial V} \varphi^\alpha\, d\sigma_\alpha\,, \tag{A.86}$$

$$\int_\Sigma \varphi^{\alpha\beta}{}_{;\beta}\, d\sigma_\alpha = \int_{\partial\Sigma} \varphi^{\alpha\beta}\, d\sigma_{\alpha\beta}\,. \tag{A.87}$$

Here ∂V and $\partial\Sigma$ are the boundaries of the four-volume V and three-dimensional hypersurface Σ. We refer to them as to the "Stokes' theorems". Relation (A.86) is usually known as the Gauss divergence theorem.

A.14 Conformal Transformations

Conformal transformations are defined as metric transformations of the type

$$g_{\alpha\beta}(x) = \Omega^2(x)\,\hat{g}_{\alpha\beta}(x)\,. \tag{A.88}$$

A tensor field $A^{\alpha_1\ldots\alpha_n}{}_{\beta_1\ldots\beta_m}$ is a field of conformal weight s if it transforms according to the formula

$$A^{\alpha_1\ldots\alpha_n}{}_{\beta_1\ldots\beta_m} = \Omega^{s-n+m}\,\hat{A}^{\alpha_1\ldots\alpha_n}{}_{\beta_1\ldots\beta_m}\,. \tag{A.89}$$

If $\hat{\nabla}_\gamma$ is the covariant derivative for a metric $\hat{g}_{\alpha\beta}$, then

$$\hat{\nabla}_\gamma A^{\alpha_1\ldots\alpha_n}{}_{\beta_1\ldots\beta_m} = \nabla_\gamma A^{\alpha_1\ldots\alpha_n}{}_{\beta_1\ldots\beta_m} \tag{A.90}$$

$$+ C^{\alpha_1}_{\gamma\sigma}\,A^{\sigma\ldots\alpha_n}{}_{\beta_1\ldots\beta_m} + \ldots - C^{\sigma}_{\gamma\beta_1}\,A^{\alpha_1\ldots\alpha_n}{}_{\sigma\ldots\beta_m} - \ldots\,,$$

where

$$C^\alpha_{\gamma\beta} = C^\alpha_{(\gamma\beta)} = -\,\Omega^{-1}\left[\delta^\alpha_\beta\,\nabla_\gamma\Omega + \delta^\alpha_\gamma\,\nabla_\beta\Omega - g_{\gamma\beta}\,g^{\alpha\sigma}\,\nabla_\sigma\Omega\right]. \tag{A.91}$$

The Weyl tensor $C_{\alpha\beta\gamma}{}^\sigma$ is invariant under conformal transformations, while the curvatures $R_{\alpha\beta\gamma}{}^\sigma$, $R_{\alpha\beta}$, and R are transformed as follows:

$$\hat{R}_{\alpha\beta\gamma}{}^\sigma = R_{\alpha\beta\gamma}{}^\sigma - 2\,\nabla_{[\alpha}C^\sigma_{\beta]\gamma} + 2C^\lambda_{\gamma[\alpha}\,C^\sigma_{\beta]\lambda}\,, \tag{A.92}$$

$$\hat{R}_{\alpha\beta} = R_{\alpha\beta} + 2\,\Omega^{-1}\,\nabla_\alpha\,\nabla_\beta\Omega + \Omega^{-1}\,g_{\alpha\beta}\,\nabla^\sigma\,\nabla_\sigma\Omega \tag{A.93}$$

$$-\,3\,\Omega^{-2}\,g_{\alpha\beta}\,g^{\gamma\sigma}\,\nabla_\gamma\Omega\,\nabla_\sigma\Omega\,,$$

$$\hat{R} = \Omega^2 R + 6\,\Omega\,g^{\gamma\sigma}\,\nabla_\gamma\,\nabla_\sigma\Omega - 12\,g^{\gamma\sigma}\,\nabla_\gamma\Omega\,\nabla_\sigma\Omega\,. \tag{A.94}$$

Generalization to n dimensions:

$$\hat{R} = \Omega^{-2}R + 2(n-1)\Omega\,\Omega_{;\alpha\beta}\,g^{\alpha\beta} - (n-1)n\,\Omega_{;\alpha}\,\Omega_\beta g^{\alpha\beta}. \tag{A.95}$$

For $n = 2$:

$$R = \Omega^{-2}\hat{R} - 2\,\frac{\Omega_{|ab}}{\Omega}\,g^{ab} + 2\,\frac{\Omega_{|a}\,\Omega_{|b}}{\Omega^2}\,g^{ab}\,,$$

where $()_{|a}$ is a two-dimensional covariant derivative.

A.15 Einstein Equations

Einstein action:

$$W[g] = \frac{c^3}{16\pi G}\left(\int_V R\sqrt{-g}\,d^4x - 2\epsilon(\boldsymbol{n})\int_{\partial V} K\sqrt{h}\,d^3y\right). \tag{A.96}$$

Here $\epsilon(\boldsymbol{n}) = n^\mu n_\mu$, and n^μ is the (outward) unit vector normal to the boundary ∂V.

Variations:

$$\delta g^{\alpha\beta} = -\,\delta g_{\mu\nu}\, g^{\mu\alpha}\, g^{\nu\beta}, \tag{A.97}$$

$$\delta|g|^{1/2} = \frac{1}{2}\,\delta g_{\mu\nu}\,|g|^{1/2}\, g^{\mu\nu}, \tag{A.98}$$

$$\delta\Gamma^\lambda_{\alpha\beta} = \frac{1}{2}\, g^{\lambda\epsilon}\left[(\delta g_{\epsilon\alpha})_{;\beta} + (\delta g_{\epsilon\beta})_{;\alpha} - (\delta g_{\alpha\beta})_{;\epsilon}\right], \tag{A.99}$$

$$\begin{aligned}\delta R^\epsilon{}_{\beta\gamma\delta} = -\,\frac{1}{2}\, g^{\epsilon\alpha}\big[&(\delta g_{\alpha\beta})_{;\gamma\delta} - (\delta g_{\alpha\beta})_{;\delta\gamma} \\ &+ (\delta g_{\alpha\gamma})_{;\beta\delta} - (\delta g_{\alpha\delta})_{;\beta\gamma} + (\delta g_{\beta\delta})_{;\alpha\gamma} - (\delta g_{\beta\gamma})_{;\alpha\delta}\big],\end{aligned} \tag{A.100}$$

$$\int dx^4\,|g|^{1/2}\,(\delta R^\alpha{}_{\beta\gamma\delta})\, S_\alpha^{\ \beta\gamma\delta} = \int dx^4\,|g|^{1/2}\,\delta g_{\mu\nu}\left[S^{\mu\nu|\alpha\beta|} - 2S^{\mu\beta|\nu\alpha|}\right]_{;\alpha\beta}, \tag{A.101}$$

$$\int dx^4\,|g|^{1/2}\,(\delta R)\, S = \int dx^4\,|g|^{1/2}\,\delta g_{\mu\nu}\left[S^{;\mu\nu} - g^{\mu\nu} S^{;\alpha}_{;\alpha} - R^{\mu\nu} S\right]. \tag{A.102}$$

Einstein equations:

$$G_{\alpha\beta} = \frac{8\pi G}{c^4}\, T_{\alpha\beta}, \tag{A.103}$$

where G is the Newtonian coupling constant, $T_{\alpha\beta}$ is the energy-momentum tensor,

$$T_{\alpha\beta} = \frac{-2c}{\sqrt{-g}}\,\frac{\delta W_m}{\delta g^{\alpha\beta}}, \tag{A.104}$$

and W_m is the action of matter. For a covariant action W_m, $T^{\alpha\beta}{}_{;\beta} = 0$.

Energy conditions. Let u^μ and l^μ be an arbitrary timelike and null vector field, respectively.

Strong energy dominance condition:

$$T_{\alpha\beta}\, u^\alpha\, u^\beta \geq \frac{1}{2}\, T^\mu_\mu\, u^\alpha\, u_\alpha. \tag{A.105}$$

Energy dominance condition: $T_{\alpha\beta}\, u^\beta$ is a non-spacelike vector.

Weak energy condition :

$$T_{\alpha\beta}\, u^\alpha\, u^\beta \geq 0\,. \tag{A.106}$$

Null energy condition :

$$T_{\alpha\beta}\, l^\alpha\, l^\beta \geq 0\,. \tag{A.107}$$

Averaged null energy condition: For every inextendable null geodesic with affine parameter λ and tangent vector l^α:

$$\int T_{\alpha\beta}\, l^\alpha\, l^\beta\, d\lambda \geq 0\,. \tag{A.108}$$

Each later energy condition is weaker than the previous one.

Conservation laws. Let ξ^μ be a Killing vector field and $\xi^{\mu\nu}$ be a Killing tensor field. If $T^{\mu\nu}$ is a symmetric tensor satisfying the condition $T^{\mu\nu}{}_{;\nu} = 0$ (the energy-momentum tensor), the quantity

$$P_\xi = \int_\Sigma T^{\mu\nu}\, \xi_\mu\, d\sigma_\nu \tag{A.109}$$

is independent of the choice of the complete Cauchy surface Σ.

If p^μ is the momentum of a freely moving particle ($p^\nu\, p^\mu_{;\mu} = 0$), then the quantities

$$P_\xi = \xi^\mu\, p_\mu\,, \quad \text{and} \quad Q_\xi = \xi_{\mu\nu}\, p^\mu\, p^\nu \tag{A.110}$$

are constant along the trajectory of the particle.

Appendix B

Spherically Symmetric Spacetimes

B.1 Spherically Symmetric Geometry

A two-dimensional sphere S^2 is a space of constant curvature, and it possesses a three parameter group $O(3)$ of isometries. The metric on S^2 can be written in the form[1]

$$d\omega^2 \equiv \omega_{XY}\, d\zeta^X\, d\zeta^Y = d\theta^2 + \sin^2\theta\, d\phi^2. \tag{B.1}$$

The most general Killing vector is a linear combination of the following three Killing vectors

$$\xi^{\mu}_{(1)}\partial_\mu = \partial_\phi\,, \tag{B.2}$$

$$\xi^{\mu}_{(2)}\partial_\mu = -(\cos\phi\,\partial_\theta - \cot\theta\,\sin\phi\,\partial_\phi)\,, \tag{B.3}$$

$$\xi^{\mu}_{(3)}\partial_\mu = \sin\phi\,\partial_\theta + \cot\theta\,\cos\phi\,\partial_\phi\,. \tag{B.4}$$

These are the usual angular momentum operators L_z, L_x and L_y which are the generators of the rotation group. Besides the continuous isometries generated by the Killing vectors, the metric (B.1) possesses discrete symmetries:

$$\phi \to -\phi\,, \qquad \theta \to \pi - \theta\,. \tag{B.5}$$

Let a, a^X, and a^X_Y be a scalar, vector, and tensor functions on S^2 which are invariant under all the isometries of a two-sphere. Then

$$a = \text{const}\,, \quad a^X = 0\,, \quad a^X_Y = \text{const}\ \delta^X_Y. \tag{B.6}$$

Spacetime is said to be spherically symmetric if it can be represented as a two-dimensional set of two-spheres and the 4-metric is invariant under the complete group

[1] Indices $X, Y, \ldots$ take values 2 and 3, while indices $A, B, \ldots$ take values 0 and 1

of isometries of two-dimensional spheres. The metric for a spherically symmetric spacetime can be written in the form

$$ds^2 \equiv g_{\mu\nu}\, dx^\mu\, dx^\nu = d\gamma^2 + r^2\, d\omega^2 \,. \tag{B.7}$$

Here $d\omega^2$ is the metric (B.1) on a unit sphere, and $d\gamma^2$ is a metric on a two-dimensional manifold M^2. If we denote by x^A $(A = 0, 1)$ coordinates on M^2, then this metric reads

$$d\gamma^2 = \gamma_{AB}(x)\, dx^A\, dx^B \,. \tag{B.8}$$

The "radius" $r = r(x)$ which enters metric (B.7) is a scalar function on M^2.

In coordinates (x^A, ζ^X) a tensor G^μ_ν in a spherically symmetric spacetime which obeys spherical symmetry

$$\mathcal{L}_{\xi_i} G^\mu_\nu = 0\,, \quad i = 1, 2, 3$$

has the following non-vanishing components

$$G^\mu_\nu = \begin{pmatrix} G^0_0 & G^0_1 & 0 & 0 \\ G^0_1 & G^1_1 & 0 & 0 \\ 0 & 0 & G^2_2 & 0 \\ 0 & 0 & 0 & G^2_2 \end{pmatrix}, \tag{B.9}$$

where $\partial G^\mu_\nu / \partial \zeta^X = 0$. If tensor G^μ_ν satisfy the conservation law $G^\mu_{\nu;\mu} = 0$, then its components obey the relation

$$G^A_{B|A} + \frac{2 r_{|A}}{r}\, G^A_B - \frac{2 r_{|B}}{r}\, G^2_2 = 0 \,.$$

B.2 Reduced Action

The variation of the Einstein-Hilbert action is

$$\delta W[\,g\,] = -\frac{1}{16\pi} \int_V d^4x\, \sqrt{-g}\, G^{\mu\nu}\, \delta g_{\mu\nu}\,, \tag{B.10}$$

where $G^{\mu\nu}$ is the Einstein tensor. The variation of the surface term vanishes for variations preserving the 3-geometry of the boundary ∂V. If we assume the spherical-symmetry ansatz for metric (B.7) and for the Einstein tensor (B.9), then

$$\delta W[\,g\,] = -\frac{1}{16\pi} \int d^2\zeta\, \sqrt{\omega} \int d^2x\, \sqrt{-\gamma}\, r^2 \left[G^{AB}\, \delta g_{AB} + 2\, G^{\theta\theta}\, \delta g_{\theta\theta} \right]. \tag{B.11}$$

The components of G^μ_ν which enter this relation do not depend on ζ^X. That is why in order to obtain the non-vanishing components of the Einstein tensor in a spherically symmetric spacetime, it is sufficient to consider variations of the metric $\delta g_{\mu\nu}$ which do not depend on ζ^X.

Denote by $W_{\rm sph}[\gamma_{AB},r]$ a reduced action, i.e., the action $W[g]$ calculated for a spherically symmetric metric (B.7). Using the relations

$$\sqrt{g}\,R \equiv \sqrt{-\gamma}\,\sqrt{\omega}\,\left[{}^{(2)}\!R[\omega] + r^2\,{}^{(2)}\!R[\gamma]\; - 4r\,{}^{(2)}\Box_\gamma r - 2\,\gamma^{AB}\,r_{|A}\,r_{|B}\right], \tag{B.12}$$

$$\int d^2\zeta\,\sqrt{\omega} = 4\pi, \tag{B.13}$$

we get

$$W_{\rm sph}[\gamma,r] \equiv \frac{1}{4}\int_{M^2} d^2x\,\sqrt{-\gamma}\,\left[{}^{(2)}\!R[\gamma]\,r^2 + 2\,\gamma^{AB}\,r_{|A}\,r_{|B} + 2\right]$$
$$-\frac{1}{2}\int_{\partial M^2} r^2\,k\,dl. \tag{B.14}$$

Here ${}^{(2)}\!R[\gamma]$ is the scalar curvature of the two-dimensional space with metric $d\gamma^2$; $(\;)_{|A}$ is a covariant derivative, and ${}^{(2)}\Box_\gamma$ is the D'Alembert operator for this metric. We denote by k the extrinsic curvature of ∂M^2, and by dl a proper distance along ∂M^2. We also note that ${}^{(2)}\!R[\omega] = 2$. Using (B.11), we get

$$\delta W_{\rm sph}[\gamma_{AB},r] = -\frac{1}{4}\int d^2x\,\sqrt{-\gamma}\,r^2\,\left[G^{AB}\,\delta\gamma_{AB}(x) + 2\,r\,G^{XY}\,\omega_{XY}\,\delta r(x)\right]. \tag{B.15}$$

The variation of the reduced action (B.14) gives the complete set of Einstein equations for spherically symmetric spacetimes:

$$r^2\,G_{AB} \equiv -2\,r\,r_{|AB} + \gamma_{AB}\left(-1 + \nabla r\cdot\nabla r + 2\,r\,{}^{(2)}\Box r\right) = 8\pi\,r^2\,T_{AB}, \tag{B.16}$$

$$G^\theta_\theta = G^\phi_\phi \equiv r^{-1}\,{}^{(2)}\Box r - \frac{1}{2}\,{}^{(2)}\!R[\gamma] = 8\pi\,T^\theta_\theta. \tag{B.17}$$

Here all the operations are defined with respect to the metric $d\gamma^2$, and $T_{\mu\nu}$ are the components of the stress-energy tensor which obeys the spherical symmetry condition.

B.3 Generalized Birkhoff's Theorem

We denote by e_{AB} and e^{AB} the two-dimensional antisymmetric tensors

$$e_{AB} = \sqrt{-\gamma}\,\epsilon_{AB}, \qquad e^{AB} = -\frac{1}{\sqrt{-\gamma}}\,\epsilon^{AB}, \tag{B.18}$$

where ϵ_{AB} and ϵ^{AB} are antisymmetric symbols ($\epsilon_{01} = 1$, $\epsilon^{01} = 1$).

Lemma 1. *Let f be a function on a two-dimensional manifold with a metric γ_{AB} which obeys the relation*

$$f_{|AB} = F\,\gamma_{AB} \tag{B.19}$$

for some function F. Then $\xi^A \equiv e^{AB} f_{|B}$ is a Killing vector.

Lemma 2. *Let ξ be a Killing vector on a two-dimensional manifold with a metric γ_{AB}. Then the equation*

$$f_{|A} = e_{AB}\,\xi^B \tag{B.20}$$

has a solution f which obeys the relation

$$f_{|AB} = F\,\gamma_{AB} \qquad F = \frac{1}{2} f^{|C}_{|C}. \tag{B.21}$$

Generalized Birkhoff's theorem: *If the stress-energy tensor generating a spherically symmetric gravitational field obeys the condition*

$$T_{AB} = \frac{1}{2} T^C_C\,\gamma_{AB}\,, \tag{B.22}$$

then the corresponding solution of the Einstein equations possesses an additional Killing vector field.

Proof: If (B.22) is satisfied, then equation (B.16) implies that $r_{|AB} \sim \gamma_{AB}$, and hence $\xi^A = e^{AB}\,r_{|B}$ is a Killing vector field for the two-dimensional metric γ_{AB}. Moreover, $\xi^A\,\partial_A r = 0$. Thus, $\xi^\mu = \delta^\mu_A\,\xi^A$ is a Killing vector field for the four-dimensional metric (B.7), which is evidently linearly independent of the Killing vectors (B.2)–(B.4).

B.4 Spherically Symmetric Vacuum Solutions

B.4.1 Schwarzschild metric

Frobenius theorem: *If a vector field ξ obeys the relation*

$$\xi_{[\alpha}\,\xi_{\beta;\gamma]} = 0\,, \tag{B.23}$$

then it is "hypersurface-orthogonal", i.e., there exists such a pair of functions t and β that

$$\xi_\alpha = \beta\,t_\alpha\,. \tag{B.24}$$

Theorem: *A spherically symmetric vacuum solution of Einstein equations is determined by one essential constant (mass M), and in the regions where $(\nabla r)^2 \neq 0$ it can be written in the form*

$$ds^2 = -\,g\,dt^2 + \frac{dr^2}{g} + r^2\,d\omega^2\,, \quad g = 1 - \frac{2M}{r}\,. \tag{B.25}$$

This solution is known as the *Schwarzschild metric.*

Proof: According to the Birkhoff's theorem,

$$\xi^A = e^{AB}\, r_{|B} \tag{B.26}$$

is a Killing vector. Its norm is

$$\xi^A\,\xi_A = e^{AB}\,e_{AC}\,r_{|B}\,r^{|C} = -\,r_{|A}\,r^{|A}\,.$$

(We used the relation $e^{AB}\,e_{AC} = -\,\delta^B_C$.) An antisymmetric tensor with three indices vanishes identically in two dimensions. Hence, (B.23) is valid for ξ_A, and according to the Frobenius theorem, vector ξ_A can be represented in the form

$$\xi_A = \beta\, t_{,A}\,. \tag{B.27}$$

In a region where $(\nabla r)^2 \neq 0$, one can use t and r as coordinates. Using (B.26) and (B.27), one gets

$$\gamma^{tr} = \gamma^{AB}\,t_{|A}\,r_{|B} = \beta^{-1}\,\gamma^{AB}\,e_{AC}\,r^{|C}\,r_{|B} = \beta^{-1}\,e^{BC}\,r_{|B}\,r_{|C} = 0\,.$$

Equation (B.27) implies

$$\xi^A\,\xi_A = \beta^2\,\gamma^{tt}\,.$$

Hence, in (t,r)-coordinates metric γ_{AB} is of the form

$$d\gamma^2 = -\,\frac{\beta^2}{g}\,dt^2 + \frac{dr^2}{g}\,, \tag{B.28}$$

where $g \equiv g^{rr}$. The Killing equation $\xi_{(0|1)} = 0$ for the metric (B.28) is

$$2\,\xi_{(0|1)} \equiv 2\,\xi_{(t,r)} - 2\,\Gamma^A_{tr}\,\xi_A = \beta' - \beta\,\frac{(g^{-1}\,\beta^2)'}{(g^{-1}\,\beta^2)} = 0\,.$$

Here $(\ \ )' = d(\ \ )/dr$. Solving this equation, one gets $\beta = c\,g$, where c is an integration constant. Using rescaling freedom $t \to C\,t$, we put $c = -1$. (A minus sign is chosen since ξ^A is a future directed vector.) The other Killing equations are satisfied identically. Finally, we have

$$ds^2 = -\,g\,dt^2 + \frac{dr^2}{g} + r^2\,d\omega^2\,. \tag{B.29}$$

Integration of the equation $G^A_A = 0$,

$$r\,{}^{(2)}\Box r + (\nabla r)^2 - 1 = 0\,, \tag{B.30}$$

allows one to find the function g. Written in (t,r)-coordinates this equation reads

$$r\,g' + g - 1 = 0\,. \tag{B.31}$$

The general solution of this equation is

$$g = 1 - \frac{2M}{r}\,, \tag{B.32}$$

where M is an integration constant. It is easy to verify that the other Einstein equations are satisfied identically.

B.4.2 Scaling properties

Consider a *scaling transformation*, i.e., a conformal transformation $ds^2 = \Omega^2\, d\tilde{s}^2$ with a constant conformal factor Ω. The gravitational action $W[\,g\,]$ transforms as $W[\,g\,] = \Omega^2\, W[\,\tilde{g}\,]$. This means that if g is a solution of the vacuum Einstein equations, then $\Omega\, g$ is also a solution. We show now that a spherically symmetric vacuum solution of the Einstein equations is unique up to a scaling transformation. For this purpose, we rewrite the Schwarzschild metric in the form

$$ds^2 = r_g^2\, d\tilde{s}^2\,, \tag{B.33}$$

$$d\tilde{s}^2 = -\,g\,d\eta^2 + \frac{dx^2}{g} + x^2\,d\omega^2\,, \tag{B.34}$$

where $g = 1 - 1/x$, $\eta = t/r_g$, and $x = r/r_g$ are dimensionless coordinates, and $r_g = 2M (\equiv 2GM/c^2)$ is the *gravitational radius.*

B.5 Kruskal Metric

B.5.1 Derivation of Kruskal metric

In order to obtain a solution which covers the complete vacuum spherically symmetric spacetime, we use coordinate-fixing conditions $g_{00} = g_{11} = 0$ and write the metric as

$$ds^2 = 2\,B\,dU\,dV + r^2\,d\omega^2\,. \tag{B.35}$$

It is convenient to use the dimensionless form of this metric:

$$d\tilde{s}^2 = r_g^{-2}\,ds^2 = 2\,b\,dU\,dV + x^2\,d\omega^2\,. \tag{B.36}$$

The non-vanishing Christoffel symbols calculated for the metric (B.36) are

$$\Gamma^U_{UU} = \frac{b_{,U}}{b}\,, \quad \Gamma^V_{VV} = \frac{b_{,V}}{b}\,, \tag{B.37}$$

and the vacuum Einstein equations (B.16) written in dimensionless form are

$$x_{|UU} = x_{|VV} = 0\,, \tag{B.38}$$

$$x\,{}^{(2)}\tilde{\Box}\,x + (\tilde{\nabla}x)^2 - 1 = 0\,. \tag{B.39}$$

Equations (B.38), written in the explicit form

$$\partial^2_U x - \frac{b_{,U}}{b}\,\partial_U x = 0\,, \quad \partial^2_V x - \frac{b_{,V}}{b}\,\partial_V x = 0\,, \tag{B.40}$$

allow first integrals

$$\partial_U x = \alpha(V)\, b\,, \qquad \partial_V x = \beta(U)\, b\,. \tag{B.41}$$

Using the coordinate freedom $U \to \tilde{U} = F(U)$ and $V \to \tilde{V} = G(V)$, one can always put $\alpha = \frac{1}{2} V$ and $\beta = \frac{1}{2} U$. Then equations (B.41) imply

$$\dot{x} = \frac{1}{2} b, \quad x = x(z), \quad b = b(z), \tag{B.42}$$

where $z = UV$, and a dot means a derivative with respect to z. Equation (B.39) now reads

$$x\left(1 + \frac{x\,\ddot{x}}{\dot{x}}\right) + z\,\dot{x} - 1 = 0\,. \tag{B.43}$$

Define $Y = (z/z') - 1$, where $(\quad)' = d(\quad)/dx$, then (B.43) implies

$$x\,Y' + Y = 0\,. \tag{B.44}$$

We have $Y = c/x$. For $c = 0$, the metric (B.36) is flat. If $c \neq 0$, one can use a scaling transformation $x \to C\,x$ to put $c = -1$. This gives

$$\frac{z'}{z} = \frac{x}{x-1}\,. \tag{B.45}$$

The sign of z is affected by the $U \to -U$ transformation. We choose the sign of z so that the solution of the above equation is

$$z = -(x-1)\exp(x-1)\,. \tag{B.46}$$

Finally, we get

$$-\,UV = (x-1)\exp(x-1)\,, \quad b = -\frac{2}{x}\exp[-(x-1)]\,. \tag{B.47}$$

This solution written for r and B gives

$$-\,UV = \left(\frac{r}{2M} - 1\right)\exp\left(\frac{r}{2M} - 1\right), \tag{B.48}$$

$$B = -\frac{16M^3}{r}\exp\left[-\left(\frac{r}{2M} - 1\right)\right]. \tag{B.49}$$

B.5.2 Relation between Kruskal and Schwarzschild metrics

The Killing vector $\xi_{(t)}$ in Kruskal coordinates is

$$\xi^{\mu}_{(t)}\,\partial_\mu = (1/4M)(V\,\partial_V - U\,\partial_U)\,. \tag{B.50}$$

In different regions of the spacetime, Kruskal coordinates (U, V) are related to Schwarz schild coordinates (t, r) as follows:
In R_+ : $U > 0, V < 0$

$$\begin{cases} U = -\left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) - \frac{t}{4M}\right], \\ V = \left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) + \frac{t}{4M}\right]. \end{cases} \tag{B.51}$$

In R_- : $U > 0, V < 0$

$$\begin{cases} U = \left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) - \frac{t}{4M}\right], \\ V = -\left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) + \frac{t}{4M}\right]. \end{cases} \tag{B.52}$$

In T_+ : $U > 0, V < 0$

$$\begin{cases} U = \left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) - \frac{t}{4M}\right], \\ V = \left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) + \frac{t}{4M}\right]. \end{cases} \tag{B.53}$$

In T_- : $U > 0, V < 0$

$$\begin{cases} U = -\left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) - \frac{t}{4M}\right], \\ V = -\left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left[\frac{1}{2}\left(\frac{r}{2M} - 1\right) + \frac{t}{4M}\right]. \end{cases} \tag{B.54}$$

The spherically symmetric vacuum spacetime is invariant under the following discrete symmetries

$$\mathbf{I}: \quad U \to -U, \quad V \to -V. \tag{B.55}$$

$$\mathbf{T}: \quad U \to -V, \quad V \to -U. \tag{B.56}$$

$$\mathbf{P}: \quad U \to V, \quad V \to U. \tag{B.57}$$

B.5.3 Kruskal spacetime as maximal analytical continuation of the Schwarzschild metric

Define $r_* \equiv r - 2M + 2M \ln|(r-2M)/2M|$ (called the *tortoise coordinate*), and define a *retarded null coordinate* u and an *advanced null coordinate* v by the relations $u = t - r_*$ and $v = t + r_*$. The Schwarzschild metric can be rewritten as

$$ds^2 = -g\,du^2 - 2\,du\,dr + r^2\,d\omega^2 \tag{B.58}$$

$$= -g\,dv^2 + 2\,dv\,dr + r^2\,d\omega^2 = -g\,du\,dv + r^2\,d\omega^2\,, \tag{B.59}$$

$g = 1 - 2M/r$. In the last expression it is assumed that r is a function of u and v defined implicitly by the relation $r_* = \frac{1}{2}(v-u)$. By introducing new coordinates

$$U = -\exp\left(\frac{-u}{4M}\right), \qquad V = \exp\left(\frac{v}{4M}\right), \tag{B.60}$$

one can rewrite the Schwarzschild metric in R_+ in the Kruskal form and use it for the construction of the maximal analytical continuation.

B.5.4 Einstein-Rosen bridge

The internal geometry of the T-invariant surface $\Sigma : U = -V$ is

$$dl^2 = \frac{dr^2}{1-2M/r} + r^2\,d\omega^2 = \Omega^4\,dl_0^2\,, \tag{B.61}$$

$$\Omega = 1 + \frac{M}{2\rho}, \quad dl_0^2 = d\rho^2 + \rho^2(d\theta^2 + \sin^2\theta\,d\phi^2)\,, \tag{B.62}$$

where ρ is defined by $r = \rho(1 + M/2\rho)^2$. The geometry of a two-dimensional section $\theta = \pi/2$ can be realized as a surface imbedded in a flat three-dimensional space:

$$dL^2 = dz^2 + dr^2 + r^2\,d\phi^2 = dr^2(1 + z'^2) + r^2\,d\phi^2\,, \tag{B.63}$$

where $z = \pm 2\sqrt{2M(r-2M)}$.

B.6 Tolman Solution

A spherically symmetric solution of the Einstein equations for a spherical distribution of dust matter (i.e., matter without pressure $p = 0$) was obtained by Tolman (1934). To derive this solution, we use comoving coordinates ($\tau = cT, R$) in which $T_{AB} = \rho c^2 \delta_A^0 \delta_B^0$ (ρ is the mass density), and the two-dimensional part $d\gamma^2$ of the spherically symmetric metric takes the form

$$d\gamma^2 = -d\tau^2 + B\,dR^2\,. \tag{B.64}$$

Here $B = B(\tau, R)$ and $r = r(\tau, R)$. Non-vanishing Christoffel symbols Γ^A_{BC} calculated for this metric are

$$\Gamma^1_{11} = B'/2B\,, \quad \Gamma^0_{11} = \dot{B}/2\,, \quad \Gamma^1_{01} = \dot{B}/2B\,. \tag{B.65}$$

Here $(\ )' = \partial_R$ and $(\ )^{\cdot} = \partial_\tau$. The integration of the equation (B.16) for the value of indices $A = 0$, $B = 1$ gives

$$B = \frac{r'^2}{1+f}\,, \tag{B.66}$$

where $f = f(R)$ is an "integration constant" obeying the condition $1 + f > 0$. The integration of the same equation for $A = 1$, $B = 1$ allows one to get

$$\dot{r}^2 = f + \frac{F}{r}\,, \tag{B.67}$$

where $F = F(R)$ is another "integration constant". By using these relations and equation (B.16) for $A = 0$ and $B = 0$, one gets

$$\frac{F'}{r'\,r^2} = 8\,\pi\,\epsilon\,, \tag{B.68}$$

where $\epsilon = G\rho/c^2$. Equation (B.67) can be integrated in parametric form.

For $f > 0$:

$$r = \frac{F}{2f}(\cosh\eta - 1)\,, \quad \tau - \tau_0(R) = \frac{F}{2f^{3/2}}(\sinh\eta - \eta)\,. \tag{B.69}$$

For $f < 0$:

$$r = \frac{F}{-2f}(1 - \cos\eta)\,, \quad \tau - \tau_0(R) = \frac{F}{2(-f)^{3/2}}(\eta - \sin\eta)\,. \tag{B.70}$$

An ambiguity $R \to \tilde{R} = \tilde{R}(R)$ in the choice of a comoving coordinate reduces three arbitrary functions $f(R)$, $F(R)$, and $\tau_0(R)$ to two invariant ones which describe the initial density distribution of dust and its radial velocity. Now we consider special cases.

Homogeneous Dust Ball ($\epsilon = \text{const}$)

$$f = -\sin^2 R\,, \quad F = a\sin^3 R\,, \quad \tau_0 = 0\,, \quad 0 \le R \le R_1\,, \tag{B.71}$$

$$r = \frac{a}{2}\sin R\,(1 - \cos\eta)\,, \quad \tau = \frac{a}{2}(\eta - \sin\eta)\,, \tag{B.72}$$

$$\epsilon = \frac{3}{\pi\,a^2\,(1 - \cos\eta)^3}\,. \tag{B.73}$$

Here a is a constant.

This solution describes part of the Robertson-Walker universe filled with the dust. At the moment of maximal expansion $\eta = \pi$ and $r = a \sin R$. For $R_1 = \pi$, we get a complete closed world. For $R_1 < \pi$, we have part of this world. The evolution of the boundary sphere is described by the equation $r = \frac{1}{2} a \sin R_1 (1 - \cos \eta)$.

Empty Spacetime ($\epsilon = 0$)

$$F = r_g = \text{const}, \quad f = -\frac{1}{R^2+1}, \quad \tau_0 = 0. \tag{B.74}$$

For this choice we get

$$r = \frac{r_g}{2}(R^2+1)(1-\cos\eta), \tag{B.75}$$

$$\tau = \frac{r_g}{2}(R^2+1)^{3/2}(\eta - \sin\eta). \tag{B.76}$$

This is a freely falling coordinate frame for the Schwarzschild geometry [Novikov (1962)]. It covers the complete spacetime, and a constant r_g coincides with the gravitational radius.

Note that the coordinate R in the above solutions allows a transformation $R \to R + b$, with b =const, which does not change the form of the metric. In order to sew the solution for the homogeneous dust ball to the empty space solution, we first use this ambiguity to rewrite f in equation (B.74) in the form $f = -[(R+b)^2+1]^{-1}$. The requirement that the functions $f(R)$ and $F(R)$ in both solutions (for dust and vacuum) takes the same value at the boundary $R = R_1$ gives

$$r_g = a \sin^3 R_1, \quad b = \cot R_1 - R_1. \tag{B.77}$$

One can verify that the other jump conditions on the boundary are identically satisfied. The proper mass M_* of the dust matter is $M_* = (3a/4)(R_1 - \frac{1}{2}\sin(2R_1))$.

Appendix C

Rindler Frame in Minkowski Spacetime

C.1 Uniformly Accelerated Motion

Many aspects of the definition and properties of black holes can be clarified and easily understood if we consider physical processes in a flat Minkowski spacetime (that is, without a gravitational field), but from the point of view of a uniformly accelerated observer. The deep reason for the possibility of clarifying physics of black holes in this way is the validity of the *equivalence principle.* According to this principle, the physical laws in a local reference frame in a gravitational field are equivalent to the physical laws in a uniformly accelerated frame. Consider a small region in the vicinity of the event horizon and use a frame which is at rest in this region. Then the equivalence principle implies that this frame has the same properties as a uniformly accelerated frame in the Minkowski spacetime. This furnishes a powerful tool for an investigation of local effects in the close vicinity of the event horizon.

In this appendix, we describe a rigid non-rotating uniformly accelerated reference frame in the Minkowski flat spacetime. In order to construct this frame, we first consider a test particle moving along a straight line. Denote by $(X^0 \equiv cT, X^1, X^2, X^3)$ the standard Cartesian coordinates in the Minkowski spacetime. We assume that the particle is moving along the X^1-axis and has constant acceleration measured by a comoving observer.[1] This means that the four-acceleration a^μ is directed along X^1 and its value $a \equiv \sqrt{a_\mu a^\mu}$ remains constant:

$$a = \text{const}. \tag{C.1}$$

The trajectory of the particle can be written in the following parametric form, with

[1] An assumption that a test particle moves with constant acceleration for an infinitely long time is physically unrealistic. For such a motion one needs to provide a particle with an infinite amount of energy. Nevertheless, the consideration of such motion is useful, and it helps one to understand better general features of a static gravitational field

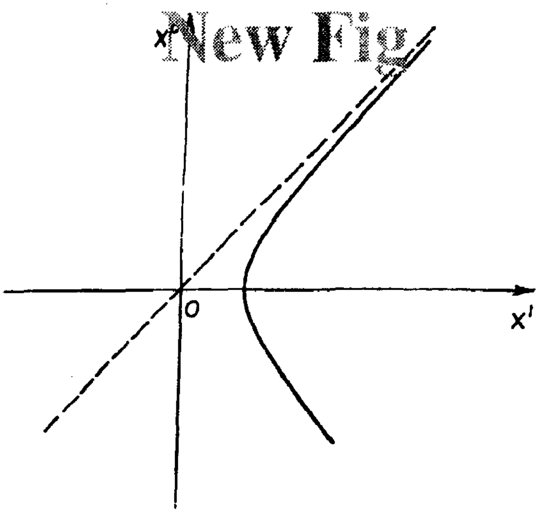


Figure C.1: A uniformly accelerated particle in Minkowski spacetime.

τ being a parameter,

$$X^0 = l_a \sinh(c\,\tau/l_a)\,, \quad X^1 = l_a \cosh(c\,\tau/l_a)\,, \quad X^2 = X_0^2\,, \quad X^3 = X_0^3\,. \quad \text{(C.2)}$$

Here $l_a = c^2/a$, and $a = (a_\mu\, a^\mu)^{1/2}$ is the value of the acceleration. The four-velocity u^μ and four-acceleration a^μ for world line (C.2) are of the form

$$u^\mu = (c\cosh(c\,\tau/l_a), c\sinh(c\,\tau/l_a), 0, 0)\,, \quad \text{(C.3)}$$

$$a^\mu = (a\sinh(c\,\tau/l_a), a\cosh(c\,\tau/l_a), 0, 0)\,. \quad \text{(C.4)}$$

These formulas show that

$$u_\mu\, u^\mu = -c^2, \quad a_\mu\, a^\mu = a^2. \quad \text{(C.5)}$$

Hence, τ is the proper time along the world line of the accelerated particle.

Curve (C.2) is a hyperbola in the (X^0-X^1)-plane (see Figure C.1). For simplicity, we choose $X_0^2 = X_0^3 = 0$. A well-known example of a uniformly accelerated motion is a motion of an electrically charged (with charge e) particle in an external homogeneous electric field directed along X^1. If the particle is initially at rest, its equation of motion coincides with (C.2). The acceleration of such a particle is $a = eE/m$, where E is the field strength, and m is its rest mass.

The null lines $X^0 = \pm\, X^1$ are asymptotes of the hyperbola. One of them ($X^0 = X^1$) coincides with the world line of a photon moving along the X^1-axis in the same direction as the accelerated particle. Quite remarkably, it never reaches the particle.

A particle moving along the X^1-axis with another (larger) value of the acceleration a and crossing $X^1 = c^2/a$ at $T = 0$ can be represented on the same picture as a hyperbola with the same asymptotes but with a vertex lying closer to the origin $X^0 = X^1 = 0$. In the limit of an infinite acceleration, the vertex tends to the origin.

C.2 Rindler Frame

We describe now a special rigid reference frame known as a *Rindler frame*,[2] constructed from world lines of particles moving with uniform acceleration along X^1-axis. Consider a family of trajectories which have the same values of coordinates X^2 and X^3. If the frame is rigid, then two trajectories of the family with different values of X^1 must have different accelerations at a given moment of time X^0. Indeed, if their accelerations are equal, then the distance between the trajectories measured in the inertial Cartesian frame is constant. Because of the Lorentz contraction, the distance measured in the moving frame will depend on the velocity of the accelerated frame at the corresponding moment of time. This time dependence of the velocity of uniformly accelerated observers results in a time dependence of the proper distance between the world lines with the same value of acceleration. In order for the frame to be rigid, the acceleration of its world lines crossing X^0 at different values X^1 must be different. Namely, the lager is the value X^1, the smaller is the acceleration.

To describe the Rindler frame, we introduce new coordinates $x^\mu = (t, \rho, x^2, x^3)$

$$\rho = \sqrt{(X^1)^2 - (X^0)^2}\,, \quad t = \frac{l_a}{c} \arctan\frac{X^0}{X^1}\,, \quad x^2 = X^2, \quad x^3 = X^3. \tag{C.6}$$

Here $l_a \equiv c^2/a$ is fixed. For fixed values of ρ, x^2, and x^3, a coordinate line of t coincides with the world line of a uniformly accelerated observer. The Rindler reference frame is formed by the world lines of these observers. The Minkowski line element

$$ds^2 = -(dX^0)^2 + (dX^1)^2 + (dX^2)^2 + (dX^3)^2 \tag{C.7}$$

in these new (Rindler) coordinates x^μ takes the form

$$ds^2 = -\frac{\rho^2}{l_a^2}\, c^2\, dt^2 + d\rho^2 + (dx^2)^2 + (dx^3)^2. \tag{C.8}$$

It is evident that the geometry of the comoving space in this Rindler frame is time independent. For a world line $\rho, x^2, x^3 =$ const, the acceleration is $a(\rho) \equiv c^2/\rho$, and $\tau = \rho t/l_a$ is a proper time. Some of these lines are shown in Figure C.2. The parameter t coincides with the proper time for the observer with $\rho = l_a$. We refer to the world line with coordinates $(l_a, 0, 0)$ as to the origin of the reference frame.

[2] Points which form a reference frame are moving along timelike curves. One may assume that these curves are world lines of observers. A frame is called *rigid* if the distance between any two lines of this congruence measured in this frame is constant.

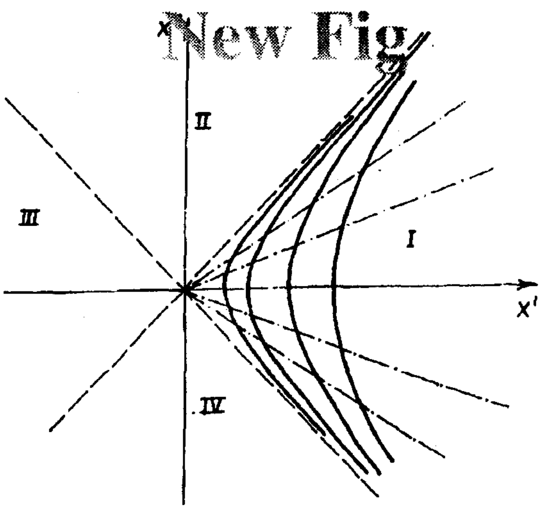


Figure C.2: A Rindler frame. World lines of Rindler observers (ρ=const) are shown by solid lines. Dashed and dotted lines represent H_t planes (t =const).

All the world lines of Rindler observers are hyperbolas with the same asymptotes $X^0 = \pm X^1$. A three-dimensional surface t = const describes the events which are simultaneous from the point of view of the comoving (Rindler) observers. This surface is a hyperplane H_t, and it obeys the equation

$$\frac{X^0}{X^1} = \tanh(ct/l_a) = \text{const}. \tag{C.9}$$

Note that all hyperplanes H_t cross one another at the same two-dimensional plane $X^0 = X^1 = 0$. The hyperplanes H_t are shown in Figure C.2 (where X^2 and X^3 dimensions are suppressed). It is evident that the proper distance between world lines of Rindler observes is time independent, so that the Rindler frame is rigid.

An important property of the Rindler frame is that it cannot be extended in the comoving space beyond $\rho = 0$ to negative values of ρ. Formally, this follows from the expression for Rindler metric (C.8) in which the metric coefficient $g_{00} = -\rho^2/l_a^2$ vanishes at $\rho = 0$. This means that the rate of proper time τ

$$d\tau = \sqrt{-g_{00}}\, dt = \frac{\rho}{l_a}\, dt \tag{C.10}$$

is infinitely slowed down as $\rho \to 0$. This indicates that there is physical peculiarity of the Rindler reference frame at $\rho = 0$. The physical reason for this peculiarity is the following. The condition of rigidity requires that the acceleration of the Rindler observer $a = c^2/\rho$ grows infinitely as $\rho \to 0$, and the reference frame fails to be rigid beyond this limit. This property of the Rindler frame resembles a similar property of a static reference frame in the spacetime of a non-rotating black hole, where the rigid frame is restricted to the region $r > r_g$. This analogy goes much deeper. As it is seen from Figure C.2, the null surface $X^0 = X^1$ which corresponds to $\rho = 0$ in Rindler coordinates is the boundary of events which can be seen by Rindler observers. The events located beyond this surface will never be seen by a uniformly accelerated observer. Thus, the plane $X^0 = X^1$ plays the role of the *event horizon* in the Rindler spacetime.

C.3 Light and Particle Propagation

The coefficients of metric (C.8) do not depend on time t. The corresponding symmetry transformations are generated by the Killing vector field

$$\xi^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial(ct)}. \tag{C.11}$$

The world lines of Rindler observers (the coordinate lines of t) coincide with the trajectories of ξ. The four-velocity u^μ of a Rindler observer is

$$u^\mu = c\xi^\mu / |\xi_\alpha \xi_\alpha|^{1/2}, \tag{C.12}$$

while his acceleration is

$$a_\mu \equiv u^\nu \nabla_\nu u_\mu = \frac{1}{2} c^2 \nabla_\mu \ln|\xi^2|. \tag{C.13}$$

For our case

$$\xi^2 = -\frac{\rho^2}{l^2}, \quad a^\mu = -\frac{c^2}{\rho}\delta^\mu_\rho. \tag{C.14}$$

A Rindler observer "feels" the inertial force of acceleration. It can be interpreted as the presence of a static gravitational field. In a static gravitational field the quantity $\alpha \equiv |\xi^2|^{1/2}$ is a *redshift factor*. To demonstrate this, consider a monochromatic wave. Its time dependence is of the form $\exp(-i\omega t)$. The proper frequency of the wave measured by an observer at a point p is

$$\omega_1 \equiv \frac{d(\text{phase})}{d(\text{proper time})} = \omega\frac{dt}{d\tau} = \frac{\omega}{\alpha(p)}. \tag{C.15}$$

This means that the frequency ω_e of the radiation emitted at a point p_e and the frequency ω_o observed at a point p_o are related as follows

$$\frac{\omega_e}{\omega_o} = \frac{\alpha_o}{\alpha_e}. \tag{C.16}$$

Any radiation emitted near the horizon has the redshift proportional to α. The redshift factor α vanishes at the horizon. Hence, the horizon is a surface of infinite redshift. This is a generic property of static spacetimes in which the event horizon always coincides with the Killing horizon (where $\xi^2 = 0$) (see Section 6.3).

Metric (C.8) was obtained by rewriting the flat spacetime metric in the Rindler coordinates. This metric is evidently a solution of the vacuum Einstein equations. Suppose that one has actually found it by solving the Einstein equations and that he is not clever enough to identify the solution with a part of the Minkowski spacetime. How can one find out that the horizon $\rho = 0$ of the solution is a regular surface and the singularity of the metric is a pure coordinate effect? First of all, it is quite easy to discover that the region ($\rho > 0, x^2, x^3$) covered by the coordinates (map) is geodesically incomplete. For example, consider a test particle freely moving in the ρ-direction. The equation of its motion[3] is ($\beta < 1$)

$$\rho = \frac{\rho_0}{\cosh(c\,t/l_a) + \beta \sinh(c\,t/l_a)}\,. \tag{C.17}$$

In the Rindler frame the motion of the particle begins at $\rho = 0$ in the infinite past $t = -\infty$. The particle propagates from the horizon until it reaches the point $\rho = \rho_0/\sqrt{1-\beta^2}$. After this the particle falls again to the horizon. It takes an infinite coordinate time t to reach the horizon. On the other hand, the proper time of this evolution remains finite. The proper time τ_{12} between any two points along the trajectory is

$$c\,\tau_{12} = \rho_0\sqrt{1-\beta^2}\left[\frac{1}{\coth(c\,t_2/l_a) + \beta} - \frac{1}{\coth(c\,t_1/l_a) + \beta}\right]. \tag{C.18}$$

The proper time for the complete motion as described in Rindler coordinates is

$$c\,\tau_{\rm tot} = \frac{2\rho_0}{\sqrt{1-\beta^2}}\,. \tag{C.19}$$

In other words, it takes a finite proper time for a freely moving particle to reach the horizon. Hence, the spacetime region I covered by the Rindler coordinates and restricted by the horizon is geodesically incomplete. Since all curvature invariants vanish identically, the spacetime at $\rho = 0$ is regular. This means that the Rindler map covers only a part of the complete spacetime and one needs to extend it beyond $\rho = 0$.

---

[3]Certainly, this and subsequent relations can be obtained by working directly in the coordinates x^μ. But it is much simpler to derive them by first writing these relations in Cartesian coordinates and then making coordinate transformation (C.8).

C.4 Maximal Analytical Extension of Rindler Spacetime

How to do this? The answer is evident: Minkowski coordinates cover the complete spacetime. It might be instructive to restore these coordinates by using the same method of analytical continuation which has been used for the construction of the Kruskal metric in Appendix B.5. For this purpose, we first introduce null coordinates

$$u = \frac{c\,t}{l_a} - \ln\left(\frac{\rho}{l_a}\right), \quad v = \frac{c\,t}{l_a} + \ln\left(\frac{\rho}{l_a}\right), \tag{C.20}$$

in which metric (C.8) takes the form

$$ds^2 = -\,l_a^2\,e^{v-u}\,du\,dv + (dx^2)^2 + (dx^3)^2. \tag{C.21}$$

The further coordinates transformation

$$V = l_a\,e^{v}, \quad U = l_a\,e^{-u} \tag{C.22}$$

ensures that the horizons ($v = -\infty$, u finite) and ($u = \infty$, v finite) are located at finite values of U and V. In these new coordinates the metric is

$$ds^2 = dU\,dV + (dx^2)^2 + (dx^3)^2. \tag{C.23}$$

The metric in (U, V)-coordinates evidently remains regular at $U = 0$ and $V = 0$ and it can be extended beyond the horizons lying at these surfaces. It can be rewritten in the form (C.7) by introducing Cartesian coordinates

$$X^0 = \frac{1}{2}(V + U), \quad X^1 = \frac{1}{2}(V - U), \quad X^2 = x^2, \quad X^3 = x^3. \tag{C.24}$$

The Killing vector field ξ^μ is globally defined as follows:

$$\xi^\mu = \frac{1}{l_a}(-U\,\partial_U + V\,\partial_V), \tag{C.25}$$

and $\xi^2 = -l_a^{-2}\,UV$. The Killing horizons are formed by two null planes $U = 0$ and $V = 0$ where $\xi^2 = 0$. These null planes divide the spacetime into four regions I, II, III, and IV (see Figure C.2). In the region I the null coordinates U and V are related to the Rindler coordinates t and ρ as follows

$$U = -\rho\,\exp(-\tau), \quad V = \rho\,\exp(\tau). \tag{C.26}$$

Similar Rindler coordinates can be introduced in other regions.

Note that the null plane $V = X^0 + X^1 = 0$ is a boundary of the set of points where no information can be received from the region I covered by the Rindler frame. Thus, this surface is analogous to the *past horizon* in the spacetime of an eternal black hole (see Appendix B.6). The value of ξ^2 is negative in I, III and positive in II, IV.

Equations (C.20) and (C.22) allow one to show that for a particle or null ray crossing the horizon $X^0 = X^1$ at a finite value of V the equation of motion in (t, ρ)-coordinates has the following universal form

$$\rho \simeq l_a e^{-t/l_a} . \tag{C.27}$$

The redshift factor has the same universal behavior $\alpha \sim \rho/l_a \sim \exp(-t/l_a)$. The quantity $1/l_a$ which enters these universal laws coincides with the limit of the redshifted acceleration a at the horizon:

$$\kappa \equiv \lim_{\rho \to 0} [a(\rho)\, \alpha(\rho)] . \tag{C.28}$$

This quantity κ is known as the *surface gravity*. Note that the origin of the Rindler frame where $\xi^2 = -1$ can be chosen arbitrarily. The value of the surface gravity depends on this choice and it changes when the origin of the frame is moved. For black holes, there exists an asymptotically flat spacetime region where $\xi^2 = \text{const}$ and the choice $\xi^2_\infty = -1$ provides a natural normalization of the surface gravity (see Section 6.3).

To conclude this appendix, we summarize the main points. The domain of a rigid Rindler reference frame in a Minkowski spacetime is restricted. The region covered by this frame (map) is incomplete. A Rindler observer cannot receive information from the regions lying beyond the event horizon. The complete spacetime manifold may be covered by another map or system of such maps. The equations of motion of particles and null rays in a Rindler reference frame have a universal asymptotic form near the horizon which is determined solely by the surface gravity of the horizon.

Note that although in this appendix we focused our attention on the analogy between the Rindler frame in a Minkowski spacetime and a local static reference frame near the horizon in the Schwarzschild spacetime, this analogy can easily be extended to local regions near a rotating (Kerr-Newman) black hole.

Appendix D

Kerr-Newman Geometry

D.1 Kerr-Newman Metric

indexKerr-Newman metric

The *Kerr-Newman metric* in *Boyer-Lindquist coordinates* $x^\mu = (t, r, \theta, \phi)$ is

$$ds^2 = -\left(1 - \frac{2Mr - Q^2}{\Sigma}\right) dt^2 - \frac{(2Mr - Q^2)2a\sin^2\theta}{\Sigma}\, dt\, d\phi$$

$$+ \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \frac{A\sin^2\theta}{\Sigma}\, d\phi^2. \tag{D.1}$$

Here we use notations

$$\Delta = r^2 - 2Mr + a^2 + Q^2, \qquad \Sigma = r^2 + a^2\cos^2\theta\,,$$

$$A = (r^2 + a^2)\Sigma + a^2(2Mr - Q^2)\sin^2\theta = (r^2 + a^2)^2 - \Delta a^2\sin^2\theta\,. \tag{D.2}$$

The vector potential of the electromagnetic field is

$$A_\mu\, dx^\mu = -\frac{Qr}{\Sigma}\,(dt - a\sin^2\theta\, d\phi)\,. \tag{D.3}$$

It is easy to verify that the metric coefficients of (D.1) obey the following relations

$$g_{33} + a\sin^2\theta\, g_{03} = (r^2 + a^2)\sin^2\theta\,, \tag{D.4}$$

$$g_{03} + a\sin^2\theta\, g_{00} = -a\sin^2\theta\,, \tag{D.5}$$

$$a\, g_{33} + (r^2 + a^2)\, g_{03} = \Delta\, a\, \sin^2\theta\,, \tag{D.6}$$

$$a\, g_{03} + (r^2 + a^2)\, g_{00} = -\Delta\,, \tag{D.7}$$

$$(g_{03})^2 - g_{00}\, g_{33} = \Delta\sin^2\theta\,, \tag{D.8}$$

$$\sqrt{-g} = \Sigma \sin\theta\,. \tag{D.9}$$

The *static limit surface* is defined by the equation $g_{00} = 0$, which implies

$$r = r_{st} = M + \sqrt{M^2 - Q^2 - a^2\cos^2\theta}\,. \tag{D.10}$$

The event horizon is defined by the equation $\Delta = 0$ and is located at

$$r = r_+ = M + \sqrt{M^2 - Q^2 - a^2}\,. \tag{D.11}$$

The surface gravity κ, the angular velocity Ω^H, and the electric potential Φ^H for the Kerr-Newman black hole are

$$\kappa = \frac{r_+ - M}{r_+^2 + a^2}, \quad \Omega^H = \frac{a}{r_+^2 + a^2}, \quad \Phi^H = \frac{Q\,r_+}{r_+^2 + a^2}\,. \tag{D.12}$$

The non-vanishing contravariant components $g^{\mu\nu}$ of the metric are

$$g^{00} = -A(\Delta\Sigma)^{-1}, \quad g^{03} = -a(\Delta\Sigma)^{-1}(2Mr - Q^2)\,,$$

$$g^{33} = (\Delta\Sigma\sin^2\theta)^{-1}(\Delta - a^2\sin^2\theta)\,, \quad g^{11} = \Delta\Sigma^{-1}\,, \quad g^{22} = \Sigma^{-1}\,. \tag{D.13}$$

D.2 Christoffel Symbol

The non-vanishing components of the *Christoffel symbol* $\Gamma_{\alpha\beta\gamma}$ for the Kerr-Newman metric in Boyer-Lindquist coordinates are

$$\Gamma_{010} = -\Gamma_{100} = p\,, \quad \Gamma_{020} = -\Gamma_{200} = -\frac{a^2 q \sin 2\theta}{\Sigma}\,,$$

$$\Gamma_{222} = -\frac{a^2\sin 2\theta}{2}\,, \quad \Gamma_{031} = \Gamma_{310} = -\Gamma_{130} = -ap\sin^2\theta\,,$$

$$\Gamma_{032} = -\Gamma_{230} = \Gamma_{320} = \frac{ab^2 q\sin 2\theta}{\Sigma}\,,$$

$$\Gamma_{111} = \frac{1}{\Delta^2}\left[(M - r)\Sigma + r\Delta\right], \quad \Gamma_{121} = -\Gamma_{211} = -\frac{a^2\sin 2\theta}{2\Delta}\,, \tag{D.14}$$

$$\Gamma_{122} = -\Gamma_{221} = -r\,, \quad \Gamma_{133} = -\Gamma_{331} = -\sin^2\theta\,(r + pa^2\sin^2\theta)\,,$$

$$\Gamma_{233} = -\Gamma_{332} = -\frac{\sin 2\theta}{2\Sigma}\left[A\left(2 + \frac{a^2\sin^2\theta}{\Sigma}\right) - b^4\right].$$

Here

$$b^2 = r^2 + a^2\,, \quad p = \Sigma^{-2}[rQ^2 + M(\Sigma - 2r^2)]\,,$$

$$q = (Q^2 - 2Mr)(2\Sigma)^{-1}\,. \tag{D.15}$$

The corresponding non-vanishing components of the Christoffel symbol $\Gamma^{\alpha}_{\beta\gamma}$ are

$$\begin{aligned}
&\Gamma^0_{01} = -\frac{p\,b^2}{\Delta}\,, \quad \Gamma^0_{13} = \frac{a\sin^2\theta}{\Delta}\,(p\,b^2 + 2qr)\,,\\
&\Gamma^0_{32} = -\frac{qa^3}{\Sigma}\sin^2\theta\sin 2\theta\,, \quad \Gamma^0_{20} = \Sigma\Gamma^2_{00} = \frac{qa^2}{\Sigma}\sin 2\theta\,,\\
&\Gamma^1_{00} = -\frac{\Delta p}{\Sigma}\,; \quad \Gamma^1_{12} = \Gamma^2_{22} = -\Delta\Gamma^2_{11} = -\frac{a^2\sin 2\theta}{2\Sigma}\,,\\
&\Gamma^1_{11} = \frac{r}{\Sigma} + \frac{M-r}{\Delta}\,, \quad \Gamma^1_{22} = -\frac{r\Delta}{\Sigma}\,, \quad \Gamma^2_{12} = \frac{r}{\Sigma}\,,\\
&\Gamma^1_{33} = -\frac{\Delta\sin^2\theta}{\Sigma}\,(r + p\,a^2\sin^2\theta)\,, \quad \Gamma^3_{01} = -\frac{ap}{\Delta}\,,\\
&\Gamma^1_{03} = \frac{a\,\Delta p\sin^2\theta}{\Sigma}\,, \quad \Gamma^2_{03} = -\frac{qab^2\sin 2\theta}{\Sigma^2}\,,\\
&\Gamma^2_{33} = -\frac{\sin 2\theta}{2\Sigma}\left[b^2 - 2a^2\sin^2\theta\, q\left(2 + \frac{a^2\sin^2\theta}{\Sigma}\right)\right],\\
&\Gamma^3_{02} = \frac{2qa}{\Sigma}\cot\theta\,, \quad \Gamma^3_{13} = \frac{r}{\Delta}(1+2q) + a^2 p\,\frac{\sin^2\theta}{\Delta}\,,\\
&\Gamma^3_{32} = \frac{\cot\theta}{\Delta}\left[(1+2q)(b^2 - 2qa^2\sin^2\theta) - 2qa^2b^2\,\Sigma^{-1}\sin^2\theta\right].
\end{aligned} \tag{D.16}$$

D.3 Symmetries

The Kerr-Newman metric has two Killing vectors $\xi^{\mu}_{(t)} = \delta^{\mu}_{t}$ and $\xi^{\mu}_{(\phi)} = \delta^{\mu}_{\phi}$

$$\begin{aligned}
&\xi^2_{(t)} = g_{tt} = -\left(1 - \frac{2Mr - Q^2}{\Sigma}\right) = -\frac{\Delta - a^2\sin^2\theta}{\Sigma}\,,\\
&\xi^2_{(\phi)} = g_{\phi\phi} = \frac{A\sin^2\theta}{\Sigma}\,, \quad \xi_{(t)}\,\xi_{(\phi)} = g_{t\phi} = -\frac{(2Mr - Q^2)\,a\sin^2\theta}{\Sigma}\,.
\end{aligned} \tag{D.17}$$

It also has a Killing tensor of the second order $\xi_{\mu\nu}$ obeying the relation

$$\xi_{\mu\nu;\lambda} + \xi_{\nu\lambda;\mu} + \xi_{\lambda\mu;\nu} = 0\,. \tag{D.18}$$

In Boyer-Lindquist coordinates its non-vanishing components are

$$\xi_{00} = a^2 \left[1 - \frac{(2Mr - Q^2)\cos^2\theta}{\Sigma}\right], \quad \xi_{11} = -\frac{a^2\cos^2\theta\,\Sigma}{\Delta},$$

$$\xi_{22} = r^2\Sigma, \quad \xi_{03} = -\frac{a\sin^2\theta}{\Sigma}\left[\Delta\, a^2\cos^2\theta + r^2(r^2+a^2)\right], \tag{D.19}$$

$$\xi_{33} = \frac{\sin^2\theta}{\Sigma}\left[r^2(r^2+a^2)^2 + \frac{1}{4}\Delta\, a^4\sin^2 2\theta\right].$$

The corresponding symmetry of the electromagnetic field (D.3) can be expressed as

$$\xi_{\mu(\nu}\, F_{\lambda)}{}^{\mu} = 0\,. \tag{D.20}$$

D.4 Motion of Test Particles

D.4.1 Integrals of motion

The equation of motion of a charged (with charge e) particle in external gravitational and electromagnetic fields is

$$m\, u^\nu\, u_{\mu;\nu} = e\, F_{\mu\nu}\, u^\nu\,, \tag{D.21}$$

where m is the mass of the particle; u^μ its four-velocity, and $F_{\mu\nu}$ is the electromagnetic field tensor. Denote by $\pi^\mu = m\,u^\mu + eA^\mu$ the generalized momentum of the particle. Let ξ^μ be a Killing vector. Since the electromagnetic field potential A_μ obeys the symmetry condition $\mathcal{L}_\xi A_\mu = 0$ in an appropriate gauge, the quantity $\pi_\mu\,\xi^\mu$ is conserved. This follows directly from the relations

$$\frac{d}{ds}(\pi_\mu\,\xi^\mu) = (\xi^\mu{}_{;\nu}\,\pi_\mu + \xi^\mu\,\pi_{\mu;\nu})\,u^\nu$$

$$= m\,\xi_{\mu;\nu}\,u^\mu\,u^\nu + e\,(\mathcal{L}_\xi A_\mu)\,u^\mu = 0\,. \tag{D.22}$$

The electromagnetic field (D.3) in the Kerr-Newman geometry obeys the symmetry relations $\mathcal{L}_{\xi_{(t)}} A_\mu = \mathcal{L}_{\xi_{(\phi)}} A_\mu = 0$, and there exist two integrals of motion: the energy

$$E = -\pi_\mu\,\xi^\mu_{(t)} = -(m\,u_0 + eA_0) \tag{D.23}$$

and the axial component of angular momentum

$$L_z = \pi_\mu\,\xi^\mu_{(\phi)} = m\,u_3 + eA_3\,. \tag{D.24}$$

In the presence of a Killing tensor $\xi_{\mu\nu}$ there exists another independent conserved quantity

$$\mathcal{K} = \xi_{\mu\nu}\, p^\mu\, p^\nu\,, \tag{D.25}$$

where $p^\mu = m\, u^\mu$ provided the electromagnetic field tensor obeys the symmetry condition (D.20). This follows from the relations

$$\begin{aligned}\frac{d\mathcal{K}}{ds} &= m^2\left(\xi_{\mu\nu;\lambda}\, u^\mu\, u^\nu\, u^\lambda + 2\,\xi_{\mu\nu}\, u^\mu{}_{;\lambda}\, u^\nu\, u^\lambda\right)\\ &= m^2\xi_{(\mu\nu;\lambda)}\, u^\mu\, u^\nu\, u^\lambda + 2\,e\,m\,\xi_{\mu\nu}\, F^\mu{}_\lambda\, u^\nu\, u^\lambda = 0\,.\end{aligned} \tag{D.26}$$

For the Kerr-Newman metric the conserved quantity $\mathcal{K}$ is

$$\mathcal{K} = m^2\left[\left(a u^0 - (r^2+a^2)u^3\right)^2 \sin^2\theta + a^2\cos^2\theta + (u_2)^2\right]. \tag{D.27}$$

Using the relations

$$a\, g^{00} - (r^2+a^2)\, g^{03} = -a\,, \qquad (r^2+a^2)\, g^{33} - a\, g^{03} = \frac{1}{\sin^2\theta}\,,$$

one can verify that

$$ap^0 - (r^2+a^2)p^3 = -ap_0 - \frac{1}{\sin^2\theta}\, p_3\,.$$

Using this formula and expressions (D.23) and(D.24) for E and L_z, one gets

$$\mathcal{K} = \left(Ea\sin\theta - \frac{L_z}{\sin\theta}\right)^2 + (p_2)^2 + m^2a^2\cos^2\theta\,. \tag{D.28}$$

Quite often instead of $\mathcal{K}$ another integral of motion $\mathcal{Q}$ related to it is used:

$$\mathcal{Q} \equiv \mathcal{K} - (Ea - L_z)^2 = L_z^2\cot^2\theta - E^2a^2\cos^2\theta + (p_2)^2 + m^2a^2\cos^2\theta\,. \tag{D.29}$$

To summarize, the equation of motion of a charged particle in the Kerr-Newman spacetime allows four integrals of motions: E, L_z, $\mathcal{K}$ (or $\mathcal{Q}$), and a trivial one, $u^\mu u_\mu = -1$. One can express the four components u^μ of the velocity as explicit functions of these integrals of motion and the coordinates r and θ. As a result, one gets the system of equations (3.6.8)–(3.6.11). For an uncharged rotating black hole these equations reduce to the system (3.4.1)–(3.4.4).

D.4.2 Hamilton-Jacobi method

To study the motion of charged particles in the Kerr-Newman spacetime one can also use the *Hamilton-Jacobi method*. Equation (D.21) is equivalent to equations

$$\frac{dx^\mu}{d\lambda}=\frac{\partial\mathcal{H}}{\partial\pi_\mu}\,,\qquad \frac{d\pi_\mu}{d\lambda}=-\frac{\partial\mathcal{H}}{\partial x^\mu} \tag{D.30}$$

for the Hamiltonian

$$\mathcal{H}=\frac{1}{2}\,g^{\mu\nu}(\pi_\mu-eA_\mu)(\pi_\nu-eA_\nu)\,. \tag{D.31}$$

The corresponding Hamilton-Jacobi equation is

$$-\frac{\partial S}{\partial\lambda}=\frac{1}{2}\,g^{\mu\nu}\left(\frac{\partial S}{\partial x^\mu}-eA_\mu\right)\left(\frac{\partial S}{\partial x^\nu}-eA_\nu\right). \tag{D.32}$$

It has the following explicit form:

$$\begin{aligned}-\frac{\partial S}{\partial\lambda} \;=\;& -\frac{1}{2\Sigma\Delta}\left((r^2+a^2)\,\frac{\partial S}{\partial t}+a\,\frac{\partial S}{\partial\phi}-eQ\,r\right)^2\\ &+\;\frac{1}{2\Sigma\sin^2\theta}\left(\frac{\partial S}{\partial\phi}+a\sin^2\theta\,\frac{\partial S}{\partial t}\right)^2+\frac{\Delta}{2\Sigma}\left(\frac{\partial S}{\partial r}\right)^2+\frac{1}{2\Sigma}\left(\frac{\partial S}{\partial\theta}\right)^2.\end{aligned} \tag{D.33}$$

The Hamilton-Jacobi equation allows separation of variables

$$S=\frac{1}{2}\,m^2\lambda-Et+L_z\phi+S_r(r)+S_\theta(\theta)\,. \tag{D.34}$$

The substitution of (D.34) into (D.33) gives

$$S_r=\int\Delta^{-1}\sqrt{\mathcal{R}}\,dr\,,\qquad S_\theta=\int\sqrt{\Theta}\,d\theta\,, \tag{D.35}$$

where

$$\mathcal{R}=[E(r^2+a^2)-L_za-eQr]^2-\Delta(m^2r^2+\mathcal{K})\,, \tag{D.36}$$

$$\Theta=\mathcal{K}-m^2a^2\cos^2\theta-\left(Ea\sin\theta-\frac{L_z}{\sin\theta}\right)^2. \tag{D.37}$$

By successively putting $\partial S/\partial\mathcal{K}$, $\partial S/\partial m^2$, $\partial S/\partial E$, and $\partial S/\partial L_z$ to zero, one gets the following set of equations

$$\int^r\frac{dr}{\sqrt{\mathcal{R}}}=\int^\theta\frac{d\theta}{\sqrt{\Theta}}\,, \tag{D.38}$$

$$\lambda=\int^r\frac{r^2}{\sqrt{\mathcal{R}}}\,dr+\int^\theta\frac{a^2\cos^2\theta}{\sqrt{\Theta}}\,d\theta\,, \tag{D.39}$$

$$t=\int^r \frac{(r^2+a^2)[E(r^2+a^2)-L_z a-eQr]}{\Delta\sqrt{\mathcal{R}}}\,dr-\int^\theta \frac{a(aE\sin^2\theta-L_z)}{\sqrt{\Theta}}\,d\theta, \tag{D.40}$$

$$\phi=\int^r \frac{a[E(r^2+a^2)-L_z a-eQr]}{\Delta\sqrt{\mathcal{R}}}\,dr-\int^\theta \frac{aE\sin^2\theta-L_z}{\sin^2\theta\sqrt{\Theta}}\,d\theta\,. \tag{D.41}$$

By differentiating these equations it is possible to show that they are equivalent to the system of first-order differential equations (3.6.8)–(3.6.11).

D.5 Stationary Congruences in the Kerr-Newman Geometry

D.5.1 Killing congruence

The Kerr-Newman metric (D.1) can be rewritten as follows:

$$ds^2=-F(dt+g_3\,d\phi)^2+h_{ij}\,dx^i\,dx^j\,, \tag{D.42}$$

where

$$F=-\xi^2_{(t)}=1-\frac{2Mr-Q^2}{\Sigma}\,,\quad g_3=\frac{g_{03}}{g_{00}}=\frac{(2Mr-Q^2)\,a\sin^2\theta}{\Delta-a^2\sin^2\theta}\,, \tag{D.43}$$

$$h_{ij}\,dx^i\,dx^j=\frac{\Sigma}{\Delta}\,dr^2+\Sigma\,d\theta^2+\frac{\Sigma\Delta\sin^2\theta}{\Delta-a^2\sin^2\theta}\,d\phi^2. \tag{D.44}$$

The four-velocity $\tilde{u}^\mu=F^{-1/2}\xi^\mu_{(t)}$ of a Killing observer associated with the Killing vector $\xi^\mu_{(t)}$ has the following components:

$$\tilde{u}^\mu=\left(\frac{\Sigma}{\Delta-a^2\sin^2\theta}\right)^{1/2}\delta^\mu_0,\quad \tilde{u}_\mu=-\left(\frac{\Delta-a^2\sin^2\theta}{\Sigma}\right)^{1/2}(1,0,0,g_3). \tag{D.45}$$

For the Killing congruence $\tilde{u}^\mu$ the rate of deformation tensor vanishes, $D_{\mu\nu}=0$, while the acceleration w_μ and the vorticity tensor $\omega_{\mu\nu}$ have the following non-vanishing components:

$$w_1=\frac{(2Mr-Q^2)r-M\Sigma}{\Sigma(\Delta-a^2\sin^2\theta)}\,,\quad w_2=-\frac{a^2(2Mr-Q^2)\sin 2\theta}{2\Sigma(\Delta-a^2\sin^2\theta)}\,, \tag{D.46}$$

$$\omega_{13}=\frac{a\sin^2\theta\,(M\Sigma-r(2Mr-Q^2))}{\Sigma^{1/2}(\Delta-a^2\sin^2\theta)^{3/2}}\,,\quad \omega_{23}=\frac{a(2Mr-Q^2)\Delta\sin 2\theta}{2\Sigma^{1/2}(\Delta-a^2\sin^2\theta)^{3/2}}\,. \tag{D.47}$$

D.5.2 Congruence of locally non-rotating observers

Another useful form of the Kerr-Newman metric is

$$ds^2 = -\alpha^2\, dt^2 + g_{ij}(dx^i + \alpha^i\, dt)(dx^j + \alpha^j\, dt)\,, \tag{D.48}$$

where

$$\alpha = \sqrt{\frac{\Sigma\Delta}{A}}, \quad \alpha^i = -\omega\delta^i_3\,, \quad \omega = \frac{a(2Mr - Q^2)}{A}. \tag{D.49}$$

$$g_{ij}\, dx^i\, dx^j = \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \frac{A\sin^2\theta}{\Sigma}\, d\phi^2. \tag{D.50}$$

Consider a stationary congruence formed by observers moving with the four-velocity

$$u_\mu = -\alpha\, \delta^t_\mu\,, \quad u^\mu = \alpha^{-1}(1, -\alpha^i)\,. \tag{D.51}$$

One has $\xi_{(\phi)\mu}\, u^\mu = 0$, so that the observers of the congruence have zero angular momentum. For this congruence, the vorticity tensor vanishes, $\omega_{\alpha\beta} = 0$. That is why the reference frame of such observers is called *locally non-rotating.* The non-vanishing components of the acceleration w_μ are

$$w_1 = \frac{[M(r^2 - a^2\cos^2\theta) - Q^2 r]\,(r^2 + a^2)^2 - (2Mr - Q^2)^2 r a^2 \sin^2\theta}{\Sigma\Delta A}\,, \tag{D.52}$$

$$w_2 = -\frac{a^2\sin 2\theta (r^2 + a^2)(2Mr - Q^2)}{2\Sigma A}\,. \tag{D.53}$$

The non-vanishing spatial components of the deformation rate tensor $D_{ij} = \alpha\, \Gamma^0_{ij}$ are

$$D_{13} = a\sin^2\theta \left(\frac{\Sigma}{\Delta A}\right)^{1/2} (p b^2 + 2qr)\,, \tag{D.54}$$

$$D_{23} = -\, a^3 q \sin^2\theta \sin 2\theta \left(\frac{\Delta}{\Sigma A}\right)^{1/2}. \tag{D.55}$$

The quantities b, p, and q which enter these relations are defined by (D.15).

One can adjoin to $\mathbf{e}^\mu_{\hat 0} = u^\mu$ three spatial unit vectors

$$\mathbf{e}^\mu_{\hat 1} = \sqrt{\frac{\Delta}{\Sigma}}\, \delta^\mu_1\,, \quad \mathbf{e}^\mu_{\hat 2} = \sqrt{\frac{1}{\Sigma}}\, \delta^\mu_2\,, \quad \mathbf{e}^\mu_{\hat 3} = \sqrt{\frac{\Sigma}{A\sin^2\theta}}\, \delta^\mu_3\,. \tag{D.56}$$

The orthonormal tetrad $\mathbf{e}_{\hat n}$ forms a local reference frame. The antisymmetric 3-tensor $\Omega_{\hat m\hat n} = e_{\hat n\alpha}\, \mathcal{F}_u\, e^\alpha_{\hat m}$ is the angular velocity tensor of rotation of this frame with respect to a locally non-rotating frame, while

$$\Omega^{\hat k} = -\frac{1}{2}\, \epsilon^{\hat k\hat m\hat n}\, \Omega_{\hat m\hat n} \tag{D.57}$$

is the angular velocity of gyroscope precession in the frame $\mathbf{e}_{\hat{m}}$. For the Kerr-Newman metric the angular velocity $\mathbf{\Omega}^{(\text{precess})}$ of gyroscope precession with respect to the frame $\mathbf{e}_{\hat{m}}$ is

$$\mathbf{\Omega}^{(\text{precess})} = \frac{A\sin\theta}{\Sigma\Delta^{1/2}}\left(\frac{\omega_{,\theta}}{\Sigma^{1/2}}\,\mathbf{e}_{\hat{r}} - \frac{\Delta^{1/2}}{\Sigma^{1/2}}\,\omega_{,r}\,\mathbf{e}_{\hat{\theta}}\right). \tag{D.58}$$

The derivation of this formula is explained in [Misner, Thorne, and Wheeler (1973)].

D.6 Algebraic Properties

The Kerr-Newman metric belongs to the Petrov type D metrics, and according to the *Goldberg-Sachs theorem* [Goldberg and Sachs (1962)], the principal null congruences are geodesic and shear-free. In Boyer-Lindquist coordinates these geodesics are defined by the equations

$$\frac{dt}{d\lambda} = \frac{r^2+a^2}{\Delta}, \quad \frac{dr}{d\lambda} = \pm 1, \quad \frac{d\theta}{d\lambda} = 0, \quad \frac{d\phi}{d\lambda} = \frac{a}{\Delta}, \tag{D.59}$$

where λ is an affine parameter. By choosing the null vectors $\boldsymbol{l}$ and $\boldsymbol{n}$ of a complex null tetrad $\boldsymbol{z}_m = (\boldsymbol{l}, \boldsymbol{n}, \boldsymbol{m}, \bar{\boldsymbol{m}})$ to be tangent to the principal null geodesics, one gets the following null tetrad proposed by Kinnersley (1969):

$$\begin{aligned} l^\mu &= \frac{1}{\Delta}\left(r^2+a^2, \Delta, 0, a\right), \\ n^\mu &= \frac{1}{2\Sigma}\left(r^2+a^2, -\Delta, 0, a\right), \\ m^\mu &= \frac{1}{\sqrt{2}\,(r+ia\cos\theta)}\left(ia\sin\theta, 0, 1, \frac{i}{\sin\theta}\right). \end{aligned} \tag{D.60}$$

The vectors of the *Kinnersley tetrad* obey the normalization conditions

$$\boldsymbol{l}\cdot\boldsymbol{n} = -1, \quad \boldsymbol{m}\cdot\bar{\boldsymbol{m}} = 1; \tag{D.61}$$

the other scalar products vanish.

The following useful representation of the covariant derivatives $\xi_{(t)\mu;\nu}$ of the Killing vector $\xi^\mu_{(t)}$

$$\xi_{(t)\mu;\nu} = -\frac{\partial F}{\partial r}\, l_{[\mu}\, n_{\nu]} + \frac{2ia\cos\theta\,(1-F)}{\Sigma}\, m_{[\mu}\,\bar{m}_{\nu]}, \tag{D.62}$$

can be used to show that the principal null vectors $\boldsymbol{l}$ and $\boldsymbol{n}$ are eigenvectors of $\xi_{(t)\mu;\nu}$ with eigenvalues $\pm\frac{1}{2}\,\partial F/\partial r$

$$\xi_{(t)\mu;\nu}\, l^\nu = \frac{1}{2}\frac{\partial F}{\partial r}\, l_\mu, \quad \xi_{(t)\mu;\nu}\, n^\nu = -\frac{1}{2}\frac{\partial F}{\partial r}\, n_\mu. \tag{D.63}$$

The function F which enters these relations is defined by (D.43).

The spin coefficients (E.5) calculated in the Kerr-Newman metric for the Kinnersley tetrad are

$$\begin{aligned}
&\varepsilon = \kappa = \lambda = \nu = \sigma = 0\,, \\
&\rho = -\frac{1}{r - ia\cos\theta}\,, \quad \tau = -\frac{ia\sin\theta}{\sqrt{2}\,\Sigma}\,, \\
&\pi = \frac{ia\sin\theta}{\sqrt{2}\,(r - ia\cos\theta)^2}\,, \quad \beta = -\frac{\bar{\rho}\cot\theta}{2\sqrt{2}}\,, \\
&\alpha = \pi - \bar{\beta}\,, \quad \mu = \frac{\rho\Delta}{2\Sigma}\,, \quad \gamma = \mu + \frac{r - M}{2\Sigma}\,.
\end{aligned} \tag{D.64}$$

D.7 Analytic Extension

Let us introduce new coordinates $(v, r, \theta, \tilde{\phi})$ related to the Boyer-Lindquist coordinates (t, r, θ, ϕ) as follows

$$dv = dt + dr_*\,, \quad dr_* = (r^2 + a^2)\,\frac{dr}{\Delta}\,, \tag{D.65}$$

$$d\tilde{\phi} = d\phi + a\,\frac{dr}{\Delta}\,. \tag{D.66}$$

The coordinate r_* is an obvious generalization of the Schwarzschild "tortoise" coordinate. For the Kerr-Newman metric, r_* is

$$r_* = r + \frac{1}{2\kappa}\ln|r - r_+| + \frac{1}{2\kappa_-}\ln|r - r_-|\,, \tag{D.67}$$

$$\kappa = (r_+ - r_-)/[2(r_+^2 + a^2)]\,, \quad \kappa_- = (r_- - r_+)/[2(r_-^2 + a^2)]\,. \tag{D.68}$$

Here $r_\pm = M \pm \sqrt{M^2 - a^2 - Q^2}$ and κ is the surface gravity of the Kerr-Newman black hole.

The coordinates $(v, r, \theta, \tilde{\phi})$ are known as the *Kerr ingoing coordinates* [Kerr (1963)]. The Kerr-Newman metric (D.1) in these coordinates takes the form

$$\begin{aligned}
ds^2 = &-\frac{\Delta}{\Sigma}\left(dv - a\sin^2\theta\,d\tilde{\phi}\right)^2 + \frac{\sin^2\theta}{\Sigma}\left[(r^2 + a^2)\,d\tilde{\phi} - a\,dv\right]^2 \\
&+ \Sigma\,d\theta^2 + 2\,dr\left(dv - a\sin^2\theta\,d\tilde{\phi}\right).
\end{aligned} \tag{D.69}$$

The Kerr ingoing coordinates are regular on the future event horizon and cover the region $I \cup II$ shown in Figure 6.4. The three-surface element of the horizon $r = r_+$, with normal vector in the inward radial direction, is [Teukolsky and Press (1974)]

$$d\Sigma^\mu = l_H^\mu \,(r_+^2 + a^2) \sin\theta \, dv \, d\theta \, d\tilde{\phi} \,, \tag{D.70}$$

where

$$l_H^\mu = (1, 0, 0, \Omega^H) \,. \tag{D.71}$$

The *Kerr outgoing coordinates* $(u, r, \theta, \bar{\phi})$, defined by

$$du = dt - dr_* \,, \quad d\bar{\phi} = d\phi - a \frac{dr}{\Delta} \tag{D.72}$$

are regular on the past horizon and cover the region $I \cup II'$ shown in Figure 6.4. The Kerr-Newman metric in these coordinates can be obtained from (D.69) by the transformation $v = -u$, $\tilde{\phi} = -\bar{\phi}$.

One can also introduce coordinates (U, V, θ, ϕ_+) which are similar to Kruskal coordinates and cover the region $I \cup I' \cup II \cup II'$ (see Figure 6.4). These coordinates are defined as

$$|U| = \exp[\kappa(r_* - t)] \,, \quad |V| = \exp[\kappa(r_* + t)] \,, \quad \phi_+ = \phi - \Omega^H t \,. \tag{D.73}$$

Here $\Omega^H = a/(r_+^2 + a^2)$ is the angular velocity of the Kerr-Newman black hole.

In these coordinates the Kerr-Newman metric takes the form

$$\begin{aligned} ds^2 = {} & \frac{G^2 a^2 \sin^2\theta}{4\kappa_+^2 \Sigma} \frac{(r - r_-)(r + r_+)}{(r^2 + a^2)(r_+^2 + a^2)} \left[\frac{\Sigma}{r^2 + a^2} + \frac{\Sigma_+}{r_+^2 + a^2} \right] (U^2 \, dV^2 + V^2 \, dU^2) \\ & + \frac{G(r - r_-)}{2\kappa_+^2 \Sigma} \left[\frac{\Sigma^2}{(r^2 + a^2)^2} + \frac{\Sigma_+^2}{(r_+^2 + a^2)^2} \right] dU \, dV \\ & + \frac{G^2 a^2 \sin^2\theta}{4\kappa_+^2 \Sigma} \frac{(r + r_+)^2}{(r_+^2 + a^2)^2} (U \, dV - V \, dU)^2 \\ & + \frac{G a \sin^2\theta}{\kappa_+ \Sigma (r_+^2 + a^2)} [\Sigma_+ (r - r_-) + (r^2 + a^2)(r + r_+)](U \, dV - V \, dU) \, d\varphi_+ \\ & + \Sigma \, d\theta^2 + \frac{A \sin^2\theta}{\Sigma} \, d\varphi_+^2 \,, \end{aligned} \tag{D.74}$$

where Σ_+ and κ_+ are the values of Σ and κ at $r = r_+$, and

$$G = \frac{r - r_+}{UV} \,. \tag{D.75}$$

In the black hole exterior where $U < 0$, $V > 0$ the coordinates (U, V) are connected with Kerr coordinates (u, v) as follows

$$U = -\exp(-\kappa u), \quad V = \exp(\kappa v). \tag{D.76}$$

If $-\infty < U, V < \infty$, these coordinates cover the regions $I \cup II \cup I' \cup II'$. A detailed derivation of the analytical continuation (D.74) and the construction of the maximal analytical continuation for the Kerr metric can be found in the book by O'Neill (1994).

Setting $U = V = 0$ reduces metric (D.74) to

$$ds^2 = 2C\, dU\, dV + \Sigma_+^2\, d\theta^2 + \frac{(r_+^2 + a^2)^2 \sin^2\theta}{\Sigma_+}\, d\phi_+^2\,, \tag{D.77}$$

where $2C = G_+\,\Sigma_+\,(r_+ - r_-)/\,[\kappa_+^2(r_+^2 + a^2)^2]$, and $G_+ = G(r_+)$ is a finite constant. By simple rescaling $U \to c\,U$, $V \to c'\,V$, one can always put $C = 1$. Metric (D.77) is evidently regular on both horizons, H^+ and H^-, V and U being the affine parameters along H^+ and H^-, respectively. It is easy to verify that

$$[\phi_+]_{H^+} = \tilde{\phi} - \Omega^H\, v\,, \quad [\phi_+]_{H^-} = \tilde{\phi} - \Omega^H\, u\,. \tag{D.78}$$

In the absence of rotation the Kerr-Newman metric reduces to the Reissner-Nordström metric (7.2.1). The analytic continuation (similar to (D.74)) of the Reissner-Nordström metric then takes the form

$$ds^2 = -2B\, dU\, dV + r^2 d\omega^2, \tag{D.79}$$

where

$$B = \frac{2r_+^4}{r^2}\,\frac{(r - r_+)(r - r_-)}{(r_+ - r_-)^2}\,\exp\left(-r_*\,\frac{r_+ - r_-}{r_+^2}\right), \tag{D.80}$$

and $r = r(U, V)$ is determined as a solution of the equation

$$UV = \left(1 - \frac{r}{r_+}\right)\exp\left[\left(r - \frac{r_-^2}{r_+ - r_-}\ln\left|\frac{r}{r_-} - 1\right|\right)\left(\frac{r_+ - r_-}{r_+^2}\right)\right]. \tag{D.81}$$

Appendix E

Newman-Penrose Formalism

E.1 Complex Null Tetrad. Spin Coefficients

In this appendix some useful formulas of the *Newman-Penrose formalism* are collected. The null-tetrad approach developed by Newman and Penrose (1962) is a very useful and powerful method for the construction of solutions of the Einstein equations and for studying physical fields propagating in the curved background. It is especially useful for studying algebraically special spaces and massless fields.

Denote by $z^\mu_m = (l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$ a *complex null tetrad* obeying the normalization conditions

$$\mathbf{l}\cdot\mathbf{n} = -1\,, \qquad \mathbf{m}\cdot\bar{\mathbf{m}} = 1\,; \tag{E.1}$$

the other scalar products vanish. The index $m = 0, 1, 2, 3$ labels the tetrad vectors. We denote by $\mathbf{z}^m$ the dual basis of 1-forms $z^m_\mu = (-n_\mu, -l_\mu, \bar{m}_\mu, m_\mu)$. The metric $g_{\mu\nu}$ can be written as

$$g_{\mu\nu} = \eta_{mn}\, z^m_\mu\, z^n_\nu = -2(l_{(\mu}\, n_{\nu)} - m_{(\mu}\,\bar{m}_{\nu)})\,. \tag{E.2}$$

The matrix η_{mn} which enters this formula, and its inverse η^{mn}, are of the form

$$\eta_{mn} = \eta^{mn} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{E.3}$$

These matrices are used for operations on the tetrad indices.

The connection coefficients $\Gamma^k{}_{mn}$ are defined as

$$\nabla_n \mathbf{z}_m = \Gamma^k{}_{mn}\, \mathbf{z}_k\,, \tag{E.4}$$

where $\nabla_n \equiv z^\nu_n \nabla_\nu$. The following 12 independent complex linear combinations of the connection coefficients are known as *spin coefficients*

$$-\kappa \equiv \Gamma_{200} = l_{\mu;\nu}\, m^\mu\, l^\nu, \qquad -\rho \equiv \Gamma_{203} = l_{\mu;\nu}\, m^\mu\, \bar{m}^\nu,$$

$$-\sigma \equiv \Gamma_{202} = l_{\mu;\nu}\, m^\mu\, m^\nu , \qquad -\tau \equiv \Gamma_{201} = l_{\mu;\nu}\, m^\mu\, n^\nu ,$$

$$\nu \equiv \Gamma_{311} = n_{\mu;\nu}\, \bar{m}^\mu\, n^\nu , \qquad \mu \equiv \Gamma_{312} = n_{\mu;\nu}\, \bar{m}^\mu\, m^\nu ,$$

$$\lambda \equiv \Gamma_{313} = n_{\mu;\nu}\, \bar{m}^\mu\, \bar{m}^\nu , \qquad \pi \equiv \Gamma_{310} = n_{\mu;\nu}\, \bar{m}^\mu\, l^\nu , \tag{E.5}$$

$$-\varepsilon \equiv \frac{1}{2}(\Gamma_{100} - \Gamma_{320}) = \frac{1}{2}(l_{\mu;\nu}\, n^\mu\, l^\nu - m_{\mu;\nu}\, \bar{m}^\mu\, l^\nu) ,$$

$$-\beta \equiv \frac{1}{2}(\Gamma_{102} - \Gamma_{322}) = \frac{1}{2}(l_{\mu;\nu}\, n^\mu\, m^\nu - m_{\mu;\nu}\, \bar{m}^\mu\, m^\nu) ,$$

$$-\gamma \equiv \frac{1}{2}(\Gamma_{231} - \Gamma_{011}) = \frac{1}{2}(l_{\mu;\nu}\, n^\mu\, n^\nu - m_{\mu;\nu}\, \bar{m}^\mu\, n^\nu) ,$$

$$-\alpha \equiv \frac{1}{2}(\Gamma_{233} - \Gamma_{013}) = \frac{1}{2}(l_{\mu;\nu}\, n^\mu\, \bar{m}^\nu - m_{\mu;\nu}\, \bar{m}^\mu\, \bar{m}^\nu) .$$

E.2 Covariant Derivatives. Ricci and Weyl Tensor

The following standard notations are used for the covariant derivatives in the directions of the null tetrad vectors:

$$D = l^\mu\, \nabla_\mu , \quad \Delta = n^\mu\, \nabla_\mu , \quad \delta = m^\mu\, \nabla_\mu , \quad \bar{\delta} = \bar{m}^\mu\, \nabla_\mu . \tag{E.6}$$

The commutators

$$[\boldsymbol{z}_m, \boldsymbol{z}_n] \equiv \nabla_m\, \boldsymbol{z}_n - \nabla_n\, \boldsymbol{z}_m = -2\Gamma^k{}_{mn}\, \boldsymbol{z}_k , \tag{E.7}$$

in this notation have the explicit form

$$\Delta D - D\Delta = (\gamma + \bar{\gamma})D + (\varepsilon + \bar{\varepsilon})\,\Delta - (\tau + \bar{\pi})\bar{\delta} - (\bar{\tau} + \pi)\delta ,$$

$$\delta D - D\delta = (\bar{\alpha} + \beta - \bar{\pi})D + \kappa\,\Delta - \sigma\bar{\delta} - (\bar{\rho} + \varepsilon - \bar{\varepsilon})\delta ,$$

$$\delta\,\Delta - \Delta\,\delta = -\bar{\nu}D + (\tau - \bar{\alpha} - \beta)\,\Delta + \bar{\lambda}\bar{\delta} + (\mu - \gamma + \bar{\gamma})\delta , \tag{E.8}$$

$$\bar{\delta}\,\delta - \delta\,\bar{\delta} = (\bar{\mu} - \mu)D + (\bar{\rho} - \rho)\,\Delta - (\bar{\alpha} - \beta)\bar{\delta} - (\bar{\beta} - \alpha)\delta .$$

The following notations are used for the null-tetrad components of the traceless part of the Ricci tensor ($Q_{\mu\nu} = R_{\mu\nu} - g_{\mu\nu}\, R/4$) and for the Weyl tensor $C_{\alpha\beta\mu\nu}$

$$\Phi_{00} = \frac{1}{2}\, Q_{\mu\nu}\, l^\mu\, l^\nu = \frac{1}{2}\, R_{00} , \quad \Phi_{01} = \bar{\Phi}_{10} = \frac{1}{2}\, Q_{\mu\nu}\, l^\mu\, m^\nu = \frac{1}{2}\, R_{02} ,$$

$$\Phi_{02} = \bar{\Phi}_{20} = \frac{1}{2}\, Q_{\mu\nu}\, m^\mu\, m^\nu = \frac{1}{2}\, R_{22} , \quad \Phi_{12} = \bar{\Phi}_{21} = \frac{1}{2}\, Q_{\mu\nu}\, n^\mu\, m^\nu = \frac{1}{2}\, R_{12} ,$$

$$\Phi_{11} = \frac{1}{4} Q_{\mu\nu}(l^\mu n^\nu + m^\mu \bar{m}^\nu) = \frac{1}{4}(R_{01} + R_{23}), \tag{E.9}$$

$$\Phi_{22} = \frac{1}{2} Q_{\mu\nu} n^\mu n^\nu = \frac{1}{2} R_{11},$$

$$\Psi_0 = C_{\alpha\beta\mu\nu} l^\alpha m^\beta l^\mu m^\nu, \qquad \Psi_1 = C_{\alpha\beta\mu\nu} l^\alpha n^\beta l^\mu m^\nu,$$

$$\Psi_2 = C_{\alpha\beta\mu\nu} l^\alpha n^\beta (l^\mu n^\nu - m^\mu \bar{m}^\nu), \tag{E.10}$$

$$\Psi_3 = C_{\alpha\beta\mu\nu} n^\alpha l^\beta n^\mu \bar{m}^\nu, \qquad \Psi_4 = C_{\alpha\beta\mu\nu} n^\alpha \bar{m}^\beta n^\mu \bar{m}^\nu.$$

The null-tetrad components of the curvature tensor can be obtained as follows:

$$R^k{}_{lmn} \equiv z^k_\alpha z^\beta_l z^\mu_m z^\nu_n R^\alpha{}_{\beta\mu\nu}$$

$$= \nabla_m \Gamma^k{}_{ln} - \nabla_n \Gamma^k{}_{lm} + \Gamma^p{}_{ln}\Gamma^k{}_{pm} - \Gamma^p{}_{lm}\Gamma^k{}_{pn} + \Gamma^p{}_{mn}\Gamma^k{}_{lp} - \Gamma^p{}_{nm}\Gamma^k{}_{lp}. \tag{E.11}$$

E.3 Newman-Penrose Equations

The following explicit form of relations (E.9)–(E.11) in terms of spin coefficients is known as the *Newman-Penrose equations*

$$D\rho - \bar{\delta}\kappa = (\rho^2 + \sigma\bar{\sigma}) + (\varepsilon + \bar{\varepsilon})\rho - \bar{\kappa}\tau - \kappa(3\alpha + \bar{\beta} - \pi) + \Phi_{00},$$

$$D\sigma - \delta\kappa = (\rho + \bar{\rho})\sigma + (3\varepsilon - \bar{\varepsilon})\sigma - (\tau - \bar{\pi} + \bar{\alpha} + 3\beta)\kappa + \Psi_0,$$

$$D\tau - \Delta\kappa = (\tau + \bar{\pi})\rho + (\bar{\tau} + \pi)\sigma + (\varepsilon - \bar{\varepsilon})\tau - (3\gamma + \bar{\gamma})\kappa + \Psi_1 + \Phi_{01},$$

$$D\alpha - \bar{\delta}\varepsilon = (\rho + \bar{\varepsilon} - 2\varepsilon)\alpha + \beta\bar{\sigma} - \bar{\beta}\varepsilon - \kappa\lambda - \bar{\kappa}\gamma + (\varepsilon + \rho)\pi + \Phi_{10},$$

$$D\beta - \delta\varepsilon = (\alpha + \pi)\sigma + (\bar{\rho} - \bar{\varepsilon})\beta - (\mu + \gamma)\kappa - (\bar{\alpha} - \bar{\pi})\varepsilon + \Psi_1,$$

$$D\gamma - \Delta\varepsilon = (\tau + \bar{\pi})\alpha + (\bar{\tau} + \pi)\beta - (\varepsilon + \bar{\varepsilon})\gamma - (\gamma + \bar{\gamma})\varepsilon + \tau\pi$$

$$- \nu\kappa + \Psi_2 + \Phi_{11} - \Lambda,$$

$$D\lambda - \bar{\delta}\pi = (\rho\lambda + \bar{\sigma}\mu) + \pi^2 + (\alpha - \bar{\beta})\pi - \nu\bar{\kappa} - (3\varepsilon - \bar{\varepsilon})\lambda + \Phi_{20},$$

$$D\mu - \delta\pi = (\bar{\rho}\mu + \sigma\lambda) + \pi\bar{\pi} - (\varepsilon + \bar{\varepsilon})\mu - \pi(\bar{\alpha} - \beta) - \nu\kappa + \Psi_2 + 2\Lambda,$$

$$D\nu - \Delta\pi = (\pi + \bar{\tau})\mu + (\bar{\pi} + \tau)\lambda + (\gamma - \bar{\gamma})\pi - (3\varepsilon + \bar{\varepsilon})\nu + \Psi_3 + \Phi_{21},$$

$$\Delta\lambda - \bar{\delta}\nu = -(\mu + \bar{\mu})\lambda - (3\gamma - \bar{\gamma})\lambda + (3\alpha + \bar{\beta} + \pi - \bar{\tau})\nu - \Psi_4, \tag{E.12}$$

$$\delta\rho - \bar{\delta}\sigma = \rho(\bar{\alpha} + \beta) - \sigma(3\alpha - \bar{\beta}) + (\rho - \bar{\rho})\tau + (\mu - \bar{\mu})\kappa - \Psi_1 + \Phi_{01},$$

$$\delta\alpha - \bar{\delta}\beta = (\mu\rho - \lambda\sigma) + \alpha\bar{\alpha} + \beta\bar{\beta} - 2\alpha\beta + \gamma(\rho - \bar{\rho}) + \varepsilon(\mu - \bar{\mu})$$

$$- \Psi_2 + \Phi_{11} + \Lambda ,$$

$$\delta\lambda - \bar{\delta}\mu = (\rho - \bar{\rho})\nu + (\mu - \bar{\mu})\pi + \mu(\alpha + \bar{\beta}) + \lambda(\bar{\alpha} - 3\beta) - \Psi_3 + \Phi_{21} ,$$

$$\delta\nu - \Delta\mu = (\mu^2 + \lambda\bar{\lambda}) + (\gamma + \bar{\gamma})\mu - \bar{\nu}\pi + (\tau - 3\beta - \bar{\alpha})\nu + \Phi_{22} ,$$

$$\delta\gamma - \Delta\beta = (\tau - \bar{\alpha} - \beta)\gamma + \mu\tau - \sigma\nu - \varepsilon\bar{\nu} - \beta(\gamma - \bar{\gamma} - \mu) + \alpha\bar{\lambda} + \Phi_{12} ,$$

$$\delta\tau - \Delta\sigma = (\mu\sigma + \bar{\lambda}\rho) + (\tau + \beta - \bar{\alpha})\tau - (3\gamma - \bar{\gamma})\sigma - \kappa\bar{\nu} + \Phi_{02} ,$$

$$\Delta\rho - \bar{\delta}\tau = -(\rho\bar{\mu} + \sigma\lambda) + (\bar{\beta} - \alpha - \bar{\tau})\tau + (\gamma + \bar{\gamma})\rho + \nu\kappa - \Psi_2 - 2\Lambda ,$$

$$\Delta\alpha - \bar{\delta}\gamma = (\rho + \varepsilon)\nu - (\tau + \beta)\lambda + (\bar{\gamma} - \bar{\mu})\alpha + (\bar{\beta} - \bar{\tau})\gamma - \Psi_3 .$$

E.4 Bianchi Identities

The Bianchi identities

$$\nabla_{[\lambda} R_{\mu\nu]\alpha\beta} = 0 , \tag{E.13}$$

form the last set of equations of the Newman-Penrose formalism. The tetrad form of this equation is

$$\nabla_{[k} R_{mn]pq} = -2R_{pqs[m} \Gamma^s{}_{nk]} + \Gamma^s{}_{p]m} R_{nk]sq} - \Gamma^s{}_{q]m} R_{nk]sp} . \tag{E.14}$$

For completeness we reproduce below an explicit form of these equations written in terms of spin coefficients

$$\bar{\delta}\Psi_0 - D\Psi_1 + D\Phi_{01} - \delta\Phi_{00} = (4\alpha - \pi)\Psi_0 - 2(2\rho + \varepsilon)\Psi_1 + 3\kappa\Psi_2$$

$$+ (\bar{\pi} - 2\bar{\alpha} - 2\beta)\Phi_{00} + 2(\varepsilon + \bar{\rho})\Phi_{01} + 2\sigma\Phi_{10} - 2\kappa\Phi_{11} - \bar{\kappa}\Phi_{02} ,$$

$$\Delta\Psi_0 - \delta\Psi_1 + D\Phi_{02} - \delta\Phi_{01} = (4\gamma - \mu)\Psi_0 - 2(2\tau + \beta)\Psi_1 + 3\sigma\Psi_2$$

$$+ (2\varepsilon - 2\bar{\varepsilon} + \bar{\rho})\Phi_{02} + 2(\bar{\pi} - \beta)\Phi_{01} + 2\sigma\Phi_{11} - 2\kappa\Phi_{12} - \bar{\lambda}\Phi_{00} ,$$

$$\bar{\delta}\Psi_3 - D\Psi_4 + \bar{\delta}\Phi_{21} - \Delta\Phi_{20} = (4\varepsilon - \rho)\Psi_4 - 2(2\pi + \alpha)\Psi_3 + 3\lambda\Psi_2$$

$$+ (2\gamma - 2\bar{\gamma} + \bar{\mu})\Phi_{20} + 2(\bar{\tau} - \alpha)\Phi_{21} + 2\lambda\Phi_{11} - 2\nu\Phi_{10} - \bar{\sigma}\Phi_{22} ,$$

$$\Delta\Psi_3 - \delta\Psi_4 + \bar{\delta}\Phi_{22} - \Delta\Phi_{21} = (4\beta - \tau)\Psi_4 - 2(2\mu + \gamma)\Psi_3 + 3\nu\Psi_2$$

$$+ (\bar{\tau} - 2\bar{\beta} - 2\alpha)\Phi_{22} + 2(\gamma + \bar{\mu})\Phi_{21} + 2\lambda\Phi_{12} - 2\nu\Phi_{11} - \bar{\nu}\Phi_{20} ,$$

$$D\Psi_2 - \bar{\delta}\Psi_1 + \Delta\Phi_{00} - \bar{\delta}\Phi_{01} + 2D\Lambda = -\lambda\Psi_0 + 2(\pi - \alpha)\Psi_1 + 3\rho\Psi_2$$
$$- 2\kappa\Psi_3 + (2\gamma + 2\bar{\gamma} - \bar{\mu})\Phi_{00} - 2(\bar{\tau} + \alpha)\Phi_{01} - 2\tau\Phi_{10} + 2\rho\Phi_{11} + \bar{\sigma}\Phi_{02}\,,$$
$$\Delta\Psi_2 - \delta\Psi_3 + D\Phi_{22} - \delta\Phi_{21} + 2\,\Delta\,\Lambda = \sigma\Psi_4 + 2(\beta - \tau)\Psi_3 - 3\mu\Psi_2 + 2\nu\Psi_1$$
$$+ (\bar{\rho} - 2\varepsilon - 2\bar{\varepsilon})\Phi_{22} + 2(\bar{\pi} + \beta)\Phi_{21} + 2\pi\Phi_{12} - 2\mu\Phi_{11} - \bar{\lambda}\Phi_{20}\,, \qquad \text{(E.15)}$$
$$D\Psi_3 - \bar{\delta}\Psi_2 - D\Phi_{21} + \delta\Phi_{20} - 2\bar{\delta}\Lambda = -\kappa\Psi_4 + 2(\rho - \varepsilon)\Psi_3 + 3\pi\Psi_2$$
$$- 2\lambda\Psi_1 + (2\bar{\alpha} - 2\beta - \bar{\pi})\Phi_{20} - 2(\bar{\rho} - \varepsilon)\Phi_{21} - 2\pi\Phi_{11} + 2\mu\Phi_{10} + \bar{\kappa}\Phi_{22}\,,$$
$$\Delta\Psi_1 - \delta\Psi_2 - \Delta\Phi_{01} + \bar{\delta}\Phi_{02} - 2\delta\Lambda = \nu\Psi_0 + 2(\gamma - \mu)\Psi_1 - 3\tau\Psi_2 + 2\sigma\Psi_3$$
$$+ (\bar{\tau} - 2\bar{\beta} + 2\alpha)\Phi_{02} + 2(\bar{\mu} - \gamma)\Phi_{01} + 2\tau\Phi_{11}\,,$$
$$D\Phi_{11} - \delta\Phi_{10} - \bar{\delta}\Phi_{01} + \Delta\Phi_{00} + 3D\Lambda = (2\gamma - \mu + 2\bar{\gamma} - \bar{\mu})\Phi_{00}$$
$$+ (\pi - 2\alpha - 2\bar{\tau})\Phi_{01} + (\bar{\pi} - 2\bar{\alpha} - 2\tau)\Phi_{10} + 2(\rho + \bar{\rho})\Phi_{11}$$
$$+ \bar{\sigma}\Phi_{02} + \sigma\Phi_{20} - \bar{\kappa}\Phi_{12} - \kappa\Phi_{21}\,,$$
$$D\Phi_{12} - \delta\Phi_{11} - \bar{\delta}\Phi_{02} + \Delta\Phi_{01} + 3\delta\Lambda = (-2\alpha + 2\bar{\beta} + \pi - \tau)\Phi_{02}$$
$$+ (\bar{\rho} + 2\rho - 2\bar{\varepsilon})\Phi_{12} + 2(\bar{\pi} - \tau)\Phi_{11} + (2\gamma - 2\bar{\mu} - \mu)\Phi_{01}$$
$$+ \bar{\nu}\Phi_{00} - \bar{\lambda}\Phi_{10} + \sigma\Phi_{21} - \kappa\Phi_{22}\,,$$
$$D\Phi_{22} - \delta\Phi_{21} - \bar{\delta}\Phi_{12} + \Delta\Phi_{11} + 3\,\Delta\,\Lambda = (\rho + \bar{\rho} - 2\varepsilon - 2\bar{\varepsilon})\Phi_{22}$$
$$+ (2\bar{\beta} + 2\pi - \bar{\tau})\Phi_{12} + (2\beta + 2\bar{\pi} - \tau)\Phi_{21} - 2(\mu + \bar{\mu})\Phi_{11}$$
$$+ \nu\Phi_{01} + \bar{\nu}\Phi_{10} - \bar{\lambda}\Phi_{20} - \lambda\Phi_{02}\,.$$

Appendix F

Wave Fields in a Curved Spacetime

F.1 Scalar Field

The action describing a free massive scalar field Φ is

$$W[\Phi] = -\frac{1}{8\pi}\int_V \left(g^{\mu\nu}\,\Phi_{,\mu}\,\Phi_{,\nu} + m^2\,\Phi^2 + \xi R\,\Phi^2\right)\sqrt{-g}\,d^4x$$

$$+\int_V J\Phi\,\sqrt{-g}\,d^4x + \frac{1}{4\pi}\,\xi\int_{\partial V} K\Phi^2\,\sqrt{h}\,d^3y\,, \tag{F.1}$$

where m is the field mass, and ξ is an arbitrary parameter. If $m = 0$, $J = 0$, and $\xi = 1/6$, the theory is conformally invariant, i.e., invariant under transformations

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \omega^2 g_{\mu\nu}\,, \qquad \Phi \to \tilde{\Phi} = \omega^{-1}\Phi\,. \tag{F.2}$$

The field equation takes the form

$$\Box\Phi - (m^2 + \xi R)\,\Phi = -\,4\pi J\,, \tag{F.3}$$

where

$$\Box \equiv \nabla^\mu\,\nabla_\mu = (-g)^{-1/2}\,\partial_\mu\left((-g)^{1/2}\,g^{\mu\nu}\,\partial_\nu\right). \tag{F.4}$$

The energy-momentum tensor is

$$T_{\mu\nu} = \frac{1}{4\pi}\left\{\Phi_{;\mu}\,\Phi_{;\nu} - \frac{1}{2}\,g_{\mu\nu}\,(\Phi_{;\alpha}\,\Phi^{;\alpha} + m^2\,\Phi^2)\right.$$

$$\left.+\xi\left[\left(R_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}\,R\right)\Phi^2 + g_{\mu\nu}\,(\Phi^2)^{;\alpha}{}_{;\alpha} - (\Phi^2)_{;\mu\nu}\right]\right\} + J\Phi\,g_{\mu\nu}\,. \tag{F.5}$$

Often (especially in quantum field theory) it is convenient to get rid of a multiplier $1/4\pi$ which enters the formulas. For this purpose, instead of Gaussian units one uses Heaviside units, which in our case is equivalent to the following transformations

$$\Phi = \sqrt{4\pi}\,\varphi\,, \quad J = \frac{1}{\sqrt{4\pi}}\,j\,. \tag{F.6}$$

In these units one has

$$W[\varphi] = -\frac{1}{2}\int_V \left(g^{\mu\nu}\,\varphi_{,\mu}\,\varphi_{,\nu} + m^2\,\varphi^2 + \xi R\,\varphi^2\right)\sqrt{-g}\,d^4x$$
$$+\int_V j\varphi\,\sqrt{-g}\,d^4x + \frac{1}{4\pi}\,\xi\int_{\partial V} K\varphi^2\,\sqrt{h}\,d^3y\,, \tag{F.7}$$

$$\Box\varphi - (m^2 + \xi R)\,\varphi = -j\,, \tag{F.8}$$

$$T_{\mu\nu} = T_{\mu\nu}[\,\varphi,\varphi\,] + j\,\varphi\,g_{\mu\nu}\,, \tag{F.9}$$

where

$$T_{\mu\nu}[\,\varphi^1,\varphi^2\,] \equiv \varphi^1_{;(\mu}\,\varphi^2_{;\nu)} - \frac{1}{2}\,g_{\mu\nu}\left(\varphi^1_{;\alpha}\,\varphi^{2;\alpha} + m^2\,\varphi^1\varphi^2\right)$$
$$+\,\xi\left[\left(R_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}\,R\right)\,\varphi^1\varphi^2 + g_{\mu\nu}\,(\varphi^1\varphi^2)^{;\alpha}{}_{;\alpha} - (\varphi^1\varphi^2)_{;\mu\nu}\right]. \tag{F.10}$$

The D'Alembert operator $\Box$ written in terms of Newman-Penrose variables reads

$$\Box \equiv -\,2\nabla_\mu\left(l^{(\mu}\,n^{\nu)} - m^{(\mu}\,\bar{m}^{\nu)}\right)\nabla_\nu = -(D - \rho - \bar{\rho})\Delta$$
$$-\,(\Delta + \mu + \bar{\mu} - \gamma - \bar{\gamma})D + (\delta + 2\beta - \tau)\bar{\delta} + (\bar{\delta} + 2\bar{\beta} - \bar{\tau})\delta\,. \tag{F.11}$$

F.2 Electromagnetic Field

The action describing the electromagnetic field A_μ is

$$W[\,A\,] = -\,\frac{1}{16\,\pi}\int_V F_{\mu\nu}\,F^{\mu\nu}\sqrt{-g}\,d^4x + \int_V A_\mu\,J^\mu\,\sqrt{-g}\,d^4x\,, \tag{F.12}$$

$$F_{\mu\nu} = 2\,A_{[\nu,\mu]}\,. \tag{F.13}$$

This action for $J = 0$ is conformally invariant, i.e., invariant under transformations

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \omega^2\,g_{\mu\nu}\,, \quad A_\mu \to \tilde{A}_\mu = A_\mu\,. \tag{F.14}$$

The Maxwell equations are

$$F^{\mu\nu}{}_{;\nu} = 4\pi J^\mu\,, \quad F_{[\mu\nu;\alpha]} = 0\,. \tag{F.15}$$

The energy-momentum tensor is

$$T_{\mu\nu} = \frac{1}{4\pi}\left(F_{\mu\alpha}\,F_\nu{}^\alpha - \frac{1}{4}\,g_{\mu\nu}\,F_{\alpha\beta}\,F^{\alpha\beta}\right). \tag{F.16}$$

In Heaviside units

$$A_\mu = \sqrt{4\pi}\,a_\mu\,, \quad F_{\mu\nu} = \sqrt{4\pi}\,f_{\mu\nu}\,, \quad J^\mu = \frac{1}{\sqrt{4\pi}}\,j^\mu, \tag{F.17}$$

the above relations take the form

$$W[\,a\,] = -\frac{1}{4}\int_V f_{\mu\nu}\,f^{\mu\nu}\sqrt{-g}\,d^4x + \int_V a_\mu\,j^\mu\,\sqrt{-g}\,d^4x\,, \tag{F.18}$$

$$f^{\mu\nu}{}_{;\nu} = j^\mu, \quad f_{\mu\nu} = 2\,a_{[\nu,\mu]}\,. \tag{F.19}$$

$$T_{\mu\nu} = T_{\mu\nu}[\,f,f\,], \quad T_{\mu\nu}[\,f^1,f^2\,] = f^1{}_{(\mu}{}^{\alpha}\,f^2_{\nu)\alpha} - \frac{1}{4}\,g_{\mu\nu}\,f^1_{\alpha\beta}\,f^{2\alpha\beta}. \tag{F.20}$$

Denote the components of the field strength $F_{\mu\nu}$ in the complex null tetrad $\mathbf{z}_m$ by

$$\Phi_0 = F_{\mu\nu}\,l^\mu\,m^\nu, \quad \Phi_1 = \frac{1}{2}\,F_{\mu\nu}(l^\mu\,n^\nu + \bar{m}^\mu\,m^\nu)\,, \quad \Phi_2 = F_{\mu\nu}\,\bar{m}^\mu\,n^\nu\,, \tag{F.21}$$

then the source-free Maxwell equations take the form

$$\begin{aligned}
D\Phi_1 - \bar{\delta}\Phi_0 &= (\pi - 2\alpha)\Phi_0 + 2\rho\Phi_1 - \kappa\Phi_2\,,\\
D\Phi_2 - \bar{\delta}\Phi_1 &= -\lambda\Phi_0 + 2\pi\Phi_1 + (\rho - 2\varepsilon)\Phi_2\,,\\
\delta\Phi_1 - \Delta\Phi_0 &= (\mu - 2\gamma)\Phi_0 + 2\tau\Phi_1 - \sigma\Phi_2\,,\\
\delta\Phi_2 - \Delta\Phi_1 &= -\nu\Phi_0 + 2\mu\Phi_1 + (\tau - 2\beta)\Phi_2\,.
\end{aligned} \tag{F.22}$$

F.3 Gravitational Perturbations

Let us consider perturbations $h_{\mu\nu}$ of the background gravitational filed $\overset{0}{g}_{\mu\nu}$. The corresponding metric $g_{\mu\nu}$ of the perturbed field is

$$g_{\mu\nu} = \overset{0}{g}_{\mu\nu} + h_{\mu\nu} \,. \tag{F.23}$$

By retaining terms up to the second order in $h_{\mu\nu}$, one gets

$$g^{\mu\nu} = \overset{0}{g}{}^{\mu\nu} - h^{\mu\nu} + h^{\mu\alpha} h_{\alpha}^{\ \nu} + \dots \,, \tag{F.24}$$

$$\sqrt{-g} = \left(-\overset{0}{g}\right)^{1/2} \left(1 + \frac{1}{2} h + \dots\right), \tag{F.25}$$

$$R_{\mu\nu} = \overset{0}{R}_{\mu\nu} + \overset{1}{R}_{\mu\nu} + \overset{2}{R}_{\mu\nu} + \dots \,, \tag{F.26}$$

$$\overset{1}{R}_{\mu\nu} = \frac{1}{2} \left(-h_{;\mu\nu} - h_{\mu\nu;\alpha}^{\quad\ \ ;\alpha} + h_{\alpha\mu;\nu}^{\quad\ \ ;\alpha} + h_{\alpha\nu;\mu}^{\quad\ \ ;\alpha}\right), \tag{F.27}$$

$$\begin{aligned} \overset{2}{R}_{\mu\nu} = \frac{1}{2} \Bigg[& \frac{1}{2} h_{\alpha\beta;\mu} h^{\alpha\beta}_{\ \ \ ;\nu} + h_{\nu}^{\ \alpha;\beta} (h_{\alpha\mu;\beta} - h_{\beta\mu;\alpha}) \\ & + h^{\alpha\beta} (h_{\alpha\beta;\mu\nu} + h_{\mu\nu;\alpha\beta} - h_{\alpha\mu;\nu\beta} - h_{\alpha\nu;\mu\beta}) \\ & - \left(h^{\alpha\beta}_{\ \ \ ;\beta} - \frac{1}{2} h^{;\alpha}\right) (h_{\alpha\mu;\nu} + h_{\alpha\nu;\mu} - h_{\mu\nu;\alpha}) \Bigg] \,, \end{aligned} \tag{F.28}$$

$$R = \overset{0}{R} + \left(\overset{1}{R} - h^{\alpha\beta} \overset{0}{R}_{\alpha\beta}\right) + \left(\overset{2}{R} - h^{\alpha\beta} \overset{1}{R}_{\alpha\beta} + h^{\alpha\mu} h_{\mu}^{\ \beta} \overset{0}{R}_{\alpha\beta}\right) + \dots \,. \tag{F.29}$$

The background metric $\overset{0}{g}_{\mu\nu}$ and its inverse $\overset{0}{g}{}^{\mu\nu}$ are used for operations with indices and the definition of the covariant derivatives.

The equations describing the gravitational perturbations can be obtained by variation of the exact Einstein equations. By introducing new variables

$$\tilde{h}_{\alpha\beta} = h_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} h \,, \tag{F.30}$$

one can write the equations for the gravitational perturbations in the form

$$\tilde{h}_{\alpha\beta}^{\quad ;\lambda}{}_{;\lambda} - \tilde{h}^{\lambda}{}_{\alpha;\beta;\lambda} - \tilde{h}^{\lambda}{}_{\beta;\alpha;\lambda} - \frac{1}{2} g_{\alpha\beta} \tilde{h}^{;\lambda}{}_{;\lambda} = 0 \,. \tag{F.31}$$

Appendix G

Wave Fields in the Kerr Metric

G.1 Teukolsky Equation

The homogeneous equations describing a wave field φ_A propagating in an external gravitational field are of the form

$$\mathcal{E}^{AB}\varphi_B = 0\,, \tag{G.1}$$

where $\mathcal{E}^{AB}$ is a covariant differential operator, and A and B represent collective tensorial indices. For example, for the electromagnetic field

$$\mathcal{E}^{\mu\nu}A_\nu = \nabla^\nu\nabla_\nu A^\mu - \nabla^\nu\nabla^\mu A_\nu\,,$$

and for gravitational perturbations

$$\mathcal{E}^{\mu\nu,\alpha\beta}\,h_{\alpha\beta} = -\,\nabla^\mu\nabla^\nu h^\alpha_\alpha - \nabla^\alpha\nabla_\alpha h^{\mu\nu} + \nabla^\alpha\nabla^\nu h^\mu_\alpha + \nabla^\alpha\nabla^\mu h^\nu_\alpha + g^{\mu\nu}\left(\nabla^\alpha\nabla_\alpha h^\beta_\beta - \nabla^\alpha\nabla^\beta h_{\alpha\beta}\right).$$

For massless fields in the Kerr metric these equations can be decoupled [Teukolsky (1973)]. This means that there exist three operators (we denote them by ${}_s\tau_A$, ${}_s\Pi^B$, and ${}_s\tilde{\Box}$) such that the following relation is valid

$${}_s\tau_A\,\mathcal{E}^{AB} = {}_s\tilde{\Box}\;{}_s\Pi^B. \tag{G.2}$$

Here $s = 0, \pm1/2, \pm1, \pm3/2, \pm2$, and $|s|$ is the spin of the field φ_A. This relation shows that the scalar ${}_s\psi = {}_s\Pi^B\varphi_B$ constructed for any solution of equation (G.1) obeys a scalar decoupled equation

$${}_s\tilde{\Box}\;{}_s\psi = 0\,. \tag{G.3}$$

Usually the covariant operator ${}_s\tilde{\Box}$ is represented in the form

$${}_s\tilde{\Box} = \Sigma^{-1}\,{}_s\Box\,. \tag{G.4}$$

(We recall that in the Kerr metric $\sqrt{-g} = \Sigma \sin\theta$.) The scalar second order differential operator ${}_s\Box$ was introduced by Teukolsky (1973). Its explicit form depends on the choice of coordinates and complex null tetrads. Using Boyer-Lindquist coordinates (D.1) and the Kinnersley tetrad (D.60), the operator ${}_s\Box$ is such that

$$\begin{aligned}
{}_s\Box\, {}_s\psi = &\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\frac{\partial^2{}_s\psi}{\partial t^2} + \frac{4aMr}{\Delta}\frac{\partial^2{}_s\psi}{\partial t\,\partial\phi} \\
&+ \left(\frac{a^2}{\Delta} - \frac{1}{\sin^2\theta}\right)\frac{\partial^2{}_s\psi}{\partial\phi^2} - \Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial_s\psi}{\partial r}\right) \\
&- \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial_s\psi}{\partial\theta}\right) - 2s\left[\frac{a(r-M)}{\Delta} + \frac{i\cos\theta}{\sin^2\theta}\right]\frac{\partial_s\psi}{\partial\phi} \\
&- 2s\left(\frac{M(r^2-a^2)}{\Delta} - r - ia\cos\theta\right)\frac{\partial_s\psi}{\partial t} + (s^2\cot^2\theta - s)\,{}_s\psi = 0\,.
\end{aligned} \tag{G.5}$$

This is the so-called *Teukolsky equation* [Teukolsky (1973)]. (We recall that $\Delta = r^2 - 2Mr + a^2$.)

A solution of the tensor field equation (G.1) can be constructed from the solutions of the related Teukolsky equation. It was demonstrated by Cohen and Kegeles (1974) for the electromagnetic field, and by Chrzanowski (1975) for gravitational perturbations. Wald (1978c) gave a simple proof of a general result which shows how to construct a solution to (G.1) once one has succeeded in deriving a decoupled equation. In order to describe the idea of this proof, let us denote by

$$(\psi, \varphi) \equiv \int \bar{\psi}_A\, \varphi^A\, d^4v \tag{G.6}$$

the scalar product of two (generally complex) tensor fields ψ_A and φ^A. Here $d^4v = \sqrt{-g}\, d^4x$ is the invariant volume element. The action $W[\varphi]$ for a real tensor field φ obeying equation (G.1) can then be written in the compact form

$$W[\varphi] = \frac{1}{2}(\varphi, \mathcal{E}\varphi)\,. \tag{G.7}$$

Further, denote by Q^* the operator which is conjugate to an operator Q with respect to the scalar product (G.6):

$$(\psi, Q\varphi) = (Q^*\psi, \varphi)\,. \tag{G.8}$$

The operator $\mathcal{E}$ in (G.7) is self-conjugate: $\mathcal{E}^* = \mathcal{E}$. By using this property, we can rewrite the equation obtained by conjugation of (G.2) in the form

$$\mathcal{E}^{AB}\, {}_s\tau_A^* = {}_s\Pi^{*B}\, {}_s\tilde{\Box}^*\,. \tag{G.9}$$

It can be shown that

$$ {}_s\tilde{\Box}^{*} = \Sigma^{-1}\ {}_{-s}\Box\,. \tag{G.10}$$

This relation shows that for any solution ${}_s\Xi$ of the scalar equation

$$ {}_s\tilde{\Box}^{*}\ {}_s\Xi = 0\,, \tag{G.11}$$

the tensor function

$$ {}_s\varphi_A = {}_s\tau_A^{*}\ {}_s\Xi \tag{G.12}$$

is a solution to the field equation (G.1). It can be shown that all solutions of equation (G.1) (up to possible gauge transformations) can be represented in this form. Thus, solutions of the scalar equation (G.11) provide complete information about the perturbing field. In the remainder of this appendix we collect the basic facts concerning the solutions of the Teukolsky equation and the procedure of reconstruction of the fields from these solutions. [For more details, see Press and Teukolsky (1973), Teukolsky and Press (1974), Chrzanowski (1975), Chrzanowski, Matzner and Sandberg (1976), Wald (1978), Candelas, Chrzanowski, and Howard (1981), Gal'tsov (1982, 1986), Futterman, Handler, and Matzner (1988).]

G.2 Separation of Variables. Spin-Weighted Spheroidal Harmonics

The coefficients of the Teukolsky equation (G.5)

$$ {}_s\Box\,{}_s\psi = 0 \tag{G.13}$$

do not depend on t and ϕ. Furthermore, the existence of a Killing tensor in the Kerr spacetimes results in the existence of an additional symmetry of the Teukolsky equation, which makes it possible to solve it by separation of variables. That is, one can use a mode decomposition

$$ {}_s\psi(t,r,\theta,\phi) = \int d\omega \sum_{\ell,m} {}_sR_{\ell m}(r,\omega)\,{}_sZ^{\omega}_{\ell m}(\theta,\phi)\,e^{-i\omega t}\,, \tag{G.14}$$

where ${}_sZ^{\omega}_{\ell m}(\theta,\phi)$ are the *spin-weighted spheroidal harmonics*

$$ {}_sZ^{\omega}_{\ell m}(\theta,\phi) = 2\pi)^{-1/2}\,{}_sS^{\omega}_{\ell m}(\theta)\,e^{im\phi}. \tag{G.15}$$

The angular problem reduces to solving the equation

$$ \frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dS}{d\theta}\right) + \left[a^2\omega^2\cos^2\theta - \frac{m^2}{\sin^2\theta} - 2a\omega s\cos\theta\right.$$

$$-\frac{2ms\cos\theta}{\sin^2\theta} - s^2\cot^2\theta + E - s^2\Bigg] S = 0\,, \tag{G.16}$$

where the function ${}_sS_{\ell m}(\theta)$ is regular on the interval $[0,\pi]$. The spin-weighted spheroidal harmonics are normalized by the condition

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \, {}_s\bar{Z}^{\omega}_{\ell' m'}(\theta,\phi)\, {}_sZ^{\omega}_{\ell m}(\theta,\phi) = \delta_{\ell\ell'}\,\delta_{mm'} \tag{G.17}$$

which means that

$$\int_0^{\pi} |\,{}_sS_{\ell m}(\theta)\,|^2 \sin\theta\, d\theta = 1\,. \tag{G.18}$$

Moreover, because of the symmetries of equation (G.16), one has

$${}_sZ^{\omega}_{\ell m}(\pi-\theta,\phi) \cong {}_{-s}Z^{\omega}_{\ell m}(\theta,\phi)\,, \tag{G.19}$$

$${}_s\bar{Z}^{\omega}_{\ell m}(\theta,\phi) \cong {}_{-s}Z^{-\omega}_{\ell -m}(\theta,\phi)\,. \tag{G.20}$$

Here and later in this appendix, we use the symbol $\cong$ to indicate that a relation is valid up to some arbitrary multiplicative constant. For the chosen fixed normalization (G.17) this constant is just a phase factor.

The required functions ${}_sS_{\ell m}(\theta)$ thus essentially follow from a Sturm-Liouville eigenvalue problem for the separation constant E. Boundary conditions of regularity should be imposed at both $\theta = 0$ and $\theta = \pi$. According to Sturm-Liouville theory, the eigenfunctions form a complete, orthogonal set on the interval $0 \le \theta \le \pi$ for each combination of s, $a\,\omega$, and m. This infinite set of eigenfunctions is labeled by ℓ.

For scalar fields ($s = 0$) the desired functions are the spheroidal wave functions that were studied in detail by Flammer (1957). For Schwarzschild black holes (where $a = 0$), the required solutions are the spin-weighted spherical harmonics ${}_sY_{\ell m}$. They can be expressed in terms of standard spherical harmonics [Goldberg *et al.* (1967)]. As we have already mentioned, the general functions are referred to as "spin-weighted spheroidal harmonics". For these, the problem was first considered by Press and Teukolsky (1973). They derived an expression for the eigenvalue E in the case when $a\,\omega$ is sufficiently small to be considered a perturbation. This expression was independently derived by Starobinsky and Churilov (1973). For rapidly rotating black holes, or higher frequencies, Press and Teukolsky resorted to numerical integration of (G.16).

A general study of the analytic properties of spin-weighted spheroidal harmonics was presented by Fackerell and Crossman (1977). They expressed the eigenfunctions as a series involving Jacobi polynomials. This series only converges for certain eigenvalues, and these lead directly to the desired separation constant for the black-hole problem. In practice, one obtains a continued fraction relation that determines the eigenvalue. A great advantage of this approach is that it can be used also for complex frequencies. Thus, it is applicable also when one considers quasinormal modes of a Kerr black hole (see section 4.4). Moreover, Fackerell and Crossman (1977) derived an expansion for the eigenvalue up to and including terms of order $(a\omega)^6$. A similar expansion, together with an alternative expansion for high frequencies, was obtained by Breuer, Ryan, and Waller (1977). Their high-frequency formula includes terms of order $(a\omega)^{-3}$. Unfortunately, there are discrepancies between the various results. This means that all these equations must be used with some care. In order to improve the situation, Seidel (1989) used the Fackerell-Crossman approach to rederive the small-$a\omega$ expansion. Having done this, he points out various misprints in the previous studies and presents a formula that should be reliable. The first two terms of this expansion are (for $s \neq 0$)

$$E \approx \ell(\ell+1) - \frac{2ma\omega s^2}{\ell(\ell+1)} + \text{ terms of order } (a\omega)^2 . \tag{G.21}$$

In most studies of rotating black holes, the angular functions have been obtained by numerical solution of (G.16), however. Examples of this include the quasinormal-mode calculations of Leaver (1985) and the study of wave scattering by Handler and Matzner (1980).

G.3 Radial Equation

The radial functions ${}_sR_{\ell m}$ also obey a second order differential equation. The exact form of this equation depends both on the choice of tetrad and the choice of coordinate system. For the Kinnersley tetrad and Boyer-Lindquist coordinates the radial equation is [Teukolsky (1973)]

$$\Delta^{-s}\frac{d}{dr}\left(\Delta^{s+1}\frac{dR}{dr}\right) + \left[\frac{K^2 - 2is(r-M)K}{\Delta} + 4ir\omega s - \lambda\right] R = 0 , \tag{G.22}$$

where

$$K \equiv (r^2+a^2)\omega - am , \qquad \lambda \equiv E - 2am\omega + a^2\omega^2 - s(s+1) . \tag{G.23}$$

By using the symmetries of the radial differential equation, we get the following relations (valid only for real frequencies)

$${}_sR_{\ell m}(r,\omega) \cong {}_s\bar{R}_{\ell -m}(r,-\omega) , \tag{G.24}$$

$${}_s\bar{R}_{\ell m}(r,\omega) \cong \Delta^{-s}\, {}_{-s}R_{\ell m}(r,\omega) . \tag{G.25}$$

An important property of the radial Teukolsky equation is that the two solutions ${}_sR_{\ell m}$ and ${}_{-s}R_{\ell m}$ are related. The exact form of these relations, known as the *Teukolsky-Starobinsky identities*, depends on the value of $|s|$. For gravitational perturbations one has

$$ {}_{+2}R_{\ell m} \cong \mathcal{D}\mathcal{D}\mathcal{D}\mathcal{D}\, {}_{-2}R_{\ell m}\,, \tag{G.26} $$

$$ {}_{-2}R_{\ell m} \cong \Delta^2\, \mathcal{D}^\dagger\, \mathcal{D}^\dagger \mathcal{D}^\dagger\, \mathcal{D}^\dagger\left(\Delta^2\, {}_{+2}R_{\ell m}\right). \tag{G.27} $$

Here

$$ \mathcal{D} = \partial_r - \frac{iK}{\Delta}\,, \quad \text{and} \quad \mathcal{D}^\dagger = \partial_r + \frac{iK}{\Delta}\,. \tag{G.28} $$

Similar relations for other values of s are given by Chandrasekhar (1983). As before, the symbol $\cong$ indicates that the constants of proportionality in these relations are not fixed. But if we require that subsequent use of both transformations (for example, first transforming ${}_2R$ into ${}_{-2}R$, and then the resultant ${}_{-2}R$ into ${}_2R$) should lead to an identity, then the product of the two corresponding constants ($\mathcal{C}$ and $\bar{\mathcal{C}}$, say) will be fixed. In the case of gravitational perturbations this leads to

$$ |\mathcal{C}|^2 = \lambda^2(\lambda+2)^2 - 8\,\omega^2\lambda[\alpha^2(5\lambda+6) - 12a^2)] + 144\,\omega^4\alpha^4 + 144\,\omega^2 M^2\,, \tag{G.29} $$

where

$$ \alpha^2 = a^2 - \frac{a\,m}{\omega}\,. \tag{G.30} $$

Solutions to the equations describing a free massless field in the exterior of a Kerr black hole can be specified by giving their asymptotic forms at null infinity and the horizon. It is convenient to introduce four sets of modes which are called *IN*-, *UP*-, *OUT*-, and *DOWN*-modes, correspondingly. Each of the modes is a solution characterized by the set $\{\ell m\omega P\}$ of quantum numbers, where the parity $P = \pm 1$. These modes are singled out by the following requirements: *IN-modes* vanish at H^- and have non-vanishing asymptotic values on $\mathcal{J}^-$. *UP-modes* vanish at $\mathcal{J}^-$ and are non-vanishing on the past horizon H^-. *OUT-modes* vanish at H^+ but not at $\mathcal{J}^+$, and finally *DOWN-modes* vanish at $\mathcal{J}^+$ and are non-vanishing on the past horizon H^+. Any two of these four solutions can be used as a complete set in the space of solutions. They are uniquely specified by the choice of boundary conditions for the corresponding radial equation (G.22). Let us discuss these conditions in more detail.

Introducing a new dependent variable

$$ {}_s\chi_{\ell m} = (r^2+a^2)^{1/2}\,\Delta^{s/2}\,{}_sR_{\ell m}\,, \tag{G.31} $$

the radial equation (G.8) can be written in the form

$$ \frac{d_s\chi_{\ell m}}{dr_*{}^2} + {}_sV_{\ell m}\,{}_s\chi_{\ell m} = 0\,. \tag{G.32} $$

Here the complex effective potential ${}_sV_{\ell m}$ is

$$ {}_sV_{\ell m}(r,\omega) = \frac{K^2 - 2is(r-M)K + \Delta(4i\,\omega s r - \lambda)}{(r^2+a^2)^2} - G^2 - \frac{dG}{dr_*}, \tag{G.33} $$

$$ G = \frac{r\Delta}{(r^2+a^2)^2} + \frac{s(r-M)}{r^2+a^2}, \tag{G.34} $$

and r_* is the "tortoise" coordinate defined by

$$ dr_* = \frac{r^2+a^2}{\Delta}\,dr. \tag{G.35} $$

In the asymptotic regions ($r_* \to \pm\infty$) the potential ${}_sV_{\ell m}$ has the form

$$ {}_sV_{\ell m}(r,\omega) = \begin{cases} \omega(\omega + 2is/r), & r_* \to \infty, \\ \varpi_s^2, & r_* \to -\infty. \end{cases} \tag{G.36} $$

Here

$$ \varpi_s = \varpi - is\kappa, \quad \varpi = \omega - m\,\Omega^H, \tag{G.37} $$

and $\kappa = (r_+ - M)/(r_+^2 + a^2)$ is the surface gravity of the Kerr black hole. Linearly independent solutions to (G.32) have the asymptotic behavior $\sim r^{\mp s}\exp(\pm i\omega r_*)$ at $r \to \infty$ and $\sim \Delta^{\pm s/2}\exp(\pm i\varpi r_*)$ at $r \to r_+$.

We define two sets of solutions ${}_s\chi^{\rm in}_{\ell m}$ and ${}_s\chi^{\rm up}_{\ell m}$ by the boundary conditions

$$ {}_{-|s|}\chi^{\rm in}_{\ell m}(r,\omega) \sim \begin{cases} {}_{-|s|}\tau^{\rm in}_{\ell m\omega}\,\Delta^{|s|/2}\,e^{-i\varpi r_*} & \text{as } r \to r_+, \\ r^{-|s|}\,e^{-i\omega r_*} + {}_{-|s|}\rho^{\rm in}_{\ell m\omega}\,r^{+|s|}\,e^{+i\omega r_*} & \text{as } r \to \infty, \end{cases} \tag{G.38} $$

and

$$ {}_{+|s|}\chi^{\rm up}_{\ell m}(r,\omega) \sim \begin{cases} {}_{+|s|}\rho^{\rm up}_{\ell m\omega}\,\Delta^{-|s|/2}\,e^{-i\varpi r_*} + \Delta^{+|s|/2}\,e^{+i\varpi r_*} & \text{as } r \to r_+, \\ {}_{+|s|}\tau^{\rm up}_{\ell m\omega}\,r^{-|s|}\,e^{+i\omega r_*} & \text{as } r \to \infty. \end{cases} \tag{G.39} $$

The choice $s = -|s|$ for *IN*-modes and $s = |s|$ for *UP*-modes is dictated by the necessity to exclude the slowest decreasing (or fastest growing) asymptotic behavior at $r \to r_+$ (*IN*-modes) or at $r \to \infty$ (*UP*-modes).

The quantities ${}_{-|s|}R^{\rm in}_{\ell m}(r,\omega)$ and ${}_{+|s|}R^{\rm up}_{\ell m}(r,\omega)$ defined by equation (G.31) are thus particular solutions of the radial equation (G.22) that satisfy the boundary conditions

$$ {}_{-|s|}R^{\rm in}_{\ell m}(r,\omega) \sim \begin{cases} {}_{-|s|}\mathcal{T}^{\rm in}_{\ell m\omega}\,\Delta^{|s|}\,e^{-i\varpi r_*} & \text{as } r \to r_+, \\ r^{-1}\,e^{-i\omega r_*} + {}_{-|s|}\mathcal{R}^{\rm in}_{\ell m\omega}\,r^{2|s|-1}\,e^{+i\omega r_*} & \text{as } r \to \infty, \end{cases} \tag{G.40} $$

and

$$
{}_{+|s|}R^{\text{up}}_{\ell m}(r,\omega) \sim \begin{cases} {}_{+|s|}\mathcal{R}^{\text{up}}_{\ell m\omega}\, \Delta^{-|s|}\, e^{-i\varpi r_*} + e^{+i\varpi r_*} & \text{as } r \to r_+\,, \\ {}_{+|s|}\mathcal{T}^{\text{up}}_{\ell m\omega}\, r^{-2|s|-1}\, e^{+i\omega r_*} & \text{as } r \to \infty\,. \end{cases} \tag{G.41}
$$

The solutions are uniquely specified by these conditions. We can interpret the solution ${}_{-|s|}R^{\text{in}}_{\ell m}(r,\omega)\, e^{-i\omega t}$ as a (unit-amplitude) spherical wave propagating *inwards* from infinity and being partially reflected back out to infinity and partially transmitted across the horizon, whereas ${}_{+|s|}R^{\text{up}}_{\ell m}(r,\omega)\, e^{-i\omega t}$ represents a (unit-amplitude) spherical wave propagating *up* from the past horizon, and being partially reflected back and partially transmitted out to infinity. The situation is depicted in Figure 4.2.

One can show that the potential ${}_sV_{\ell m}$ for λ defined by (G.33) with E fixed possesses the symmetry

$$
{}_s\bar{V}_{\ell m}(r,\omega) = {}_{-s}V_{\ell m}(r,\omega)\,. \tag{G.42}
$$

That is, if ${}_s\chi_{\ell m}$ is a solution of equation (G.32), then ${}_{-s}\bar{\chi}^{\text{in}}_{\ell m}$ is a solution to the same equation. This means that one can define the remaining two solutions (*OUT-modes* and *DOWN-modes*) in terms of the known ones:

$$
{}_{|s|}\chi^{\text{out}}_{\ell m}(r,\omega) = {}_{-|s|}\bar{\chi}^{\text{in}}_{\ell m}(r,\omega)\,, \tag{G.43}
$$

$$
{}_{-|s|}\chi^{\text{down}}_{\ell m}(r,\omega) = {}_{|s|}\bar{\chi}^{\text{up}}_{\ell m}(r,\omega)\,. \tag{G.44}
$$

The corresponding radial functions ${}_sR^{\text{out}}_{\ell m}$ and ${}_sR^{\text{down}}_{\ell m}$ are given by (G.31), while ${}_s\psi^{\text{out}}$ and ${}_s\psi^{\text{down}}$ are defined by (G.14).

G.4 Massless Scalar Field

The case of a massless scalar field is technically the simplest one. Consider a scalar field φ obeying the equation

$$
\Box\varphi = 0\,. \tag{G.45}
$$

It is easy to verify that

$$
{}_{s=0}\Box = -\Sigma\,\Box\,, \tag{G.46}
$$

so that the solutions to equation (G.45) simply coincide with the solutions of the Teukolsky equation for $s=0$.

To make the consideration more concrete, we fix from now on the phase factors that enter (G.20) and (G.24) as follows [Chrzanowski (1975)]:

$$
{}_s\bar{Z}^{\omega}_{\ell m}(\theta,\phi) = (-1)^m\, {}_{-s}Z^{-\omega}_{\ell -m}(\theta,\phi)\,, \quad {}_sR_{\ell m}(r,\omega) = (-1)^m\, {}_s\bar{R}_{\ell -m}(r,-\omega)\,. \tag{G.47}
$$

We define

$$\varphi^{\rm in}(\ell,m,\omega;x) = N_0^{\rm in}\,{}_0Z^{\omega}_{\ell m}(\theta,\phi)\,{}_0R^{\rm in}_{\ell m}(r,\omega)\,e^{-i\omega t}, \tag{G.48}$$

$$\varphi^{\rm up}(\ell,m,\omega;x) = N_0^{\rm up}\,{}_0Z^{\omega}_{\ell m}(\theta,\phi)\,{}_0R^{\rm up}_{\ell m}(r,\omega)\,e^{-i\omega t}. \tag{G.49}$$

Relations (G.47) imply

$$\bar{\varphi}^{\rm in}(\ell,m,\omega;x) = \varphi^{\rm in}(\ell,-m,-\omega;x)\,, \tag{G.50}$$

$$\bar{\varphi}^{\rm up}(\ell,m,\omega;x) = \varphi^{\rm up}(\ell,-m,-\omega;x)\,, \tag{G.51}$$

provided

$$\bar{N}^{\rm in}(m,\omega) = N^{\rm in}(-m,-\omega)\,, \quad \bar{N}^{\rm up}(m,\omega) = N^{\rm up}(-m,-\omega)\,. \tag{G.52}$$

We define the inner product for a pair ψ and ϕ of complex solutions of (G.45) as

$$<\psi,\phi> = i\int_{\Sigma} d\sigma^{\mu}\left(\bar{\psi}\,\phi_{;\mu} - \phi\,\bar{\psi}_{;\mu}\right). \tag{G.53}$$

Σ is a complete spacelike hypersurface in the exterior Kerr spacetime which passes through the intersection of the horizons $H^-\cup H^+$ and spatial infinity I^0, and $d\sigma^{\mu}$ is a future-directed normal to Σ. One can choose t=const as the equation of Σ. In the limit $t\to-\infty$ this surface coincides with $H^-\cup\mathcal{J}^-\cup\tilde{\Sigma}$, where $\tilde{\Sigma}$ is the space-region of the infinite past. For massless fields wavepacket-like solutions vanish on $\tilde{\Sigma}$, so that one can write

$$\int_{\Sigma} = \int_{H^-} + \int_{\mathcal{J}^-}\,.$$

The integrals over H^- and $\mathcal{J}^-$ can be evaluated explicitly since one knows the form of the radial functions ${}_0R^{\rm in}_{\ell m}$, ${}_0R^{\rm up}_{\ell m}$ at both the horizon and spatial infinity (cf. (G.40) and (G.41)). In a similar way one can move the surface Σ to the future and rewrite the integral in terms of integration over $H^+\cup\mathcal{J}^+$.

It is easy to verify that the coefficients $N_0^{\rm in}$ and $N_0^{\rm up}$ in (G.48) and (G.49) can be chosen so that

$$<\varphi^{\Lambda}(\ell,m,\omega;x),\varphi^{\Lambda'}(\ell',m',\omega';x)> = \pi^{\Lambda}\,\delta_{\Lambda\Lambda'}\,\delta_{\ell\ell'}\,\delta_{mm'}\,\delta(\omega-\omega')\,, \tag{G.54}$$

where $\Lambda\in\{\rm in, up\}$, and π^{Λ} is the sign-function

$$\pi^{\rm in}(\omega,m) = \begin{cases} +1 & \text{if } \omega>0\,, \\ -1 & \text{if } \omega<0\,, \end{cases} \qquad \pi^{\rm up}(\omega,m) = \begin{cases} +1 & \text{if } \varpi>0\,, \\ -1 & \text{if } \varpi<0\,. \end{cases} \tag{G.55}$$

Solutions (G.48) and (G.49) for $(-\infty<\omega<\infty)$ form the complete set of modes. This set can be used to decompose any solution of the scalar field equation (G.45) in the exterior of the Kerr black hole. Instead of this set, it is more convenient to use another one that is defined as

$$\{\varphi^\Lambda(x), \bar{\varphi}^\Lambda(x)\}_{\ell m\omega} \tag{G.56}$$

where $\Lambda \in \{\text{in}, \text{up}\}$, $\omega > 0$ for *IN*-modes, and $\varpi > 0$ for *UP*-modes. The equivalence of the two sets of modes follows from (G.50) and (G.51). For the basis (G.56)

$$< \varphi^\Lambda(\ell, m, \omega; x), \varphi^{\Lambda'}(\ell', m', \omega'; x) > = \delta_{\Lambda\Lambda'}\, \delta_{\ell\ell'}\, \delta_{mm'}\, \delta(\omega - \omega')\,; \tag{G.57}$$

the other inner products vanish. A solution φ of the field equation (G.45) can be written as

$$\begin{aligned} \varphi(x) &= \sum_J \left[c_J\, \varphi(J; x) + \bar{c}_J\, \bar{\varphi}(J; x)\right] &\text{(G.58)}\\ &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \int_0^\infty d\omega \left[c_{\text{in},\ell m\omega}\, \varphi^{\text{in}}(\ell m\omega; x) + \bar{c}_{\text{in},\ell m\omega}\, \bar{\varphi}^{\text{in}}(\ell m\omega; x)\right] \\ &+ \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \int_{m\Omega^H}^\infty d\omega \left[c_{\text{up},\ell m\omega}\, \varphi^{\text{up}}(\ell m\omega; x) + \bar{c}_{\text{up},\ell m\omega}\, \bar{\varphi}^{\text{up}}(\ell m\omega; x)\right]. &\text{(G.59)} \end{aligned}$$

Here J is a shorthand for $\{\Lambda, \ell, m, \omega\}$, and Λ takes values "in" and "up".

Another useful complete set of modes (out-basis) can be constructed by using the solutions

$$\varphi^{\text{out}}(\ell, m, \omega; x) = N_0^{\text{out}}\, {}_0Z^\omega_{\ell m}(\theta, \phi)\, {}_0\bar{R}^{\text{in}}_{\ell m}(r, \omega)\, e^{-i\omega t}, \tag{G.60}$$

$$\varphi^{\text{down}}(\ell, m, \omega; x) = N_0^{\text{down}}\, {}_0Z^\omega_{\ell m}(\theta, \phi)\, {}_0\bar{R}^{\text{up}}_{\ell m}(r, \omega)\, e^{-i\omega t}. \tag{G.61}$$

These solutions obey relations similar to (G.50)–(G.51). The set

$$\{\varphi^\Lambda(x), \bar{\varphi}^\Lambda(x)\}_{\ell m\omega} \tag{G.62}$$

where $\Lambda \in \{\text{out}, \text{down}\}$, $\omega > 0$ for *OUT*-modes, and $\varpi > 0$ for *DOWN*-modes obey the normalization conditions (G.57) and can be used to obtain a decomposition of the solution φ similar to (G.58).

G.5 Electromagnetic Field

The procedure of construction of the bases for higher spins is similar to the one used for the scalar massless field. Once solutions to the Teukolsky equation are known, one can use them to obtain solutions of the field equations. The basic relations for the electromagnetic field are collected in this section. The gravitational perturbations are considered in the next section. [For more details, see Chrzanowski (1975), Wald (1978), Candelas, Chrzanowski, and Howard (1981), Gal'tsov (1982, 1986), Futterman, Handler, and Matzner (1988).]

For the electromagnetic field propagating in the exterior of the eternal Kerr black hole we define a complete set of complex modes as

$$\{a_\mu^\Lambda(x), \bar{a}_\mu^\Lambda(x)\}_{\ell m \omega P}, \tag{G.63}$$

where $\Lambda \in \{\mathrm{in}, \mathrm{up}\}$ and $P = \pm 1$. The explicit form of the modes was obtained by Chrzanowski (1975).

When $\Lambda = \mathrm{in}$, one has ($\omega > 0$)

$$a_\mu^{\mathrm{in}}(\ell, m, \omega, P; x) =$$

$$N_1^{\mathrm{in}} \left\{\Theta_\mu \ {}_{+1}Z_{\ell m}^\omega(\theta, \phi) + P\bar{\Theta}_\mu \ {}_{-1}Z_{\ell m}^\omega(\theta, \phi)\right\} \ {}_{-1}R_{\ell m}^{\mathrm{in}}(r, \omega)\, e^{-i\omega t} \tag{G.64}$$

(in the *ingoing radiation gauge* $a_\nu l^\nu = 0$).

When $\Lambda = \mathrm{up}$, one has ($\varpi > 0$)

$$a_\mu^{\mathrm{up}}(\ell, m, \omega, P; x) =$$

$$N_1^{\mathrm{up}} \left\{\Upsilon_\mu \ {}_{-1}Z_{\ell m}^\omega(\theta, \phi) + P\bar{\Upsilon}_\mu \ {}_{+1}Z_{\ell m}^\omega(\theta, \phi)\right\} \ {}_{+1}R_{\ell m}^{\mathrm{up}}(r, \omega)\, e^{-i\omega t} \tag{G.65}$$

(in the *outgoing radiation gauge* $a_\nu n^\nu = 0$). Here, N_1^Λ are constants that will be fixed later, and Θ_μ, Υ_μ are the first-order differential operators:

$$\Theta_\mu = -l_\mu(\bar{\delta} + 2\bar{\beta} + \bar{\tau}) + \bar{m}_\mu(D + \bar{\rho}),$$

$$\Upsilon_\mu = \bar{\rho}^{\,-2}\left[n_\mu(\delta + \bar{\pi} - 2\bar{\alpha}) - m_\mu(\Delta + \bar{\mu} - 2\bar{\gamma})\right].$$

The fact that the "in"- and "up"-modes (G.64) and (G.65) are expressed in different gauges does not cause any difficulty because only gauge-independent quantities constructed from them are actually used.

Using relations (G.47), one can verify that

$$\bar{a}_\mu^{\mathrm{in}}(\ell, m, \omega; x) = P a_\mu^{\mathrm{in}}(\ell, -m, -\omega; x), \tag{G.66}$$

$$\bar{a}_\mu^{\mathrm{up}}(\ell, m, \omega; x) = P a_\mu^{\mathrm{up}}(\ell, -m, -\omega; x), \tag{G.67}$$

provided N_1^{in} and N_1^{up} obey (G.52).

Having constructed mode solutions for the vector potential a_μ, one can derive the corresponding expansions for the scalar field-strengths. The operators ${}_{\pm 1}\Pi^{\mu\nu}$ relating ${}_{\pm 1}\psi$ to the potential a_μ, i.e.,

$${}_{\pm 1}\psi = {}_{\pm 1}\Pi^\mu\, a_\mu, \tag{G.68}$$

have the following form

$$_1\Pi^\mu = (\delta - \beta - \bar{\alpha} - 2\tau + \bar{\pi})\, l^\mu - (D - 2\rho - \bar{\rho})\, m^\mu ,$$

$$_{-1}\Pi^\mu = \rho^{-2}\left[(\Delta + \gamma - \bar{\gamma} + 2\mu + \bar{\mu})\, \bar{m}^\mu - (\bar{\delta} + \alpha + \bar{\beta} + 2\pi - \bar{\tau})\, n^\mu\right].$$

The spin coefficients for the Kerr metric are given by (D.64). The mode solutions for *DOWN*- and *OUT*-modes are defined similarly.

In the application of the mode decomposition to the quantum theory it is convenient to fix the constants N_1^Λ appearing in (G.64) and (G.65) by demanding that the mode set satisfies the orthonormality condition

$$< a^\Lambda(\ell, m, \omega, P; x), a^{\Lambda'}(\ell', m', \omega', P'; x) > = \delta_{\Lambda\Lambda'}\, \delta_{\ell\ell'}\, \delta_{mm'}\, \delta(\omega - \omega')\, \delta_{PP'}\, ; \quad \text{(G.69)}$$

the other inner products vanish. Here, the inner product $< \psi, \phi >$ for complex vector fields ψ_α and $\phi_\alpha$is defined as follows:

$$< \psi, \phi > = 2i \int_\Sigma \mathrm{d}\sigma^\mu \left(\bar{\psi}_{[\mu,\alpha]}\, \phi^\alpha - \phi_{[\mu,\alpha]}\, \bar{\psi}^\alpha\right). \quad \text{(G.70)}$$

By using Stokes' formula (A.87), it is easy to verify that for solutions that decrease at infinity rapidly enough this definition of the inner product is identical to equation (2.40) of [Candelas, Chrzanowski, and Howard (1981)]. Σ is a complete spacelike hypersurface in the exterior Kerr spacetime, and $d\sigma^\mu$ is the future-directed normal to Σ. As in the case of the electromagnetic field, the surface of integration Σ can be moved either to the infinite future or to the infinite past, so that the integration is performed either over $H^- \cup \mathcal{J}^-$ or over $H^+ \cup \mathcal{J}^+$, where the asymptotics of the fields are known.

An arbitrary electromagnetic wave propagating in the exterior of the Kerr metric can be expanded in terms of the modes (G.64) and (G.65)

$$a_\mu(x) = \sum_J \left\{c_J\, a_\mu(J; x) + \bar{c}_J\, \bar{a}_\mu(J; x)\right\}. \quad \text{(G.71)}$$

Here J is a shorthand for $\{\Lambda, \ell, m, \omega, P\}$, and Λ takes values "in" and "up". A similar decomposition is valid if one use the set of *OUT*- and *DOWN*-modes.

G.6 Gravitational Perturbations

For gravitational perturbations propagating in the exterior of the eternal Kerr black hole we define a complete set of complex modes as

$$\left\{h^\Lambda_{\mu\nu}(x), \bar{h}^\Lambda_{\mu\nu}(x)\right\}_{\ell m\omega P} \quad \text{(G.72)}$$

where $\Lambda \in \{\text{in}, \text{up}\}$ and $P = \pm 1$.

The explicit form of the modes for $\Lambda = \text{in}$ is $(\omega > 0)$

$$h^{\text{in}}_{\mu\nu}(\ell, m, \omega, P; x) =$$

$$N_2^{\text{in}} \left\{ \Theta_{\mu\nu}\ {}_{+2}Z^{\omega}_{\ell m}(\theta,\phi) + P\bar{\Theta}_{\mu\nu}\ {}_{-2}Z^{\omega}_{\ell m}(\theta,\phi) \right\}\ {}_{-2}R^{\text{in}}_{\ell m}(r,\omega)\, e^{-i\omega t} \tag{G.73}$$

(in the *ingoing radiation gauge* $h_{\mu\nu}\, l^{\nu} = 0,\ h_{\nu}{}^{\nu} = 0$).

When $\Lambda = \text{up}$, we have $(\varpi > 0)$

$$h^{\text{up}}_{\mu\nu}(\ell, m, \omega, P; x) =$$

$$N_2^{\text{up}} \left\{ \Upsilon_{\mu\nu}\ {}_{-2}Z^{\omega}_{\ell m}(\theta,\phi) + P\bar{\Upsilon}_{\mu\nu}\ {}_{+2}Z^{\omega}_{\ell m}(\theta,\phi) \right\}\ {}_{+2}R^{\text{up}}_{\ell m}(r,\omega)\, e^{-i\omega t} \tag{G.74}$$

(in the *outgoing radiation gauge* $h_{\mu\nu}\, n^{\nu} = 0,\ h_{\nu}{}^{\nu} = 0$).

Here, N_2^{Λ} are constants which will be fixed later, and $\Theta_{\mu\nu}$, $\Upsilon_{\mu\nu}$ are the second-order differential operators [Chrzanowski (1975)]:

$$\begin{aligned}
\Theta_{\mu\nu} \;=\; & -l_{\mu}\, l_{\nu}(\bar{\delta} + \alpha + 3\bar{\beta} - \bar{\tau})(\bar{\delta} + 4\bar{\beta} + 3\bar{\tau}) \\
& - \bar{m}_{\mu}\, \bar{m}_{\nu}(D - \bar{\rho})(D + 3\bar{\rho}) \\
& + l_{(\mu}\, \bar{m}_{\nu)}[(D + \rho - \bar{\rho})(\bar{\delta} + 4\bar{\beta} + 3\bar{\tau}) \\
& + (\bar{\delta} + 3\bar{\beta} - \alpha - \pi - \bar{\tau})(D + 3\bar{\rho})]\,,
\end{aligned}$$

$$\begin{aligned}
\Upsilon_{\mu\nu} \;=\; & \bar{\rho}^{\,-4}\{-n_{\mu} n_{\nu}(\delta - 3\bar{\alpha} - \beta + 5\bar{\pi})(\delta - 4\bar{\alpha} + \bar{\pi}) \\
& - m_{\mu}\, m_{\nu}(\Delta + 5\bar{\mu} - 3\bar{\gamma} + \gamma)(\Delta + \bar{\mu} - 4\bar{\gamma}) \\
& + n_{(\mu}\, m_{\nu)}[(\delta + 5\bar{\pi} + \beta - 3\bar{\alpha} + \tau)(\Delta + \bar{\mu} - 4\bar{\gamma}) \\
& + (\Delta + 5\bar{\mu} - \mu - 3\bar{\gamma} - \gamma)(\delta - 4\bar{\alpha} + \bar{\pi})]\}\,.
\end{aligned}$$

The fact that the *IN*- and *UP*-modes (G.73) and (G.74) are expressed in different gauges does not cause any difficulty because only gauge independent quantities constructed from them are actually used. Chrzanowski's mode set comprises a complete set of solutions to the perturbed field equations for the Kerr spacetime.

Using relations (G.47), one can verify that

$$\bar{a}^{\text{in}}_{\mu}(\ell, m, \omega; x) = P a^{\text{in}}_{\mu}(\ell, -m, -\omega; x)\,, \tag{G.75}$$

$$\bar{a}^{\text{up}}_{\mu}(\ell, m, \omega; x) = P a^{\text{up}}_{\mu}(\ell, -m, -\omega; x)\,, \tag{G.76}$$

provided N_1^{in} and N_1^{up} obey (G.52).

Having constructed mode solutions for the metric perturbations, one can derive the corresponding expansions for the Weyl scalars. The operators ${}_{\pm 2}\Pi^{\mu\nu}$ relating ${}_{\pm 2}\psi$ to the metric perturbations $h_{\mu\nu}$, i.e.,

$${}_{\pm 2}\psi = {}_{\pm 2}\Pi^{\mu\nu}\, h_{\mu\nu}\,, \tag{G.77}$$

have the following form

$$\begin{aligned}
{}_2\Pi^{\mu\nu} = & -\frac{1}{2}\{(D-\bar{\rho})(D-\bar{\rho})\, m^\mu m^\nu \\
& +(\delta+\bar{\pi}-3\beta-\bar{\alpha})(\delta+\bar{\pi}-2\beta-2\bar{\alpha})\, l^\mu l^\nu \\
& -[(D-\bar{\rho})(\delta+2\bar{\pi}-2\beta) \\
& +(\delta+\bar{\pi}-3\beta-\bar{\alpha})(D-2\bar{\rho})]\, l^{(\mu} m^{\nu)}\},
\end{aligned}$$

$$\begin{aligned}
{}_{-2}\Pi^{\mu\nu} = & -\frac{1}{2}\rho^{-4}\{(\Delta+\bar{\mu}+3\gamma-\bar{\gamma})(\Delta+\bar{\mu}+2\gamma-2\bar{\gamma})\, \bar{m}^\mu \bar{m}^\nu \\
& +(\bar{\delta}-\bar{\tau}+3\alpha+\bar{\beta})(\bar{\delta}-\bar{\tau}+2\alpha+2\bar{\beta})\, n^\mu n^\nu \\
& -[(\Delta+\bar{\mu}+3\gamma-\bar{\gamma})(\bar{\delta}-2\bar{\tau}+2\alpha) \\
& +(\bar{\delta}-\bar{\tau}+3\alpha+\bar{\beta})(\Delta+2\bar{\mu}+2\gamma)]\, n^{(\mu} \bar{m}^{\nu)}\}.
\end{aligned}$$

In the application of the mode decomposition to the quantum theory it is convenient to fix the constants N^Λ appearing in (G.73) and (G.74) by demanding that the mode set satisfies the orthonormality condition

$$< h^\Lambda(\ell,m,\omega,P;x), h^{\Lambda'}(\ell',m',\omega',P';x) > = \delta_{\Lambda\Lambda'}\,\delta_{\ell\ell'}\,\delta_{mm'}\,\delta(\omega-\omega')\,\delta_{PP'}\,; \quad \text{(G.78)}$$

the other inner products vanish. Here $<\psi,\phi>$ is the inner product defined for complex symmetric tensor fields $\psi_{\alpha\beta}$ and $\phi_{\alpha\beta}$ as [Candelas, Chrzanowski, and Howard (1981)]:

$$< \psi,\phi > = -i\int_\Sigma \mathrm{d}\sigma^\mu \left(\bar{\tilde{\psi}}^{\alpha\beta}\,\tilde{\phi}_{\alpha\beta;\mu} - \phi^{\alpha\beta}\,\bar{\tilde{\psi}}_{\alpha\beta;\mu} + 2\tilde{\phi}_{\alpha\mu}\,\bar{\tilde{\psi}}^{\alpha\beta}{}_{;\beta} - 2\bar{\tilde{\psi}}_{\alpha\mu}\,\tilde{\phi}^{\alpha\beta}{}_{;\beta}\right). \quad \text{(G.79)}$$

Here $\tilde{\psi}_{\alpha\beta} = \psi_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\,\psi^\mu_\mu$ and $\tilde{\phi}_{\alpha\beta} = \phi_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\,\phi^\mu_\mu$. Σ is a complete spacelike hypersurface in the exterior Kerr spacetime, and $d\sigma^\mu$ is the future-directed normal to Σ. As in the case of the electromagnetic field, the surface of integration Σ can be moved either to the infinite future or to the infinite past, so that the integration is performed either over $H^-\cup\mathcal{J}^-$ or $H^+\cup\mathcal{J}^+$, where the asymptotics of the fields are known.

An arbitrary gravitational perturbation in the exterior of the Kerr black hole can be expanded in terms of the modes (G.73) and (G.74). Thus, we can write

$$h_{\mu\nu}(x) = \sum_J \{a_J\, h_{\mu\nu}(J;x) + \bar{a}_J\, \bar{h}_{\mu\nu}(J;x)\}\,, \quad \text{(G.80)}$$

where J is a shorthand for $\{\Lambda,\ell,m,\omega,P\}$, and Λ takes values "in" and "up". A similar decomposition is valid if one use *OUT*- and *DOWN*-modes.

Appendix H

Quantum Fields in Kerr Spacetime

H.1 Quantum Theory in an External Field

In this section we briefly describe the general scheme of quantization of a free Bose field in a given (not necessarily gravitational) external field.[1]

The general expression for the action describing a system of real Bose fields $\varphi_A(x)$ $(A = 1, \ldots, M)$ interacting with an arbitrary, given external field $g_Y(x)$ $(Y = 1, \ldots, Q)$ is written in the form

$$W[\varphi] = -\frac{1}{2}\int \left[\varphi_{A,\mu}\, P^{AB\mu\nu}\, \varphi_{B,\nu} + \varphi_A\, N^{AB\mu}\, \varphi_{B,\mu} + \varphi_A\, T^{AB}\, \varphi_B\right] d^4x\,, \tag{H.1}$$

where $P^{AB\mu\nu} = P^{(AB)(\mu\nu)}$, $N^{AB\mu} = N^{[AB]\mu}$, $T^{AB} = T^{(AB)}$ are real functions of the external field g_Y and its derivatives. Recall that summation is carried out over repeated indices (including the indices A and B). The variation of this action with respect to the dynamic variables φ_A results in the following field equations:

$$D^{AB}\,\varphi_B \equiv [P^{AB\mu\nu}\,\partial_\mu\,\partial_\nu - (N^{AB\mu} - P^{AB\mu\nu}{}_{,\nu})\,\partial_\mu - (T^{AB} + \frac{1}{2}\,N^{AB\mu}{}_{,\mu})]\,\varphi_B = 0. \tag{H.2}$$

An arbitrary pair of functions φ^1_A and φ^2_A obeys the relation

$$\varphi^2_A\, D^{AB}\,\varphi^1_B - \varphi^1_A\, D^{AB}\,\varphi^2_B = [\,B^\mu(\varphi^1, \varphi^2)\,]_{,\mu}\,, \tag{H.3}$$

where

$$B^\mu(\varphi^1, \varphi^2) = \varphi^2_A\, P^{AB\mu\nu}\,\varphi^1_{B,\nu} - \varphi^1_A\, P^{AB\mu\nu}\,\varphi^2_{B,\nu} + \varphi^1_A\, N^{AB\mu}\,\varphi^2_B\,. \tag{H.4}$$

Hence, it can be shown, using Gauss' theorem (A.86), that the expression

$$B(\varphi^1, \varphi^2) = \int_\Sigma B^\mu(\varphi^1, \varphi^2)\, d\sigma_\mu \tag{H.5}$$

[1] A more detailed presentation of this theory, and also of the theory of Fermi fields in curved spacetime, can be found in DeWitt (1965, 1975), Birrell and Davies (1982), and Wald (1994).

calculated for an arbitrary pair of solutions φ^1 and φ^2 of equation (H.2) is independent of the choice of the complete Cauchy surface.[2] The antisymmetric bilinear form B, defined on the space of solutions of equations (H.2), will be called the *canonical form* corresponding to this equation. Quite often, instead of the canonical form B, a scalar product

$$< \varphi, \psi > = iB(\bar{\varphi}, \psi) \tag{H.6}$$

connected with it is used.

In quantum theory, the field $\hat{\varphi}_A(x)$ is treated as the operator solution of equation (H.2). The canonical commutation relations imposed on this operator are formulated in a standard manner. Let

$$\pi^A(x) = \frac{\delta W}{\delta \varphi_{A,0}(x)} = -P^{AB0\nu}\varphi_{B,\nu} + \frac{1}{2}N^{AB0}\varphi_B \tag{H.7}$$

be the momentum of the field φ_A; then

$$[\hat{\varphi}_A(\mathbf{x}, x^0), \hat{\varphi}_B(\mathbf{x}', x^0)] = 0, \quad [\hat{\pi}^A(\mathbf{x}, x^0), \hat{\pi}^B(\mathbf{x}', x^0)] = 0,$$

$$[\hat{\varphi}_A(\mathbf{x}, x^0), \hat{\pi}^B(\mathbf{x}', x^0)] = i\,\delta^B_A\,\delta^3(\mathbf{x}, \mathbf{x}'). \tag{H.8}$$

A simple check shows that the canonical commutation relations (H.8) are equivalent to the relations

$$[B(\varphi^1, \hat{\varphi}), B(\varphi^2, \hat{\varphi})] = iB(\varphi^1, \varphi^2) \tag{H.9}$$

provided the equality is satisfied for an arbitrary pair φ^1 and φ^2 of classical solutions of system (H.2).

The commutation relations written in the form (H.9) have a larger range of applicability than (H.8). Thus, they are valid for a degenerate theory [i.e., when $\det(P^{AB00}) = 0$]. The presence of constraints in the degenerate theory dictates a change in the rules (H.8) of standard canonical quantization. In a theory with gauge invariance a solution is defined up to a gauge transformation

$$\varphi'_B = \varphi_B + R^i_B\,\lambda_i, \tag{H.10}$$

where R^i_B are differential operators and λ_i are arbitrary functions. In other words, for arbitrary functions λ_i the following identity is valid

$$\tilde{D}^{AB}R^i_B\,\lambda_i = 0. \tag{H.11}$$

It can be shown that the scalar product $< \varphi, \psi >$ for any two solutions φ and ψ of the equation (H.2) is gauge invariant.

[2]If the space is non-compact, the solutions φ^1 and φ^2 are assumed to fall off sufficiently rapidly at infinity.

H.2 Vacuum. Many-Particle States

To define a "particle", consider a set of complex solutions of equation (H.2) and choose in this space a basis; that is, a complete system of linearly independent solutions. It is convenient to use a basis $\{\varphi_J, \bar{\varphi}_J\}$ which consists of complex-conjugate solutions $\{\varphi_A(J;x), \bar{\varphi}_A(J;x)\}$ and satisfies the *normalization conditions:*

$$< \bar{\varphi}_J, \varphi_{J'} > = < \bar{\varphi}_J, \varphi_{J'} > = 0, \quad < \varphi_J, \varphi_{J'} > = \delta_{JJ'}. \tag{H.12}$$

Here the indices $J, J', \dots$ label the basic solutions. Each solution of equations (H.2), satisfying the imposed boundary conditions, admits a decomposition with respect to this basis. Thus, for the field operator $\hat{\varphi}$, we can write

$$\hat{\varphi}_A(x) = \sum_J \left[\hat{a}_J \, \varphi_A(J;x) + \hat{a}_J^\dagger \, \bar{\varphi}_A(J;x) \right], \tag{H.13}$$

where

$$\hat{a}_J = < \varphi_J, \hat{\varphi} >, \quad \hat{a}_J^\dagger = - < \bar{\varphi}_J, \hat{\varphi} > . \tag{H.14}$$

If the field $\hat{\varphi}$ is Hermitian ($\hat{\varphi}_A^\dagger = \hat{\varphi}$), then $\hat{a}_J^\dagger = (\hat{a}_J)^\dagger$. The time-independent operators $\hat{a}_J^\dagger$ and $\hat{a}_J$ are called the operators of *creation* and *annihilation* of a particle in the mode $\varphi_A(J;x)$. These operators satisfy the following commutation relations:

$$[\hat{a}_J, \hat{a}_{J'}] = \left[\hat{a}_J^\dagger, \hat{a}_{J'}^\dagger\right] = 0, \quad \left[\hat{a}_J, \hat{a}_{J'}^\dagger\right] = \delta_{JJ'}. \tag{H.15}$$

This is readily verified if the relation (H.9) and normalization conditions (H.12) are used.

The vacuum state $|0\rangle$ corresponding to a given choice of basis $\{\varphi_J, \bar{\varphi}_J\}$ is defined by the condition

$$\hat{a}_J|0\rangle = 0. \tag{H.16}$$

The n-particle state $|J_1, \dots, J_n\rangle$ with one particle in each of the modes $\varphi_A(J_1;x)$, $\dots$, $\varphi_A(J_n;x)$ is obtained from the vacuum by applying to it the appropriate set of creation operators:

$$|J_1, \dots, J_n\rangle = \hat{a}_{J_1}^\dagger \dots \hat{a}_{J_n}^\dagger |0\rangle . \tag{H.17}$$

These multiparticle basis states constitute the eigenstates of the operator $\hat{n}_J = \hat{a}_J^\dagger \, \hat{a}_J$ of the number of particles in the Jth mode,

$$\hat{n}_J|J_1, \dots, J_n\rangle = n_J|J_1, \dots, J_n\rangle , \quad n_J = \sum_{k=1}^{n} \delta_{JJ_k} , \tag{H.18}$$

and satisfying the following orthonormalization and completeness conditions:

$$\langle J_1, \dots, J_n | J_1', \dots, J_m' \rangle = 0 \quad \text{if} \quad n \neq m ,$$

$$\langle J_1, \dots, J_n | J'_1, \dots, J'_n \rangle = \sum_{\substack{\text{over all permutation} \\ (J_1, \dots, J_n)}} \delta_{J_1 J'_1} \dots \delta_{J_n J'_n}, \tag{H.19}$$

$$\hat{I} = |0\rangle\langle 0| + \sum_{n=1}^{\infty} |J_1, \dots, J_n\rangle \frac{1}{n!} \langle J_1, \dots, J_n| . \tag{H.20}$$

Here $\hat{I}$ is the identity operator, and summation is carried out over all sets $\{J_1, \dots, J_n\}$.

Obviously, the choice of basis $\{\varphi_J, \bar{\varphi}_J\}$ and of the related definition of a "particle" is far from unambiguous. The formal construction given above becomes physically meaningful only if we succeed in describing clearly a set of attributes sufficient for distinguishing between the vacuum (or one-particle) state and all other allowed quantum states of the system. In the long run, this problem reduces to describing a detector capable of recording the particles. According to the quantum theory of measurement, this instrument is described by an Hermitian operator whose eigenvectors are the states corresponding to a certain number of particles recorded by this instrument.

All necessary concepts can be defined in the asymptotic in- and out-regions in the standard scheme of the theory, where the external field is "switched off" in the distant past and distant future. In each of these regions, one has to operate with the free Bose field theory in flat spacetime. Because of the time-invariance of the Minkowski space, the energy $\hat{H}$ is conserved in this regions. One also has

$$\left[\hat{H}, \hat{\varphi}\right] = \frac{1}{i} \xi^\mu \partial_\mu \hat{\varphi} . \tag{H.21}$$

Here $\xi^\mu = \delta^\mu_{(T)}$ is the Killing vector field generating the time translations. The vacuum state in each of the asymptotic regions is defined as the lowest eigenstate of the energy operator $\hat{H}$. This choice unambiguously corresponds to choosing functions with positive frequency with respect to the time coordinate T as basis modes:

$$\xi^\mu \partial_\mu \varphi_A(J; x) = -i \omega_J \varphi_A(J, x) . \tag{H.22}$$

H.3 S-Matrix

In order to distinguish between two bases consisting of the solutions $\{\varphi_J, \bar{\varphi}_J\}$, for which relations (H.22) act as the asymptotic boundary conditions in the future (out-basis) and in the past (in-basis), we equip the basis functions with a superscript "out" and "in", respectively. Likewise, we make use of these additional indices in order to distinguish between the quantities defined via these bases. For instance,

$$\hat{\varphi} = \sum_J \left(\hat{a}_{\text{in}, J} \varphi_J^{\text{in}} + \hat{a}^\dagger_{\text{in}, J} \bar{\varphi}_J^{\text{in}} \right) = \sum_J \left(\hat{a}_{\text{out}, J} \varphi_J^{\text{out}} + \hat{a}^\dagger_{\text{out}, J} \bar{\varphi}_J^{\text{out}} \right), \tag{H.23}$$

$$\hat{a}_{\text{in}, J} |0; \text{in}\rangle = 0, \quad \hat{a}_{\text{out}, J} |0; \text{out}\rangle = 0 . \tag{H.24}$$

In general, the in- and out-bases do not coincide because the in- and out-basis functions satisfy different boundary conditions. The coefficients of the matrices $A_{JJ'}$ and $B_{JJ'}$, which relate the in- and out-bases,

$$\varphi_J^{\text{out}} = \sum_{J'} \left(A_{JJ'}\, \varphi_{J'}^{\text{in}} + B_{JJ'}\, \bar{\varphi}_{J'}^{\text{in}} \right), \tag{H.25}$$

are known as *Bogoliubov transformation* coefficients. Using the normalization conditions (H.12), we find

$$A_{JJ'} = < \varphi_{J'}^{\text{in}}, \varphi_J^{\text{out}} >, \quad B_{JJ'} = - < \bar{\varphi}_{J'}^{\text{in}}, \varphi_J^{\text{out}} > . \tag{H.26}$$

In order to simplify the expressions, it is convenient to use matrix notation. We denote by $\hat{\mathbf{a}}$ and $\hat{\mathbf{a}}^\dagger$ rows constructed from $\hat{a}_J$ and $\hat{a}_J^\dagger$, the index J enumerating the elements in the row. (In the general case J takes infinitely many values.) We use notations $\mathbf{\Phi}$ and $\bar{\mathbf{\Phi}}$ for column, constructed from φ_J and $\bar{\varphi}_J$, respectively. We also denote by $\mathbf{A}$ and $\mathbf{B}$ the matrices with matrix elements given by (H.26), and by C the following block-matrix:

$$\mathsf{C} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \bar{\mathbf{B}} & \bar{\mathbf{A}} \end{pmatrix}. \tag{H.27}$$

Using this matrix notation, we have

$$\hat{\varphi} = \hat{\mathbf{a}}_{\text{in}}\, \mathbf{\Phi}^{\text{in}} + \hat{\mathbf{a}}_{\text{in}}^\dagger\, \bar{\mathbf{\Phi}}^{\text{in}} = \hat{\mathbf{a}}_{\text{out}}\, \mathbf{\Phi}^{\text{out}} + \hat{\mathbf{a}}_{\text{out}}^\dagger\, \bar{\mathbf{\Phi}}^{\text{out}},$$

$$\begin{pmatrix} \Phi^{\text{out}} \\ \bar{\Phi}^{\text{out}} \end{pmatrix} = \mathsf{C} \begin{pmatrix} \Phi^{\text{in}} \\ \bar{\Phi}^{\text{in}} \end{pmatrix}, \quad \begin{pmatrix} \Phi^{\text{in}} \\ \bar{\Phi}^{\text{in}} \end{pmatrix} = \mathsf{C}^{-1} \begin{pmatrix} \Phi^{\text{out}} \\ \bar{\Phi}^{\text{out}} \end{pmatrix}, \tag{H.28}$$

$$(\hat{\mathbf{a}}_{\text{out}}, \hat{\mathbf{a}}_{\text{out}}^\dagger) = (\hat{\mathbf{a}}_{\text{in}}, \hat{\mathbf{a}}_{\text{in}}^\dagger)\ \mathsf{C}^{-1}, \quad (\hat{\mathbf{a}}_{\text{in}}, \hat{\mathbf{a}}_{\text{in}}^\dagger) = (\hat{\mathbf{a}}_{\text{out}}, \hat{\mathbf{a}}_{\text{out}}^\dagger)\ \mathsf{C}.$$

Relations (H.26) yield the following expression for the coefficients of matrix C^{-1} inverse to C:

$$\mathsf{C}^{-1} = \begin{pmatrix} \mathbf{A}^+ & -\mathbf{B}' \\ -\mathbf{B}^+ & \mathbf{A}' \end{pmatrix}. \tag{H.29}$$

Primes in the above formulas denote transposition, and + stands for Hermitian conjugation of matrices: $(\ )^+ = \overline{(\ )}'$. The conditions $\mathsf{C}\mathsf{C}^{-1} = \mathsf{C}^{-1}\mathsf{C} = \mathsf{I}$ imply that the following equalities are true:

$$\mathbf{A}\,\mathbf{A}^+ - \mathbf{B}\,\mathbf{B}^+ = \mathbf{I}, \quad \mathbf{A}^+\mathbf{A} - \mathbf{B}'\,\bar{\mathbf{B}} = \mathbf{I},$$

$$(\mathbf{A}^{-1}\,\mathbf{B})' = \mathbf{A}^{-1}\,\mathbf{B}, \quad (\bar{\mathbf{B}}\,\mathbf{A}^{-1})' = \bar{\mathbf{B}}\,\mathbf{A}^{-1} \tag{H.30}$$

(here $\mathbf{I}$ is the identity matrix).

The *S-matrix* operator relating the in- and out-states is defined by the relations

$$\hat{\mathbf{a}}_{\rm in}\,\hat{S} = \hat{S}\,\hat{\mathbf{a}}_{\rm out} \qquad \hat{\mathbf{a}}^{\dagger}_{\rm in}\,\hat{S} = \hat{S}\,\hat{\mathbf{a}}^{\dagger}_{\rm out}\,. \tag{H.31}$$

It can be verified that this operator is unitary ($\hat{S}\,\hat{S}^{+} = \hat{I}$), has the property

$$\hat{S}|J_1,\ldots,J_n;\mathrm{out}\rangle = |J_1,\ldots,J_n;\mathrm{in}\rangle\,, \tag{H.32}$$

and admits the following representation:

$$\hat{S} = \sum_n |J_1,\ldots,J_n;\mathrm{in}\rangle\frac{1}{n!}\langle J_1,\ldots,J_n;\mathrm{out}|\,. \tag{H.33}$$

If we substitute the last of the relations (H.28) into (H.31) and solve the resulting equation, the operator $\hat{S}$ can be expressed in terms of $\{\hat{\mathbf{a}}_{\rm out},\hat{\mathbf{a}}^{\dagger}_{\rm out}\}$. The corresponding solution admits the following representation [Berezin (1965), DeWitt (1965)]:

$$\hat{S} = e^{iW_0} N_{\rm out}\left[\exp\left(\frac{1}{2}\,\hat{\mathbf{a}}_{\rm out}\,\mathbf{\Lambda}\,\hat{\mathbf{a}}'_{\rm out} + \hat{\mathbf{a}}^{\dagger}_{\rm out}\,(\mathbf{M}-\mathbf{I})\,\hat{\mathbf{a}}'_{\rm out} + \frac{1}{2}\,\hat{\mathbf{a}}^{\dagger}_{\rm out}\,\mathbf{V}\,\hat{\mathbf{a}}^{+}_{\rm out}\right)\right], \tag{H.34}$$

where $N_{\rm out}[\ldots]$ denotes the operation of normal ordering[3] with respect to the out-operators, $\{\hat{a}_{{\rm out},J},\hat{a}^{\dagger}_{{\rm out},J}\}$, and

$$\mathbf{\Lambda} = \mathbf{A}^{-1}\,\mathbf{B}\,, \quad \mathbf{V} = -\,\bar{\mathbf{B}}\,\mathbf{A}^{-1}\,, \quad \mathbf{M} = \mathbf{A}^{-1\prime},$$
$$e^{iW_0} = \theta\left[\,\det(\mathbf{A}^{+}\mathbf{A})\,\right]^{-1/4}, \quad |\,\theta\,| = 1\,. \tag{H.35}$$

Here $\hat{\mathbf{a}}_{\rm out}\,\mathbf{\Lambda}\,\hat{\mathbf{a}}'_{\rm out} \equiv \Sigma_{J,J'}\,\hat{a}_{{\rm out},\,J}\,\Lambda_{JJ'}\,\hat{a}_{{\rm out},\,J'}$, etc. Equalities (H.30) imply that the matrices $\mathbf{\Lambda}$ and $\mathbf{V}$ are symmetric:

$$\mathbf{\Lambda}' = \mathbf{\Lambda}\,, \quad \mathbf{V}' = \mathbf{V}. \tag{H.36}$$

The result (H.34) is fundamental for the quantum theory in an external field that we are discussing here. Namely, it states that it is possible to calculate explicitly the S-matrix operator containing exhaustive information on quantum effects of creation, scattering, and absorption of particles in the external field provided we know the Bogoliubov transformation coefficients determined by solving the classical equations (H.2). It can be shown that the matrices $\mathbf{V}$, $\mathbf{M}$, and $\mathbf{\Lambda}$ in (H.35) are directly related to probability amplitudes of the elementary processes of creation, scattering, and annihilation in the external field:

$$\begin{aligned}\langle J,J';\mathrm{out}|0;\mathrm{in}\rangle &= e^{iW_0}\,V_{JJ'}\,,\\ \langle J;\mathrm{out}|J';\mathrm{in}\rangle &= e^{iW_0}\,M_{JJ'}\,,\\ \langle 0;\mathrm{out}|J,J';\mathrm{in}\rangle &= e^{iW_0}\,\Lambda_{JJ'}\,.\end{aligned} \tag{H.37}$$

[3] As a result of this operation, all creation operators are placed to the left of annihilation operators in the expansion of the appropriate expression into a series in creation and annihilation operators. The same expression (H.34) describes the S-matrix in the case of Fermi fields as well. The matrices $\mathbf{\Lambda}$ and $\mathbf{V}$ are then antisymmetric. The general expression for them in terms of Bogoliubov coefficients is given in the monographs of Berezin (1965) and DeWitt (1965).

H.4 Massless Fields Quantization in Kerr Spacetime

H.4.1 *IN*-, *UP*-, *OUT*-, and *DOWN*-modes

In this section we construct convenient in- and out-bases for quantum massless fields in the spacetime of an eternal Kerr black hole. For this purpose, we use mode decompositions for the scalar massless, electromagnetic and gravitational perturbations discussed in Appendix G.

The field operator solution $\hat{\varphi}_A(x)$ for these fields in the black hole exterior can be written in the form

$$\hat{\varphi}_A(x) = \sum_\Lambda \sum_J \left[\hat{a}_{\Lambda J}\, \varphi_A^\Lambda(J;x) + \hat{a}^\dagger_{\Lambda J}\, \bar{\varphi}_A^\Lambda(J;x)\right]. \tag{H.38}$$

Here $\varphi_J^\Lambda = \varphi_A^\Lambda(J;x)$ are basic solutions, J denoting the complete set of quantum numbers $\{\ell, m, \omega, P\}$ enumerating the modes. Λ takes values either "in" and "up", or "out" and "down". We recall also that, according to our agreement, modes $\varphi_J^{\rm in}$ and $\varphi_J^{\rm out}$ are defined for $\omega > 0$, while the modes $\varphi_J^{\rm up}$ and $\varphi_J^{\rm down}$ are defined for $\varpi > 0$.

The basic solutions are normalized as follows

$$< \varphi_{J'}^{\Lambda'}, \varphi_J^\Lambda > = \delta_{\Lambda\Lambda'}\, \delta_{JJ'}, \quad < \bar{\varphi}_{J'}^{\Lambda'}, \bar{\varphi}_J^\Lambda > = -\delta_{\Lambda\Lambda'}\, \delta_{JJ'}; \tag{H.39}$$

the other inner products vanish. Here

$$\delta_{JJ'} \equiv \delta_{PP'}\, \delta_{\ell\ell'}\, \delta_{mm'}\, \delta(\omega - \omega'). \tag{H.40}$$

The normalization conditions and canonical commutation relations (H.15) imply that the operators of creation and annihilation

$$\hat{a}_{\Lambda J} = < \varphi_J^\Lambda, \hat{\varphi} >, \quad \hat{a}^\dagger_{\Lambda J} = - < \bar{\varphi}_J^\Lambda, \hat{\varphi} > \tag{H.41}$$

obey the standard commutation relations

$$\left[\hat{a}_{\Lambda J}, \hat{a}^\dagger_{\Lambda' J'}\right] = \delta_{\Lambda\Lambda'}\, \delta_{JJ'}; \tag{H.42}$$

the other commutators vanish.

Any of the four mode solutions, *IN*, *UP*, *OUT*, and *DOWN*, with the same quantum numbers J can be represented as a linear combination of any two of them. We shall need the expressions for $\varphi_J^{\rm in}$ and $\varphi_J^{\rm up}$ in terms of the $\varphi_J^{\rm out}$ and $\varphi_J^{\rm down}$ modes. The only problem in deriving these relations is connected with superradiant modes. For these modes $\omega\,\varpi < 0$, and hence the decomposition of the positive frequency in-mode (say $\varphi_J^{\rm in}$) must contain a negative frequency out-mode (say $\varphi_J^{\rm down}$). In order to be able to deal with this type of situation, we introduce the following notation:

$$\tilde{\varphi}_A^{\rm down}(\ell, m, \omega, P; x) = \begin{cases} \varphi_A^{\rm down}(\ell, m, \omega, P; x) & \text{if } \varpi > 0, \\ \bar{\varphi}_A^{\rm down}(\ell, -m, -\omega, P; x) & \text{if } \varpi < 0, \end{cases} \tag{H.43}$$

$$\tilde{\varphi}^{\rm up}_A(\ell,m,\omega,P;x) = \begin{cases} \varphi^{\rm up}_A(\ell,m,\omega,P;x) & \text{if } \varpi > 0\,, \\ \bar{\varphi}^{\rm up}_A(\ell,-m,-\omega,P;x) & \text{if } \varpi < 0\,. \end{cases} \tag{H.44}$$

The functions $\tilde{\varphi}^{\rm up}_J$ and $\tilde{\varphi}^{\rm down}_J$ obey the normalization relations

$$< \tilde{\varphi}^{\rm up}_{J'}, \tilde{\varphi}^{\rm up}_J > = \delta_{JJ'}\,, \tag{H.45}$$

$$< \tilde{\varphi}^{\rm down}_{J'}, \tilde{\varphi}^{\rm down}_J > = \delta_{JJ'}\,, \tag{H.46}$$

which are valid for all (positive and negative) values of ϖ.

By using these notations, we can write the sought decomposition as

$$\varphi^{\rm in}_A(J;x) = R_J\,\varphi^{\rm out}_A(J;x) + T_J\,\tilde{\varphi}^{\rm down}_A(J;x)\,, \tag{H.47}$$

$$\tilde{\varphi}^{\rm up}_A(J;x) = t_J\,\varphi^{\rm out}_A(J;x) + r_J\,\tilde{\varphi}^{\rm down}_A(J;x)\,. \tag{H.48}$$

The quantities on both sides of these relations are defined for $\omega > 0$. Substituting (H.47) and (H.48) into the normalization conditions (H.39) for the *IN*- and *UP*-modes, one gets

$$|\,R_J|^2 + \sigma_J\,|\,T_J|^2 = 1\,, \tag{H.49}$$

$$|\,r_J|^2 + \sigma_J\,|\,t_J|^2 = 1\,, \tag{H.50}$$

$$R_J\,\bar{t}_J + \sigma_J\,T_J\,\bar{r}_J = 0\,. \tag{H.51}$$

Here $\sigma_J = \text{sign}(\varpi_J)$. These relations imply

$$|\,R_J|^2 = |\,r_J|^2\,, \quad |\,T_J|^2 = |\,t_J|^2\,. \tag{H.52}$$

H.4.2 *DN*-modes

Definition

For the construction of the sought bases we need another set of solutions. Let us introduce in the spacetime of the eternal Kerr black hole coordinates (U,V,θ,ϕ_+) given by (D.73). The metric (D.74) in these coordinates is invariant under the following discrete symmetry transformation

$$\mathbf{J}:\ U \to -U\,, \quad V \to -V\,. \tag{H.53}$$

The operator $\mathbf{J}$ is an isometry that maps the region I onto the region I' (see e.g., Figure 11.1). Denote by $\varphi^*_A = \mathbf{J}^*\varphi_A$ a tensor that is obtained from φ_A by means of the operation of "dragging" by the map $\mathbf{J}$ (see Appendix A.7). If φ_A is a solution of the covariant equations with coefficients determined by the metric g, then φ_A^* is again a solution of the same equations. We define[4]

$$\varphi_A^{\rm dn}(J;x) = [\,\bar{\varphi}_A^{\rm up}(J;x)\,]^* . \tag{H.54}$$

These new modes obey the normalization conditions ($\varpi > 0$)

$$< \varphi_{J'}^{\rm dn}, \varphi_J^{\rm dn} > = \delta_{JJ'} ; \tag{H.55}$$

the other inner products vanish.

Scalar massless field

Consider the behavior of UP-modes of a massless scalar field near the past horizon H^-. Equation (G.48) implies

$$\varphi^{\rm up}(\ell,m,\omega;x)\,|_{H^-} \sim \frac{1}{\sqrt{r_+^2+a^2}}\,\frac{1}{\sqrt{4\pi\varpi}}\;{}_0Z_{\ell m}^{\omega}\left(\theta,\bar{\phi}\right)\exp(-i\omega u) . \tag{H.56}$$

Using the relation

$$\bar{\omega}\,u - m\,\bar{\phi} = \varpi\,u - m\,\tilde{\phi}_H \tag{H.57}$$

which is valid on the horizon H^-, one can rewrite (H.56) in the form

$$\varphi^{\rm up}(\ell,m,\omega;x)\,|_{H^-} \sim \alpha_{\ell,m,\omega}\left(\theta,\tilde{\phi}_H\right)\exp(-i\varpi u) . \tag{H.58}$$

Using relation (D.76), we can rewrite (H.58) as

$$\varphi^{\rm up}\,|_{H^-} \sim \alpha_{\ell,m,\omega}\left(\theta,\tilde{\phi}_H\right)\exp\left[i\varpi\kappa^{-1}\ln(-U)\right]\vartheta(-U) , \tag{H.59}$$

where U is an affine parameter along H^-. We include the Heaviside step function $\vartheta(-U)$ to make explicit the fact that $\varphi^{\rm up}$ is defined and non-vanishing only for $U < 0$. The map **J** restricted to H^- is equivalent to the transformation $U \to -U$. Hence, we have

$$\varphi^{\rm dn}\,|_{H^-}(\ell,m,\omega;x) \sim \bar{\alpha}_{\ell,m,\omega}\left(\theta,\tilde{\phi}_H\right)\exp[-i\varpi\kappa^{-1}\ln(U)]\,\vartheta(U) . \tag{H.60}$$

The modes $\varphi_J^{\rm dn}$ are defined for $\varpi > 0$. It is convenient to use solutions $\tilde{\varphi}_J^{\rm dn}$ defined as

$$\tilde{\varphi}_A^{\rm dn}(\ell,m,\omega,P;x) \;=\; \begin{cases} \varphi_A^{\rm dn}(\ell,m,\omega,P;x) & \text{if } \varpi > 0 , \\[2mm] \bar{\varphi}_A^{\rm dn}(\ell,-m,-\omega,P;x) & \text{if } \varpi < 0 . \end{cases} \tag{H.61}$$

[4] Index "dn" used for the new modes is chosen to recall that these modes are obtained from "up" modes by means of the inversion transformation (see footnote on page 371).

Define

$$c_J = [1 - \exp(-2\pi\varpi\,\sigma_J/\kappa)]^{-1/2}, \qquad s_J = [\exp(2\pi\varpi\,\sigma J/\kappa) - 1]^{-1/2}. \tag{H.62}$$

Note that $\sigma_J \varpi = |\varpi|$, so that one has $c_J = c_{\bar{J}}$ and $s_J = s_{\bar{J}}$, where $J = \{\ell, m, \omega, P\}$ and $\bar{J} = \{\ell, -m, -\omega, P\}$.

The following two functions on H^-

$$\varphi_J^{\rm n} = c_J\,\varphi_J^{\rm dn} + s_J\,\bar{\varphi}_J^{\rm up}, \tag{H.63}$$

$$\varphi_J^{\rm p} = c_J\,\varphi_J^{\rm up} + s_J\,\bar{\varphi}_J^{\rm dn}, \tag{H.64}$$

for $\varpi > 0$ are of positive frequency with respect to the affine parameter U along H^-. This follows from the following lemma.

Lemma: *For positive real p and arbitrary real q one has*

$$\int_{-\infty}^{\infty} {\rm d}U\, e^{-ipU}\left[e^{-iq\ln U}\,\vartheta(U) + e^{-\pi q}\,e^{-iq\ln(-U)}\,\vartheta(-U)\right] = 0\,. \tag{H.65}$$

Proof: Define[5]

$$F(p,q) = \int_0^{\infty} {\rm d}U\, e^{ipU}\, e^{-iq\ln U}. \tag{H.66}$$

Suppose $p > 0$, and compare $F(p,q)$ with $F(-p,q)$. In order to do this, we extend $\ln U$ to the complex plane. We assume that the branch cut lies on the negative real axis. This permits us to make all the necessary contour rotations we shall use. Put $U = -ix$ and rotate the contour of integration so as to make x real and positive. Using the relation $\ln(ix) = i\pi/2 + \ln x$, we get

$$F(p,q) = ie^{\pi q/2}\int_0^{\infty} {\rm d}x\, e^{-px}\, e^{-iq\ln x}. \tag{H.67}$$

A similar representation for $F(-p,q)$ can be obtained if we substitute in the corresponding integral $U = -ix$ and rotate the contour so that x becomes real and positive. As a result, we have

$$F(-p,q) = -ie^{-\pi q/2}\int_0^{\infty} {\rm d}x\, e^{-px}\, e^{-iq\ln x}. \tag{H.68}$$

Comparing (H.67) and (H.68), we get

$$F(p,q) = -\,e^{\pi q}\,F(-p,q)\,. \tag{H.69}$$

Equality (H.65) follows immediately since the quantity standing on the left-hand side of (H.65) is identical to $F(-p,q) + e^{-\pi q}F(p,q)$.

One can introduce modes $\tilde{\varphi}_J^p$ and $\tilde{\varphi}_J^n$ by relations similar to (H.43), (H.44), and (H.61). These modes are defined for $\omega > 0$ and are of positive frequency with respect to U.

[5]Note that this integral does not converge absolutely. Nevertheless, in all the applications where we consider wavepackets, an additional integration over q which is required to construct the wavepacket makes the integral convergent.

Electromagnetic field

The *UP*-modes (G.65) for the electromagnetic field in the outgoing gauge

$$a_\mu n^\mu = 0 \tag{H.70}$$

near the past horizon have the following form [Candelas, Chrzanowski, and Howard (1981)]

$$\begin{aligned} &a_\mu^{\rm up}(\ell, m, \omega; x)\,|_{H^-} \\ &\sim \frac{1}{\sqrt{r_+^2+a^2}}\,\frac{1}{\sqrt{4\pi\varpi}}\left[\frac{\rho}{\bar\rho}\;{}_{-1}Z^\omega_{\ell m}\left(\theta,\tilde\phi\right) m_\mu + P\,\frac{\bar\rho}{\rho}\;{}_{+1}Z^\omega_{\ell m}\left(\theta,\tilde\phi\right)\bar m_\mu\right]\exp(-i\omega u) \\ &+ (n_\mu \text{ terms})\,. \end{aligned} \tag{H.71}$$

One can use the allowed freedom in gauge transformations to exclude on H^- the terms proportional to n_μ. In fact, on the past horizon H^- which in the (U, V, θ, ϕ) coordinates is described by the equation $V = 0$, one has $n_\mu \sim V_{,\mu}$. Let us make the gauge transformation $a_\mu \to a_\mu + \lambda_{,\mu}$, with $\lambda = V f(U, \theta, \phi)$. This transformation, restricted to H^-, affects only those terms of $a_\mu^{\rm up}$ which are proportional n_μ. Since $f(U, \theta, \phi)$ is an arbitrary function these terms can be put equal to zero. As a result, we can write $a_\mu^{\rm up}\,|_{H^-}$ in the form

$$\begin{aligned} a_\mu^{\rm up}(\ell, m, \omega; x)\,|_{H^-} &\sim \left[\alpha_{\ell,m,\omega}\left(\theta,\tilde\phi_H\right) m_\mu + \beta_{\ell,m,\omega}\left(\theta,\tilde\phi_H\right)\bar m_\mu\right]\exp(-i\varpi u) \\ &= \left[\alpha_{\ell,m,\omega}\left(\theta,\tilde\phi_H\right) m_\mu + \beta_{\ell,m,\omega}\left(\theta,\tilde\phi_H\right)\bar m_\mu\right]\exp\left[i\varpi\kappa^{-1}\ln(-U)\right]\vartheta(-U)\,. \end{aligned} \tag{H.72}$$

DN-modes for the electromagnetic fields are defined by the requirements that the corresponding solutions vanish on $\mathcal{J}^-$ in the space I', and in a specially chosen gauge have the following form on H^-

$$\begin{aligned} &a_\mu^{\rm dn}(\ell, m, \omega; x)\,|_{H^-} \\ &\sim \left[\bar\alpha_{\ell,m,\omega}\left(\theta,\tilde\phi_H\right)\bar m_\mu + \bar\beta_{\ell,m,\omega}\left(\theta,\tilde\phi_H\right) m_\mu\right]\exp\left[-i\varpi\kappa^{-1}\ln(U)\right]\vartheta(U)\,. \end{aligned} \tag{H.73}$$

The dependence on the affine parameter U for fixed values of θ and $\tilde\phi_H$ (i.e., along the generator of the horizon) for $a_\mu^{\rm up}$ and $a_\mu^{\rm dn}$ is the same as for the corresponding scalar modes (H.59) and (H.60). That is why one can use the arguments given above to show that the electromagnetic modes $\varphi_J^{\rm p}$ and $\varphi_J^{\rm n}$ defined by (H.63) and (H.64) are (for $\varpi > 0$) of positive frequency with respect to U.

Gravitational perturbations

The *UP*-modes (G.74) for the gravitational perturbations in the outgoing gauge

$$h_{\mu\nu} n^\nu = 0\,, \quad h^\mu_\mu = 0 \tag{H.74}$$

near the past horizon have the following form [Candelas, Chrzanowski, and Howard (1981)]:

$$h^{\rm up}_{\mu\nu}(\ell, m, \omega; x)\,|_{H^-}$$

$$\sim C\left[\frac{\rho^2}{\bar\rho^2}\;{}_{-2}Z^\omega_{\ell m}\left(\theta, \tilde\phi\right) m_\mu m_\nu + P\,\frac{\bar\rho^2}{\rho^2}\;{}_{+2}Z^\omega_{\ell m}\left(\theta, \tilde\phi\right) \bar m_\mu \bar m_\nu\right] \exp(-i\omega u)$$

$$+\ (\text{terms in } n_\mu)\,. \tag{H.75}$$

One can use the allowed freedom in gauge transformations to exclude on H^- the terms containing n_μ [for details, see e.g., Frolov (1978c)]. As a result, we can write $h^{\rm up}_{\mu\nu}\,|_{H^-}$ in the form

$$h^{\rm up}_{\mu\nu}(\ell, m, \omega; x)\,|_{H^-} \sim \left[\alpha_{\ell,m,\omega}(\theta, \tilde\phi_H)\, m_\mu m_\nu + \beta_{\ell,m,\omega}(\theta, \tilde\phi_H)\, \bar m_\mu \bar m_\nu\right] \exp(-i\varpi u)$$

$$= \left[\alpha_{\ell,m,\omega}(\theta, \tilde\phi_H)\, m_\mu m_\nu + \beta_{\ell,m,\omega}(\theta, \tilde\phi_H)\, \bar m_\mu \bar m_\nu\right] \exp[\,i\varpi\kappa^{-1}\ln(-U)]\,\vartheta(-U). \tag{H.76}$$

DN-modes for the gravitational perturbations are defined by the requirements that the corresponding solutions vanish at $\mathcal{J}^-$ in the space I' and in a specially chosen gauge have the following form on H^-:

$$h^{\rm dn}_{\mu\nu}(\ell, m, \omega; x)\,|_{H^-}$$

$$\sim \left[\bar\alpha_{\ell,m,\omega}(\theta, \tilde\phi_H)\, \bar m_\mu \bar m_\nu + \bar\beta_{\ell,m,\omega}(\theta, \tilde\phi_H)\, m_\mu m_\nu\right] \exp[-i\varpi\kappa^{-1}\ln(U)]\,\vartheta(U). \tag{H.77}$$

The dependence on the affine parameter U for fixed values of θ and $\tilde\phi_H$ (i.e., along the generator of the horizon) for $h^{\rm up}_{\mu\nu}$ and $h^{\rm dn}_{\mu\nu}$ is the same as for the corresponding scalar modes (H.59) and (H.60). That is why one can use the arguments given above to show that the modes $\varphi^{\rm p}_J$ and $\varphi^{\rm n}_J$ for the gravitational perturbations defined by (H.63) and (H.64) are (for $\varpi > 0$) of positive frequency with respect to U.

H.5 Wald's Bases, Bogoliubov Transformation, and S-matrix

In the spacetime of the eternal Kerr black hole we choose for the in-basis the following sets of positive-frequency solutions

$$\mathbf{\Phi}_J^{\mathrm{in}\,(\sigma_J>0)} = \begin{pmatrix} \varphi_J^{\mathrm{in}} \\ \varphi_J^{\mathrm{n}} \\ \varphi_J^{\mathrm{p}} \end{pmatrix}, \quad \mathbf{\Phi}_J^{\mathrm{in}\,(\sigma_J<0)} = \begin{pmatrix} \varphi_J^{\mathrm{in}} \\ \varphi_{\tilde{J}}^{\mathrm{n}} \\ \varphi_{\tilde{J}}^{\mathrm{p}} \end{pmatrix}. \tag{H.78}$$

Here $J = \{\ell, m, \omega, P\}$, $\tilde{J} = \{\ell, -m, -\omega, P\}$, and $\sigma_J = \mathrm{sign}(\varpi_J)$. The modes are defined for $\omega > 0$. Note that the basis functions are different for non-superradiant ($\sigma_J > 0$) and superradiant ($\sigma_J < 0$) modes.

Likewise, we form the out-basis by sets of functions

$$\mathbf{\Phi}_J^{\mathrm{out}\,(\sigma_J>0)} = \begin{pmatrix} \varphi_J^{\mathrm{out}} \\ \varphi_J^{\mathrm{down}} \\ \varphi_J^{\mathrm{dn}} \end{pmatrix}, \quad \mathbf{\Phi}_J^{\mathrm{out}\,(\sigma_J<0)} = \begin{pmatrix} \varphi_J^{\mathrm{out}} \\ \varphi_{\tilde{J}}^{\mathrm{down}} \\ \varphi_{\tilde{J}}^{\mathrm{dn}} \end{pmatrix}. \tag{H.79}$$

Similar bases (for a non-rotating black hole) were introduced by Wald (1975). In what follows we refer to them as to *Wald's bases.*

We write the equations relating $\mathbf{\Phi}_J^{\mathrm{in}}$ to $\mathbf{\Phi}_J^{\mathrm{out}}$ in the following matrix form:

$$\mathbf{\Phi}_J^{\mathrm{in}} = \mathbf{A}_J^{+} \, \mathbf{\Phi}_J^{\mathrm{out}} - \mathbf{B}'_J \, \bar{\mathbf{\Phi}}_J^{\mathrm{out}}. \tag{H.80}$$

Transformation matrices $\mathbf{A}$ and $\mathbf{B}$ which relate the sets of in-basis and out-basis modes are easily determined from the relations

$$\varphi_A^{\mathrm{in}}(J;x) = R_J \, \varphi_A^{\mathrm{out}}(J;x) + T_J \, \tilde{\varphi}_A^{\mathrm{down}}(J;x), \tag{H.81}$$

$$\tilde{\varphi}_A^{\mathrm{d}}(J;x) = s_J \, \bar{t}_J \, \bar{\varphi}_A^{\mathrm{out}}(J;x) + s_J \, \bar{r}_J \, \bar{\tilde{\varphi}}_A^{\mathrm{down}}(J;x) + c_J \, \tilde{\varphi}_A^{\mathrm{dn}}(J;x), \tag{H.82}$$

$$\tilde{\varphi}_A^{\mathrm{p}}(J;x) = c_J \, t_J \, \varphi_A^{\mathrm{out}}(J;x) + c_J \, r_J \, \tilde{\varphi}_A^{\mathrm{down}}(J;x) + s_J \, \bar{\tilde{\varphi}}_A^{\mathrm{dn}}(J;x). \tag{H.83}$$

The first of these relations is just equation (H.47), while the last two relations follow from (H.48), (H.63), and (H.64). Note that both sides of these equalities are defined for $\omega > 0$. By comparing (H.80) and (H.81)–(H.83), we obtain the Bogoliubov transformation matrices $\mathbf{A}$ and $\mathbf{B}$ [Frolov (1986)]

$$\mathbf{A}^{\sigma>0} = \begin{pmatrix} \overline{R} & 0 & c\,\bar{t} \\ \overline{T} & 0 & c\,\bar{r} \\ 0 & c & 0 \end{pmatrix}, \qquad \mathbf{A}^{\sigma<0} = \begin{pmatrix} \overline{R} & s\,\bar{t} & 0 \\ 0 & 0 & c\,r \\ 0 & c & 0 \end{pmatrix}, \tag{H.84}$$

$$\mathbf{B}^{\sigma>0} = \begin{pmatrix} 0 & -s\,\bar{t} & 0 \\ 0 & -s\,\bar{r} & 0 \\ 0 & 0 & -s \end{pmatrix}, \quad \mathbf{B}^{\sigma<0} = \begin{pmatrix} 0 & 0 & -c\,\bar{t} \\ -T & -s\,r & 0 \\ 0 & 0 & -s \end{pmatrix}. \tag{H.85}$$

In these relations we omit the index J. It is understood that they are valid for arbitrary J.

The field operator $\hat{\varphi}$ can be decomposed in terms of in- and out-bases

$$\hat{\varphi} = \sum_J \left[\hat{\mathbf{a}}_{\mathrm{in},J}\,\mathbf{\Phi}_J^{\mathrm{in}} + \hat{\mathbf{a}}_{\mathrm{in},J}^{\dagger}\,\bar{\mathbf{\Phi}}_J^{\mathrm{in}}\right] = \sum_J \left[\hat{\mathbf{a}}_{\mathrm{out},J}\,\mathbf{\Phi}_J^{\mathrm{out}} + \hat{\mathbf{a}}_{\mathrm{out},J}^{\dagger}\,\bar{\mathbf{\Phi}}_J^{\mathrm{out}}\right], \tag{H.86}$$

where we define

$$\hat{\mathbf{a}}_{\mathrm{in}} = (\hat{a}_{\mathrm{in}}, \hat{a}_{\mathrm{n}}, \hat{a}_{\mathrm{p}})\,, \quad \hat{\mathbf{a}}_{\mathrm{in}}^{\dagger} = \left(\hat{a}_{\mathrm{in}}^{\dagger}, \hat{a}_{\mathrm{n}}^{\dagger}, \hat{a}_{\mathrm{p}}^{\dagger}\right), \tag{H.87}$$

$$\hat{\mathbf{a}}_{\mathrm{out}} = (\hat{a}_{\mathrm{out}}, \hat{a}_{\mathrm{down}}, \hat{a}_{\mathrm{dn}})\,, \quad \hat{\mathbf{a}}_{\mathrm{out}}^{\dagger} = \left(\hat{a}_{\mathrm{out}}^{\dagger}, \hat{a}_{\mathrm{down}}^{\dagger}, \hat{a}_{\mathrm{dn}}^{\dagger}\right). \tag{H.88}$$

Using relations (H.34) and (H.35), we obtain the S-matrix

$$S = \prod_J S_J\,,$$

$$\hat{S}_J = e^{iW_0} N_{\mathrm{out}}\left[\exp\left(\frac{1}{2}\,\hat{\mathbf{a}}_{\mathrm{out}}\,\mathbf{\Lambda}\,\hat{\mathbf{a}}_{\mathrm{out}}' + \hat{\mathbf{a}}_{\mathrm{out}}^{\dagger}\,(\mathbf{M}-\mathbf{I})\,\hat{\mathbf{a}}_{\mathrm{out}}' + \frac{\mathbf{1}}{\mathbf{2}}\,\hat{\mathbf{a}}_{\mathrm{out}}^{\dagger}\mathbf{V}\,\hat{\mathbf{a}}_{\mathrm{out}}^{+}\right)\right]. \tag{H.89}$$

We omitted indices J on the right-hand side of (H.89) and denoted

$$\hat{\mathbf{a}}_{\mathrm{out}}' = \begin{pmatrix} \hat{a}_{\mathrm{out}} \\ \hat{a}_{\mathrm{down}} \\ \hat{a}_{\mathrm{dn}} \end{pmatrix}, \quad \hat{\mathbf{a}}_{\mathrm{out}}^{+} = \begin{pmatrix} \hat{a}_{\mathrm{out}}^{\dagger} \\ \hat{a}_{\mathrm{down}}^{\dagger} \\ \hat{a}_{\mathrm{dn}}^{\dagger} \end{pmatrix}. \tag{H.90}$$

The matrices $\mathbf{\Lambda} = \mathbf{A}^{-1}\mathbf{B}$, $\mathbf{V} = -\bar{\mathbf{B}}\,\mathbf{A}^{-1}$, and $\mathbf{M} = \mathbf{A}^{-1\prime}$ for non-superradiant ($\sigma > 0$) and superradiant ($\sigma < 0$) modes are [Frolov (1986)]

$$\mathbf{\Lambda}^{\sigma>0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -s/c \\ 0 & -s/c & 0 \end{pmatrix}, \quad \mathbf{\Lambda}^{\sigma<0} = \begin{pmatrix} 0 & 0 & t/(c\,r) \\ 0 & 0 & -s/c \\ t/(c\,r) & -s/c & 0 \end{pmatrix},$$

$$\mathbf{V}^{\sigma>0} = \begin{pmatrix} 0 & 0 & s\,t/c \\ 0 & 0 & s\,r/c \\ s\,t/c & s\,r/c & 0 \end{pmatrix}, \quad \mathbf{V}^{\sigma<0} = \begin{pmatrix} 0 & t/r & 0 \\ t/r & 0 & s/(c\,r) \\ 0 & s/(c\,r) & 0 \end{pmatrix}, \tag{H.91}$$

$$\mathbf{M}^{\sigma>0} = \begin{pmatrix} R & 0 & t/c \\ T & 0 & r/c \\ 0 & 1/c & 0 \end{pmatrix}, \quad \mathbf{M}^{\sigma<0} = \begin{pmatrix} 1/\bar{R} & 0 & 0 \\ 0 & 0 & 1/(c\,r) \\ -s\,\bar{t}/(c\,\bar{R}) & 1/c & 0 \end{pmatrix}.$$

Sometimes it is more convenient to redefine the operators of creation and annihilation for superradiant N- and P-modes as

$$\hat{b}_{\mathrm{n},\ell m\omega P} = \hat{a}_{\mathrm{n},\ell -m -\omega P}\,, \quad \hat{b}^{\dagger}_{\mathrm{n},\ell m\omega P} = \hat{a}^{\dagger}_{\mathrm{n},\ell -m -\omega P}\,, \tag{H.92}$$

$$\hat{b}_{\mathrm{p},\ell m\omega P} = \hat{a}_{\mathrm{p},\ell -m -\omega P}\,, \quad \hat{b}^{\dagger}_{\mathrm{p},\ell m\omega P} = \hat{a}^{\dagger}_{\mathrm{p},\ell -m -\omega P}\,. \tag{H.93}$$

If we also assume that the operators $\hat{b}$ and $\hat{b}^{\dagger}$ coincide with $\hat{a}$ and $\hat{a}^{\dagger}$ for all IN-modes and for non-superradiant N- and P-modes, then the first of the relations (H.86) written in in terms of these operators takes the form

$$\begin{aligned}\hat{\varphi}(x) = \sum_{P=\pm 1}\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\Bigg\{\int_0^{\infty} \mathrm{d}\omega \left[\hat{b}_{\mathrm{in},\ell m\omega P}\,\varphi^{\mathrm{in}}(\ell m\omega P;x) + \hat{b}^{\dagger}_{\mathrm{in},\ell m\omega P}\,\bar{\varphi}^{\mathrm{in}}(\ell m\omega P;x)\right] \\ + \int_{m\,\Omega^H}^{\infty} \mathrm{d}\omega \left[\hat{b}_{\mathrm{n},\ell m\omega P}\,\varphi^{\mathrm{n}}(\ell m\omega P;x) + \hat{b}^{\dagger}_{\mathrm{n},\ell m\omega P}\,\bar{\varphi}^{\mathrm{n}}(\ell m\omega P;x)\right. \\ \left. + \hat{b}_{\mathrm{p},\ell m\omega P}\,\varphi^{\mathrm{p}}(\ell m\omega P;x) + \hat{b}^{\dagger}_{\mathrm{p},\ell m\omega P}\,\bar{\varphi}^{\mathrm{p}}(\ell m\omega P;x)\right]\Bigg\}.\end{aligned} \tag{H.94}$$

One can also define new out-operators of creation and annihilation $\{\hat{\mathbf{b}}_{\mathrm{out}}, \hat{\mathbf{b}}^{\dagger}_{\mathrm{out}}\}$, and write the decomposition of $\hat{\varphi}$ in terms of these operators similarly to (H.94).

H.6 Averaging over "Non-observable" States

For the black hole problem, out-particles can be naturally divided into two categories. Particles in OUT-modes can reach a distant observer. We call them "visible". Particles in $DOWN$- and DN-modes fall into the black hole and cannot be registered by a distance observer. We call them "invisible". Averaging a quantum state of the complete system over states of the "invisible" particles gives a density matrix describing the state of the "visible" particles.

To simplify formulas, we use the following notations. Denote by $\{\hat{\beta}_{\alpha}, \hat{\beta}^{\dagger}_{\alpha}\}$ operators of annihilation and creation of the "visible" particles and by $\{\hat{b}_a, \hat{b}^{\dagger}_a\}$ operators of annihilation and creation of the "invisible" particles. Indices α and a enumerate modes of the "visible" and "invisible" particles, respectively. Corresponding vacuum states $|0;\beta\rangle$ and $|0;b\rangle$, that is, the states without "visible" and "invisible" particles, are defined as

$$\hat{\beta}_{\alpha}|0;\beta\rangle = 0\,, \quad \hat{b}_a|0;b\rangle = 0\,. \tag{H.95}$$

For the black hole problem,

$$\hat{\beta}_{\alpha} \Leftrightarrow \hat{a}_{\mathrm{out},J}\,, \quad \hat{b}_a \Leftrightarrow \{\hat{a}_{\mathrm{down},J}, \hat{a}_{\mathrm{dn},J}\}\,, \tag{H.96}$$

$$|0;\beta\rangle|0;b\rangle = |0;\text{out}\rangle\,. \tag{H.97}$$

A vector $|\Psi\rangle$ in the space of out-states admits the following decomposition:

$$|\Psi\rangle = \sum_{k,\,m}\sum_{\substack{\alpha_1,\dots,\alpha_k\\ a_1,\dots,a_m}} \Psi_{\alpha_1,\dots,\alpha_k;a_1,\dots,a_m}|\alpha_1,\dots,\alpha_k;\beta\rangle|a_1,\dots,a_m;b\rangle\,, \tag{H.98}$$

where

$$|\alpha_1,\dots,\alpha_k;\beta\rangle|a_1,\dots,a_m;b\rangle = \hat\beta^\dagger_{\alpha_1}\dots\hat\beta^\dagger_{\alpha_k}\hat b^\dagger_{a_1}\dots\hat b^\dagger_{a_m}|0;\text{out}\rangle\,. \tag{H.99}$$

For the expectation value $\langle\Psi|\hat F|\Psi\rangle$ of an operator $\hat F$ that depends only on the "visible" states $\hat F = F(\hat\beta^\dagger_\alpha,\hat\beta_\alpha)$, we find

$$\langle\Psi|\hat F|\Psi\rangle = \sum_{\substack{k,\,m\\ k',\,m'}}\sum_{\substack{\alpha_1,\dots,\alpha_k\\ \alpha'_1,\dots,\alpha'_{k'}\\ a_1,\dots,a_m\\ a'_1,\dots,a'_{m'}}} \bar\Psi_{\alpha'_1,\dots,\alpha'_{k'};a'_1,\dots,a'_{m'}}\,\Psi_{\alpha_1,\dots,\alpha_k;a_1,\dots,a_m} \tag{H.100}$$

$$\times\langle\alpha'_1,\dots,\alpha'_{k'};\beta|\langle a'_1,\dots,a'_{m'};b|F(\hat\beta^\dagger_\alpha,\hat\beta_\alpha)|a_1,\dots,a_m;b\rangle|\alpha_1,\dots,\alpha_k;\beta\rangle,$$

Since the operator $\hat F$ is independent of $\hat b_a$ and $\hat b^\dagger_a$ and the states $|a_1,\dots,a_m;b\rangle$ satisfy the normalization conditions

$$\langle a'_1,\dots,a'_{m'};b|a_1,\dots,a_m;b\rangle = \delta_{mm'}\sum_{\substack{\text{over all permutation}\\ (a_1,\dots,a_m)}}\delta_{a'_1a_1}\dots\delta_{a'_{m'}a_m}\,, \tag{H.101}$$

we can rewrite (H.100) in the form

$$\langle\Psi|\hat F|\Psi\rangle = \sum_{k,\,k'}\sum_{\substack{\alpha_1,\dots,\alpha_k\\ \alpha'_1,\dots,\alpha'_{k'}}} R_{\alpha_1,\dots,\alpha_k;\alpha'_1,\dots,\alpha'_{k'}}\langle\alpha'_1,\dots,\alpha'_{k'};\beta|\hat F|\alpha_1,\dots,\alpha_k;\beta\rangle$$

$$= \text{Tr}_\beta\left(\hat\rho\hat F\right), \tag{H.102}$$

where

$$R_{\alpha_1,\dots,\alpha_k;\alpha'_1,\dots,\alpha'_{k'}} = \sum_m m!\sum_{a_1,\dots,a_m}\bar\Psi_{\alpha'_1,\dots,\alpha'_{k'};a_1,\dots,a_m}\,\Psi_{\alpha_1,\dots,\alpha_k;a_1,\dots,a_m}\,, \tag{H.103}$$

$$\hat\rho = \sum_{k,\,k'}\sum_{\substack{\alpha_1,\dots,\alpha_k\\ \alpha'_1,\dots,\alpha'_{k'}}}|\alpha_1,\dots,\alpha_k;\beta\rangle R_{\alpha_1,\dots,\alpha_k;\alpha'_1,\dots,\alpha'_{k'}}\langle\alpha'_1,\dots,\alpha'_{k'};\beta|\,, \tag{H.104}$$

and Tr_β denotes the operator of trace over "visible" particle states. It is readily verified, using relation (H.102) for the identity operator $\hat F = \hat I$, that the density matrix $\hat{\rho}$ for the normalized state $|\Psi\rangle(\langle\Psi|\Psi\rangle = 1)$ satisfies the normalization condition $\mathrm{Tr}_\beta(\hat{\rho}) = 1$.

The density matrix $\hat{\rho}$ is *independent* of the definition of the concept of particle for the "invisible" states. Let us make a Bogoliubov transformation in the subspace of solutions corresponding to the "invisible" particles and denote by $\hat{U}$ a unitary operator generating this transformation, then one has

$$\hat{U}|\alpha_1,\ldots,\alpha_k;\beta\rangle|a_1,\ldots,a_m;b\rangle$$
$$= \sum_{\substack{m' \\ a'_1,\ldots,a'_{m'}}} U_{a_1,\ldots,a_m;a'_1,\ldots,a'_{m'}}|\alpha_1,\ldots,\alpha_k;\beta\rangle|a'_1,\ldots,a'_{m'};b\rangle\,, \tag{H.105}$$

$$\sum_{\substack{m' \\ a'_1,\ldots,a'_{m'}}} U_{a_1,\ldots,a_m;a'_1,\ldots,a'_{m'}}\,\bar{U}_{a''_1,\ldots,a''_{m''};a'_1,\ldots,a'_{m'}}$$
$$= \delta_{mm''}\,\delta_{a_1a''_1}\cdots\,\delta_{a_ma''_{m''}}\,. \tag{H.106}$$

The coefficients $\Psi_{\alpha_1,\ldots,\alpha_k;a_1,\ldots,a_m}$ of expansion (H.98) are transformed as follows:

$$\Psi'_{\alpha_1,\ldots,\alpha_k;a'_1,\ldots,a'_{m'}} = \sum_{\substack{m \\ a_1,\ldots,a_m}} \Psi_{\alpha_1,\ldots,\alpha_k;a_1,\ldots,a_m}\,U_{a_1,\ldots,a_m;a'_1,\ldots,a'_{m'}}\,, \tag{H.107}$$

while the coefficients of the matrix $R_{\alpha_1,\ldots,\alpha_k;\alpha'_1,\ldots,\alpha'_{k'}}$ remain unaltered as a result of the unitarity condition (H.106).

Appendix I

Quantum Oscillator

I.1 Action

The action for an oscillator of unit mass and frequency ω is

$$S = \int dt\, L\,, \quad L = \frac{1}{2}\left(\frac{dx}{dt}\right)^2 - \frac{1}{2}\,\omega^2\, x^2\,. \tag{I.1}$$

The equation of motion

$$\frac{d^2x}{dt^2} + \omega^2\, x = 0 \tag{I.2}$$

has a solution

$$x(t) = x_0 \cos\omega t + \frac{p_0}{\omega}\, \sin\omega t\,, \tag{I.3}$$

where $x_0 = x(0)$ and $p_0 = \dot{x}(0)$.

The Euclidean version of action (I.1) is

$$S_E = -i\, S|_{t=-i\tau} = \frac{1}{2}\int d\tau \left[\left(\frac{dx}{d\tau}\right)^2 + \omega^2\, x^2\right]. \tag{I.4}$$

The equation of motion for the Euclidean "oscillator"

$$\frac{d^2x}{d\tau^2} - \omega^2\, x = 0 \tag{I.5}$$

has a solution

$$x(\tau) = x_0 \cosh\omega\tau + \frac{p_0}{\omega}\, \sinh\omega\tau\,. \tag{I.6}$$

The value of the Euclidean action calculated for a solution connecting an "initial" point x_1 at $\tau = 0$ and a "final" point x_2 at $\tau = \beta$ is

$$S_E(x_1, 0; x_2, \beta) = \frac{\omega}{2}\left[\frac{\cosh\omega\beta}{\sinh\omega\beta}(x_1^2 + x_2^2) - \frac{2x_1x_2}{\sinh\omega\beta}\right]$$

$$= \frac{\omega}{4}\left[\tanh\left(\frac{\omega\beta}{2}\right)(x_1 + x_2)^2 + \coth\left(\frac{\omega\beta}{2}\right)(x_1 - x_2)^2\right]. \quad (I.7)$$

I.2 Quantization and Representations

The standard canonical quantization requires that

$$[\hat{p}, \hat{x}] = -i, \quad (I.8)$$

where p is the momentum $p = \partial L/\partial\dot{x}$. The eigenvalues E

$$\hat{H}\,\Psi = E\,\Psi \quad (I.9)$$

of the Hamiltonian

$$\hat{H} = \frac{1}{2}\left(\hat{p}^2 + \omega^2\hat{x}^2\right) \quad (I.10)$$

are discrete $E_n = (n + 1/2)\,\omega$, and the corresponding eigenvectors $|\Psi_n\rangle$ in the x-representation $\Psi_n(x) = \langle x|\Psi_n\rangle$ are

$$\Psi_n(x) = \left(\frac{\omega}{\pi}\right)^{1/4}(2^n n!)^{-1/2}\,e^{-\omega x^2/2}\,H_n(\omega^{1/2}x). \quad (I.11)$$

Here Hermite polynomials $H_n(x)$ can be obtained by using the following generating function

$$e^{2zx - z^2} = \sum_{n=0}^{\infty}\frac{z^n}{n!}\,H_n(x). \quad (I.12)$$

The operator of creation $\hat{a}^\dagger$ and the operator of annihilation $\hat{a}$ are defined as

$$\hat{a}^\dagger = \frac{1}{\sqrt{2\omega}}(\omega\hat{x} - i\hat{p}), \quad \hat{a} = \frac{1}{\sqrt{2\omega}}(\omega\hat{x} + i\hat{p}). \quad (I.13)$$

They obey the commutation relations

$$[\hat{a}, \hat{a}^\dagger] = 1, \quad [\hat{a}, \hat{a}] = [\hat{a}^\dagger, \hat{a}^\dagger] = 0. \quad (I.14)$$

Hamiltonian (I.10) written in terms of the operators of creation and annihilation is

$$\hat{H} = \omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right). \quad (I.15)$$

The representation

$$\hat{a}^{\dagger} = z\,, \quad \hat{a} = \frac{d}{dz}\,, \tag{I.16}$$

when we do not specify the nature of the variable z, clearly satisfies the commutation relations. The problem is to find the inner product for the functions $f(z)$ which the operators $\hat{a}^{\dagger}$ and $\hat{a}$ act on such that those operators are mutually adjoint. In the representation proposed by Segal and Bargmann one considers polynomials $f(z) = \langle z|f\rangle$ in the complex variable z and defines the scalar product as

$$\langle f_1|f_2\rangle = \int d\mu_z\, \overline{f_1(z)}\, f_2(z)\, e^{-|z|^2}\,, \tag{I.17}$$

where

$$d\mu_z = \frac{dz\, d\bar{z}}{2\pi i} = \frac{dx\, dy}{\pi}\,, \quad z = x + iy\,. \tag{I.18}$$

In this representation a vector

$$|\Psi\rangle = \sum_n \Psi_n|n\rangle = \sum_n \frac{\Psi_n}{\sqrt{n!}} \left(\hat{a}^{\dagger}\right)^n |0\rangle \tag{I.19}$$

is represented by a function[1]

$$\Psi(z) = \langle z|\Psi\rangle = \sum_n \frac{\Psi_n}{\sqrt{n!}}\, z^n\,. \tag{I.20}$$

The following relation (the "resolution of unity") is valid

$$I = \int d\mu_z\, e^{-|z|^2} |z\rangle\langle z|\,. \tag{I.21}$$

Here the integration extends over the entire complex plane. One also has

$$\langle z|z'\rangle = e^{\bar{z}\, z'}\,. \tag{I.22}$$

In the Segal-Bargmann representation an operator

$$\hat{A} = \sum_{n,m} |n\rangle\, A_{nm}\, \langle m| \tag{I.23}$$

is represented by a kernel

$$A(z,\bar{\zeta}) = \sum_{n,m} A_{nm}\, \frac{z^n}{\sqrt{n!}}\, \frac{\bar{\zeta}^m}{\sqrt{m!}}\,. \tag{I.24}$$

[1]This representation is similar to the coherent state representation, and differs from the latter by normalization factor $\exp(-|z|^2/2)$.

An entire function of two complex variables z and z' is uniquely defined by its value on the submanifold $z = \bar{z}'$. For this reason, one can also use $A(z, \bar{z})$ as a function representing the operator $\hat{A}$.

The action of $\hat{A}$ on an arbitrary function $f(z)$ is given by

$$(Af)(z) = \int d\mu_\zeta \, A\left(z, \bar{\zeta}\right) f(\zeta) \, e^{-|\zeta|^2}, \tag{I.25}$$

and the product of two operators is

$$(A_1 A_2)(z, \bar{z}) = \int d\mu_\zeta \, A_1\left(z, \bar{\zeta}\right) A_2(\zeta, \bar{z}) \, e^{-|\zeta|^2}. \tag{I.26}$$

Another useful representation of the operator $\hat{A}$ is the normal form representation

$$\hat{A} = \sum_{n,m} K_{nm} \left(\hat{a}^\dagger\right)^n (\hat{a})^m. \tag{I.27}$$

One can associate the following function

$$K(z, \bar{\zeta}) = \sum_{n,m} K_{nm} \, z^n \, \bar{\zeta}^m. \tag{I.28}$$

The two functions A and K connected with operator $\hat{A}$ are related as

$$A(z, \bar{z}) = e^{|z|^2} K(z, \bar{z}). \tag{I.29}$$

For this representation one has

$$\langle z|n\rangle = \frac{z^n}{\sqrt{n!}}. \tag{I.30}$$

One can also write

$$\langle z|x\rangle = \sum_n \langle z|n\rangle \langle n|x\rangle. \tag{I.31}$$

Using this relation and relations (I.11), (I.12), and (I.30), one gets

$$\langle z|x\rangle = \left(\frac{\omega}{\pi}\right)^{1/4} \exp\left[-\frac{\omega}{2} x^2 + \sqrt{2\omega}\, x\, z - \frac{1}{2} z^2\right]. \tag{I.32}$$

The following Gaussian integrals are often useful for calculations

$$\int d\mu_z \exp\left[-\frac{1}{2}(z, \bar{z}) \begin{pmatrix} A & B \\ B' & C \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \end{pmatrix} + (f_1, f_2) \begin{pmatrix} z \\ \bar{z} \end{pmatrix}\right]$$

$$= \left[\det \begin{pmatrix} B' & C \\ A & B \end{pmatrix}\right]^{-1/2} \exp\left[\frac{1}{2} (f_1, f_2) \begin{pmatrix} A & B \\ B' & C \end{pmatrix}^{-1} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}\right]. \tag{I.33}$$

Since the bilinear form in the exponent is symmetric, one has $B' = B$.

I.3 Quantum Oscillator at Finite Temperature

The density matrix $\hat{\rho}$ for a quantum oscillator at the temperature T is

$$\hat{\rho} = Z^{-1}(\beta)\, e^{-\beta\omega\hat{a}^{\dagger}\hat{a}}\,, \tag{I.34}$$

where $\beta = T^{-1}$ is the inverse temperature. Function $Z(\beta)$ can be found from the normalization condition $\mathrm{Tr}\hat{\rho} = 1$

$$Z(\beta) = \frac{1}{\exp(\beta\,\omega/2) - \exp(-\beta\,\omega/2)}\,. \tag{I.35}$$

The density matrix $\hat{\rho}$ can be identically rewritten as

$$\hat{\rho} = e^{\beta F}\, e^{-\beta\hat{H}}\,, \tag{I.36}$$

where $\hat{H}$ is the Hamiltonian of the oscillator and

$$F(\beta) = \beta^{-1}\ln[1 - \exp(-\beta\,\omega)] + \frac{1}{2}\,\omega \tag{I.37}$$

is the free energy. The entropy S defined by the response of the free energy on the change of the temperature

$$dF = -\,S\,dT \tag{I.38}$$

is of the form

$$S = \beta^2\,\frac{dF}{d\beta} = \frac{\beta\,\omega}{\exp(\beta\,\omega) - 1} - \ln[1 - \exp(-\beta\,\omega)]\,. \tag{I.39}$$

Simple calculations give the following expression for the density matrix in x-representation

$$\rho(x, x') = \langle x|\,\hat{\rho}\,|x'\rangle = \rho_0 \exp S_E(x_1, 0; x_2, \beta)$$

$$\equiv \rho_0 \exp\left\{\frac{\omega}{4}\left[\tanh\left(\frac{\omega\beta}{2}\right)(x_1 + x_2)^2 + \coth\left(\frac{\omega\beta}{2}\right)(x_1 - x_2)^2\right]\right\}. \tag{I.40}$$

Here $S_E(x_1, 0; x_2, \beta)$ is the Euclidean action for the oscillator (I.7) and

$$\rho_0 = \left[\frac{\omega}{\pi}\tanh\left(\frac{\beta\,\omega}{2}\right)\right]^{1/2}. \tag{I.41}$$

I.4 Two Mode Coupled Oscillators

The above-described formalism can be easily generalized to the case of a finite number of oscillators. Let i be an index enumerating oscillators, so that

$$\hat{H} = \sum_i \hat{H}_i\,, \qquad \hat{H}_i = \omega_i \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2}\right). \tag{I.42}$$

Here $\hat{a}_i^\dagger$ and $\hat{a}_i$ are operators of creation and annihilation for ith oscillator. All the relations of Section I.2 are easily generalized to the multi-dimensional case. In particular, in the Segal-Bargmann representation states are described by polynomials of n-complex variables $z = (z_1, z_2, \dots z_n)$. Relation (I.33) remains also valid if one understands z, $\bar{z}$, f_1, f_2 as n-dimensional objects, and A, B, and C as $n\times n$-matrices.

Consider a special type of canonical transformations for two oscillators which is of special interest for our purposes.

$$\hat{b}_1 = \cosh\sigma\,\hat{a}_1 + \sinh\sigma\,\hat{a}_2^\dagger\,,$$

$$\hat{b}_2 = \sinh\sigma\,\hat{a}_1^\dagger + \cosh\sigma\,\hat{a}_2\,. \tag{I.43}$$

Operators $\hat{b}_i$ and $\hat{b}_i^\dagger$ obey the same canonical commutation relations (I.14) as operators $\hat{a}_i$ and $\hat{a}_i^\dagger$.

The inverse transformation is of the form

$$\hat{a}_1 = \cosh\sigma\,\hat{b}_1 - \sinh\sigma\,\hat{b}_2^\dagger\,,$$

$$\hat{a}_2 = -\sinh\sigma\,\hat{b}_1^\dagger + \cosh\sigma\,\hat{b}_2\,. \tag{I.44}$$

Denote by $|0_a\rangle$ a vacuum state annihilated by the operators $\hat{a}_i$

$$\hat{a}_1|0_a\rangle = \hat{a}_2|0_a\rangle = 0\,. \tag{I.45}$$

Denote by $\Psi(z_1, z_2) = \langle z_1, z_2|\Psi\rangle$ a function that corresponds to this vector in the Segal-Bargmann representation based on $\hat{b}_i$ operators. Using relations (I.44), one has the following system of equations which determines $\Psi(z_1, z_2)$

$$\frac{\partial\Psi}{\partial z_1} - \tanh\sigma\, z_2\,\Psi = 0\,,$$

$$\frac{\partial\Psi}{\partial z_2} - \tanh\sigma\, z_1\,\Psi = 0\,. \tag{I.46}$$

The solution is

$$\Psi(z_1, z_2) = \frac{1}{\cosh\sigma}\, e^{\tanh\sigma\, z_1 z_2}\,. \tag{I.47}$$

An integration constant is found from the normalization condition $\langle\Psi|\Psi\rangle = 1$.

We shall also need an expression for state $|\Psi\rangle$ in the x-representation

$$\Psi(x_1, x_2) = \langle x_1, x_2|\Psi\rangle\,. \tag{I.48}$$

Using resolution of unity (I.21), one gets

$$\Psi(x_1, x_2) = \int d\mu_{z_1} \int d\mu_{z_2}\, e^{-|z_1|^2-|z_2|^2} \langle x_1|z_1\rangle\langle x_2|z_2\rangle\Psi(z_1, z_2)\,. \tag{I.49}$$

Using relation (I.31) and calculating Gaussian integrals with the help of (I.33), one gets

$$\Psi(x_1, x_2) = \left(\frac{\omega}{\pi}\right)^{1/2} \exp\left[-\omega\left(\frac{1}{2} + \sinh^2\sigma\right)(x_1^2 + x_2^2) + \omega\sinh 2\sigma\; x_1x_2\right]. \tag{I.50}$$

In the book we use the following special choice of the parametrization of canonical transformations (I.43)

$$\cosh\sigma = \frac{1}{\sqrt{1 - \exp(-\beta\,\omega)}}\,, \qquad \sinh\sigma = \frac{1}{\sqrt{\exp(\beta\,\omega) - 1}}\,. \tag{I.51}$$

The function $\Psi(x_1, x_2)$ in this parametrization takes the form

$$\Psi(x_1, x_2) = \left(\frac{\omega}{\pi}\right)^{1/2} \exp\left\{-\frac{\omega}{2}\left[\frac{\cosh(\frac{\beta\omega}{2})}{\sinh(\frac{\beta\omega}{2})}(x_1^2 + x_2^2) - \frac{2x_1x_2}{\sinh(\frac{\beta\omega}{2})}\right]\right\}. \tag{I.52}$$

Bibliography

ABRAHAMS, A. M., D. BERNSTEIN, D. HOBILL, E. SEIDEL, and L. SMARR: 1992, *Phys. Rev.* **D45**, 3544.

ABRAHAMS, A. M. and G. B. COOK: 1994, *Phys. Rev.* **D50**, R2364.

ABRAHAMS, A. M., G. B. COOK, S. L. SHAPIRO, and S. A. TEUKOLSKY: 1994, *Phys. Rev.* **D49**, 5153.

ABRAHAMS, A. M. and C. R. EVANS: 1992, *Phys. Rev.* **D46**, R4117.

ABRAHAMS, A. M. and C. R. EVANS: 1993, *Phys. Rev. Lett.* **70**, 2980.

ABRAHAMS, A. M. and R. H. PRICE: 1996a, *Phys. Rev.* **D53**, 1963.

ABRAHAMS, A. M. and R. H. PRICE: 1996b, *Phys. Rev.* **D53**, 1972.

ABRAMOVICI, A., W. E. ALTHOUSE, R. W. P. DREVER, Y. GURSEL, S. KAWAMURA, F. J. RAAB, D. SHOEMAKER, L. SIEVERS, R. E. SPERO, K. S. THORNE, R. E. VOGT, R. WEISS, S. E. WHITCOMB, M. E. ZUCKER: 1992, *Science* **256**, 325.

ABRAMOWICZ, M. A.: 1990, *Mon. Not. RAS* **245**, 733.

ABRAMOWICZ, M. A.: 1992, *Mon. Not. RAS* **256**, 710.

ABRAMOWICZ, M. A., X.-M. CHEN, M. GRANATH, and J.-P. LASOTA: 1996, *Astrophys. J.*, **471**, 762.

ABRAMOWICZ, M. A., X. CHEN, S. KATO, J.-P. LASOTA, and O. REGEV: 1995, *Astrophys. J.* **438**, L37.

ABRAMOWICZ, M. A., B. CZERNY, J. P. LASOTA, and E. SZUZKIEVICZ: 1988, *Astrophys. J.* **332**, 646.

ABRAMOWICZ, M. A. and P. LASOTA: 1986, *Am. J. Phys.* **54**, 936.

ABRAMOWICZ, M. A. and A. R. PRASANNA: 1990, *Mon. Not. RAS* **245**, 720.

ABRAMOWICZ, M. A., E. SZUSKIEWICZ, and F. H. WALLINDER: 1989, in F. Meyer, W. J. Duschl, J. Frank and E. Meyer-Hofmeister (eds.), *Theory of Accretion Discs*, NATO ASI Series, Vol. **290**, Kluwer, p.141.

ADLER, S. L.: 1982, *Rev. Mod. Phys.* **54**, 729.

AFFLECK, I. K., O. ALVAREZ, and N. S. MANTON: 1982, *Nucl. Phys.* **B197**, 509.

AFFLECK, I. K. and N. S. MANTON: 1982, *Nucl. Phys.* **B194**, 38.

AHARONOV, Y., A. CASHER, and S. NUSSINOV: 1987, *Phys. Lett.* **B191**, 51.

AICHELBURG, P. C. and P. U. SEXL: 1971, *Gen. Rel. & Grav.* **2**, 303.

AITCHISON, I. J. R.: 1985, *Contemp. Phys.* **26**, 333.

AKHMEDOV, E. T.: 1997, E-print hep-th/9711153.

ALCOCK, C., E. FARHI, and A. V. OLINTO: 1986, *Astrophys. J.* **310**, 261.

ALIEV, A. I., D. V. GAL'TSOV, and A. A. SOKOLOV: 1980, *Izv. Vuzov, Fizika* No. 3, 7.

AMAN, E. G.: 1983, Preprint Nuclear Research Institute, USSR Academy of Sciences, P-0272, Moscow.

ANDERSON, A. and R. H. PRICE: 1991, *Phys. Rev.* **D43**, 3147.

ANDERSON, P. R.: 1990, *Phys. Rev.* **D41**, 1152.

ANDERSON, P. R., W. A. HISCOCK, and D. A. SAMUEL: 1995, *Phys. Rev.* **D51**, 4337.

ANDERSON, W. G.: 1994, *Phys. Rev.* **D50**, 4786.

ANDERSON, W. G. and W. ISRAEL: 1997, unpublished.

ANDERSSON, N.: 1992, *Proc. R. Soc. London* **439**, 47.

ANDERSSON, N.: 1993a, *Proc. R. Soc. London* **442**, 427.

ANDERSSON, N.: 1993b, *Class. Quantum Grav.* **10**, L61.

ANDERSSON, N.: 1994, *Class. Quantum Grav.* **11** 3003.

ANDERSSON, N.: 1995a, *Phys. Rev.* **D51**, 353.

ANDERSSON, N.: 1995b, *Phys. Rev.* **D52**, 1808.

ANDERSSON, N.: 1997, *Phys. Rev.* **D55**, 468.

ANDERSSON, N., M. E. ARAÚJO, and B. F. SCHUTZ: 1993a, *Class. Quantum Grav.* **10** 735.

ANDERSSON, N., M. E. ARAÚJO, and B. F. SCHUTZ: 1993b, *Class. Quantum Grav.* **10** 757.

ANDERSSON, N., M. E. ARAÚJO, and B. F. SCHUTZ: 1994, *Phys. Rev.* **D49**, 2703.

ANDERSSON, N and S. LINNÆUS: 1992, *Phys. Rev.* **D46**, 4179.

ANDERSSON, N. and H. ONOZAWA: 1996, *Phys. Rev.* **D54**, 7470.

ANDERSSON, N. and K.-E. THYLWE: 1994, *Class. Quantum Grav.* **11**, 2991.

ANNINOS, P., D. BERNSTEIN, S. BRANDT, D. HOBILL, E. SEIDEL, and L. SMARR: 1994, *Phys. Rev.* **D50**, 3801.

ANNINOS, P., D. BERNSTEIN, S. BRANDT, J. LIBSON, J. MASSÓ, E. SEIDEL, L. SMARR, W.-M. SUEN, and P. WALKER: 1995, *Phys. Rev. Lett.* **74**, 630.

ANNINOS, P., K. CAMARDA, J. MASSÓ, E. SEIDEL, W.-M. SUEN, and J. TOWNS: 1995, *Phys. Rev.* **D52**, 2059.

ANNINOS, P., G. DAUES, J. MASSÓ, E. SEIDEL, and W.-M. SUEN: 1995, *Phys. Rev.* **D51**, 5562.

ANNINOS, P., C. DEWITT-MORETTE, R. A. MATZNER, P. YIOUTAS, and T. R. ZHANG: 1992, *Phys. Rev.* **D46**, 4477.

ANNINOS, P., D. HOBILL, E. SEIDEL, L. SMARR, and W.-M. SUEN: 1993, *Phys. Rev. Lett.* **71**, 2851.

ANNINOS, P., D. HOBILL, E. SEIDEL, L. SMARR, and W.-M. SUEN: 1995, *Phys. Rev.* **D52**, 2044.

ANNINOS, P., R. H. PRICE, J. PULLIN, E. SEIDEL, and W.-M. SUEN: 1995, *Phys. Rev.* **D52**, 4462.

APOSTOLATOS, T., D. KENNEFICK, A. ORI, and E. POISSON: 1993, *Phys. Rev.* **D47**, 5376.

ARAÚJO, M. E., D. NICHOLSON, and B. F. SCHUTZ: 1993, *Class. Quantum Grav.* **10**, 1127

ARAKI, H. and E. H. LIEB: 1970, *Commun. Math. Phys.* **18**, 160.

AREF'EVA, I. YA., K. S. VISWANATHAN, and I. V. VOLOVICH: 1995, *Nucl. Phys.* **B452**, 346.

ARNOLD, V. I.: 1990, *Singularities of caustics and wave fronts* (Kluwer Academic Publishers, Dordrecht).

ARNOWITT, R., S. DESER, and C. W. MISNER: 1963, in L. Witten (ed.), *Gravitation, an Introduction to Current Research*, N. Y.-London, p. 227.

ARONS, J.: 1983, *Astrophys. J.* **266**, 215.

ARTEMOVA, J., G. BISNOVATYI-KOGAN, G. BJÖRNSSON, and I. NOVIKOV: 1996, *Astrophys. J.* **456**, 119.

BAADE, W. and F. ZWICKY: 1934, *Proc. Nat. Acad. Sci.* **20**, 254.

BACHELOT, A. and A. MOTET-BACHELOT: 1992, in *Proceedings of the IV International Conference on Hyperbolic Problems*, Vieweg (Taosmina).

BAHCALL, S., B. W. LYNN, and S. B. SELIPSKY: 1990, *Astrophys. J.* **362**, 251.

BALASIN, H. and H. NACHBAGAUER: 1995, *Class. Quantum Grav.* **12**, 707.

BALASIN, H. and H. NACHBAGAUER: 1996, *Class. Quantum Grav.* **13**, 731.

BALBINOT, R.: 1984, *Phys. Lett.* **B136**, 337.

BALBINOT, R. and R. BERGAMINI: 1982, *Nuovo Cim.* **B68**, 104.

BALBINOT, R., R. BERGAMINI, and B. GIORGINI: 1982, *Nuovo Cim.* **B71**, 27.

BALBINOT, R. and M. R. BROWN: 1984, *Phys. Lett.* **A100**, 80.

BALBINOT, R. and E. POISSON: 1990, *Phys. Rev.* **D41**, 395.

BALEK, V., J. BIČÁK, and Z. STUCHLÍK: 1989, *Bull. Astron. Ist. Czechosl.* **40**, 133.

BANACH, R. and J. S. DOWKER: 1979a, *J. Phys.* **A12**, 2255.

BANACH, R. and J. S. DOWKER: 1979b, *J. Phys.* **A12**, 2545.

BAÑADOS, M., M. HENNEAUX, C. TEITELBOIM, and J. ZANELLI: 1993, *Phys. Rev.* **D48**, 1506.

BAÑADOS, M., C. TEITELBOIM, and J. ZANELLI: 1992, *Phys. Rev. Lett.* **69**, 1849.

BAÑADOS, M., C. TEITELBOIM, and J. ZANELLI: 1994, *Phys. Rev. Lett.* **72**, 957.

BANKS, T.: 1995, *Nucl. Phys. Proc. Suppl.* **41**, 21.

BANKS, T., L. SUSSKIND, and M. E. PESKIN: 1984, *Nucl. Phys.* **B244**, 125.

BARDEEN, J. M.: 1968, *Bull. Am. Phys. Soc.* **13**, 41.

BARDEEN, J. M.: 1981, *Phys. Rev. Lett.* **46**, 382.

BARDEEN, J. M., B. CARTER, and S. W. HAWKING: 1973, *Commun. Math. Phys.* **31**, 161.

BARDEEN, J. M. and J. A. PETTERSON: 1975, *Astrophys. J.* **195**, L65.

BARDEEN, J. M. and T. PIRAN: 1983, *Phys. Rep.* **196**, 205.

BARDEEN, J. M. and W. H. PRESS: 1973, *J. Math. Phys.* **14**, 7.

BARDEEN, J. M., W. H. PRESS, and S. A. TEUKOLSKY: 1972, *Astrophys. J.* **178**, 347.

BARRABÈS, C., B. BOISSEAU, and W. ISRAEL: 1995, *Mon. Not. RAS* **276**, 432.

BARRABÈS, C. and V. FROLOV: 1996, *Phys. Rev.* **D53**, 3315.

BARRABÈS, C., A. GRAMAIN, E. LESIGNE, and P. S. LETELIER: 1992, *Class. Quantum Grav.* **9**, L105.

BARRABÈS, C. and W. ISRAEL: 1991, *Phys. Rev.* **D43**, 1129.

BARRABÈS, C., W. ISRAEL, and P. S. LETELIER: 1991, *Phys. Lett.* **A160**, 41.

BARRABÈS, C., W. ISRAEL, and E. POISSON: 1990, *Class. Quantum Grav.* **7**, L273.

BARROW, I. D.: 1982, *Cosmology and elementary particle physics*, in *Fundamentals of Cosmic Physics*, **7**, Gordon and Breach.

BARROW, I. D. and I. SILK: 1983, *The Left Hand of Creation*, Basic Books, N. Y.

BARROW, J. D., BURD, A. B., and D. LANCASTER: 1986, *Class. Quantum Grav.* **3**, 551.

BARROW, J. D., E. J. COPELAND, and A. R. LIDDLE: 1992, *Phys. Rev.* **D46**, 645.

BARTNIK, R. and J. MCKINNON: 1988, *Phys. Rev. Lett.* **61**, 141.

BARVINSKY, A. O., V. P. FROLOV, and A. I. ZELNIKOV: 1995, *Phys. Rev.* **D51**, 1741.

BAYM, G. and C. J. PETHICK: 1979, *Ann. Rev. Astr. Ap.* **17**, 415.

BEGELMAN, M. C., R. D. BLANDFORD, and M. I. REES: 1984, *Rev. Mod. Phys.* **56**, 255.

BEGELMAN, M. C. and D. L. MEIER: 1982, *Astrophys. J.* **253**, 873.

BEGELMAN, M. and M. REES: 1996, *Gravity's Fatal Attraction: Black Holes in the Universe.* (Scientific Americal Library) W. H. Freeman.

BEKENSTEIN, J. D.: 1972, *Lett. Nuovo Cim.* **4**, 737.

BEKENSTEIN, J. D.: 1972a, *Phys. Rev. Lett.* **28**, 452.

BEKENSTEIN, J. D.: 1972b, *Phys. Rev.* **D5**, 1239.

BEKENSTEIN, J. D.: 1972c, *Phys. Rev.* **D5**, 2403.

BEKENSTEIN, J. D.: 1973a, *Phys. Rev.* **D7**, 949.

BEKENSTEIN, J. D.: 1973b, *Phys. Rev.* **D7**, 2333.

BEKENSTEIN, J. D.: 1974, *Phys. Rev.* **D9**, 3292.

BEKENSTEIN, J. D.: 1975, *Ann. Phys.* **91**, 75.

BEKENSTEIN, J. D.: 1980, *Physics Today* **33**, 24.

BEKENSTEIN, J. D.: 1982, *Phys. Rev.* **D26**, 950.

BEKENSTEIN, J. D.: 1983, *Phys. Rev.* **D27**, 2262.

BEKENSTEIN, J. D.: 1995, *Phys. Rev.* **D51**, R6608.

BEKENSTEIN, J. D.: 1996, *Talk given at 2nd International Sakharov Conference on Physics, Moscow, Russia, 20-23 May 1996*, E-print gr-qc/9605059.

BEKENSTEIN, J. D. and A. MEISELS: 1977, *Phys. Rev.* **D15**, 2775.

BEKENSTEIN, J. D. and V. F. MUKHANOV: 1995, *Phys. Lett.* **B360**, 7.

BELINSKY, V. A., E. M. LIFSHITZ, and I. M. KHALATHNIKOV: 1970, *Usp. Fiz. Nauk* **102**, 463.

BENDER, P., E. BLOOM, L. COMINSKY, H. FORD, A. HARMON, P. MICHELSON, I. NOVIKOV, K. POSTNOV, J. SCARGLE, J. SWANK, J. C. WHEELER, and K. WOOD: 1995, in: E. W. Kolb and R. D. Peccei (eds.), *Particles and Nuclear Astrophysics and Cosmology in the Next Millenium*, Proceedings of the 1994 Snowmass Summer Study, p.377.

BEREZIN, F. A.: 1965, *Method of Secondary Quantization*: Nauka, Moscow. (1966, Academic Press, N. Y.)

BEREZIN, V. A.: 1980, Preprint NRI, P-0183, Moscow.

BEREZIN, V.: 1990, *Phys. Lett.* **B241**, 194.

BEREZIN, V.: 1997, *Phys. Rev.* **D55**, 2139.

BEREZIN, V. A., V. A. KUZMIN, and I. I. TKACHEV: 1988, *Plys. Lett.* **B207**, 397.

BERNSTEIN, D., D. HOBILL, E. SEIDEL, and L. SMARR: 1994a, *Phys. Rev.* **D50**, 3760.

BERNSTEIN, D., D. HOBILL, E. SEIDEL, L. SMARR, and J. TOWNS: 1994b, *Phys. Rev.* **D50**, 5000.

BESKIN, V. S.: 1997, *Physics - Uspekhi* **40**, 659.

BESKIN, V. S., A. V. GUREVICH, and YA. N. ISTOMIN: 1983, *Zh. Eksp. Teor. Fiz.* **85**, 401.

BESKIN, V. S., A. V. GUREVICH, and YA. N. ISTOMIN: 1993, *Physics of Pulsar Magnetosphere*, Cambridge University Press, Cambridge.

BESKIN, V. S., YA. N. ISTOMIN, and V. I. PARIEV: 1992, in J. Roland, H. Sol, and G. Pelletier (eds.), *Extragalactic radio sources - from Beams to Jets.* Proc. 7th A. I. P. Meeting, Cambridge University Press, Cambridge.

BESKIN, V. S., YA. N. ISTOMIN, and V. I. PARIEV: 1993, *Sov. Astron.* **36**, 642.

BESKIN, V. S. and L. M. MALISHKIN: 1996, *Astronomy Lett.* **22**, 475.

BESKIN, V. S. and V. I. PARIEV: 1993, *Sov. Phys. Uspekhi* **36**, 529.

BESKIN, V. S. and U. N. PIDOPRIGORA: 1995, *Zh. Eksp. Teor. Fiz.* **107**, 1025.

BIČÁK, J.: 1972, *Gen. Rel. & Grav.* **3**, 331.

BIČÁK, J.: 1980a, *Proc. R. Soc. London* **A371**, 429.

BIČÁK, J.: 1980b, *Gen. Rel. & Grav.* **12**, 195.

BIČÁK, J.: 1982, in R. Ruffini (ed.), *The Proceedings of the Second Marcel Grossmann Meeting on General Relativity*, North-Holland Publishing Company, p.277.

BIČÁK, J.: 1983, in B. Bertotti, F. de Felice, A. Pascolini (eds.), *10th Int. Conf. on General Relativity and Gravitation*, Padova, **2**, p. 688.

BIČAK, J. and L. DVOŘÁK: 1976, *Gen. Rel. & Grav.* **7**, 959.

BIČÁK, J. and L. DVOŘÁK: 1980, *Phys. Rev.* **D22**, 2933.

BIČÁK, J. and C. HOENSELAERS: 1985, *Phys. Rev.* **D31**, 2476.

BIČÁK, J. and V. JANIŠ: 1985, *Mon. Not. RAS* **212**, 899.

BIČÁK, J., O. SEMERÁK, and P. HADRAVA: 1993, *Mon. Not. RAS* **263**, 545.

BIČÁK, J., Z. STUCHLÍK, and V. BALEK: 1989, *Bull. Astron. Ist. Czechosl.* **40**, 65.

BIČÁK, J. and Z. STUCHLÍK: 1976, *Mon. Not. RAS* **175**, 381.

BIRKHOFF, G. D.: 1923, *Relativity and Modern Physics*, Harvard Univ. Press.

BIRRELL, N. D. and P. C. W. DAVIES: 1982, *Quantum Fields in Curved Space*, Cambridge University Press, Cambridge, England.

BISHOP, N. T.: 1982, *Gen. Rel. & Grav.* **14**, 717.

BISHOP, N. T.: 1984, *Gen. Rel. & Grav.* **16**, 589.

BISNOVATYI-KOGAN, G. S.: 1989, *Physical Problems of the Theory of Stellar Evolution*, Nauka, Moscow (in Russian).

BISNOVATYI-KOGAN, G. S., V. N. LUKASH, and I. D. NOVIKOV: 1980, in *Proc. of the Fifth European Regional Meeting in Astronomy*, Liege, Belgium, G. 1.1.

BIZON, P.: 1990, *Phys. Rev. Lett.* **64**, 2844.

BIZON, P. and T. CHMAJ: 1992, *Phys. Lett.* **B297**, 55.

BJÖRNSSON, G. and R. SVENSSON: 1991, *Astrophys. J.* **371**, L69.

BJÖRNSSON, G. and R. SVENSSON: 1992, *Astrophys. J.* **394**, 500.

BLANCHET, L., T. DAMOUR, B. R. IYER, C. M. WILL, AND A. G. WISEMAN: 1995, *Phys. Rev. Lett.* **74**, 3515.

BLANDFORD, R. D.: 1976, *Mon. Not. RAS* **176**, 465.

BLANDFORD, R. D.: 1979, in C. Hazard and S. Mitton (eds.), *Active Galactic Nuclei*, Cambridge Univ. Press, p. 241.

BLANDFORD, R. D.: 1987, in S. W. Hawking and W. Israel (eds.), *300 Years of Gravitation*, Cambridge University Press, p.199.

BLANDFORD, R. D.: 1990, in R. D. Blandford, H. Netzer, and L. Woltier (eds.), *Active Galacttic Nuclei*, Spring Verlag, Berlin.

BLANDFORD, R. D. and D. G. PAYNE: 1982, *Mon. Not. RAS* **199**, 883.

BLANDFORD, R. D. and M. J. REES: 1978, *Phys. Sci.* **17**, 265.

BLANDFORD, R. D. and K. S. THORNE: 1979, in S. W. Hawking and W. Israel (eds.), *General Relativity, An Einstein Centenary Survey*, Cambridge Univ. Press, p. 454.

BLANDFORD, R. D. and R. L. ZNAJEK: 1977, *Mon. Not. RAS* **179**, 433.

BLEYER, U., V. P. FROLOV, and A. I. ZEL'NIKOV: 1987, *Ann. Physik* **44**, 371.

BLOME, H. J and B. MASHHOON: 1984, *Phys. Lett.* **A100**, 231.

BOCHAROVA, N. M., K. A. BRONNIKOV, and V. N. MEL'NIKOV: 1970, *Vestnik MGU* **1**, 706.

BOCHKAREV, N. G., R. A. SUNYAEV, T. S. KRUZINA, A. M. CHEREPASHCHUK, and N. I. SHAKURA: 1988, *Sov. Astron. J.* **32**, 405.

BOERSMA, S. and T. DRAY: 1995a, *J. Math. Phys.* **36**, 1378.

BOERSMA, S. and T. DRAY: 1995b, *J. Math. Phys.* **36**, 1394.

BOERSMA, S. and T. DRAY: 1995c, *Gen. Rel. & Grav.* **27**, 319.

BOGORODSKY, A. F.: 1962, *Einstein's Field Equations and Their Application to Astronomy*, Kiev Univ.

BOGOVALOV, S. V.: 1989, *Soviet Astron Lett.* **15**, 469.

BOLASHENKO, P. A. and V. P. FROLOV: 1984, in *Abstracts of the USSR Conference on the Modern Theoretical and Experimental Problems of Relativity and Gravitation*, Moscow, **1**, 102.

BOLASHENKO, P. A. and V. P. FROLOV: 1986, *P. N. Lebedev Physics Inst. Trudy* **169**, 159; *Proceedings of the Lebedev Physics Institute of the Academy of Sciences of the USSR*, **169**: *The Physical Effects in the Gravitational Field of Black Holes* (ed. M. A. Markov), 1987, Nova Science Publishers, Commack.

BOLASHENKO, P. A. and V. P. FROLOV: 1987, *Kratkie Soobshcheniya po Fizike* (Soviet Physics — Lebedev Inst. Reports), No. 12, p. 42.

BOLASHENKO, P. A. and V. P. FROLOV: 1988, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Fourth Seminar on Quantum Gravity*, Moscow, World Scientific, Singapore, p. 654.

BOLASHENKO, P. A. and V. P. FROLOV: 1989, *Proceedings of the P. N. Lebedev Physics Institute of the Academy of Sciences of the USSR*, **197**: *Quantum Theory and Gravitation* (ed. M. A. Markov), 1989, Nova Science Publishers, Commack.

BOLOTOVSKY, B. M. and S. M. STOLYAROV: 1980, in *Annual Einstein Memorial volume* Nauka, Moscow.

BOLTON, C. T.: 1972a, *Nature* **235**, 271.

BOLTON, C. T.: 1972b, *Nature Phys. Sci.* **240**, 124.

BOLTON, C. T.: 1975, *Astrophys. J.* **200**, 269.

BOMBELLI, L., R. KOUL, J. LEE, and R. SORKIN: 1986, *Phys. Rev.* **D34**, 373.

BONANNO, A., S. DROZ, W. ISRAEL, and S. MORSINK: 1994, *Can. J. Phys* **72** 755-759. (erratum: *Can. J. Phys* **73**, 251 (1995)).

BONANNO, A., S. DROZ, W. ISRAEL, and S. MORSINK: 1995, *Proc. R. Soc. London* **A450** 553.

BONDI, H.: 1952, *Mon. Not. RAS* **112**, 195

BONDI, H. and T. GOLD: 1955, *Proc. R. Soc. London* **A229**, 416.

BONDI, H. and F. HOYLE: 1944, *Mon. Not. RAS* **104**, 273.

BONDI, H., M. G. J. VAN DER BURG, and A. W. K. METZNER: 1962, *Proc. R. Soc. London* **A269**, 21.

BONNOR, W. B.: 1953, *Proc. Phys. Soc. London*, **A67**, 225.

BONNOR, W. B. and P. C. VAIDYA: 1970, *Gen. Rel. & Grav.* **1**, 127.

BOSE, S. K.: 1975, *J. Math. Phys.* **16**, 772.

BOULWARE, D. G.: 1975a, *Phys. Rev.* **D11**, 1404.

BOULWARE, D. G.: 1975b, *Phys. Rev.* **D12**, 350.

BOULWARE, D. G.: 1976, *Phys. Rev.* **D13**, 2169.

BOULWARE, D. G.: 1980, *Ann. Phys.* **124**, 169.

BOULWARE, D. G.: 1992, *Phys. Rev. D* **46**, 4421.

BOUSSO, R. and S. W. HAWKING: 1996, *Phys. Rev.* **D54**, 6312.

BOVYN, M.: 1983, *Phys. Rev.* **D28**, 703.

BOWEN, J., J. RAUBER, and J. W. YORK JR.: 1984, *Class. Quantum Grav.* **1**, 591.

BOWEN, J. and J. W. YORK JR.: 1980, *Phys. Rev.* **D21**, 2047.

BOWICK, M. J., S. B. GIDDINGS, J. A. HARVEY, G. T. HOROWITZ, AND A. STROMINGER: 1988, *Phys. Rev. Lett.* **61**, 2823.

BOYER, R. H.: 1969, *Proc. R. Soc. London* **A311**, 245.

BOYER, R. H. and R. W. LINDQUIST: 1967, *J. Math. Phys.* **8**, 265.

BRADEN, H. W., J. D. BROWN, B. F. WHITING, and J. W. YORK, JR.: 1990, *Phys. Rev.* **D42**, 3376.

BRADEN, H. W., B. F. WHITING, and J. W. YORK JR.: 1987, *Phys. Rev.* **D36**, 3614.

BRADY, P. R.: 1994, *Class. Quantum Grav.* **11**, 1255.

BRADY, P. R.: 1995, *Phys. Rev.* **D51**, 4168.

BRADY, P. and C. M. CHAMBERS: 1995, *Phys. Rev.* **D51**, 4177.

BRADY, P. and J. SMITH: 1995, *Phys. Rev. Letters* **75**, 1256.

BRANDERBURG, A., Å. NORDLUND, R. F. STEIN, and U. TORKELSON: 1995, *Astrophys. J.* **446**, 741.

BRANDERBURG, A., Å. NORDLUND, R. F. STEIN, and U. TORKELSON: 1996, *Astrophys. J.* **458**, L45.

BRANDT, S. R. and E. SEIDEL: 1995a, *Phys. Rev.* **D52**, 856.

BRANDT, S. R. and E. SEIDEL: 1995b, *Phys. Rev.* **D52**, 870.

BRECKENRIDGE, J. C., D. A. LOWE, R. C. MYERS, A. W. PEET, A. STROMINGER, AND C. VAFA: 1996, *Phys. Lett* **B381**, 423.

BRECKENRIDGE, J. C., R. C. MYERS, A. W. PEET, and C. VAFA: 1997, *Phys. Lett.* **B391**, 93.

BREITENLOHNER, P., P. FORGÁCS, and D. MAISON: 1992, *Nucl. Phys.* **B383**, 357.

BREUER, R. A.: 1975, *Gravitational Perturbation Theory and Synchrotron Radiation*, Springer Lect. Notes in Physics, no **44**.

BREUER, R. A., M. P. RYAN JR., and S. WALLER: 1977, *Proc. R. Soc. London* **A358**, 71.

BRILL, D. R.: 1996, *Lecture given at "Raychaudhuri session", ICGC-95 conference, Puna (India), 1995*; E-print gr-qc/9607026.

BRILL, D. R., P. L. CHRZANOWSKI, C. M. PEREIRA, E. D. FACKERELL, AND J. R. IPSER: 1972, *Phys. Rev* **D5**, 1913.

BRILL, D. R. and R. W. LINDQUIST: 1963, *Phys. Rev.* **131**, 471.

BRILLOUIN, L.: 1956, *Science and Information Theory*, Academic Press, N. Y.

BRONNIKOV, K. A. and YU. N. KIREEV: 1978, *Phys. Lett.* **A67**, 95.

BROUT, R., S. MASSAR, R. PARENTANI, and P. SPINDEL: 1995, *Phys. Rev.* **D52**, 4559.

BROWDY, S. and G. J. GALLOWAY: 1995, *J. Math. Phys.* **36**, 4952.

BROWN, G. and H. E. BETHE: 1994, *Astrophys. J.* **423**, 659.

BROWN, J. D.: 1995, *Phys. Rev.* **D52**, 7011.

BROWN, J. D., G. L. COMER, E. A. MARTINEZ, J. MELMED, B. F. WHITING, AND J. W. YORK, JR.: 1990, *Class. Quantum Grav.* **7**, 1433.

BROWN, J. D., E. A. MARTINEZ, J. W. YORK, JR.: 1991, *Phys. Rev. Lett.* **66**, 2281.

BROWN, J. D., and J. W. YORK JR.: 1993a, *Phys. Rev.* **D47**, 1407.

BROWN, J. D., and J. W. YORK JR.: 1993b, *Phys. Rev.* **D47**, 1420.

BROWN, M. R.: 1984, *J. Math. Phys.* **25**, 136.

BROWN, M. R. and A. C. OTTEWILL: 1985, *Phys. Rev.* **D31**, 2514.

BROWN, M. R., A. C. OTTEWILL, and D. N. PAGE: 1986, *Phys. Rev.* **D33**, 2840.

BUNTING, G.: 1983, *Proof of the Uniqueness Conjecture for Black Holes*, PhD Thesis, Dept. of Math., Univ. of New England, Armilade, Australia.

BUNTING, G. L. and A. K. M. MASOOD-UL-ALAM: 1987, *Gen. Rel. & Grav.* **19**, 174.

BURBIDGE, G. R.: 1972, *Comm. Astrophys. Space. Sci.* **4**, 105.

CADERNI, N. and M. CALVANI: 1979, *Phys. Lett.* **A71**, 1.

ČADEŽ, A.: 1974, *Ann. Phys.* **83**, 449.

CALDWELL, R. R., H. A. CHAMBLIN, and G. W. GIBBONS: 1996, *Phys. Rev.* **53**, 7103.

CALLAN, C. G., D. FRIEDAN, E. J. MARTINEC, and M. J. PERRY: 1985, *Nucl. Phys.* **B262**, 593.

CALLAN, C. G., JR. and J. M. MALDACENA: 1996, *Nucl. Phys.* **B472**, 591.

CALLAN, C. and F. WILCZEK: 1995, *Phys. Lett.* **B333**, 55. L2597

CALLEN, H. B.: 1960, *Thermodynamics*, Wiley, New York.

CALVANI, M.: 1980, *Nuovo Cim.* **A58**, 364.

CAMENZIND, M: 1986a, *Astron. Astrophys.* **156,** 137.

CAMENZIND, M: 1986b, *Astron. Astrophys.* **162,** 32.

CAMENZIND, M: 1987, *Astron. Astrophys.* **184,** 341.

CAMENZIND, M: 1990, *Rev. Mod. Astron.* **3,** 234.

CAMPBELL, B. A., M. J. DUNCAN, N. KALOPER, and K. A. OLIVE: 1990, *Phys. Lett.* **B251**, 34.

CAMPBELL, B. A., N. KALOPER, and K. A. OLIVE: 1991, *Phys. Lett.* **B263**, 364.

CAMPBELL, B. A., N. KALOPER, and K. A. OLIVE: 1992, *Phys. Lett.* **B285**, 199.

CANDELAS, P.: 1980, *Phys. Rev.* **D21**, 2185.

CANDELAS, P., P. CHRZANOWSKI, and K. W. HOWARD: 1981, *Phys. Rev.* **D24**, 297.

CANDELAS, P. and K. W. HOWARD: 1984, *Phys. Rev.* **D29**, 1618.

CANDELAS, P. and B. P. JENSEN: 1986, *Phys. Rev.* **D33**, 1596 (1986).

CAPPER, D. M. and M. J. DUFF: 1974, *Nuovo Cim.* **A23**, 173.

CARLINI, A., V. FROLOV, M. MENSKY, I. NOVIKOV, and H. SOLENG: 1995, *Int. J. Mod. Phys.* **D4**, 557.

CARLINI, A. and I. NOVIKOV: 1996, *Int. J. Mod. Phys.* **D5**, 445.

CARLIP, S.: 1995, *Class. Quantum Grav.* **12**, 2853.

CARLITZ, R. D. and R. S. WILLEY: 1987, *Phys. Rev.* **D36**, 2336.

CARR, B. J.: 1983, in M. A. Markov and P. C. West (eds.), *Quantum Gravity*, Plenum Press, N. Y., p. 337.

CARR, B. J.: 1992, *Report at Zeldovich memorial meeting*, Moscow (unpublished).

CARR, B. J., J. H. GILBERT, and J. E. LIDSEY: 1994, *Phys. Rev.* **D50**, 4853.

CARR, B. J. and S. W. HAWKING: 1974, *Mon. Not. RAS* **168**, 399.

CARTER, B.: 1966a, *Phys. Lett.* **21**, 423.

CARTER, B.: 1966b, *Phys. Rev.* **141**, 1242.

CARTER, B.: 1968a, *Phys. Rev.* **174**, 1559.

CARTER, B.: 1968b, *Commun. Math. Phys.* **10**, 280.

CARTER, B.: 1969, *J. Math. Phys.* **10**, 70.

CARTER, B.: 1971, *Phys. Rev. Lett.* **26**, 331.

CARTER, B.: 1973a, in C. DeWitt and B. S. DeWitt (eds.), *Black Holes*, Gordon and Breach, New York.

CARTER, B.: 1973b, *Commun. Math. Phys.* **30**, 261.

CARTER, B.: 1974, *Phys. Rev. Lett.* **33**, 558.

CARTER, B.: 1976, in R. Ruffini (ed.), *Proceedings of the First Marcel Grossmann Meeting on General Relativity*, North-Holland Amsterdam.

CARTER, B.: 1977, *Phys. Rev.* **D16**, 3395.

CARTER, B.: 1979, in S. W. Hawking and W. Israel (eds.), *General Relativity, An Einstein Centenary Survey*, Cambridge Univ. Press.

CARTER, B.: 1985, *Commun. Math. Phys.* **99**, 563.

CARTER, B.: 1987, in A. Coley, C. Dyer, T. Tupper (eds.), *Proceedings of the 2nd Canadian Conference on General Relativity and Relativistic Astrophysics*, Toronto, World Scientific, Singapore.

CARTER, B. and V. P. FROLOV: 1989, *Class. Quantum Grav.* **6**, 569.

CARTER, B., V. P. FROLOV, and O. HEINRICH: 1991, *Class. Quantum Grav.* **8**, 135.

CASSIDY, M. J. and S. W. HAWKING: 1997, E-print hep-th/9709066.

CHAN, K. C. K. and R. B. MANN: 1995, *Class. Quantum Grav.* **12**, 1609.

CHANDRASEKHAR, S.: 1931, *Astrophys. J.* **74**, 81.

CHANDRASEKHAR, S.: 1975, *Proc. R. Soc. London* **A343**, 289.

CHANDRASEKHAR, S.: 1976a, *Proc. R. Soc. London* **A348**, 39.

CHANDRASEKHAR, S.: 1976b, *Proc. R. Soc. London* **A349**, 571.

CHANDRASEKHAR, S.: 1979a, *Proc. R. Soc. London* **A365**, 453.

CHANDRASEKHAR, S.: 1979b, in S. W. Hawking and W. Israel (eds.), *General Relativity, An Einstein Centenary Survey*, Cambridge Univ. Press, p. 370.

CHANDRASEKHAR, S.: 1980, *Proc. R. Soc. London* **A369**, 425.

CHANDRASEKHAR, S.: 1983, *The Mathematical Theory of Black Holes*, Clarendon Press, Oxford.

CHANDRASEKHAR, S.: 1984, *Proc. R. Soc. London* **A392**, 1.

CHANDRASEKHAR, S. and S. DETWEILER: 1975a, *Proc. R. Soc. London* **A344**, 441.

CHANDRASEKHAR, S. and S. DETWEILER: 1975b, *Proc. R. Soc. London* **A345**, 145.

CHANDRASEKHAR, S. and S. DETWEILER: 1976, *Proc. R. Soc. London* **A350**, 165.

CHANDRASEKHAR, S. and S. DETWEILER: 1977, *Proc. R. Soc. London* **A352**, 325.

CHANDRASEKHAR, S. and J. B. HARTLE: 1982, *Proc. R. Soc. London* **A384**, 301.

CHANDRASEKHAR, S. and B. S. XANTHOPOULOS: 1979, *Proc. R. Soc. London* **A367**, 1.
CHASE, J. E.: 1970, *Commun. Math. Phys.* **19**, 276.
CHEREPASHCHUK, A. M.: 1996, *Russian Physics - Uspekhi*, **39**, 759.
CHIBA, T., T. NAKAMURA, K. NAKAO, and M. SASAKI: 1994, *Class. Quantum Grav.* **11**, 431.
CHIBA, T. and J. SODA: 1996, *Prog. Theor. Phys.* **96**, 567.
CHING, E. S. C., P. T. LEUNG, W. M. SUEN, and K. YOUNG: 1995a, *Phys. Rev. Lett.* **74**, 2414.
CHING, E. S. C., P. T. LEUNG, W. M. SUEN, and K. YOUNG: 1995b, *Phys. Rev.* **D52**, 2118.
CHING, E. S. C., P. T. LEUNG, W. M. SUEN, and K. YOUNG: 1995c, *Phys. Rev. Lett.* **74**, 4588.
CHING, E. S. C., P. T. LEUNG, W. M. SUEN, and K. YOUNG: 1995d, *Wave-propagation in gravitational systems: Completeness of quasinormal modes* (Preprint
CHITRE, D. M.: 1975, *Phys. Rev.* **D11**, 760.
CHITRE, D. M.: 1976, *Phys. Rev.* **D13**, 2713.
CHO, Y. M. and P. G. O. FREUND: 1975, *Phys. Rev.* **D12**, 1588.
CHODOS, A. and S. DETWEILER: 1982, *Gen. Rel. & Grav.* **14**, 879.
CHOPTUIK, M. W.: 1993, *Phys. Rev. Lett.* **70**, 9.
CHOPTUIK, M. W., T. CHMAJ, and P. BIZON: 1996, *Phys. Rev. Lett.* **77**, 424.
CHRISTENSEN, S. M.: 1976, *Phys. Rev.* **D14**, 2490.
CHRISTENSEN, S. M.: 1978, *Phys. Rev.* **D17**, 946.
CHRISTENSEN, S. M. and S. A. FULLING: 1977, *Phys. Rev.* **D15**, 2088.
CHRISTODOULOU, D.: 1970, *Phys. Rev. Lett.* **25**, 1596.
CHRISTODOULOU, D.: 1984, *Commun. Math. Phys.* **93**, 171.
CHRZANOWSKI, P. L.: 1975, *Phys. Rev.* **D11**, 2042.
CHRZANOWSKI, P. L., R. A. MATZNER, V. D. SANDBERG, and M. P. RYAN JR: 1976, *Phys. Rev.* **D14**, 317.
CHRZANOWSKI, P. L. and C. W. MISNER: 1974, *Phys. Rev.* **D10**, 1701.
CLARKE, C. J. S.: 1973, *Commun. Math. Phys.* **32**, 205.
CLARKE, C. J. S.: 1975, *Gen. Rel. & Grav.* **6**, 35.
CLARKE, C. J. S.: 1976, *Commun. Math. Phys.* **49**, 17.
CLINE, D. B. and W. HONG: 1992, *Astrophys. J.* **401**, L57.
CLINE, D. B. and W. HONG: 1994, *BAAS*, **185**, 116.
COHEN, J. M. and L. S. KEGELES: 1974, *Phys. Rev.* **D10**, 1070.
COHEN, J. M. and R. M. WALD: 1971, *J. Math. Phys.* **12**, 1845.
COLEMAN, S. and S. HUGHES: 1993, *Phys. Lett.* **B309**, 246.
COLEMAN, S., J. PRESKILL, and F. WILCZEK: 1991, *Phys. Rev. Lett.* **67**, 1975.
COLEMAN, S., J. PRESKILL, and F. WILCZEK: 1992, *Nucl. Phys.* **B378**, 175.
COLPI, M., L. MARACHI, and P. TREVES: 1984, *Astrophys. J.* **280**, 319
COPSON, E. T.: 1928, *Proc. R. Soc. London* **A118**, 184.

CORLEY, S. and T. JACOBSON: 1996, *Phys. Rev.* **D54**, 1568.

COUCH, W. E. and E. T. NEWMAN: 1973, *J. Math. Phys.* **14**, 285.

COURANT, R. and D. HILBERT: 1962, *Methods of Mathematical Physics,* **2**: *Partial Differential Equations*, Interscience, New York, London.

COWIE, L. L., J. P. OSTRIKER, and A. A. STARK: 1978, *Astrophys. J.* **226**, 1041.

COWLEY, A. P.: 1992, *Ann. Rev. Astron. Astrophys.* **30**, 287.

COWLEY, A. P., D. CRAMPTON, J. B. HUTCHINGS, R. REMILLARD, AND J. E. PENFOLD: 1983, *Astrophys. J.* **272**, 118.

COWLEY, A. P., P. C. SCHMIDTKE, D. CRAMPTON, and J. B. HUTCHINGS: 1990, *Astrophys. J.* **350**, 288.

COYNE, D. G.: 1993, in S. Kalara and D. V. Nanopoulos (eds.), *International symposium on black holes, membranes, wormholes and superstrings*, (Sigapore, World Scientific Publishing Co.), p. 159.

CUNNINGHAM, C. T., R. H. PRICE, and V. MONCRIEF: 1978, *Astrophys. J.* **224**, 643.

CUNNINGHAM, C. T., R. H. PRICE, and V. MONCRIEF: 1979, *Astrophys. J.* **230**, 870.

CUTLER, C., T. A. APOSTOLATOS, L. BILDSTEN, L. S. FINN, E. E. FLANNAGAN, D. KENNEFICK, D. M. MARKOVIC, A. ORI, E. POISSON, G. J. SUSSMAN, AND K. S. THORNE: 1993, *Phys. Rev. Lett* **70**, 1984.

CUTLER, C., L. S. FINN, E. POISSON, and G. J. SUSSMAN: 1993, *Phys. Rev.* **D47**, 1511.

CUTLER, C., D. KENNEFICK, and E. POISSON: 1994, *Phys. Rev.* **D50**, 3816.

DADHICH, N.: 1983, *Phys. Lett.* **A98**, 103.

DAMOUR, T.: 1978, *Phys. Rev.* **D18**, 3598.

DAMOUR, T.: 1979, Thèse de Doctorat d'état, Univ. Paris VI.

DAMOUR, T.: 1982, in R. Ruffini (ed.), *Proceedings of the Second Marcel Grossmann Meeting on General Relativity*, North-Holland, Amsterdam.

DAMOUR, T.: 1983, in N. Deruelle and T. Piran (eds.), *Gravitational Radiation*, North-Holland, Amsterdam, p. 59.

DAMOUR, T., N. DERUELLE, and R. RUFFINI: 1976, *Lett. Nuovo Cim.* **15**, 257.

DAMOUR, T. and R. RUFFINI: 1975, *Phys. Rev. Lett.* **35**, 463.

DARWIN, C.: 1959, *Proc. R. Soc. London A* **249**, 180.

DAVIES, P. C. W.: 1976, *Proc. R. Soc. London* **A351**, 129.

DAVIES, P. C. W.: 1977, *Proc. R. Soc. London* **A353**, 499.

DAVIES, P. C. W., S. A. FULLING, and W. G. UNRUH: 1976, *Phys. Rev.* **D13**, 2720.

DAVIS, M., R. RUFFINI, W. H. PRESS, and R. H. PRICE: 1971, *Phys. Rev. Lett.* **27**, 1466.

DAVIS, M., R. RUFFINI, and J. TIOMNO: 1972, *Phys. Rev.* **D5**, 2932.

DE FELICE, F.: 1981, *Nuovo Cim.* **65**, 224.

DE FELICE, F.: 1991, *Mon. Not. RAS* **252**, 197.

DE FELICE, F.: 1994, *Class. Quantum Grav.* **11**, 1283.
DE LA CRUZ, V., J. E. CHASE, and W. ISRAEL: 1970, *Phys. Rev. Lett.* **24**, 423.
DE LA CRUZ, V. and W. ISRAEL: 1967, *Nuovo Cim.* **A51**, 744.
DE OLIVEIRA, H. P.: 1996, E-print gr-qc/9605008.
DE VILLIERS, J.-P. and V. FROLOV: 1997, E-print gr-qc/9711045.
D'EATH, P. D.: 1975a, *Phys. Rev.* **D11**, 1387.
D'EATH, P. D.: 1975b, *Phys. Rev.* **D12**, 2183.
D'EATH, P. D.: 1978, *Phys. Rev.* **D18**, 990.
D'EATH, P. D.: 1979, in J. Ehlers (ed.), *Isolated Gravitating Systems in General Relativity*, Enrico Fermi Summer School Course LXVII, North-Holland, Amsterdam.
DEMERS, J. G., R. LAFRANCE, and R.C. MYERS: 1995, *Phys.Rev.* **D52**, 2245.
DEMIANSKI, M. and L. P. GRISHCHUK: 1974, *Gen. Rel. & Grav.* **5**, 673.
DEMIANSKI, M. and I. D. NOVIKOV: 1982, *Gen. Rel. & Grav.* **14**, 439.
DESER, S. and R. JACKIW: 1988, *Commun. Math. Phys.* **118**, 495.
DESER, S., R. JACKIW, and G. 'T HOOFT: 1984, *Ann. Phys.* **152**, 220.
DETWEILER, S. L.: 1976, *Proc. R. Soc. London* **A349**, 217.
DETWEILER, S. L.: 1977, *Proc. R. Soc. London* **A352**, 381.
DETWEILER, S. L.: 1978, *Astrophys. J.* **225**, 687.
DETWEILER, S. L.: 1979, in L. Smarr (ed.), *Sources of Gravitational Radiation*, Cambridge Univ. Press, p.211.
DETWEILER, S. L.: 1980a, *Phys. Rev.* **D22**, 2323.
DETWEILER, S. L.: 1980b, *Ap. J.* **239**, 292.
DETWEILER, S. L. and J. R. IPSER: 1973, *Astrophys. J.* **185**, 675.
DETWEILER, S. L. and E. SZEDENITZ: 1979, *Astrophys. J.* **231**, 211.
DEUTSCH, D.: 1992, Phys. Rev. D **44**, 3197.
DEWITT, B. S.: 1965, *Dynamical Theory of Groups and Fields*, Gordon and Breach, New York.
DEWITT, B. S.: 1975, *Phys. Rep.* **C19**, 297.
DEWITT, B. S. and R. W. BREHME: 1960, *Ann. Phys.* **9**, 220.
DEWITT-MORETTE, C. and B. L. NELSON: 1984, *Phys. Rev.* **D29**, 663.
DHURAUDHAR, S. V. and N. DADHICH: 1984a, *Phys. Rev.* **D29**, 2712.
DHURAUDHAR, S. V. and N. DADHICH: 1984b, *Phys. Rev.* **D30**, 1625.
DIENER, P., V. P. FROLOV, A. M. KHLOPOV, I. D. NOVIKOV, and C. J. PETHICK: 1997, *Astrophys. J.* **479**, 164.
DIRAC, P. A. M.: 1931, *Proc. R. Soc. London* **A133**, 60.
DOBIASCH, P. and D. MAISON: 1982, *Gen. Rel. & Grav.* **14**, 231.
DOKUCHAYEV, V. I.: 1987, *Zh. Eksp. Teor. Fiz.* **92**, 1921.
DOLGOV, A. D.: 1980a, *Zh. Eksp. Teor. Fiz.* **79**, 337.
DOLGOV, A. D.: 1980b, *Yad. Fiz.* **32**, 1606.
DOLGOV, A. D.: 1981, *Phys. Rev.* **D24**, 1042.
DOLGOV, A. D. and I. B. KHRIPLOVICH: 1997, E-print hep-th/9703042.

DOLGOV, A. D., I. B. KHRIPLOVICH, A. I. VAINSTEIN, and V. I. ZAKHAROV: 1989, *Nucl. Phys.* **B315**, 138.

DOLGOV, A. D., I. B. KHRIPLOVICH, and V. I. ZAKHAROV: 1987, *JETP Lett.* **45**, 651.

DOLGOV, A. D., I. B. KHRIPLOVICH, and V. I. ZAKHAROV: 1988a, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Fourth Seminar on Quantum Gravity*, Moscow, World Scientific, Singapore, p. 673.

DOLGOV, A. D., I. B. KHRIPLOVICH, and V. I. ZAKHAROV: 1988b, *ibid.* p. 679.

DOLGOV, A. D. and YA. B. ZEL'DOVICH: 1980, *Usp. Fiz. Nauk* **130**, 559.

DOLGOV, A. D. and YA. B. ZEL'DOVICH: 1981, *Rev. Mod. Phys.* **53**, 1.

DONETZ, E. E. and D. V. GAL'TSOV: 1993, *Phys. Lett.* **B302**, 411.

DOROSHKEVICH, A. G. and I. D. NOVIKOV: 1978, *Zh. Eksp. Teor. Fiz.* **74**, 3.

DOROSHKEVICH, A. G., I. D. NOVIKOV, and A. G. POLNAREV: 1972, *Zh. Eksp. Teor. Fiz.* **63**, 1538; 1973, *Sov. Phys. JETP* **36**, 816.

DOROSHKEVICH, A. G., YA. B. ZEL'DOVICH, and I. D. NOVIKOV: 1965, *Zh. Eksp. Teor. Fiz.* **49**, 170.

DOWKER, H. F., J. P. GAUNTLETT, S. B. GIDDINGS, and G. T. HOROWITZ: 1994, *Phys. Rev.* **D50**, 2662.

DOWKER, J. S.: 1972, *J. Phys.* **A5**, 936.

DOWKER, J. S.: 1978, *Phys. Rev.* **D18**, 1856.

DOWKER, J. S.: 1986, *Phys. Rev.* **D33**, 3150.

DOWKER, J. S.: 1987, *Phys. Rev.* **D36**, 3742.

DOWKER, J. S. and R. BANACH: 1978, *J. Phys.* **A11**, 2255.

DRAY, T. and G. 'T HOOFT: 1985, *Commun. Math. Phys.* **99**, 613.

DRESLER, A.: 1989, in D. E. Osterbrock and J. S. Miller (eds.), *IAU Symposium 134, Active Galactic Nuclei*, Kluwer, Dordrecht, p.217.

DROZ, S., M. HEUSLER, and N. STRAUMANN: 1991, *Phys. Lett.* **B268**, 371.

DROZ, S., W. ISRAEL, and S. MORSINK: 1996, *Physics World* **9**, (January), 34.

DUFF, M. J.: 1977, *Nucl. Phys.* **B125**, 334.

DWIVEDI, I. H. and P. S. JOSHI: 1989, *Class. Quantum Grav.* **6**, 1599.

DWIVEDI, I. H. and P. S. JOSHI: 1991, *Class. Quantum Grav.* **8**, 1339.

DWIVEDI, I. H. and P. S. JOSHI: 1992, *Class. Quantum Grav.* **9**, L69.

DYMNIKOVA, I. G.: 1982, Preprint FTI-95, Leningrad.

DYMNIKOVA, I. G.: 1986, *Usp. Fiz. Nauk* **148**, 393.

DYSON, F.: 1976, *Institute for Advanced Study Preprint* (unpublished).

EARDLEY, D. M.: 1974, *Phys. Rev. Lett.* **33**, 442.

EARDLEY, D. M., E. W. HIRSCHMANN, and J. H. HORNE: 1995, *Phys. Rev.* **D52**, R5397.

EARDLEY, D. M., G. T. HOROWITZ, D. A. KASTOR, and J. TRASCHEN: 1995, *Phys. Rev. Lett.* **75**, 3390.

EARDLEY, D. M. and W. H. PRESS: 1975, *Ann. Rev. Astron. Astrophys.* **13**, 381.

EARDLEY, D. and L. SMARR: 1979, *Phys. Rev.* **D19**, 2239.

ECHEVERRIA, F., G. KLINKHAMMER, and K. S. THORNE: 1991, *Phys. Rev.* **D44**, 1077.

EDDINGTON, A. S.: 1924, *Nature* **113**, 192.

EDDINGTON, A. S.: 1926, *The Internal Constitution of the Stars*, Cambridge: Cambridge University Press.

EDELSTEIN, L. A. and C. V. VISHVESHWARA: 1970, *Phys. Rev.* **D1**, 3514.

EINSTEIN, A., L. INFELD, and B. HOFFMAN: 1938, *Ann. Math.* **39**, 65.

EINSTEIN, A. and N. ROSEN: 1935, *Phys. Rev.* **48**, 73.

EISENSTAEDT, J.: 1982, *Archive for History of Exact Science* **27**, 157.

ELLIS, G. F. R. and A. R. KING: 1974, *Commun. Math. Phys.* **38**, 119.

ELLIS, J., J. S. HAGELIN, D. V. NANOPOULOS, and M. SREDNICKI: 1984, *Nucl. Phys.* **B241**, 381.

ELLIS, J., E. MAVROMATOS, and D. V. NANOPOULOS: 1992, *Phys. Lett.* **B293**, 142.

ELSTER, T.: 1982a, *Phys. Lett.* **A89**, 125.

ELSTER, T.: 1982b, *Phys. Lett.* **A93**, 58.

ELSTER, T.: 1983a, *J. Phys.* **A16**, 989.

ELSTER, T.: 1983b, *Phys. Lett.* **A94**, 205.

ELSTER, T.: 1984, *Class. Quantum Grav.* **1**, 43.

EMPARAN, R.: 1995, *Phys. Rev. Lett.* **75**, 3386.

ERNST, F. J.: 1968a, *Phys. Rev.* **167**, 1175.

ERNST, F. J.: 1968b, *Phys. Rev.* **168**, 1415.

ERNST, F. J.: 1976a, *J. Math. Phys.* **17**, 54.

ERNST, F. J.: 1976b, *J. Math. Phys.* **17**, 515.

ERNST, F. J. and W. J. WILD: 1976, *J. Math. Phys.* **17**, 182.

EVANS, C. R. and J. S. COLEMAN: 1994, *Phys. Rev. Lett.* **72**, 1782.

FABBRICHESI, M., R. PETTORINO, G. VENEZIANO, and G. A. VILKOVISKII: 1994, *Nucl. Phys.* **B419**, 147.

FABIAN, A. C.: 1994, *Astrophys. J. Suppl.*, **92**, 555.

FACKERELL, E. D. and R. G. CROSSMANN: 1977, *J. Math. Phys.* **18**, 1849.

FARHI, E. and A. H. GUTH: 1987, *Phys. Lett.* **B183**, 149.

FARHI, E., A. H. GUTH, and J. GUVEN: 1990, *Nucl. Phys.* **B339**, 417.

FAWCETT, M. S.: 1983, *Commun. Math. Phys.* **89**, 103.

FAWCETT, M. S. and B. WHITING: 1982, in M. J. Duff and C. J. Isham (eds.), *Quantum Structure of Space and Time*, Cambridge Univ. Press.

FERMI, E.: 1921, *Nuovo Cim.* **22**, 176.

FERRARI, V. and B. MASHHOON: 1984a, *Phys. Rev. Lett.* **52**, 1361.

FERRARI, V. and B. MASHHOON: 1984b, *Phys. Rev.* **D30**, 295.

FERRARI, V. and R. RUFFINI: 1981, *Phys. Lett.* **B98**, 381.

FEYNMAN, R. P. and A. R. HIBBS: 1965, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New-York.

FINKELSTEIN, D.: 1958, *Phys. Rev.* **110**, 965.

FLAMM, L.: 1916, *Phys. Z.* **17**, 448.

FLAMMER, C.: 1957, *Spheroidal wave functions*, Stanford University Press.

FLANAGAN, E.: 1991, *Phys. Rev.* **D46**, 1429.

FLANAGAN, E. and A. ORI: 1996, *Phys. Rev* **D53** R1754.

FLANAGAN, E. and R. M. WALD: 1996, *Phys. Rev.* **D54**, 6233.

FOLACCI, A.: 1992, *Phys. Rev.* **D46**, 2726.

FORD, H. C., R. J. HARMS, Z. I. TSVETANOV, G. F. HARTIG, L. L. DRESSEL, G. A. KRISS, R. C. BOHLIN, A. F. DAVIDSEN, B. MARGON, and A. K. KOCHHAR: 1994, *Astrophys. J.* **435**, L27.

FORD, K. W. and J. A. WHEELER: 1959a, *Ann. Phys.* **7**, 259.

FORD, K. W. and J. A. WHEELER: 1959b, *Ann. Phys.* **7**, 287.

FORD, L. H. and T. A. ROMAN: 1990, *Phys. Rev.* **D41**, 3662.

FORD, L. H. and T. A. ROMAN: 1992, *Phys. Rev.* **D46**, 1328.

FORD, L. H. and T. A. ROMAN: 1993, *Phys. Rev.* **D48**, 776.

FORD, L. H. and T. A. ROMAN: 1995, *Phys. Rev.* **D51**, 4277.

FORD, L. H. and T. A. ROMAN: 1996, *Phys. Rev.* **D53**, 5496.

FRÖMAN, N., P. O. FRÖMAN, N. ANDERSSON, and A. HÖKBACK: 1992, *Phys. Rev.* **D45**, 2609.

FREDENHAGEN, K. and R. HAAG: 1990, *Commun. Math. Phys.* **127**, 273.

FRIEDLANDER, F. G.: 1975, *The Wave Equation in Curved Space-Time*, Cambridge University Press, Cambridge, England.

FRIEDMAN, J. L. and J. R. IPSER: 1987, *Astrophys. J.* **314**, 594.

FRIEDMAN, J., M. S. MORRIS, I. D. NOVIKOV, F. ECHEVERRIA, G. KLINKHAMMER, K. S. THORNE, and U. YURTSEVER: 1990, *Phys. Rev.* **D42** , 1915.

FRIEDMAN, J. L., N. PAPASTAMATIOU, and J. Z. SIMON: 1992a, *Phys. Rev.* **D46**, 4442.

FRIEDMAN, J. L., N. PAPASTAMATIOU, and J. Z. SIMON: 1992b, *Phys. Rev.* **D46**, 4456.

FRIEDMAN, J. L., K. SCHLEICH, and D. M. WITT: 1993, *Phys. Rev. Lett.* **71**, 1486.

FRIEDMAN, J. L. and B. F. SCHUTZ: 1974, *Phys. Rev. Lett.* **32**, 243.

FRIEDRICH, H. and J. M. STEWART: 1983, *Proc. R. Soc. London* **A385**, 345.

FROLOV, A.: 1997, *Phys. Rev.* **D56**, 6433.

FROLOV, V. P.: 1973, PhD Thesis, P. N. Lebedev Institute of Physics.

FROLOV, V. P.: 1974, *Zh. Eksp. Teor. Fiz.* **66**, 813.

FROLOV, V. P.: 1976a, *P. N. Lebedev Inst. of Physics Trudy*, **96**, 72.

FROLOV, V. P.: 1976b, *Usp. Fiz. Nauk* **118**, 473.

FROLOV, V. P.: 1978a, *Gen. Rel. & Grav.* **9**, 569.

FROLOV, V. P.: 1978b, in *Annual Einstein Memorial Volume 1975-1976*, Nauka, Moscow, p. 82.

FROLOV, V.P.: 1978c, *Fortschr. Phys.* **26**, 563.

FROLOV, V. P.: 1979, *Gen. Rel. & Grav.* **10**, 833.

FROLOV, V. P.: 1981, *Phys. Rev. Lett.* **46**, 1349.

FROLOV, V. P.: 1982, *Phys. Rev.* **D26**, 954.

FROLOV, V. P.: 1983a, in M. A. Markov, V. I. Man'ko, and A. E. Shabad (eds.), *Group Theoretical Methods in Physics*, Proceedings of the second Zvenigorod Seminar, USSR, 24-26 Nov. 1982, **2**, Gordon and Breach.

FROLOV, V. P.: 1983b, *Introduction to Black Hole Physics*, Znaniye, Moscow.

FROLOV, V. P.: 1983c, in M. A. Markov and P. C. West (eds.), *Quantum Gravity*, Plenum Press, N. Y.-London, p. 303.

FROLOV, V. P.: 1986, *P. N. Lebedev Institute of Physics. Trudy* **169**, 3; *Proceedings of the Lebedev Physics Institute of the Academy of Sciences of the USSR*, **169**: *The Physical Effects in the Gravitational Field of Black Holes* (ed. M. A. Markov) 1987, Nova Science Publishers, Commack.

FROLOV, V. P.: 1991, *Phys. Rev.* **D43**, 3878.

FROLOV, V. P.: 1992, *Phys. Rev.* **D46**, 5383.

FROLOV, V.: 1995, *Phys. Rev. Lett.* **74**, 3319.

FROLOV, V.: 1996, in N. Sánchez and A. Zichichi (eds.), *String Gravity and Physics at the Planck Energy Scale*, Kluwer Acad. Publ., p.187.

FROLOV, V. and D. FURSAEV: 1997a, *Phys. Rev.* **D56**, 2212.

FROLOV, V. and D. FURSAEV: 1997b, E-print hep-th/9705207.

FROLOV, V., D. FURSAEV, and A. ZELNIKOV: 1996a, *Phys. Lett.* **B382**, 220.

FROLOV, V., D. FURSAEV, and A. ZELNIKOV: 1996b, *Phys. Rev.* **D54**, 2711.

FROLOV, V. P., D. V. FURSAEV, and A. I. ZELNIKOV: 1996c, *Nucl. Phys.* **B486**, 339.

FROLOV, V. P. and A. D. GARCIA: 1983, *Phys. Lett.* **A99**, 421.

FROLOV, V., S. HENDY, and A. L. LARSEN: 1996, *Phys. Rev.* **D54**, 5093.

FROLOV, V. P., A. M. KHLOPOV, I. D. NOVIKOV, and C. J. PETHICK: 1994, *Astrophys. J.* **432**, 680.

FROLOV, V. P. and A. L. LARSEN: 1995, *Nucl. Phys.* **B449**, 149.

FROLOV, V. P., M. A. MARKOV, and V. F. MUKHANOV: 1989, *Phys. Lett.* **B216**, 272.

FROLOV, V. P., M. A. MARKOV, and V. F. MUKHANOV: 1990, *Phys. Rev.* **D41**, 383.

FROLOV, V. P. and I. D. NOVIKOV: 1990, *Phys. Rev.* **D42**, 1057.

FROLOV, V. and I. NOVIKOV: 1993a, *Phys. Rev.* **D48**, 1607.

FROLOV, V. and I. NOVIKOV: 1993b, *Phys. Rev.* **D48**, 4545.

FROLOV, V. P. and D. N. PAGE: 1993, *Phys. Rev. Lett.* **71**, 3902.

FROLOV, V. P. and N. SÁNCHEZ: 1986, *Phys. Rev.* **D33**, 1604.

FROLOV, V. P. and E. M. SEREBRIANY: 1987, *Phys. Rev.* **D35**, 3779.

FROLOV, V. P., V. D. SKARZHINSKY, A. I. ZELNIKOV, and O. HEINRICH: 1989, *Phys. Lett.* **B224**, 255.

FROLOV, V. P. and K. S. THORNE: 1989, *Phys. Rev.* **D39**, 2125.

FROLOV, V. P. and G. A. VILKOVISKY: 1979, Preprint IC/79/69, International Center for Theoretical Physics, Trieste, Italy.

FROLOV, V. P. and G. A. VILKOVISKY: 1981, *Phys. Lett.* **B106**, 307.

FROLOV, V. P. and G. A. VILKOVISKY: 1982, in R. Ruffini (ed.), *Proceedings of the Second Marsel Grossmann Meeting on General Relativity*, North-Holland, Amsterdam, p. 455.

FROLOV, V. P. and G. A. VILKOVISKY: 1983, in M. A. Markov and P. C. West (eds.), *Quantum Gravity*, Plenum Press, N. Y.-London p. 267.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1980, in *Abstracts of Contributed Papers of the 9th Intern. Conf. on General Relativity and Gravitation*, Jena, DDR, **3**, 555.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1982, *Phys. Lett.* **B115**, 372.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1983, *Phys. Lett.* **B123**, 197.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1984, *Phys. Rev.* **D29**, 1057.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1985a, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Third Seminar on Quantum Gravity*, Moscow, World Scientific, Singapore, p. 313.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1985b, *Phys. Rev.* **D32**, 3150.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1987, *Phys. Rev.* **D35**, 3031.

FROLOV, V. P. and A. I. ZEL'NIKOV: 1988, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Fourth Seminar on Quantum Gravity*, Moscow, World Scientific, Singapore, p. 568.

FROLOV, V. P., A. I. ZELNIKOV, and U. BLEYER: 1987, *Ann. Physik* **44**, 371.

FRONSDAL, C.: 1959, *Phys. Rev.* **116**, 778.

FRY, J. N., K. A. OLIVE, and M. S. TURNER: 1980a, *Phys. Rev.* **D22**, 2953.

FRY, J. N., K. A. OLIVE, and M. S. TURNER: 1980b, *Phys. Rev.* **D22**, 2977.

FRY, J. N., K. A. OLIVE, and M. S. TURNER: 1980c, *Phys. Rev. Lett.* **45**, 2074.

FUKUGITA, M. and E. L. TURNER: 1996, *Astrophys. J.* **460**, L81.

FULLING, S. A.: 1977a, *J. Phys.* **A10**, 917.

FULLING, S. A.: 1977b, *Phys. Rev.* **D15**, 2411.

FULLING, S. A.: 1989, *Aspects of quantum field theory in curved spacetime*, Cambridge Univ. Press.

FULLING, S. A. and S. N. M. RUIJSENAARS: 1987, *Phys. Rep.* **152**, 135.

FURSAEV, D. V.: 1995, *Phys. Rev.* **D51**, 5352.

FURSAEV, D. V. and S. N. SOLODUKHIN: 1995, *Phys. Rev.* **D52**, 2133.

FUTTERMAN, J. A. H., F. A. HANDLER, and R. A. MATZNER: 1988, *Scattering from Black Holes*, Cambridge Univ. Press, Cambridge.

GAISER, B. D. and R. V. WAGONER: 1980, *Astrophys. J.* **240**, 648.

GALLOWAY, G. J.: 1983, *J. Phys.* **A16**, 1435.

GAL'TSOV, D. V.: 1980, in *Group-Theoretical Methods in Physics*, Nauka, Moscow, **2**, p. 183.

GAL'TSOV, D. V.: 1982, *J. Phys.* **A15**, 3737.

GAL'TSOV, D. V.: 1986, *Particles and Fields in the Vicinity of Black Holes*, MGU, Moscow (in Russian).
GAL'TSOV, D. V., E. E. DONETS, and M. YU. ZOTOV: 1997, *JETP Lett.* **65**, 895.
GAL'TSOV, D. V., A. A. ERSHOV, and M. YU. MOROZOV: 1988, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Fourth Seminar on Quantum Gravity*, Moscow, World Scientific, Singapore, p. 629.
GALTSOV, D. V. and A. A. MATIUKHIN: 1992, *Class. Quantum Grav.* **9**, 2039.
GAL'TSOV, D. V. and V. I. PETUKHOV: 1978, *Zh. Eksp. Teor. Fiz.* **74**, 801.
GAL'TSOV, D. V., V. I. PETUKHOV, and A. N. ALIEV: 1984a, *Phys. Lett.* **A105**, 346.
GAL'TSOV, D. V., G. V. POMERANTSEVA, and G. A. CHIZHOV: 1983, *Izv. Vuzov, Fizika* No. 8, 75.
GAL'TSOV, D. V., G. V. POMERANTSEVA, and G. A. CHIZHOV: 1984b, *Izv. Vuzov, Fizika* No. 8, 81.
GANNON, D.: 1975, *J. Math. Phys.* **16**, 2364.
GANNON, D.: 1976, *Gen. Rel. & Grav.* **7**, 219.
GARFINKLE, D.: 1995, *Phys. Rev.* **D51**, 5558.
GARFINKLE, D., S. B. GIDDINGS, and A. STROMINGER: 1994, *Phys. Rev.* **49**, 958.
GARFINKLE, D., G. HOROWITZ, and A. STROMINGER: 1991, *Phys. Rev.* **D43**, 3140; 1992 **D45**, 3888(E).
GARFINKLE, D. and A. STROMINGER: 1991, *Phys. Lett.* **B256**, 146.
GEGENBERG, J., G. KUNSTATTER, and D. LOUIS-MARTINEZ: 1995, *Phys. Rev.* **D51**, 1781.
GERLACH, U. H.: 1974, *Phys. Rev. Lett.* **32**, 1023.
GERLACH, U. H.: 1975, *Phys. Rev.* **D11**, 2762.
GERLACH, U. H.: 1976, *Phys. Rev.* **D14**, 1479.
GEROCH, R.: 1967, *J. Math. Phys.* **8**, 782.
GEROCH, R. P.: 1968, *J. Math. Phys.* **9**, 450.
GEROCH, R.: 1969, *Commun. Math. Phys.* **13**, 180.
GEROCH, R.: 1971, *J. Math. Phys.* **12**, 918.
GEROCH, R. P. and J. B. HARTLE: 1982, *J. Math. Phys.* **23**, 680.
GEROCH, R. P., E. H. KRONHEIMER, and R. PENROSE: 1972, *Proc. R. Soc. London* **A327**, 545.
GEROCH, R. P., C. B. LIANG, and R. M. WALD: 1982, *J. Math. Phys.* **23**, 432.
GEROCH, R. and J. TRASCHEN: 1987, *Phys. Rev.* **D36**, 1017.
GIACCONI, R., H. GURSKY, F. R. PAOLINI, and B. B. ROSSI: 1962, *Phys. Rev. Lett.* **9**, 439.
GIACCONI, R., S. MURRAY, H. GURSKY, E. KELLOGG, E. SCHREIER, AND H. TANNANBAUM: 1972, *Astrophys. J.* **178**, 281.
GIBBONS, G. W.: 1972, *Commun. Math. Phys.* **27**, 87.
GIBBONS, G. W.: 1975, *Commun. Math. Phys.* **44**, 245.
GIBBONS, G. W.: 1976, *Mon. Not. RAS* **177**, 37P.

GIBBONS, G. W.: 1977, *Phys. Rev.* **D15**, 3530.
GIBBONS, G. W.: 1982, *Nucl. Phys.* **B207**, 337.
GIBBONS, G. W.: 1984, *Aspects of supergravity*, Preprint, D. A. M. T. P. Cambridge.
GIBBONS, G. W.: 1986, in A. Jadczyk (ed.), *Fields and Geometry 1986*, Proceedings of the 22nd Karpacz Winter School of Theoretical Physics, World Scientific, Singapore.
GIBBONS, G. W. and S. W. HAWKING: 1977, *Phys. Rev.* **D15**, 2752.
GIBBONS, G. W. and C. M. HULL: 1982, *Phys. Lett.* **B109**, 190.
GIBBONS, G. W. and R. E. KALLOSH: 1995, *Phys. Rev.* **D51**, 2839.
GIBBONS, G. W. and K. MAEDA: 1988, *Nucl. Phys.* **B298**, 741.
GIBBONS, G. W. and M. J. PERRY: 1976, *Phys. Rev. Lett.* **36**, 985.
GIBBONS, G. W. and M. J. PERRY: 1978, *Proc. R. Soc. London* **A358**, 467.
GIBBONS, G. W. and D. L. WILTSHIRE: 1986, *Ann. Phys.* **167**, 201.
GIDDINGS, S. B. and L. THORLACIUS: 1995, in E. W. Kolb and R. D. Peccei (eds.), *Particle and Nuclear Astrophysics and Cosmology in the Next Millenium*, Proceedings of the 1994 Snowmass Summer Study, World Sci., p.430.
GIES, D. R. and C. T. BOLTON: 1986, *Astrophys. J.* **304**, 371.
GINZBURG, V. L.: 1964, *Dokl. Akad. Nauk USSR* **156**, 43.
GINZBURG, V. L.: 1975, *Pis'ma v Zh. Eksp. Teor. Fiz.* **22**, 514.
GINZBURG, V. L. and I. M. FRANK: 1947, *Soviet Phys. Doklady* **56**, 583.
GINZBURG, V. L. and V. P. FROLOV: 1976, *Pis'ma v Astr. Zh.* **2**, 474.
GINZBURG, V. L. and V. P. FROLOV: 1987, *Usp. Fiz. Nauk* **153**, 633.
GINZBURG, V. L. and L. M. OZERNOI: 1964, *Zh. Eksp. Teor. Fiz.* **47**, 1030.
GLASS, E. N. and A. HARPAZ: 1981, *Phys. Rev.* **D24**, 3038.
GLEISER, R. J., C. O. NICASIO, R. H. PRICE, and J. PULLIN: 1995, *Second order perturbations of a Schwarzschild black hole* (preprint).
GLENDENNING, N. K.: 1997, *Compact Stars*, Astronomy and Astrophysics Library, Springer.
GNEDIN, M. L. and N. Y. GNEDIN: 1993, *Class. Quantum Grav.*, **10**, 1083.
GÖDEL, K.: 1949, *Rev. Mod. Phys.* **21**, 447.
GOLDBERG, J. N., A. J. MACFARLANE, E. T. NEWMAN, F. RORLICH, and E. C. G. SUDARSHAN: 1967, *J. Math. Phys.* **8**, 2155.
GOLDBERG, J. N. and R. SACHS: 1962, *Acta Phys. Polonica* **22**, Suppl., 13.
GOLDREICH, P. and W. H. JULIAN: 1969, *Astrophys. J.* **157**, 869.
GOLDWIRTH, D. S., M. S. PERRY, T. PIRAN, and K. S. THORNE: 1994, *Phys. Rev.* **D49**, 3951.
GRAVES, J. C. and D. R. BRILL: 1960, *Phys. Rev.* **120**, 1507.
GREENE, B. R., S. D. MATHUR, and C. M. O'NEILL: 1993, *Phys. Rev.* **D47**, 2242.
GRIB, A. A., S. G. MAMAEV, and V. M. MOSTEPANENKO: 1980, *Quantum Effects in High-Intensity External Fields*, Atomizdat, Moscow.
GRILLO, G.: 1991, *Class. Quantum Grav.* **8**, 739.
GRISHCHUK, L. P. and YU. V. SIDOROV: 1990, *Phys. Rev.* **D42**, 3413.

GUINN, J. W., C. M. WILL, Y. KOJIMA, and B. F. SCHUTZ: 1990, *Class. Quantum Grav.* **7**, L47.

GULLSTRAND, A.: 1922, *Arkiv. Mat. Astron. Fys.* **16**, (8), 1.

GUNDLACH, C.: 1995, *Phys. Rev. Lett.* **75**, 3214.

GUNDLACH, C.: 1997, *Phys. Rev.* **D55**, 695.

GUNDLACH, C., R. H. PRICE, and J. PULLIN: 1994a, *Phys. Rev.* **D49**, 883.

GUNDLACH, C., R. H. PRICE, and J. PULLIN: 1994b, *Phys. Rev.* **D49**, 890.

GUNTER, D. L.: 1980, *Philos. Trans. R. Soc.* **A296**, 497.

GUREVICH, A. V. and YA. N. ISTOMIN: 1985, *Zh. Eksp. Teor. Fiz.* **89**, 3.

GÜRSEL, Y., I. D. NOVIKOV, V. SANDBERG, and A. A. STAROBINSKY: 1979b, *Phys. Rev.* **D20**, 1260.

GÜRSEL, Y., V. SANDBERG, I. D. NOVIKOV, and A. A. STAROBINSKY: 1979a, *Phys. Rev.* **D19**, 413.

HAARDT, F. and L. MARASCHI: 1993, *Astrophys. J.*, **413**, 507.

HADAMARD, J.: 1923, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press, New Haven.

HAEHNELT, M. and M. REES: 1993, *Mon. Not. RAS* **263**, 168.

HAGIHARA, Y.: 1931, *Jap. J. Astr. Geoph.* **8**, 67.

HAHN, S. and R. W. LINDQUIST: 1964, *Ann. Phys.* **29**, 304.

HÁJÍČEK, P., B. S. KAY, and K. V. KUCHAŘ: 1992, *Phys. Rev.* **D46**, 5439.

HÁJÍČEK, P. and W. ISRAEL: 1980, *Phys. Lett.* **A80**, 9.

HALZEN, F., E. ZAS, J. H. MACGIBBON, and T. C. WEEKES: 1991, *Nature* **353**, 807.

HAMADÉ, R. S., J. H. HORNE, and J. M. STEWART: 1996, *Class. Quantum Grav.* **13**, 2241.

HAMADÉ, R. S. and J. M. STEWART: 1996, *Class. Quantum Grav.* **13**, 497.

HAMBLI, N. and C. P. BURGESS: 1996, *Phys. Rev.* **D53**, 5717.

HANDLER, F. A. and R. A. MATZNER: 1980, *Phys. Rev.* **D22**, 2331.

HANNI, R. S.: 1977, *Phys. Rev.* **D16**, 933.

HANNI, R. S. and R. RUFFINI: 1973, *Phys. Rev.* **D8**, 3259.

HARMS, R. J., H. C. FORD, Z. I. TSVETANOV, G. F. HARTIG, L. L. DRESSEL, G. A. KRISS, R. BOHLIN, A. F. DAVIDSEN, B. MARGON, and A. K. KOCHHAR: 1994, *Astrophys. J.* **435**, L35.

HARRISON, E.: 1970, *Phys. Rev.* **D1**, 2726.

HARTLE, J. B.: 1971, *Phys. Rev.* **D3**, 2938.

HARTLE, J. B.: 1972, in J. Klauder (ed.), *Magic without Magic, John Archibald Wheeler*, Freeman, San Francisco, p.259.

HARTLE, J. B.: 1973, *Phys. Rev.* **D8**, 1010.

HARTLE, J. B.: 1974, *Phys. Rev.* **D9**, 2749.

HARTLE, J.: 1994, *Phys. Rev.* **D49**, 6543.

HARTLE, J.: 1995, *Phys. Rev.* **D51**, 1800.

HARTLE, J. B. and S. W. HAWKING: 1972, *Commun. Math. Phys.* **26**, 87.

HARTLE, J. B. and S. W. HAWKING: 1976, *Phys. Rev.* **D13**, 2188.
HARTLE, J. B. and S. W. HAWKING: 1983, *Phys. Rev.* **D28**, 2960.
HARTLE, J. B. and B. L. HU: 1978, *Phys. Rev.* **D20**, 1772.
HARTLE, J. B. and B. L. HU: 1980, *Phys. Rev.* **D21**, 2756.
HARTLE, J. B. and D. C. WILKINS: 1974, *Commun. Math. Phys* **38**, 47.
HASSAN, S. and A. SEN: 1992, *Nucl. Phys.* **B375**, 103.
HASSLACHER, B. and E. MOTTOLA: 1981, *Phys. Lett.* **B99**, 221.
HASWELL, C. A., E. L. ROBINSON, K. HORNE, R. STIENING, and T. M. C. ABBOTT: 1993, *Astrophys. J.* **411**, 802.
HASWELL, C. A. and A. W. SHAFTER: 1990, *Astrophys. J. Lett.* **359**, L47.
HAUGAN, M. P., S. L. SHAPIRO, and I. WASSERMAN: 1982, *Astrophys. J.* **257**, 283.
HAWKING, S. W.: 1971a, *Mon. Not. RAS* **152**, 75.
HAWKING, S. W.: 1971b, *Phys. Rev. Lett.* **26**, 1344.
HAWKING, S. W.: 1972a, *Commun. Math. Phys.* **25**, 152.
HAWKING, S. W.: 1972b, *Commun. Math. Phys.* **25**, 167.
HAWKING, S. W.: 1973, *Commun. Math. Phys.* **33**, 323.
HAWKING, S. W.: 1974, *Nature* **248**, 30.
HAWKING, S. W.: 1975, *Commun. Math. Phys.* **43**, 199.
HAWKING, S. W.: 1976a, *Phys. Rev.* **D13**, 191.
HAWKING, S. W.: 1976b, *Phys. Rev.* **D14**, 2460.
HAWKING, S. W.: 1978, *Nucl. Phys.* **B144**, 349.
HAWKING, S. W.: 1979, in S. W. Hawking and W. Israel (eds.), *General Relativity, An Einstein Centenary Survey*, Cambridge Univ. Press.
HAWKING, S. W.: 1981, *Commun. Math. Phys.* **80**, 421.
HAWKING, S.: 1982, *Commun. Math. Phys.* **87**, 395.
HAWKING, S. W.: 1984, in M. A. Markov and P. C. West (eds.), *Quantum Gravity, Proceedings of the Second Seminar on Quantum Gravity*, Moscow, Plenum Press, New York, London, p. 19.
HAWKING, S. W.: 1988a, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Fourth Seminar on Quantum Gravity*, Moscow World Scientific, Singapore, p. 125.
HAWKING, S. W.: 1988b, *Phys. Rev.* **D37**, 904.
HAWKING, S. W.: 1989, *Phys. Lett.* **B231**, 237.
HAWKING, S. W.: 1990a, *Mod. Phys. Lett.* **A5**, 145.
HAWKING, S. W.: 1990b, *Mod. Phys. Lett.* **A5**, 453.
HAWKING, S. W.: 1992, *Phys. Rev.* **D46**, 603.
HAWKING, S. W.: 1996, *Phys. Rev.* **D53**, 3099.
HAWKING, S. W. and G. F. ELLIS: 1973, *The Large-Scale Structure of Spacetime*, Cambridge Univ. Press.
HAWKING, S. W. and J. B. HARTLE: 1972, *Commun. Math. Phys.* **27**, 283.
HAWKING, S. W., G. T. HOROWITZ, and S. F. ROSS: 1995, *Phys. Rev.* **51**, 4302.

HAWKING, S. W. and C. J. HUNTER: 1996, *Class. Quantum Grav.* **13**, 2735.
HAWKING, S. W., I. MOSS, and J. STEWART: 1982, *Phys. Rev.* **D26**, 2681.
HAWKING, S. W., D. N. PAGE, and C. N. POPE: 1979, *Phys. Lett.* **B86**, 175.
HAWKING, S. W., D. N. PAGE, and C. N. POPE: 1980, *Nucl. Phys.* **B170**, 283.
HAWKING, S. W. and R. PENROSE: 1970, *Proc. R. Soc. London* **A314**, 529.
HAWKING, S. W. and S. F. ROSS: 1995, *Phys. Rev. Lett.* **75**, 3382.
HAYAKAWA, S. and M. MATSUOKO: 1964, *Progr. Theor. Phys. Suppl.* **30**, 204.
HAYASHI, K. and T. SAMURA: 1994, *Phys. Rev.* **D50**, 3666.
HAYWARD, G.: 1993, *Phys. Rev.* **D47**, 3275.
HAYWARD, G. and J. LOUKO: 1990, *Phys. Rev.* **D42**, 4032.
HEADING, J.: 1977, *J. Phys.* **A10**, 885.
HECKLER, A. F.: 1997, *Phys. Rev. Lett.* **78**, 3430.
HECKLER, A. F., E. W. KOLB: 1996, *Astrophys. J.* **472** L85.
HEUSLER, M.: 1992, *J. Math. Phys.* **33**, 3497.
HEUSLER, M.: 1995, *Class. Quantum Grav.* **12**, 779.
HEUSLER, M.: 1996a, *Black Hole Uniqueness Theorems*, Cambridge Univ. Press.
HEUSLER, M.: 1996b, *Class. Quantum Grav.* **12**, 2021.
HEUSLER, M., S. DROZ, and N. STRAUMANN: 1991, *Phys. Lett.* **B271**, 61.
HEYVAERTS, J. and C. NORMAN: 1989, *Astrophys. J.* **347**, 1055.
HIROTANI, K., M. TAKAHACHI, S.-YA. NITTA, and A. TOMIMATSU: 1992, *Astrophys. J.* **386**, 455.
HIRSCHMANN, E. W. and D. M. EARDLEY: 1995a, *Phys. Rev.* **D51**, 4198.
HIRSCHMANN, E. W. and D. M. EARDLEY: 1995b, *Phys. Rev.* **D52**, 5850.
HIRSCHMANN, E. W. and D. M. EARDLEY: 1997, *Phys. Rev.* **D56**, 4696.
HISCOCK, W. A.: 1977, *Phys. Rev.* **D16**, 2673.
HISCOCK, W. A.: 1981, *Phys. Rev.* **D23**, 2813.
HISCOCK, W. A.: 1983, *Phys. Rev. Lett.* **50**, 1734.
HISCOCK, W. A.: 1987, *Phys. Rev.* **D35**, 1161.
HOD, S. and T. PIRAN: 1997, *Phys. Rev.* **D55**, 3485.
HORIUCHI, S., L. MESTEL, and I. OKAMOTO: 1995, *Mon. Not. RAS* **275**, 1160.
HOROWITZ, G. T.: 1992, In *Proceedings of 1992 Trieste Spring School on String Theory and Quantum Gravity*, p.59; E-print hep-th/9210119.
HOROWITZ, G. T.: 1996, *Talk given at Pacific Conference on Gravitation and Cosmology*, Seul, Korea, 1-6 Feb.; E-print gr-qc/9604051.
HOROWITZ, G.: 1997, E-print gr-qc/9704072.
HOROWITZ, G. T., D. A. LOWE, and J. M. MALDACENA: 1996, *Phys. Rev. Lett.* **77**, 430.
HOROWITZ, G. T. and A. STROMINGER: 1996, *Phys. Rev. Lett.* **77**, 2368.
HOWARD, K. W.: 1984, *Phys. Rev.* **D30**, 2532.
HOWARD, K. W. and P. CANDELAS: 1984, *Phys. Rev. Lett.* **53**, 403.
HOYLE, F. and W. A. FOWLER: 1963, *Nature* **197**, 533.

HOYLE, F., W. A. FOWLER, G. BURBIDGE, and E. M. BURBIDGE: 1964, *Astrophys. J.* **139**, 909.

HOYLE, F. and R. A. LYTTLETON: 1939, *Proc. Camb. Phil. Soc.* **35**, 405.

HSU, S. D. U.: 1990, *Phys. Lett.* **B251**, 343.

HU, B. L. and L. PARKER: 1978, *Phys. Rev.* **D17**, 3292.

HUET, P. and M. PESKIN: 1995, *Nucl. Phys.* **B434**, 3.

HUGHES, S. A., C. R. KEETON II, P. WALKER, K. T. WALSH, S. L. SHAPIRO, AND S. A. TEUKOLSKY: 1994, *Phys. Rev.* **D49**, 4004.

HUT, P.: 1977, *Mon. Not. RAS* **180**, 379.

IPSER, J. R.: 1971, *Phys. Rev. Lett.* **27**, 529.

ISAACSON, R. A.: 1968a, *Phys. Rev.* **166**, 1263.

ISAACSON, R. A.: 1968b, *Phys. Rev.* **166**, 1272.

ISHAM, C. J.: 1995, *Plenary session lecture given at the GR14 conference, Florence, August 1995*; E-print gr-qc/9510063.

ISRAEL, W.: 1967, *Phys. Rev.* **164**, 1776.

ISRAEL, W.: 1968, *Commun. Math. Phys.* **8**, 245.

ISRAEL, W.: 1971, *Gen. Rel. & Grav.* **2**, 53.

ISRAEL, W.: 1973, *Lett. Nuovo Cim.* **6**, 267.

ISRAEL, W.: 1976, *Phys. Lett.* **A57**, 107.

ISRAEL, W.: 1983, *Sci. Prog. Oxf.* **68**, 333,

ISRAEL, W.: 1984, *Found. Phys.* **14**, 1049.

ISRAEL, W.: 1985, *Can. J. Phys.* **63**, 34.

ISRAEL, W.: 1986a, *Phys. Rev. Let.* **56**, 789.

ISRAEL, W.: 1986a, *Phys. Rev. Lett.* **56**, 789.

ISRAEL, W.: 1986b, *Phys. Rev. Let.* **57**, 397.

ISRAEL, W.: 1986b, *Can. J. Phys.* **64**, 120.

ISRAEL, W.: 1987, in S. W. Hawking and W. Israel, (eds.), *300 Years of Gravity*, Cambridge Univ. Press., p.199.

ISRAEL, W.: 1997, in B. Jones and D. Marković (eds.), *Relativistic Astrophysics*, Cambridge University Press, p.173.

ISRAEL, W. and K. A. KHAN: 1964, *Nuovo Cim.* **33**, 331.

IYER, B. R. and A. KUMAR: 1978, *Phys. Rev.* **D18**, 4799.

IYER, S.: 1987, *Phys. Rev.* **D35**, 3632.

IYER, S. and C. M. WILL: 1987, *Phys. Rev.* **D35**, 3621.

IYER, V. and R. M. WALD: 1994, *Phys. Rev.* **D50**, 846.

IYER, V. and R. M. WALD: 1995, *Phys. Rev.* **D52**, 4430.

JACOBSON, T.: 1991, *Phys. Rev.* **D44**, 1731.

JACOBSON, T.: 1993, *Phys. Rev.* **D48**, 728.

JACOBSON, T.: 1994, E-print gr-qc/9404039.

JACOBSON, T.: 1995, in: *Mexican School on Gravitation 1994* p.87; E-print hep-th/9510026.

JACOBSON, T.: 1996, *Phys. Rev.* **D53**, 7082.

JACOBSON, T., G. KANG, and R. C. MYERS: 1994, *Phys. Rev.* **D49**, 6587.

JACOBSON, T. and S. VENKATARAMANI: 1995, *Class. Quantum Grav.***12**, 1055.

JANO, K. and S. BOCHNER: 1953, *Curvature and Betti Numbers*, Princeton Univ. Press.

JENSEN, B. P., J. G. MCLAUGHLIN, and A. C. OTTEWILL: 1988, *Class. Quantum Grav.* **5**, L187.

JENSEN, B. P., J. G. MCLAUGHLIN, and A. C. OTTEWILL: 1991, *Phys. Rev.* **D43**, 4142.

JENSEN, B. P., J. G. MCLAUGHLIN, and A. C. OTTEWILL: 1992, *Phys. Rev.* **D45**, 3002.

JENSEN, B. P., J. G. MCLAUGHLIN, and A. C. OTTEWILL: 1995, *Phys. Rev.* **D51**, 5676.

JENSEN, B. P. and A. C. OTTEWILL: 1989, *Phys. Rev.* **D39**, 1130.

JOHNSON, C. V., R. R. KHURI, and R. C. MYERS: 1996, *Phys. Lett.* **B378**, 78.

JOHNSTON, H. M. and S. R. KULKARNI: 1990, in *Accretion-Powered Compact Binaries*, Cambridge, Cambridge University Press.

JOHNSTON, H. M., S. R. KULKARNI, and J. B. OKE: 1989, *Astrophys. J.* **345**, 492.

JOHNSTON, M. and R. RUFFINI: 1974, *Phys. Rev.* **D10**, 2324.

JOHNSTON, M., R. RUFFINI, and F. ZERILLI: 1973, *Phys. Rev. Lett.* **31**, 1317.

JORDAN, P., J. EHLERS, and R. SACHS: 1961, *Akad. Wiss. Mainz. Abh. Math. Naturwiss.* Kl, No. 1, p. 2.

JOSHI, P. S.: 1993, *Global Aspects in Gravitation and Cosmology,* Clarendon Press, Oxford.

KABURAKI, O. and I. OKAMOTO: 1991, *Phys. Rev.* **D43**, 340.

KABURAKI, O., I. OKAMOTO, and J. KATZ: 1993, *Phys. Rev.* **D47**, 2234.

KALLOSH, R., A. LINDE, T. ORTIN, A. PEET, and A. VAN PROEYEN: 1992, *Phys. Rev.* **D46**, 5278.

KALLOSH, R., J. RAHMFELD, and W. K. WONG: 1997, E-print hep-th/9706048.

KALUZA, T.: 1921, *Sitz. Preuss. Akad. Wiss.* **K1**, 966.

KARDASHEV, N. S., I. G. MITROFANOV, and I. D. NOVIKOV: 1984, *Astron. Zh.* **61**, 1113.

KARDASHEV, N. S. and I. D. NOVIKOV: 1983, in G. O. Abell and G. Chincarini (eds.), *Early Evolution of the Universe and Its Present Structure (IAU Symp. 104, Crete, 1982)*, Reidel Publ. Company, p.463.

KARDASHEV, N. S., I. D. NOVIKOV, A. G. POLNAREV, and B. E. STERN: 1983, *Astron. Zh.* **60**, 209: in M. Burns *et al.* (eds.), *Positron-Electron Pairs in Astrophysics*, Goddard Space Flight Center.

KATES, R. E.: 1980, *Phys. Rev.* **D22**, 1853.

KATES, R. E.: 1981, *Ann. Phys.* **132**, 1.

KAWALER, S. D.: 1997, in the book: S. D. Kawaler, I. Novikov, G. Srinivasan, *Stellar Remnants*, Saas-Fee Advanced Course 25, Springer, page 1.

KAY, B. S., M. J. RADZIKOWSKI, and R. M. WALD: 1997, *Commun. Math. Phys.* **183**, 533.

KAY, B. S. and R. M. WALD: 1987, *Class. Quantum. Grav.* **4**, 893.

KAY, B. S. and R. M. WALD: 1991, *Phys. Rep.* **207**, 49.

KENNEL, C. F., F. S. FUJIMURA, and I. OKAMOTO: 1983, *Geophys. Astrophys. Fluid Dyn.* **26**, 147.

KERR, R. P.: 1963, *Phys. Rev. Lett.* **11**, 237.

KESTIN, J.: 1968, *A Course of Thermodynamics*, Vol.2, Blaisdell, Waltham, Mass.

KHLOPOV, M. YU. and A. G. POLNAREV: 1980, *Phys. Lett.*, **B97**, 383.

KIEM, Y.: 1994, E-print hep-th/9407100.

KIHARA, M. and A. TOMIMATSU: 1982, *Progr. Theor. Phys.* **67**, 349.

KIM, S. W. and K. S. THORNE: 1991, *Phys. Rev.* **D43**, 3929.

KING, A. R.: 1975, *Phys. Rev.* **D 11**, 763.

KING, A. R. and J. P. LASOTA: 1977, *Astr. Astrophys.* **58**, 175.

KINNERSLEY, W.: 1969, *J. Math. Phys.* **10**, 1195.

KINNERSLEY, W. and M. WALKER: 1970, *Phys. Rev.* **D2**, 1359.

KIPPENHAHN, R. and A. WEIGERT: 1990, *Stellar Structure and Evolution*, Springer-Verlag.

KLAUDER, J. R. and E. C. G. SUDARSHAN: 1968, *Fundamentals of Quantum Optics*, W. A. Benjamin, N. Y.

KLEIHAUS, B., J. KUNZ, and A. SOOD: 1996, *Phys. Rev.* **D54**, 5070.

KLEIHAUS, B., J. KUNZ, and A. SOOD: 1997, E-print hep-th/9705179.

KLEIN, O.: 1926, *Z. Phys.* **37**, 895.

KLEIN, O.: 1961, in *Werner Heisenberg und die Physik unserer Zeit*, Braunschweig, Vieweg, s. 58.

KLINKHAMMER, G.: 1991, *Phys. Rev.* **D43**, 2542.

KLINKHAMMER, G.: 1992, *Phys. Rev.* **D46**, 3388.

KLINKHAMMER, G. and K. S. THORNE: (unpublished).

KODAMA, H.: 1979, *Progr. Theor. Phys.* **62**, 1434.

KODAMA, H.: 1980, *Progr. Theor. Phys.* **63**, 1217.

KODAMA, H., M. SASAKI, and K. SATO: 1982, *Progr. Theor. Phys.* **68**, 1979.

KOIKE, T., T. HARA, and S. ADACHI: 1995, *Phys. Rev. Lett.* **74**, 5170.

KOIKE, T. and T. MISHIMA: 1995, *Phys. Rev.* **D51**, 4045.

KOJIMA, Y. and T. NAKAMURA: 1983a, *Phys. Lett.* **A96**, 335.

KOJIMA, Y. and T. NAKAMURA: 1983b, *Phys. Lett.* **A99**, 37.

KOJIMA, Y. and T. NAKAMURA: 1984a, *Progr. Theor. Phys.* **71**, 79.

KOJIMA, Y. and T. NAKAMURA: 1984b, *Progr. Theor. Phys.* **72**, 494.

KOKKOTAS, K. D.: 1991, *Class. Quantum Grav.* **8**, 2217.

KOKKOTAS, K. D.: 1993, *Nouvo Cim.* **B108**, 991.

KOKKOTAS, K. D. and B. F. SCHUTZ: 1988, *Phys. Rev.* **D37** 3378.

KOLB, E. W. and M. S. TURNER: 1983, *Ann. Rev. Nucl. Part. Sci.* **33**, 645.

KORMENDY, J.: 1993, in J. E. Beckman, H. Netzer, and L. Colina (eds.), *The Nearest Active Galaxies*, Consejo Superior de Investigaciones Cientificas, Madrid, p.197.

KORMENDY, J., R. BENDER, J. MAGORRIAN, S. TREMAINE, K. GEBHARDT, D. RICHSTONE, A. DRESLER, S. M. FABER, C. GRILLMAIR, and T. R. LAUER: 1997, E-print astro-ph/9703188.

KORMENDY, J. and D. RICHSTONE: 1992, *Astrophys. J.* **393**, 559.

KORMENDY, J. and D. RICHSTONE: 1995, *Ann. Rev. Astron. Astroph.* **33**, 581.

KOVETZ, A. and T. PIRAN: 1975a, *Lett. Nuovo Cim.* **12**, 39.

KOVETZ, A. and T. PIRAN: 1975b, *Lett. Nuovo Cim.* **12**, 560.

KRAMER, D., H. STEPHANI, M. MACCALLUM, and E. HERLT: 1980, *Exact Solutions of the Einstein Field Equations*, Deutscher Verlag Wiss., Berlin.

KRASNIKOV, S. V.: 1996, *Phys. Rev.* **D54**, 7322.

KRAUSS, L. M. and H. LIU: 1997, *Nucl. Phys.* **B491**, 365.

KRAUSS, L. M. and F. WILCZEK: 1989, *Phys. Rev. Lett.* **62**, 1221.

KROLIK, J. H. and R. A. LONDON: 1983, *Astrophys. J.* **267**, 371.

KRORI, K. D., S. CHAUDHURI, and S. DOWERAH: 1983, *Can. J. Phys.* **61**, 1192.

KRORI, K. D., S. CHAUDHURI, and S. DOWERAH: 1984, *J. Math. Phys.* **25**, 607.

KRUSKAL, M. D.: 1960, *Phys. Rev.* **119**, 1743.

KUCHAŘ, K.: 1994, *Phys. Rev.* **D50**, 3961.

KUIPER, L., J. VAN PARADIJS, and M. VAN DER KLIS: 1988, *Astron. Astrophys.* **203**, 79.

KULKARNI, A. D.: 1984, *J. Math. Phys.* **25**, 1028.

KULKARNI, A. D., L. SHEPLEY, and J. W. YORK JR.: 1983, *Phys. Lett.* **A96**, 228.

KUNDT, W. and M. TRÜMPER: 1966, *Z. Phys.* **192**, 419.

KÜNZLE, H. P.: 1991, *Class. Quantum Grav.* **8**, 2283.

KÜNZLE, H. P.: 1994, *Commun. Math. Phys.* **162**, 371.

KÜNZLE, H. P. and A. K. MASOOD-UL-ALAM: 1990, *J. Math. Phys.* **31**, 928.

KURODA, Y.: 1984a, *Progr. Theor. Phys.* **71**, 1422.

KURODA, Y.: 1984b, *Progr. Theor. Phys.* **72**, 63.

LAFLAMME, R.: 1989a, *Physica* **A158**, 58.

LAFLAMME, R.: 1989b, *Nucl. Phys.* **B324**, 233.

LAKE, K.: 1979, *Phys. Rev.* **D19**, 421.

LAKE, K. and C. HELLABY: 1981, *Phys. Rev.* **D24**, 3019.

LAMB, F. K.: 1991, in D. L. Lambert (ed.), *"Frontiers of Stellar Evolution"*, Astronomical Society of the Pacific, p.299.

LANDAU, L. D.: 1932, *Physikalische Zeitschrift Sowjetunion* **1**, 285.

LANDAU, L. D. and E. M. LIFSHITZ: 1959, *Fluid Mechanics*, Pergamon Press, London.

LANDAU, L. D. and E. M. LIFSHITZ: 1975, *The Classical Theory of Fields*, Pergamon Press, Oxford.

LAUER, T. R. *et al.*: 1992, *Astronom. J.* **103**, 703.

LAVRELASHVILI, G.: 1997, E-print gr-qc/9701049.

LAVRELASHVILI, G. V., V. A. RUBAKOV, and P. G. PINYAKOV: 1988, in M. A. Markov, V. A. Berezin, and V. P. Frolov (eds.), *Quantum Gravity, Proceedings of the Fourth Seminar on Quantum Gravity*, Moscow, World Scientific, Singapore, p. 139.

LEAHY, D. A. and W. G. UNRUH: 1979, *Phys. Rev.* **D19**, 3509.

LÉAUTÉ, B. and B. LINET: 1976, *Phys. Lett.* **A58**, 5.

LÉAUTÉ, B. and B. LINET: 1982, *J. Phys.* **A15**, 1821.

LEAVER, E. W.: 1985, *Proc. R. Soc. London* **A402**, 285.

LEAVER, E. W.: 1986a, *J. Math. Phys.* **27**, 1238.

LEAVER, E. W.: 1986b, *Phys. Rev.* **D34**, 384.

LEAVER, E. W.: 1990, *Phys. Rev.* **D41**, 2986.

LEE, C. W.: 1976, *Commun. Math. Phys.* **51**, 157.

LEE, C. H.: 1976, *J. Math. Phys.* **17**, 1226.

LEE, C. H.: 1977, *Phys. Lett.* **B68**, 152.

LEE, K.-Y., V. P. NAIR, and E. WEINBERG: 1992a, *Phys. Rev. Lett.* **68**, 1100.

LEE, K.-Y., V. P. NAIR, and E. WEINBERG: 1992b, *Phys. Rev.* **D45**, 2751.

LEE, K.-Y., V. P. NAIR, and E. WEINBERG: 1992c, *Gen. Rel. & Grav.* **24**, 1203.

LEE, K. and E. J. WEINBERG: 1991, *Phys. Rev.* **D44**, 3159.

LEMAÎTRE, G: 1933, *Ann. Soc. Bruxelles* **A53**, 51.

LEWIS, G. N. and M. RANDAL: 1961, *Thermodynamics*, 2nd edition, Mc Graw-Hill, New-York.

LI, ZH.-Y., T. CHIUEN, and M. C. BEGELMAN: 1992, *Astrophys. J.* **394**, 459.

LIEBLING, S. L. and M. W. CHOPTUIK: 1996, *Phys. Rev. Lett.* **77**, 1424.

LIFSHITZ, E. M.: 1946, *Zh. Eksp. Teor. Fiz.* **16**, 587.

LINDQUIST, R.: 1963, *J. Math. Phys.* **4**, 938.

LINET, B.: 1976, *J. Phys.* **A9**, 1081.

LINET, B.: 1977a, *C. R. Acad. Sc. Paris* **A284**, 215.

LINET, B.: 1977b, *C. R. Acad. Sc. Paris* **A284**, 1167.

LINET, B.: 1979, *J. Phys.* **A12**, 839.

LIPUNOV, V. M.: 1987, *Astrophysics of Neutron Stars*, Nauka, Moscow (in Russian); 1992, Springer-Verlag, Berlin (English translation).

LIPUNOV, V. M., K. A. POSTNOV, E. M. PROKHOROV, and E. YU. OSMINKIN: 1994, *Astrophys. J. Lett.* **423**, L121.

LIPUNOV, V. M., K. A. POSTNOV, and M. E. PROKHOROV: 1996, in R. A. Sunyaev (ed.), *Review of Astrophysics and Space Science*, Harward Acad. Publ., **17**, 1.

LIU, H.: 1995, *Class. Quantum Grav.* **12**, 543.

LIU, H. and B. MASHHOON: 1995, On the spectrum of oscillations of a Schwarzschild black hole (preprint).

LOHIYA, D.: 1982, *J. Phys.* **A15**, 1815.

LONSDALE, S., and I. MOSS: 1988, *Nucl. Phys.* **B298**, 693.

LOSSEV, A. and I. D. NOVIKOV: 1992, *Class. Quantum Grav.* **9**, 2309.

LOUIS-MARTINEZ, D. and G. KUNSTATTER: 1994, *Phys. Rev.* **D49**, 5227.

LOUSTO, C.: 1993, *Nucl. Phys.* **B410**, 155; Erratum, ibid: 1995 **B449**, 433.

LOUSTO, C.: 1996, E-print gr-qc/9601006.

LOUSTO, C. O. and N. SÁNCHEZ: 1992, *Nucl. Phys.* **B383**, 377.

LOVELACE, R. V. E.: 1976, *Nature* **262**, 649.

LOVELACE, R. V. E., J. MACAUSLAN, and M. BURNS: 1979, in *Proceedings of La Jolla Institute Workshop on Particle Acceleration Mechanism in Astrophysics*, American Inst. of Phys., N. Y.

LUCKOCK, H. C. and I. MOSS: 1986, *Phys. Lett.* **B176**, 314.

LUDVIGSEN, M. and J. A. G. VICKERS: 1983, *J. Phys.* **A16**, 1169.

LUMINET, J. P. and J. A. MARCK: 1985, *Mon. Not. RAS* **212**, 57.

LYNDEN-BELL, D.: 1969, *Nature* **233**, 690.

LYUTIKOV, M.: 1994, *Phys. Rev.* **D49**, 4041.

MACDONALD, D. A.: 1984, *Mon. Not. RAS* **211**, 313.

MACDONALD, D. A. and W. M. SUEN: 1985, *Phys. Rev.* **D32**, 848.

MACDONALD, D. A. and K. S. THORNE: 1982, *Mon. Not. RAS* **198**, 345.

MACGIBBON, J. H.: 1987, *Nature* **329**, 308.

MACGIBBON, J. H. and B. J. CARR: 1991, *Astrophys. J.* **371**, 447.

MADSEN, J.: 1994, in B. Sinha, Y. P. Viyogi, and S. Raha (eds.), *Proceedings of the 2nd Int. Conf. on Physics and Astrophysics of Quark-Gluon Plasma. Calcutta, 1993*, World Scientific, p.186.

MAEDA, K., K. SATO, M. SASAKI, and H. KODAMA: 1982, *Phys. Lett.* **B108**, 98.

MAISON, D.: 1996, *Phys. Lett.* **B366**, 82.

MAJUMDAR, B. and N. PANCHAPAKESAN: 1989, *Phys. Rev.* **D40**, 2568.

MALDACENA, J. M.: 1996, *Black Holes in String Theory*, PhD thesis; E-print hep-th/9607235.

MALDACENA, J. M. and A. STROMINGER: 1996, *Phys. Rev. Lett.* **77**, 428.

MANDAL, G., A. M. SENGUPTA, and S. R. WADIA: 1991, *Mod. Phys. Lett.* **6**, 1685.

MANN, R. B. and S. F. ROSS: 1995, *Phys. Rev.* **D52**, 2254.

MARCK, J. A.: 1983, *Proc. R. Soc. London* **A385**, 431.

MARKOV, M. A.: 1965, *Suppl. Progr. Theor. Phys.* Extra Number, p. 85.

MARKOV, M. A.: 1966, *Zh. Eksp. Teor. Fiz.* **51**, 879.

MARKOV, M. A.: 1970, *Ann. Phys.* **59**, 109.

MARKOV, M. A.: 1971, *Cosmology and Elementary Particles*, preprint IC/71/33, Trieste, Italy.

MARKOV, M. A.: 1973, *Usp. Fiz. Nauk* **111**, 3.

MARKOV, M. A.: 1974, in C. DeWitt-Morette (ed.), *Gravitational Radiation and Gravitational Collapse*, D. Reidel, Dordrecht, p. 106.

MARKOV, M. A.: 1976, *On the Nature of Matter*, Nauka, Moscow.

MARKOV, M. A.: 1981a, *On the Upper Limit of the Cosmic Rays Energy Spectrum* (DUMAND Type Experiments), Preprint P-0197 INR, Moscow.

MARKOV, M. A.: 1981b, *Maximon-Type Scenario of the Universe*, Preprint P-0207 INR, Moscow.

MARKOV, M. A.: 1981c, *On the maximon and the concept of elementary particles*, Preprint P-0208 INR, Moscow.

MARKOV, M. A.: 1982, *JETP Letters* **36**, 266.

MARKOV, M. A.: 1984, *Ann. Phys.* **155**, 333.

MARKOV, M. A. and V. P. FROLOV: 1970, *Teor. Mat. Fiz.* **3**, 3.

MARKOV, M. A. and V. P. FROLOV: 1972, *Teor. Mat. Fiz.* **13**, 41.

MARKOV, M. A. and V. P. FROLOV: 1979, *Pis'ma v Zh. Eksp. Teor. Fiz.* **29**, 372.

MARKOV, M. A. and V. F. MUKHANOV: 1985, *Nuovo Cim.* **B86**, 97.

MARKOV, M. A. and I. M. ZHELEZNYKH: 1981, in J. Learned (ed.), *Proc. of the 1979 DUMAND Summer Workshop at Khabarovsk and Lake Baikal*, Hawaii DUMAND Center, Univ. of Hawaii, p. 177.

MARTELLINI, M. and A. TREVES: 1977, *Phys. Rev.* **D15**, 3060.

MASEVICH, A. G. and A. V. TUTUKOV: 1988, *Stellar evolution: theory and observations*, Nauka, Moscow (in Russian).

MASHHOON, B.: 1973, *Astrophys. J. Lett.* **181**, L65.

MASHHOON, B.: 1985, *Phys. Rev* **D31**, 290.

MASOOD-UL-ALAM, A. K. M.: 1992, *Class. Quantum Grav.* **9**, L53.

MASOOD-UL-ALAM, A. K. M.: 1993, *Class. Quantum Grav.* **10**, 2649.

MASSIEU, M. J.: 1869, *C. R. Acad. Sci. Paris* **69**, 858.

MATZNER, R. A.: 1968, *J. Math. Phys.* **9** 163.

MATZNER, R. A., C. DEWITT-MORETTE, B. NELSON, and T.-R. ZHANG: 1985, *Phys. Rev.* **D31**, 1869.

MATZNER, R. A. and M. P. RYAN JR.: 1977, *Phys. Rev.* **D16**, 1636.

MATZNER, R. A. and M. P. RYAN JR.: 1978, *Ap. J. Suppl.* **36**, 451.

MATZNER, R. A., H. E. SEIDEL, S. L. SHAPIRO, L. SMARR, W.-M. SUEN, S. A. TEUKOLSKY, and J. WINICOUR: 1995, *Science* **270**, 941.

MATZNER, R. A., N. ZAMORANO, and V. D. SANDBERG: 1979, *Phys. Rev.* **D19**, 2821.

MAYO, A. E. and J. D. BEKENSTEIN: 1996, *Phys. Rev.* **D54**, 5059.

MAZEH, T., J. VAN PARADIJS, E. P. J. VAN DEN HEUVEL, and G. J. SAVONIJE: 1986, *Astron. Astrophys.* **157**, 113.

MAZUR, P. O.: 1982, *J. Phys.* **A15**, 3173.

MAZUR, P. O.: 1984, *Gen. Rel. & Grav.* **16**, 211.

MCCARTHY, P. J.: 1972a, *J. Math. Phys.* **13**, 1837.

MCCARTHY, P. J.: 1972b, *Proc. R. Soc. London* **A330**, 517.

MCCARTHY, P. J.: 1973, *Proc. R. Soc. London* **A333**, 317.

MCCARTHY, P. J. and M. CRAMPIN: 1973, *Proc. R. Soc. London* **A335**, 301.

MCCLINTOCK, J. E.: 1988, in M. Kafatos (ed.), *Supermassive Black Holes*, Cambridge University Press, Cambridge, p.1.

McClintock, J. E.: 1992, in Y. Tanaka and K. Koyama (eds.), *Frontiers of X-ray Astronomy*, Tokyo: Universal Academic Press, p.333.

McClintock, J. E. and R. A. Remillard: 1990, *Bull. Am. Astron. Soc.* **21**, 1206.

McNamara, J. M.: 1978a, *Proc. R. Soc. London* **A358**, 499.

McNamara, J. M.: 1978b, *Proc. R. Soc. London* **A364**, 121.

Melvin, M. A.: 1964, *Phys. Lett.* **8**, 65.

Melvin, M. A.: 1965, *Phys. Rev.* **139**, B225.

Mensky, M. and I. Novikov: 1996a, *Int. J. Mod. Phys.* **D5**, 179.

Mensky, M. and I. Novikov: 1996b, *Int. J. Mod. Phys.* **D5**, 1.

Mestel, L.: 1968, *Mon. Not. RAS* **138**, 359.

Mestel, L. and S. Shibata: 1994, *Mon. Not. RAS* **271**, 621.

Michel, F. C.: 1969, *Astrophys. J.* **158**, 727.

Michel, F. C.: 1973, *Astrophys. J.* **180**, 207.

Michel, F. C.: 1991, *Theory of Neutron Star Magnetospheres.* University of Chicago Press. Chicago.

Mikheeva, E. V. and I. D. Novikov: 1993, Phys. Rev. D **47**, 1432.

Miller, J. G.: 1979, *J. Math. Phys.* **20**, 1345.

Mineshige, S. and J. C. Wheeler: 1989, *Astrophys. J.* **343**. 241.

Misner, C. W.: 1960, *Phys. Rev.* **118**, 1110.

Misner, C. W.: 1963, *Ann. Phys.* **24**, 102.

Misner, C. W.: 1972, *Phys. Rev. Lett.* **28**, 994.

Misner, C. W.: 1978, *Phys. Rev.* **D18**, 4510.

Misner, C. W., K. S. Thorne, and J. A. Wheeler: 1973, *Gravitation*, Freeman, San Francisco.

Misner, C. W. and J. A. Wheeler: 1957, *Ann. Phys.* **2**, 525.

Misra, R. M.: 1977, *Progr. Theor. Phys.* **58**, 1205.

Miyoshi, M., J. Moran, J. Herrnstein, L. Greenhill, N. Nakai, P. Diamond, and M. Inoue: 1995, Nature **373**, 127.

Moncrief, V.: 1974a, *Ann. Phys.* **88**, 323.

Moncrief, V.: 1974b, *Phys. Rev.* **D9**, 2707.

Moncrief, V.: 1974c, *Phys. Rev.* **D10**, 1057.

Moncrief, V.: 1975, *Phys. Rev.* **D12**, 1526.

Moreschi, O. M. and S. Dain: 1996, *Phys. Rev.* **D53**, R1745.

Morgan, D.: 1991, *Phys. Rev.* **D43**, 3144.

Morris, M. S., K. S. Thorne, and U. Yurtsever: 1988, *Phys. Rev. Lett.* **61**, 1446.

Moscoso, M. D. and J. C. Wheeler: 1994, in A. W. Shafter (ed.), *Interacting Binary Stars*, Astronomical society of the Pacific, Conference Series, **56**, p.100

Moss, I. G.: 1985, *Phys. Rev.* **D32**, 1333.

Mukhanov, V. F.: 1986, *Sov. Phys. JETP Lett.* **44**, 63.

Mukhanov, V. and R. Brandenberger: 1992, *Phys. Rev. Lett.* **68**, 1969.

MUKHANOV, V., A. WIPF, and A. ZELNIKOV: 1994, *Phys. Lett.* **B332**, 283.

MÜLLER, R. and C. LOUSTO: 1994, *Phys. Rev.* **D49**, 1922.

MÜLLER ZUM HAGEN, H.: 1970, *Proc. Camb. Phil. Soc.* **67**, 415.

MÜLLER ZUM HAGEN, H., D. C. ROBINSON, and H. SEIFERT: 1973, *Gen. Rel. & Grav.* **4**, 53.

MÜLLER ZUM HAGEN, H., D. C. ROBINSON, and H. J. SEIFERT: 1974, *Gen. Rel. & Grav.* **5**, 61.

MYERS, R. C.: 1987, *Nucl. Phys.* **B289**, 701.

MYERS, R. C. and M. J. PERRY : 1986, *Annals Phys.* **172**, 304.

MYSAK, L. and G. SZEKERES: 1966, *Can. J. Phys.* **44**, 617.

NADEJIN, D. K., I. D. NOVIKOV, and A. G. POLNAREV: 1977, in *Abstracts of contributed papers GR8*, Waterloo, Canada, p. 382.

NADEJIN, D. K., I. D. NOVIKOV, and A. G. POLNAREV: 1978, *Astron. Zh.* **55**, 216.

NAKAMURA, T. and K. OOHARA: 1983, *Phys. Lett.* **A98**, 403.

NAKAMURA, T., K. OOHARA, and Y. KOJIMA: 1987, *Progr. Theor. Phys. Suppl.* **90**.

NAKAMURA, T. and M. SASAKI: 1981a, *Phys. Lett.* **B106**, 69.

NAKAMURA, T. and M. SASAKI: 1981b, *Phys. Lett.* **A89**, 185.

NAKAMURA, T. and H. SATO: 1976, *Phys. Lett.* **B61**, 371.

NAKAMURA, T., S. L. SHAPIRO, and S. A. TEUKOLSKY: 1988, *Phys. Rev.* **D38**, 2972.

NARAYAN, R.: 1997, in J. N. Bahcall and J. P. Ostriker (eds.), *Unsolved problems in Astrophysics*, Princeton University Press, page 301.

NARAYAN, R., J. E. MCCLINTOCK, and I. YI: 1996, *Astrophys. J.*, **457**, 821.

NARAYAN, R. and I. YI: 1994, *Astrophys. J.* **428**, L13.

NASELSKY, P. D. and A. G. POLNAREV: 1985, *Astr. Zh.* **29**, 487.

NE'EMAN, Y.: 1965, *Astrophys. J.* **141**, 1303.

NERNST, W.: 1906, *Nachr. kgl. Ges. Wiss. Góttingen Math.-physik. Kl.*, Heft 1.

NEWMAN, E. and R. PENROSE: 1962, *J. Math. Phys.* **3**, 566.

NEWMAN, E. and T. UNTI: 1963, *J. Math. Phys.* **4**, 1467.

NEWMAN, R. P.: 1986a, *Class. Quantum Grav.* **3**, 527.

NEWMAN, R. P.: 1986b, in P. G. Bergmann and V. De Sabbata (eds,), *Topological Properties and Global Structure of Space-Time*, Plenum, New-York, p.153.

NITTA, S., M. TAKAHASHI, and A. TOMIMATSU: 1991, *Phys. Rev.* **D44**, 2295.

NITYANANDA, R. and R. NARAYAN: 1981, *Phys. Lett.* **A82**, 1.

NOLLERT, H. P.: 1993, *Phys. Rev.* **D47** , 5253.

NOLLERT, H. P. and B. G. SCHMIDT: 1992, *Phys. Rev.* **D45** , 2617.

NORDSTRÖM, G.: 1918, *Proc Kon. Ned. Akad. Wet.* **20**, 1238.

NOVIKOV, I. D.: 1961, *Astron. Zh.* **38**, 564.

NOVIKOV, I. D.: 1962a, *Vestnik MGU* Ser. III, No. **5**, 90.

NOVIKOV, I. D.: 1962b, *Vestnik MGU* Ser. III, No. **6**, 61.

NOVIKOV, I. D.: 1963, *Astron. Zh.* **40**, 772.
NOVIKOV, I. D.: 1964a, *Soobshch. GAISH* No. **132**, 3, 43.
NOVIKOV, I. D.: 1964b, *Astron. Zh.* **41**, 1075.
NOVIKOV, I. D.: 1966a, *Pis'ma v Zh. Eksp. Teor. Fiz.* **3**, 223.
NOVIKOV, I. D.: 1966b, *Astron. Zh.* **43**, 911.
NOVIKOV, I. D.: 1969, *Zh. Eksp. Teor. Fiz.* **57**, 949.
NOVIKOV, I. D.: 1970, *Zh. Eksp. Teor. Fiz.* **59**, 262.
NOVIKOV, I. D.: 1974, in *Astrophysics and Gravitation*, Proceedings of the 16th Solvay Conf. on Physics, 1973, edition de l'Universite de Bruxelles, p.317.
NOVIKOV, I. D.: 1983, *Evolution of the Universe*, Cambridge University Press, p.169.
NOVIKOV, I. D.: 1989, *JETP* **95**, 769.
NOVIKOV, I. D.: 1990, *Black Holes and the Universe*, Cambridge Univ. Press.
NOVIKOV, I. D.: 1992, *Phys. Rev.* **D45**, 1989.
NOVIKOV, I. D.: 1993, in N. Sánchez and A. Zichichi (eds.), *The Proceedings of the International Workshop on Theoretical Physics, String Quantum Gravity and Physics at the Planck Energy Scale*, Erice, Italy 21-28 June 1992, World Scientific, Singapore, p.480.
NOVIKOV, I. D.: 1997, *Physics of Black Holes*, TAC preprint; to appear in the Proceedings of the Symposium "Non-Linear Phenomena in Accreation Discs around Black Holes", Laugarvata, Iceland.
NOVIKOV, I. D. and V. P. FROLOV: 1989, *Physics of Black Holes*, Kluwer Acad. Publ.; Russion version: 1986, *Fizika Chernych Dyr*, Nauka, Moscow.
NOVIKOV, I. D. and L. M. OZERNOY: 1964, Publication of the P. N. Lebedev Phys. Inst., A-17.
NOVIKOV, I. D. and A. G. POLNAREV: 1979, in L. L. Smarr (ed.), *Sources of Gravitational Radiation* Proceedings of the Battelle Seattle Workshop, July 1978, Cambridge Univ. Press, Cambridge, p.173.
NOVIKOV, I. D. and A. G. POLNAREV: 1980, *Astron. Zh.* **57**, 250.
NOVIKOV, I. D., A. G. POLNAREV, A. A. STAROBINSKY, and YA. B. ZEL'DOVICH: 1979, *Astron. Astrophys.* **80**, 104.
NOVIKOV, I. D. and A. A. STAROBINSKY: 1980a, in *Abstracts of contributed papers of the 9th Intern. Conf. on General Relativity and Gravitation*, Jena, DDR, p. 268.
NOVIKOV, I. D. and A. A. STAROBINSKY: 1980b, Publications of the Space Research Institute, Pr-585, Moscow.
NOVIKOV, I. D. and A. A. STAROBINSKY: 1980c, *Zh. Eksp. Teor. Fiz.* **78**, 3.
NOVIKOV, I. D. and B. E. STERN: 1986, in G. Giuricin *et al.* (eds.), *Structure and Evolution of Active Galactic Nuclei*, D. Reidel, Dordrecht, p. 149.
NOVIKOV, I. D. and K. S. THORNE: 1973, in C. Dewitt and B. S. DeWitt (eds.), *"Black holes, Les Astres Occults"*, Gordon and Breach Science Publishers, p.343.
NOVIKOV, I. D. and YA. B. ZEL'DOVICH: 1966, *Nuovo Cim. Suppl.* **4**, 810; addendum 2.

OHTA, T. and T. KIMURA: 1982, *Progr. Theor. Phys.* **68**, 1175.
OKAMOTO, I.: 1978, *Mon. Not. RAS* **185**, 69.
OKAMOTO, I.: 1992, *Mon. Not. RAS* **152**, 192.
O'NEILL, B.: 1994, *The Geometry of Kerr Black Holes*, A. K. Peters, Wellesley.
ONOZAWA, H., T. MISHIMA, T. OKAMURA, and H. ISHIHARA: 1996, *Phys. Rev.* **D53**, 7033.
OOHARA, K.: 1984, *Progr. Theor. Phys.* **71**, 738.
OOHARA, K. and T. NAKAMURA: 1983a, *Phys. Lett.* **A98**, 407.
OOHARA, K. and T. NAKAMURA: 1983b, *Progr. Theor. Phys.* **70**, 757.
OOHARA, K. and T. NAKAMURA: 1984, *Progr. Theor. Phys.* **71**, 91.
OOHARA, K. and H. SATO: 1981, *Progr. Theor. Phys.* **65**, 1891.
OPPENHEIMER, J. R. and H. SNYDER: 1939, *Phys. Rev.* **56**, 455.
OPPENHEIMER, J. R. and G. VOLKOFF: 1939, *Phys. Rev.* **55**, 374.
ORI, A.: 1991, *Phys. Rev. Lett.* **67**, 789.
ORI, A.: 1992, *Phys. Rev. Lett.* **68**, 2117.
ORI, A.: 1997, *Gen. Rel. & Grav.* **29**, 881.
ORI, A., and T. PIRAN: 1990, *Phys. Rev.* **D42**, 1068.
ORTIZ, M. E.: 1992, *Phys. Rev.* **D45**, R2586.
OSHIRO, Y., K. NAKAMURA, and A. TOMIMATSU: 1994, *Prog. Theor. Phys.* **91**, 1265.
OSTRIKER, J. P., R. MCCRAY, R. WEAVER, and A. YAHIL: 1976, *Astrophys. J.* **208**, L61.
PACZYNSKI, B: 1974, *Astron. Astrophys.* **34**, 161.
PACZYNSKI, B.: 1983, *Astrophys. J.* **273**, L81.
PAGE, D. N.: 1976a, *Phys. Rev.* **D13**, 198.
PAGE, D. N.: 1976b, *Phys. Rev.* **D14**, 3260.
PAGE, D. N.: 1976c, *Phys. Rev.* **D14**, 1509.
PAGE, D. N.: 1977, *Phys. Rev.* **D16**, 2402.
PAGE, D. N.: 1980, *Phys. Rev. Lett.* **44**, 301.
PAGE, D. N.: 1982, *Phys. Rev.* **D25**, 1499.
PAGE, D. N.: 1983, *Phys. Rev. Lett.* **50**, 1013.
PAGE, D. N.: 1993a, *Sci. Am.* **267** (August), 10.
PAGE, D. N.: 1993b, in N. Sánchez and A. Zichichi (eds.), *String Quantum Gravity and Physics at the Planck Energy Scale*. Proceedings of the NATO International Workshop on Theoretical Physics, Erice, Italy, 21-28 June 1992, World Scientific, Singapore, 417.
PAGE, D. N.: 1994, in R. B. Mann and R. G. McLenaghan (eds.), *Proceedings of the 5th Canadian Conference on General Relativity and Relativistic Astrophysics, University of Waterloo, 13-15 May, 1993*, World Scientific, Singapore, p.1.
PAGE, D. N.: 1997, E-print gr-qc/9706029.
PAGE, D. and K. S. THORNE: 1974, *Astrophys. J.* **191**, 499.
PAINLEVÉ, P.: 1921, *Comp. Rend. Acad. Sci. (Paris)* **173**, 677.

PANANGADEN, P. and R. M. WALD: 1977, *Phys. Rev.* **D16**, 929.

PAPADOPOULOS, D. and B. C. XANTHOPOULOS: 1984, *Nuovo Cim.* **B83**, 113.

PARENTANI, R., J. KATZ, and I. OKAMOTO: 1995, *Class. Quantum Grav.* **12**, 1663.

PARK, M. G.: 1990a, *Astrophys. J.* **354**, 64.

PARK, M. G.: 1990b, *Astrophys. J.* **354**, 83.

PARK, S. I. and E. L. VISHNIAK: 1989, *Astrophys. J.* **337**, 78.

PARK, S. I. and E. L. VISHNIAK: 1990, *Astrophys. J.* **353**, 1.

PELEG, Y., S. BOSE, and L. PARKER: 1997, *Phys. Rev.* **D55**, 4525.

PELLETIER, G. and R. PUDRITZ: 1992, *Astrophys. J.* **394**, 117.

PENROSE, R.: 1963, *Phys. Rev. Lett.* **10**, 66.

PENROSE, R.: 1964, in C. DeWitt and B. DeWitt (eds.), *Relativity, Groups, and Topology*, N. Y.-London, p. 565.

PENROSE, R.: 1965a, *Phys. Rev. Lett.* **14**, 57.

PENROSE, R.: 1965b, *Proc. R. Soc. London* **A284**, 159.

PENROSE, R.: 1968, *Structure of Space-Time* in C. M. DeWitt and J. A. Wheeler (eds.), *Battelle Rencontres*, Benjamin, N. Y.

PENROSE, R.: 1969, *Rev. Nuovo Cim.* **1**, 252.

PENROSE, R.: 1972, *Sci. Am.* **226**, No. 5, 38.

PENROSE, R.: 1973, *Ann. Phys.* **224**, 125.

PENROSE, R.: 1974, in Seminar at Cambridge Univ., Cambridge Univ.

PENROSE, R.: 1976, in M. Cahen and M. Flato (eds.), *Differential Geometry and Relativity*, D. Reidel, Dordrecht.

PENROSE, R.: 1978, in N. R. Leibovitz *et al.* (eds.), *Theoretical Principles in Astrophysics and General Relativity*, Univ. of Chicago Press, Chicago, p. 217.

PENROSE, R.: 1979, in S. W. Hawking and W. Israel (eds.), *General Relativity, an Einstein Centenary Survey*, Cambridge Univ. Press, Cambridge.

PEŘINA, J., Z. HRADIL, and B. JURČO: 1994, *Quantum Optics and Fundamentals of Physics*, Kluwer Acad. Publ., Dodrecht.

PERJÉS, Z.: 1993, *Nucl. Phys.* **B403**, 809.

PETERSON, B. M.: 1997, *Active Galactic Nuclei*, Cambridge Univ. Press.

PETRICH, L. I., S. L. SHAPIRO, and I. WASSERMAN: 1985, *Astrophys. J. Suppl.* **58**, 297.

PHINNEY, S.: 1983a, in A. Ferrari and A. Pacholczyk (eds.), *Astrophysical Jets.*, Proc. of International Workshop, D. Reidel Publishing Company, Dordrecht.

PHINNEY, S.: 1983b, PhD Thesis, Cambridge.

PIKEL'NER, S. B.: 1961, *Foundations of Cosmical Electrodynamics*, Nauka, Moscow (1964, NASA).

PIRAN, T. and J. SHAHAM: 1977, *Phys. Rev.* **D16**, 1615.

PIRAN, T., J. SHAHAM, and J. KATZ: 1975, *Astrophys. J. Lett.* **196**, L107.

PIRANI, F. A. E.: 1959, *Proc. R. Soc. London* **A252**, 96.

PIRANI, F. A. E.: 1964, in *Lectures on General Relativity*, Brandeis Summer Institute in Theoretical Physics, **1**, p. 249.

PLANCK, M.: 1911, *Thermodynamik*, 3d ed., Veit and Co., Leipzig, p.279; English translation: *Treatise on Thermodynamis*, Reprint Dover, New York, 1945.

POISSON, E.: 1993a, *Phys. Rev.* **D47**, 1497.

POISSON, E.: 1993b, *Phys. Rev.* **D48**, 1860.

POISSON, E.: 1995, *Phys. Rev.* **D52**, 5719.

POISSON, E. and W. ISRAEL: 1988, *Class. Quantum Grav.* **5**, L201.

POISSON, E. and W. ISRAEL: 1990, *Phys. Rev.* **D41**, 1796.

POISSON, E. and M. SASAKI: 1995, *Phys. Rev.* **D51**, 5753.

POLCHINSKI, J.: 1989, *Nucl. Phys.* **B325**, 619.

POLITZER, H. D.: 1992, *Phys. Rev. D* **46**, 4470.

POLLARD, D.: 1983, *J. Phys.* **A16**, 565.

POLNAREV, A. G.: 1994, *Astron. and Astrophys. Transactions* **5**, 35.

POLNAREV, A. G. and M. YU. KHLOPOV: 1981, *Astr. Zh* **58**, 706.

POLNAREV, A. G. and M. YU. KHLOPOV: 1985, *Usp. Fiz. Nauk* **145**, 369.

POLNAREV, A. G. and M. REES: 1994, Astron Astrophys. **283**, 301.

POLNAREV, A. G. and R. ZEMBOWICZ: 1991, *Phys. Rev.* **D43**, 1106.

POLYAKOV, A. M.: 1974, *JETP Lett.* **20**, 194.

POLYAKOV, A. M.: 1974, *Pis'ma v Zh. Eksp. Teor. Fiz.* **20**, 194.

POLYAKOV, A. M.: 1975, Zh. Eksp. Teor. Fiz. **41**, 988.

POPLAVSKY, R. P.: 1981, *Thermodynamics of Information Processes*, Nauka, Moscow

PRENDERGAST, K. H. and G. R. BURBIDGE: 1968, *Astrophys. J.* **151**, L83.

PRESKILL, J.: 1992, In *Houston 1992, Proceedings, Black Holes, Membranes, Wormholes, and Superstrings*, p.22; E-print hep-th/9209058.

PRESS, W. H.: 1971, *Ap. J* **170**, L105.

PRESS, W. H.: 1972, *Astrophys. J.* **175**, 243.

PRESS, W. H.: 1972, *Astrophys. J.* **175**, 243.

PRESS, W. H. and S. A. TEUKOLSKY: 1972, *Nature* **238**, 211.

PRESS, W. H. and S. A. TEUKOLSKY: 1973, *Astrophys. J.* **185**, 649.

PRICE, R. H.: 1972a, *Phys. Rev.* **D5**, 2419.

PRICE, R. H.: 1972b, *Phys. Rev.* **D5**, 2439.

PRICE, R. H. and J. PULLIN: 1994, *Phys. Rev. Lett.* **72**, 3297.

PRICE, R. H. and K. S. THORNE: 1986, *Phys. Rev.* **D33**, 915.

PRINGLE, J. E. and M. J. REES: 1972, *Astron. Astrophys.* **21**, 1.

PUDRITZ, R.: 1981a, *Mon. Not RAS* **195**, 881.

PUDRITZ, R.: 1981b, *Mon. Not RAS* **195**, 887.

PUNSLY, B.: 1991, *Astrophys. J.* **372**, 424.

PUNSLY, B. and F. C. CORONITI: 1990a, *Astrophys. J.* **350**, 518.

PUNSLY, B. and F. C. CORONITI: 1990b, *Astrophys. J.* **354**, 583.

REDMOUNT, I. H.: 1984, Preprint, Kyoto Univ. RIFP-584.

REES, M. J.: 1982, in G. R. Reigler and R. D. Blandford (eds.), *Proceedings of AIP Conference, 'The Galactic Center'*, Caltech, N. Y., p. 166.

REES, M.: 1990a, Science **247**, N4944, 16 February, p.817.

REES, M.: 1990b, Scientific American, November, p.26.

REES, M.: 1996, in O. Lahv, E. Terlevich, and R. J. Terlevich (eds.). *Gravitational Dynamics*, Cambridge Univ. Press, p.103.

REES, M. J., M. C. BEGELMAN, R. D. BLANDFORD, and E. PHINNEY: 1982, *Nature* **295**, 17.

REGGE, T. and J. A. WHEELER: 1957, *Phys. Rev.* **108**, 1063.

REISSNER, H.: 1916, *Ann. Phys.* **50**, 106.

REUTER, M.: 1992, *Class. Quantum Grav.* **9**, 751.

RINDLER, W.: 1966, *Am. J. Phys.* **34**, 1174.

ROBINSON, D. C.: 1975, *Phys. Rev. Lett.* **34**, 905.

ROBINSON, D. C.: 1977, *Gen. Rel. & Grav.* **8**, 695.

ROMAN, T. A.: 1986, *Phys. Rev.* **D33**, 3526.

ROMAN, T. A.: 1988, *Phys. Rev.* **D37**, 546.

ROMAN, T. A. and P. G. BERGMANN: 1983, *Phys. Rev.* **D28**, 1265.

RUBACK, P.: 1988, *Class. Quantum Grav.* **5**, L155.

RUDERMAN, M. and P. G. SUTHERLAND: 1975, *Astrophys. J.* **196**, 51.

RUFFINI, R.: 1973, *Phys. Rev.* **D7**, 972.

RUFFINI, R.: 1979, in *Astrofisica e Cosmologia Gravitazione, Quanti e Relativita*, Giunti Barbera, Firenze.

RUFFINI, R. and M. SASAKI: 1981, *Prog. Theor. Phys.* **66**, 1627.

RUFFINI, R. and J. A. WHEELER: 1971a, *Physics Today* **24**, 30.

RUFFINI, R. and J. A. WHEELER: 1971b, in *Proceedings of the Conference on Space Physics*, European Space Research Organization, Paris.

RUFFINI, R. and J. R. WILSON: 1975, *Phys. Rev.* **D12**, 2959.

RYLOV, YU. A.: 1961, *Zh. Eksp. Teor. Fiz.* **40**, 1955.

SACHS, R. K.: 1961, *Proc. R. Soc. London* **A264**, 309.

SACHS, R. K.: 1962, *Phys. Rev.* **128**, 2851.

SACHS, R. K.: 1964, in C. DeWitt and B. DeWitt (eds.), *Relativity, Groups, and Topology*, N. Y.-London.

SAKHAROV, A. D.: 1968, *Soviet Phys. Doklady* **12**, 1040.

SAKHAROV, A. D.: 1976, *Theor. Math. Phys. (USSR)* **23**, 435.

SAKURAI, T.: 1985, *Astron. Astrophys.* **152**. 121.

SALPETER, E.: 1964, *Astrophys. J.* **140**, 796.

SÁNCHEZ, N.: 1976, *J. Math. Phys.* **17**, 688.

SÁNCHEZ, N.: 1977, *Phys. Rev.* **D16**, 937.

SÁNCHEZ, N.: 1978a, *Phys. Rev.* **D18**, 1030.

SÁNCHEZ, N.: 1978b, *Phys. Rev.* **D18**, 1798.

SÁNCHEZ, N.: 1997, E-print hep-th/9711068.

SASAKI, M.: 1994, *Prog. Theor. Phys* **92**, 17.

SASAKI, M. and T. NAKAMURA: 1981, *Phys. Lett.* **A87**, 85.

SASAKI, M. and T. NAKAMURA: 1982a, *Phys. Lett.* **A89**, 68.

SASAKI, M. and T. NAKAMURA: 1982b, *Progr. Theor. Phys.* **67**, 1788.

SASAKI, M. and T. NAKAMURA: 1990, *Gen. Rel. & Grav.* **22**, 1351.

SATO, K.: 1983, *Multiple Kerr Metric Solutions of Einstein's Field Equations*, Preprint RIFP-527, Kyoto Univ.

SATO, K., M. SASAKI, H. KODAMA, and K. MAEDA: 1981, *Progr. Theor. Phys.* **65**, 1443.

SCHILD and MAEDER: 1985, *Astron. Astrophys.* **143**, L7.

SCHMIDT, B. G.: 1971, *Gen. Rel. & Grav.* **1**, 269.

SCHULMAN, L. S.: 1968, *Phys. Rev.* **176**, 1558.

SCHULTZ, A. L. and R. H. PRICE: 1985, *Astrophys. J.* **291**, 1.

SCHUMACHER, B., W. A. MILLER, and W. H. ZUREK: 1992, *Phys. Rev.* **D46**, 1416.

SCHUTZ, B. F. and C. M. WILL: 1985, *Ap. J.* **291**, L33.

SCHVARTSMAN, V. F.: 1971, *Astr. Zh.* **15**, 377.

SCHWARZSCHILD, K.: 1916, *Sitzber. Deut. Akad. Wiss.* Berlin, K1. Math.-Phys. Tech., s. 189.

SCHWARZSHILD, M.: 1958, *Structure and Evolution of the Stars*, Princeton, N. J.: Princeton University Press.

SCHWINGER, J. S.: 1951, *Phys. Rev.* **82**, 664.

SCHWINGER, J. S.: 1954, *Phys. Rev.* **94**, 1362.

SCIAMA, D. W.: 1976, *Vistas Astron.* **19**, 385.

SCIAMA, D. W.: 1981, in C. J. Isham, R. Penrose, and D. W. Sciama (eds.), *Quantum Gravity* 2, A Second Oxford Symposium, Oxford Univ. Press, Oxford.

SCIAMA, D. W., P. CANDELAS, and D. DEUTSCH: 1981, *Advances in Phys.* **30**, 327.

SEIDEL, E.: 1989, *Class. Quantum Grav.* **6**, 1057.

SEIDEL, E.: 1991, *Phys. Rev.* **D44**, 950.

SEIDEL, E. and S. IYER: 1990, *Phys. Rev.* **41**, 374.

SEIDEL, E. and W.-M. SUEN: 1992, *Phys. Rev. Lett.* **69**, 1845.

SEMERÁK, O.: 1994, *Astron. Astrophys.* **291**, 679.

SEN, A.: 1991, *Phys. Lett.* **B271**, 295.

SEN, A.: 1992a, *Phys. Lett.* **B274**, 34.

SEN, A.: 1992b, *Phys. Rev. Lett.* **69**, 1006.

SEXL, P. U.: 1975, *Acta Phys. Austriaca* **B42**, 303.

SHAKURA, N. I.: 1972, *Astr. Zh.* **16**, 756.

SHAKURA, N. I. and R. A. SUNYAEV: 1973, *Astron. Astrophys.* **24**, 337.

SHAPIRO, S. L.: 1973a, *Astrophys. J.* **180**, 531.

SHAPIRO, S. L.: 1973b, *Astrophys. J.* **185**, 69.

SHAPIRO, S. L., A. P. LIGHTMAN, and D. M. EARDLEY: 1976, *Astrophys. J.*, **204**, 187.

SHAPIRO, S. L. and S. A. TEUKOLSKY: 1983, *Black Holes, White Dwarfs, and Neutron Stars. The Physics of Compact Objects.*, Wiley-Interscience, New York.

SHAPIRO, S. and S. TEUKOLSKY: 1992, *Phys. Rev.* **D45**, 2739.

SHAPIRO, S. L., S. A. TEUKOLSKY, and J. WINICOUR: 1995, *Phys. Rev.* **D52**, 6982.

SHAPIRO, S. L. and I. WASSERMAN: 1982, *Astrophys. J.* **260**, 838.

SHARP, N. A.: 1979, *Gen. Rel. & Grav.* **10**, 659.

SHELLARD, E. P. S. and A. VILENKIN: 1994, Cosmic Strings., Cambridge Univ. Press, Cambridge.

SHIBATA, M.: 1993a, *Phys. Rev.* **D48**, 663.

SHIBATA, M.: 1993b, *Prog. Theor. Phys.* **90**, 595.

SHIBATA, M.: 1994, *Phys. Rev.* **D50**, 6297.

SHIBATA, M., M. SASAKI, H. TAGOSHI, and T. TANAKA: 1995, *Phys. Rev.* **D51**, 1646.

SHKLOVSKY, I. S.: 1967a, *Astrophys. J.* **148**, L1.

SHKLOVSKY, I. S.: 1967b, *Astron. Zh.* **44**, 930.

SHLOSSMAN, I., M. BEGELMAN, and J. FRANK: 1990, Nature **345**, 679.

SHVARTSMAN, V. F.: 1971, *Zh. Eksp. Teor. Fiz.* **60**, 881.

SIBGATULLIN, N. R.: 1973, *Dokl. Akad. Nauk USSR* **209**, 815.

SIBGATULLIN, N. R.: 1974, *Zh. Eksp. Teor. Fiz.* **66**, 1187.

SIBGATULLIN, N. R.: 1984, *Oscillations and Waves in Strong Gravitational and Electromagnetic Fields*, Nauka, Moscow.

SIBGATULLIN, N. R. and G. A. ALEKSEEV: 1974, *Zh. Eksp. Teor. Fiz.* **67**, 1233.

SIMON, W.: 1984, *J. Math. Phys.* **25**, 1035.

SIMON, W.: 1985, *Gen. Rel. & Grav.* **17**, 761.

SIMONE, L. E. and C. M. WILL: 1992, *Class. Quantum Grav.* **9** 963.

SKYRME, T. H. R.: 1961, *Proc. R. Soc. London* **A260**, 127.

SMARR, L.: 1973a, *Phys. Rev. Lett.* **30**, 71.

SMARR, L.: 1973b, *Phys. Rev.* **D7**, 289.

SMARR, L.: 1977, *Phys. Rev.* **D15**, 2069.

SMARR, L.: 1979, in L. Smart (ed.), *Sources of Gravitational Radiation*, Cambridge Univ. Press, Cambridge, p.245.

SMARR, L., A. ČADEŽ, B. DEWITT, and K. EPPLEY: 1976, *Phys. Rev.* **D14**, 2443.

SMARR, L. and J. W. YORK JR.: 1978, *Phys. Rev.* **D17**, 2529.

SMITH, A. G., and C. M. WILL: 1980, *Phys. Rev.* **D22**, 1276.

SOKOLOV, D. D. and A. A. STAROBINSKY: 1977, *Soviet Phys. Doclady* **22**, 312.

SOKOLOWSKI, L. M. and P. MAZUR: 1980, *Phys. Rev.* **D13**, 1113.

SREDNICKI, M.: 1993, *Phys. Rev. Lett.* **71**, 666.

SRINIVASAN, G.: 1997, in the book: S. D. Kawaler, I. Novikov, G. Srinivasan, *Stellar Remnants*, Saas-Fee Advanced Course 25, Springer, page 97.

STARK, R. F. and T. PIRAN: 1985, *Phys. Rev. Lett* **55**, 891.

STAROBINSKY, A. A.: 1973, *Zh. Eksp. Teor. Fiz.* **64**, 48 (*Sov. Phys. JETP* **37**, 28).

STAROBINSKY, A. A. and S. M. CHURILOV: 1973, *Zh. Eksp. Teor. Fiz.* **65**, 3.

STEINMÜLLER, B., A. R. KING, and J. P. LASOTA: 1975, *Phys. Lett.* **A51**, 191.

STEWART, J. M.: 1975, *Proc. R. Soc. London* **A344**, 65.

STEWART, J. M. and M. WALKER: 1973, in *Springer Tracts in Modern Physics*, **69**, Springer-Verlag, Heidelberg, p. 69.
STROMINGER, A.: 1995, *Les Houches Lectures on Black Holes*, E-print hep-th/9501071
STROMINGER, A. and C. VAFA: 1996, *Phys. Lett.* **B379**, 99.
STURROCK, P.: 1971, *Astrophys. J.* **164**, 529.
SUDARSKY, D.: 1995, *Class. Quantum Grav.* **12**, 579.
SUDARSKY, D. and R. M. WALD: 1992, *Phys. Rev.* **D46**, 1453.
SULLIVAN, B. T. and W. ISRAEL: 1980, *Phys. Lett* **79A**, 371
SUN, Y. and R. H. PRICE: 1988, *Phys. Rev.* **D38**, 1040.
SUSHKOV, S. V.: 1995, *Class. Quantum Grav.* **12**, 1685.
SUSSKIND, L. and J. UGLUM: 1994, *Phys. Rev.* **D50**, 2700.
SVENSSON, R.: 1994, *Astrophys. J.* Supplement Series, **92**, 585.
SYNGE, J. L.: 1950, *Proc. Roy. Irish. Acad.* **A53**, 83.
SYNGE, J. L.: 1960, *Relativity: The General Theory*, North-Holland, Amsterdam.
SZEKERES, G.: 1960, *Publ. Mat. Debrecent* **7**, 285.
SZILLARD, L.: 1929, *Z. Phys.* **53**, 840.
'T HOOFT, G.: 1974, *Nucl. Phys.* **B79**, 276.
'T HOOFT, G.: 1985, *Nucl. Phys.* **B256**, 727.
'T HOOFT, G.: 1988, *Commun. Math. Phys.* **117**, 685.
TACHIZAWA, T., K. MAEDA, and T. TORII: 1995, *Phys. Rev.* **D51**, 4054.
TAGOSHI, H.: 1995, *Progr. Theor. Phys.* **93** 307.
TAGOSHI, H. and T. NAKAMURA: 1994, *Phys. Rev.* **D49**, 4016.
TAGOSHI, H. and M. SASAKI: 1994, *Prog. Theor. Phys* **92**, 745.
TAKAHACHI, M.: 1991, *Publ. Astron. Soc. Japan.* **43**, 569.
TAKAHASHI, M., S. NITTA, Y. TATEMATSU, and A. TOMIMATSU: 1990, *Astrophys. J.* **386**, 455.
TAMBURINO, L. A. and J. H. WINICOUR: 1966, *Phys. Rev.* **150**, 1039.
TANAKA, T., M. SHIBATA, M. SASAKI, H. TAGOSHI, and T. NAKAMURA: 1993, *Prog. Theor. Phys* **90**, 65.
TANGHERLINI, F. R.: 1963, *Nuovo Cim.* **77**, 636.
TASHIRO, Y. and H. EZAWA: 1981, *Prog. Theor. Phys* **66**, 1612.
TAYLOR, J. H.: 1994, *Rev. Mod. Phys.* **66**, 711.
TEITELBOIM, C.: 1972a, *Nuovo Cim.* **113**, 397.
TEITELBOIM, C.: 1972b, *Phys. Rev.* **D5**, 2941.
TEITELBOIM, C.: 1972c, *Nuovo Cim.* **113**, 326.
TERNOV, I. M., V. R. KHALILOV, and G. A. CHIZHOV: 1975, *Zh. Eksp. Teor. Fiz.* **68**, 377.
TERNOV, I. M., V. R. KHALILOV, G. A. CHIZHOV, and A. B. GAINA: 1978, *Izv. Vuzov, Fizika*, No. 9, 109.
TEUKOLSKY, S. A.: 1972, *Phys. Rev. Lett.* **29**, 1114.
TEUKOLSKY, S. A.: 1973, *Astrophys. J.* **185**, 635.
TEUKOLSKY, S. A. and W. H. PRESS: 1974, *Astrophys. J.* **193**, 443.

THIRRING, H. and J. LENSE: 1918, *Phys. Z.* **19**, 156.

THORNE, K. S.: 1972, in J. Klauder (ed.), *Magic without Magic: John Archibald Wheeler*, W. H. Freeman, San Francisco, p. 231.

THORNE, K. S.: 1976, *General Relativity Astrophysics*, Preprint, Caltech, OAP-462.

THORNE, K. S.: 1978, in N. R. Lebovitz, W. H. Reid and P. O. Vandervoort (eds.) *Theoretical Principles in Astrophysics and Relativity* (Univ. Chicago Press), p. 149

THORNE, K. S.: 1985, *Black Holes: The Membrane Viewpoint*, Preprint, Caltech, GRP-031.

THORNE, K. S.: 1986, in S. L. Shapiro and S. A. Teukolsky (eds.), *Highlights of Modern Astrophysics*, Wiley, N. Y.

THORNE, K. S.: 1987, in S. W. Hawking and W. Israel (eds.), *300 Years of Gravitation*, Cambridge University Press, p. 330.

THORNE, K. S.: 1993, in R. J. Gleiser, C. N. Kozameh, and O. M. Moreschi (eds.), *GR13: General Relativity and Gravitation 1992 – Proceedings of the 13th International Conference on General Relativity and Gravitation, Cordoba, Argentina, 1992*, Bristol Institute of Physics, p.295.

THORNE, K. S.: 1994a, in M. Sasaki (ed.), *Relativistic Cosmology*, Proceedings of the Eighth Nishinomiya-Yukawa Memorial Symposium, Universal Academy Press, Kyoto, p. 67.

THORNE, K. S.: 1994b, *Black holes and time warps: Einstein's outrageous legacy*, W. W. Norton and Co., New York.

THORNE, K.: 1995, in: E. W. Kolb and R. D. Peccei (eds.), Particles and Nuclear Astrophysics and Cosmology in the Next Millenium, Proceedings of the 1994 Snowmass Summer Study, p.160.

THORNE, K. S.: 1997a, E-print gr-qc/9704042.

THORNE, K. S.: 1997b, E-print gr-qc/9706079.

THORNE, K. S.: 1997c, in B. Jones and D. Marković (eds.), *Relativistic Astrophysics*, Cambridge Univ. Press, p.259.

THORNE, K. S. and R. D. BLANDFORD: 1982, in D. Heeschen and C. Wade (eds.), *Extragalactic Radio Sources*, p. 255.

THORNE, K. S. and J. B. HARTLE: 1985, *Phys. Rev.* **D31**, 1815.

THORNE, K. S. and D. A. MACDONALD: 1982, *Mon. Not. RAS* **198**, 339.

THORNE, K. S. and R. H. PRICE: 1975, *Astrophys. J.* **195**, L101.

THORNE, K. S., R. H. PRICE, and D. A. MACDONALD: 1986, *Black Holes: The Membrane Paradigm*, Yale Univ. Press, New Haven.

TIPLER, F. J.: 1978a, *Phys. Rev.* **D17**, 2521.

TIPLER, F. J.: 1978b, *J. Diff. Equ.* **30**, 165.

TIPLER, F. J., C. J. S. CLARKE, and G. F. R. ELLIS: 1980, in A. Held (ed.), *General Relativity and Gravitation: One Hundred Years After the Birth of Albert Einstein*, Plenum Press, New York, vol.2.

TOD, K. P.: 1992, *Class. Quantum Grav.* **9**, 1581.

TOLMAN, R. G.: 1934, *Proc. Nat. Acad. Sci. US* **20**, 169.

TOMBOULIS, E.: 1980, *Phys. Lett.* **B97**, 77.

TOMIMATSU, A.: 1983, *Progr. Theor. Phys.* **70**, 385.

TOMIMATSU, A.: 1984, *Phys. Lett.* **A103**, 374.

TOMIMATSU, A. and M. KIHARA: 1982, *Progr. Theor. Phys.* **67**, 1406.

TORII, T. and K. MAEDA: 1993, *Phys. Rev.* **D48**, 1643.

TORII, T., K. MAEDA, and T. TACHIZAWA: 1995, *Phys. Rev.* **D51**, 1510.

TREMAINE, S.: 1995, *Astronom. J.* **110**, 628.

TREVES, A., T. BELLONI, R. H. D. CORBET, K. EBISAWA, R. FALOMO *et al.*: 1990, *Astrophys. J.* **364**, 266.

TUTUKOV, A. V. and A. M. CHEREPASHCHUK: 1993, *Astr. Zh.* **70**, 307.

UMEZAWA, H., H. MATSUMOTO, and Y. TAKAHASHI: 1981, *Thermo Field Dynamics and Condensed States* (North-Holland, Amsterdam).

UMEZAWA, H. and Y. TAKAHASHI: 1975, *Collective Phenomena* **2**, 55.

UNRUH, W. G.: 1973, *Phys. Rev. Lett.* **31**, 1265.

UNRUH, W. G.: 1974, *Phys. Rev.* **D10**, 3194.

UNRUH, W. G.: 1976a, *Proc. R. Soc. London* **A348**, 447.

UNRUH, W. G.: 1976b, *Phys. Rev.* **D14**, 870.

UNRUH, W. G.: 1976c, *Phys. Rev.* **D14** 3251.

UNRUH, W. G.: 1977, *Phys. Rev.* **D15**, 365.

UNRUH, W. G.: 1981, *Phys. Rev. Lett.* **46**, 1351.

UNRUH, W. G.: 1995, *Phys. Rev.* **D51**, 2827.

UNRUH, W. G. and R. M. WALD: 1982, *Phys. Rev.* **D25**, 942.

UNRUH, W. G. and R. M. WALD: 1983a, *Phys. Rev.* **D27**, 2271.

UNRUH, W. G. and R. M. WALD: 1983b, *Gen. Rel. & Grav.* **15**, 195.

UNRUH, W. G. and R. M. WALD: 1984, *Phys. Rev.* **D29**, 1047.

UNRUH, W. G. and R. M. WALD: 1995, *Phys. Rev.* **D52**, 2176.

VAIDYA, P. C.: 1951, *Proc.of the Indian Acad. Sci.* **A33**, 264.

VAIDYA, P. C.: 1953, *Nature* **171**, 260.

VALTONEN, M. J.: 1996, *Mon. Not. RAS* **278**, 186.

VAN DEN HEUVEL, E. P. J.: 1983, in W. H. G. Lewin and E. P. J. van den Heuvel (eds.), *Accretion-Driven Stellar X-ray Sources*, Cambridge University Press, Cambridge, p.303.

VAN DEN HEUVEL, E. P. J. and G. M. H. J. HABETS: 1984, *Nature* **309**, 598.

VAN DER MAREL, R. P.: 1995, in I. Appenzeller (ed.), *Highlights of Astronomy* **10**, p.527.

VAN NIEUWENHUIZEN, P., D. WILKINSON, and M. J. PERRY: 1976, *Phys. Rev.* **D13**, 778.

VAN STOCKUM, W. J.: 1937, *Proc. Roy. Soc. Edin.* **57**, 135.

VENEZIANO, G.: 1993, in M. C. Bento, O. Bertolami, J. M. Mourão, and R. F. Picken (eds.), *Classical and Quantum Gravity*, Proceedings of the First Iberian Meeting on Gravity, World Scientific, p.134.

VILENKIN, A.: 1979a, *Phys. Rev.* **D20**, 373.
VILENKIN, A.: 1979b, *Phys. Rev.* **D20**, 1807.
VILENKIN, A.: 1982, *Nucl. Phys.* **B196**, 240.
VISHVESHWARA, C. V.: 1968, *J. Math. Phys.* **9**, 1319.
VISHVESHWARA, C. V.: 1970a, *Nature* **227**, 936.
VISHVESHWARA, C. V.: 1970b, *Phys. Rev.* **D1**, 2870.
VISSER, M.: 1995, *Lorentzian Wormholes: From Einstein to Hawking* (American Inst. of Physics Press).
VISSER, M.: 1997, *Phys. Rev.* **D55**, 5212.
VLADIMIROV, YU. S.: 1982, *Frames of Reference in Gravitation Theory*, Energoizdat, Moscow.
VOLKOV, M. S. and D. V. GAL'TSOV: 1989, *JETP Lett.* **50**, 346.
VOLKOV, M. S. and D. V. GAL'TSOV: 1990, *Sov. J. Nucl. Phys.* **51**, 747.
VOLKOV, M. S. and N. STRAUMANN: 1997, *Phys. Rev. Lett.* **79**, 1428.
VOLOVICH, I. V., V. A. ZAGREBNOV, and V. P. FROLOV: 1976, *Teor. Mat. Fiz.* **29**, 191.
VOLOVICH, I. V., V. A. ZAGREBNOV, and V. P. FROLOV: 1978, *Sov. J. Part. Nucl.* **9**, 147.
WALD, R. M.: 1972, *Phys. Rev.* **D6**, 406.
WALD, R. M.: 1973, *J. Math. Phys* **14**, 1453.
WALD, R. M.: 1974a, *Ann. Phys.* **82**, 548.
WALD, R. M.: 1974b, *Phys. Rev.* **D10**, 1680.
WALD, R. M.: 1974c, *Astrophys. J.* **191**, 231.
WALD, R. M.: 1975, *Commun. Math. Phys.* **45**, 9.
WALD, R. M.: 1977, *Commun. Math. Phys.* **54**, 1.
WALD, R. M.: 1978a, *Phys. Rev.* **D17**, 1477.
WALD, R. M.: 1978b, *Ann. Phys.* **110**, 472.
WALD, R. M.: 1978c, *Phys. Rev. Lett.* **41**, 203.
WALD, R. M.: 1979a, *J. Math. Phys.* **20**, 1056.
WALD, R. M.: 1979b, *Phys. Rev.* **D20**, 1271.
WALD, R. M.: 1980, *J. Math. Phys.* **21**, 218.
WALD, R. M.: 1984, *General Relativity*, Univ. of Chicago Press, Chicago.
WALD, R. M.: 1990, *J. Math. Phys.* **31**, 725.
WALD, R. M.: 1992, in V. De Sabbata and Z. Zhang (eds.), *Black Hole Physics*, NATO ASI Series, Kluwer Academic Publ., p.55.
WALD, R. M.: 1993, *Phys. Rev.* **D48**, R3427.
WALD, R.M.: 1994, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, Univ. of Chicago Press, Chicago.
WALD, R. M.: 1997, *Phys. Rev.* **D56**, 6467.
WALD, R. M. and U. YURTSEVER: 1991, *Phys. Rev.* **D44**, 403.
WALKER, M.: 1970, *J. Math. Phys.* **11**, 2280.
WALKER, M. and R. PENROSE: 1970, *Commun. Math. Phys.* **18**, 265.

WALLINDER, F. H.: 1993, *Comments on Astrophys.* **16**, 331.

WANDELL, A., A. YAHIL, and M. MILGROM: 1984, *Astrophys. J.* **282**, 53.

WARNER, N. P.: 1982, *Commun. Math. Phys.* **86**, 419.

WAUGH, B. and K. LAKE: 1988, *Phys. Rev.* **D38**, 1315.

WEBER, E. J. and L. DAVIS: 1967, *Astrophys. J.* **148**, 217.

WEBSTER, B. L. and P. MURDIN: 1972, *Nature* **235**, 37-38.

WEYL, H.: 1917, *Ann. Physik* **54**, 117.

WHEELER, J. A.: 1955, *Phys. Rev.* **97**, 511.

WHEELER, J. A.: 1962, *Geometrodynamics*, Academic Press, New York.

WHEELER, J. A.: 1968, *Am. Sci.* **59**, **1**.

WHEELER, J. C.: 1997, in B. Jones and D. Marković (eds.), *Relativistic Astrophysics*, Cambridge Univ. Press, p.211.

WHEELER, J. C., SOON-WOOK KIM, and M. MOSCOSO: 1993, in *"Cataclysmic Variables and Related Physics"*, 2nd Technion Haifa Conference, Eilat, Israel. Annals of the Israel Physical Society, **10**, p.180.

WHITING, B. F.: 1989, J. Math. Phys. **30**, 1301.

WHITING, B. F. and J. W. YORK, JR.: 1988, *Phys. Rev. Lett.* **61**, 1336.

WHITT, B.: 1988, *Phys. Rev.* **D38**, 3000.

WILD, W. J. and R. M. KERNS: 1980, *Phys. Rev.* **D21**, 332.

WILD, W. J., R. M. KERNS, and W. J. DRISH JR.: 1981, *Phys. Rev.* **D23**, 829.

WILKINS, D.: 1972, *Phys. Rev.* **D15**, 814.

WILL, C. M.: 1994, in M. Sasaki (ed.), *Relativistic Cosmology*, Proceedings of the Eighth Nishinomiya-Yukawa Memorial Symposium, Universal Academy Press, Kyoto, p. 83.

WILSON, A. H.: 1957, *Thermodynamics and Statistical Mechanics*, Cambridge Univ. Press, Cambridge, Chap.7.

WILTSHIRE, D. L.: 1986, *Phys. Lett.* **B169**, 36.

WILTSHIRE, D. L.: 1988, *Phys. Rev.* **D38**, 2445.

WITTEN, E.: 1988, *Nucl. Phys.* **B311**, 46.

WITTEN, E.: 1989, *Nucl. Phys.* **B323**, 113.

WITTEN, E.: 1991, *Phys. Rev.* **D44**, 314.

XANTHOPOULOS, B. C.: 1983, *Proc. R. Soc. London* **A388**, 117.

YAMAZAKI, M.: 1983a, *Phys. Rev. Lett.* **50**, 1027.

YAMAZAKI, M.: 1983b, *Progr. Theor. Phys.* **69**, 503.

YANG, P. S.: 1979, *Phys. Rev.* **D20**, 834.

YASSKIN, P. B.: 1975, *Phys. Rev.* **D12**, 2212.

YI, P.: 1995, *Phys. Rev.* **51**, 2813.

YODZIS, P., H. J. SEIFERT, and H. MÜLLER ZUM HAGEN: 1973, *Cammun. Math. Phys.* **34**, 135.

YODZIS, P., H. J. SEIFERT, and H. MÜLLER ZUM HAGEN: 1974, *Commun. Math. Phys.* **37**, 29.

YORK, J. W., JR.: 1983, *Phys. Rev.* **D28**, 2929.

YORK, J. W., JR.: 1985, *Phys. Rev.* **D31**, 775.

YORK, J. W., JR.: 1986, *Phys. Rev.* **D33**, 2092.

YORK, J. W., JR.: 1989, in C. Evans, L. Finn, and D. Hobill (eds.), *Frontiers in Numerical Relativity*, Cambridge Univ. Press, Cambridge, England.

YOUNG, P. J.: 1976, *Phys. Rev.* **D14**, 3281.

YURTSEVER, U.: 1988, *Phys. Rev.* **D38**, 1731.

YURTSEVER, U.: 1993, *Class. Quantum Grav.* **10**, L17.

ZAKHAROV, A. F.: 1986, *Sov. Phys. JEPT* **64**, 1.

ZAKHAROV, A. F.: 1989, *Sov. Phys. JEPT* **68**, 217.

ZANNIAS, T.: 1984, *Phys. Rev.* **D30**, 1161.

ZASLAVSKII, O. B.: 1991, *Phys. Rev.* **D43**, 605.

ZAUMEN, W. T.: 1974, Nature **247**, 530.

ZEL'DOVICH, YA. B.: 1962a, *Zh. Eksp. Teor. Fiz.* **42**, 641.

ZEL'DOVICH, YA. B.: 1962b, *Zh. Eksp. Teor. Fiz.* **42**, 1667.

ZEL'DOVICH, YA. B.: 1962c, *Zh. Eksp. Teor. Fiz.* **43**, 1037.

ZEL'DOVICH, YA. B.: 1964, *Sov. Phys. Doklady* **9**, 195.

ZEL'DOVICH, YA. B.: 1970, *Pis'ma v Zh. Eksp. Teor. Fiz.* **12**, 443.

ZEL'DOVICH, YA. B.: 1971, *Sov. Phys. JETP Lett.* **14**, 180.

ZEL'DOVICH, YA. B.: 1972, *Sov. Phys. JETP* **35**, 1085.

ZEL'DOVICH, YA. B.: 1976, *Pis'ma v Zh. Eksp. Teor. Fiz.* **24**, 29.

ZEL'DOVICH, YA. B.: 1977a, *Sov. Phys. JETP* **45**, 9.

ZEL'DOVICH, YA. B.: 1977b, *Sov. Phys. Usp.* **20**, 945.

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1964a, *Dokl. Akad. Nauk USSR* **155**, 1033.

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1964b, *Sov. Phys. Doklady* **158**, 811.

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1966, *Astron. Zh.* **43**, 758.

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1967a, *Astr. Zh.* **10**, 602.

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1967b, *Relativistic Astrophysics*, Nauka, Moscow (in Russian).

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1971a,*Theory of Gravitatian and Stellar Evolution*, Nauka, Moscow (in Russian).

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1971b, *Relativistic Astrophysics*, **1**, *Stars and Relativity*, Univ. of Chicago Press, Chicago.

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1975, *The Structure and Evolution of the Universe*, Nauka, Moscow (in Russian).

ZEL'DOVICH, YA. B. and I. D. NOVIKOV: 1983, *Relativistic Astrophysics*, **2**, *The Structure and Evolution of the Universe*, Univ. of Chicago Press, Chicago.

ZEL'DOVICH, YA. B., I. D. NOVIKOV, and A. A. STAROBINSKY: 1974, *Zh. Eksp. Teor. Fiz.* **66**, 1897.

ZEL'DOVICH, YA. B. and A. A. STAROBINSKY: 1971, *Zh. Eksp. Teor. Fiz.* **61**, 2161.

ZEL'MANOV, A. L.: 1956, *Dokl. Akad. Nauk USSR* **107**, 815.

ZEL'NIKOV, A. I. and V. P. FROLOV: 1982, *Zh. Eksp. Teor. Fiz.* **82**, 321.

ZERILLI, F. J.: 1970a, *Phys. Rev.* **D2**, 2141.
ZERILLI, F. J.: 1970b, *Phys. Rev. Lett.* **24**, 737.
ZERILLI, F. J.: 1974, *Phys. Rev.* **D9**, 860.
ZHANG, T. R. and C. DEWITT-MORETTE: 1984, *Phys. Rev. Lett.* **52**, 2313.
ZHUK, A. I. and V. P. FROLOV: 1981, *Kratk. Soobshch. po Fiz.* No. **9**, 25.
ZNAJEK, R. L.: 1978, *Mon. Nat. RAS* **185**, 833.
ZOUROS, T. J. M. and D. M. EARDLEY: 1979, *Ann. Phys.* **118**, 139.
ZUREK, W. H.: 1982, *Phys. Rev. Lett.* **49**, 1683.
ZUREK, W. H. and K. S. THORNE: 1985, *Phys. Rev. Lett.* **54**, 2171.
ZWICKY, F.: 1958, *Handbuch der Physik* **51**, 766.

Index

Index