DUALITY, MATHEMATICS

The term *duality* used in our daily life means the sort of harmony of two opposite or *complementary* parts by which they integrate into a whole. Inner beauty in natural phenomena is bound up with duality, which has always been a rich source of inspiration in human knowledge through the centuries. The theory of duality is a vast subject, significant in art and natural science. Mathematics lies at its root. By using abstract languages, a common mathematical structure can be found in
3 many physical theories. This structure is independent of the physical contents of the system and exists in wider classes of problems in engineering and sciences (see Ref. 1).

According to Tonti (2), for every physical theory we can identify (a) some *configuration variables* that describe the state of the system and (b) some *source variables* that describe the source of the phenomenon. The displacement vector in continuum mechanics and the scalar potential in electrostatics are examples of configuration variables. Forces and charges are examples of source variables. Besides these two types of quantities, there are also some paired (i.e., one-to-one) *intermediate variables* that describe the internal (or constitutive) properties of the system, such as velocity and momentum in dynamics, electrical field intensity and the flux density in electrostatics, and the two electromagnetic tensors in electromagnetism.

Let $\mathcal{U}$ and $\mathcal{U}^*$ be, respectively, the real vector spaces of configuration variables and source variables, and let $\mathcal{V}$ and $\mathcal{V}^*$ denote the paired intermediate variable spaces. By introducing a so-called *geometric operator* $\Lambda : \mathcal{U} \to \mathcal{V}$ and an equilibrium operator $B : \mathcal{V}^* \to \mathcal{U}^*$, such that the duality relation between $\mathcal{V}$ and $\mathcal{V}^*$ is a one-to-one mapping, the *primal system* $\mathcal{S}_p := \{\mathcal{U}, \mathcal{V}; \Lambda\}$ and the *dual system* $\mathcal{S}_d := \{\mathcal{U}^*, \mathcal{V}^*; B\}$ are linked into a whole system $\mathcal{S} = \mathcal{S}_p \cup \mathcal{S}_d$. The system is called *geometrically linear* if Λ is linear. In this case, B is the adjoint operator of Λ. If Λ is an $m \times n$ matrix, then $\mathcal{S}$ is a finite-dimensional *algebraic system*. Optimization in such systems is known as *mathematical programming*. If Λ is a continuous (partial) differential operator, then $\mathcal{S}$ is an infinite-dimensional (*partial*) *differential system*, and optimization problems fall into the *calculus of variations*. It is shown in Ref. 3 that under certain conditions, if there is a theorem in the primal system $\mathcal{S}_p$, then in the dual system $\mathcal{S}_d$ there exists a complementary theorem and vice versa. If there is a theorem defined on the whole system $\mathcal{S}$, then exchanging the dual elements in this theorem leads to another parallel theorem. Generally speaking, the *theory of duality* is the study of the intrinsic relations between the primal system $\mathcal{S}_p$ and the dual system $\mathcal{S}_d$.

In the theory of optimization, let $P : \mathcal{U} \to \mathbb{R}$ and $P^* : \mathcal{V}^* \to \mathbb{R}$ be real-valued functions. If $P(\boldsymbol{u}) \geq P^*(\boldsymbol{v}^*)$ for all vectors $(\boldsymbol{u}, \boldsymbol{v}^*)$ in the *Cartesian product space* $\mathcal{U} \times \mathcal{V}^*$, then an infimum of P and a supremum of P^* exist and $\inf P(\boldsymbol{u}) \geq \sup P^*(\boldsymbol{v}^*)$. Under certain conditions we have $\inf P(\boldsymbol{u}) = \sup P^*(\boldsymbol{v}^*)$. A statement of this type is called a *duality theorem.*

Let $L(\boldsymbol{u}, \boldsymbol{v}^*): \mathscr{U} \times \mathscr{V}^* \to \mathbb{R}$ be a so-called *Lagrangian form.* Under certain conditions we have $\inf_u \sup_{v*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \sup_{v*} \inf_u L(\boldsymbol{u}, \boldsymbol{v}^*)$. Such a statement is called a *minimax theorem.* In convex optimization, the minimax theorem is connected to a *saddle-point theorem.* The main purpose of the theory of duality in mathematical optimization is to make inquiries about a corresponding pair of optimization problems, namely, (a) the primal problem to find $\bar{\boldsymbol{u}}$ such that $P(\bar{\boldsymbol{u}}) = \inf_u P(\boldsymbol{u})$ and (b) the dual problem to find $\bar{\boldsymbol{v}}^*$ such that $P^*(\bar{\boldsymbol{v}}^*) = \sup_{v*} P^*(\boldsymbol{v}^*)$ and to discover relations between corresponding duality, minimax, and theorems of critical points. In numerical analysis, the primal problem provides only upper-bound approaches to the solution. However, the dual problem will give a lower bound of solutions. The numerical methods to find the *primal–dual solution* $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ in each iteration are known as *primal–dual methods.* In finite element analysis of boundary value problems, such methods as *mixed/hybrid methods* have been studied extensively by engineers for more than 30 years. In the past decade, *primal–dual algorithms* have emerged as the most important and useful algorithms for mathematical programming (4).

Duality in natural science is amazingly beautiful. It has excellent theoretical properties, powerful practical applications, and pleasing relationships with the existing fundamental theories. In geometrically linear systems, the common mathematical structure and theorems take particularly symmetric forms. The duality theory has been well studied for both convex problems (see Refs. 5–8) and nonconvex problems (see Refs. 9 and 10). However, in *geometrically nonlinear systems,* where Λ is a nonlinear operator, such symmetry is broken. The duality theory in these systems was studied in Ref. 11. An interesting *triality theorem* in nonconvex systems has been discovered recently in Refs. 12 and 13, which can be used either to solve some nonlinear variational problems or to develop algorithms for numerical solutions in nonconvex, *nonsmooth optimization* (see Ref. 3).

FRAMEWORK AND CANONIC EQUATIONS

Let $\mathscr{U}$, $\mathscr{U}^*$ and $\mathscr{V}$, $\mathscr{V}^*$ be two pairs of *real vector spaces,* finite- or infinite-dimensional, and let $(*, *): \mathscr{U} \times \mathscr{U}^* \to \mathbb{R}$ and $\langle *, * \rangle: \mathscr{V} \times \mathscr{V}^* \to \mathbb{R}$ be certain *bilinear forms.* We say that these two bilinear forms put the paired spaces $\mathscr{U}$, $\mathscr{U}^*$ and $\mathscr{V}$, $\mathscr{V}^*$ *in duality,* respectively. Let the *geometric operator* Λ be a continuous linear transformation from $\mathscr{U}$ to $\mathscr{V}$. The *equilibrium operator* B in geometrically linear system is simply the adjoint operator $\Lambda^*: \mathscr{V}^* \to \mathscr{U}^*$ defined by

$$\langle \Lambda \boldsymbol{u}, \boldsymbol{v}^* \rangle = (\boldsymbol{u}, \Lambda^* \boldsymbol{v}^*) \qquad \forall \boldsymbol{u} \in \mathscr{U}, \boldsymbol{v}^* \in \mathscr{V}^* \tag{1}$$

Thus the two paired dual spaces $\mathscr{U}$, $\mathscr{U}^*$ and $\mathscr{V}$, $\mathscr{V}^*$ are linked, respectively, by a so-called *geometrical* (*or definition*) *equation:*

$$\boldsymbol{v} = \Lambda \boldsymbol{u} \tag{2}$$

and an *equilibrium equation:*

$$\boldsymbol{u}^* = \Lambda^* \boldsymbol{v}^* \tag{3}$$

In calculus of variations, if Λ is a gradient-like operator, its adjoint Λ^* should be a divergence-like operator, Eq. (1) is then the well-known Gauss–Green formula.

The readers who are interested primarily in the finite-dimensional case will not need knowledge of *convex analysis* in what follows. Instead, they can simply interpret $\mathscr{U} = \mathscr{U}^* = \mathbb{R}^n$, $\mathscr{V} = \mathscr{V}^* = \mathbb{R}^m$, with $(\boldsymbol{u}, \boldsymbol{u}^*)$ and $\langle \boldsymbol{v}, \boldsymbol{v}^* \rangle$ as the ordinary inner products on the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. In this case, we can of course identify Λ with a certain $m \times n$ matrix $A = \{a_{ij}\}$, and

$$\langle A\boldsymbol{u}, \boldsymbol{v}^* \rangle = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ji} u_i v_j^* = \sum_{i=1}^{n} u_i \left(\sum_{j=1}^{m} a_{ji} v_j^* \right) = (\boldsymbol{u}, A^* \boldsymbol{v}^*)$$

So the adjoint of $\Lambda = A$ is a transpose matrix $A^* = A^T$.

A subset $\mathscr{C} \subset \mathscr{U}$ is said to be a *convex set* if for any given $\theta \in [0, 1]$, we have

$$\theta \boldsymbol{u}_1 + (1 - \theta) \boldsymbol{u}_2 \in \mathscr{C} \qquad \forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathscr{C}$$

By a *convex function* $F: \mathscr{U} \to \overset{\leftrightarrow}{\mathbb{R}} := [-\infty, +\infty]$ we shall mean that for any given $\theta \in [0, 1]$, we obtain

$$F(\theta \boldsymbol{u}_1 + (1 - \theta) \boldsymbol{u}_2) \le \theta F(\boldsymbol{u}_1) + (1 - \theta) F(\boldsymbol{u}_2) \qquad \forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathscr{U} \tag{4}$$

F is *strictly convex* if the inequality is strict. The *indicator function* of a subset $\mathscr{C} \subset \mathscr{U}$ is defined by

$$\Psi_C(\boldsymbol{u}) = \begin{cases} 0 & \text{if } \boldsymbol{u} \in \mathscr{C} \\ +\infty & \text{otherwise} \end{cases} \tag{5}$$

which plays an important role in constrained optimization. This is a convex function if and only if $\mathscr{C}$ is convex.

A function F on $\mathscr{U}$ is said to be *proper* if $F(\boldsymbol{u}) > -\infty \ \forall \boldsymbol{u} \in \mathscr{U}$ and $F(\boldsymbol{u}) < +\infty$ for at least one $\boldsymbol{u}$. Conversely, given a convex function F defined on a nonempty convex set $\mathscr{C}$, one can set $\bar{F}(\boldsymbol{u}) = F(\boldsymbol{u}) + \Psi_{\mathscr{C}}(\boldsymbol{u})$. In this way one can relax the constraint $\boldsymbol{u} \in \mathscr{C}$ on F to get a proper function $F(\boldsymbol{u}) + \Psi_{\mathscr{C}}(\boldsymbol{u})$ on the whole space $\mathscr{U}$.

A function F on $\mathscr{U}$ is *lower semicontinuous* (l.s.c.) if

$$\lim_{\boldsymbol{u}_n \to \boldsymbol{u}} \inf F(\boldsymbol{u}_n) \ge F(\boldsymbol{u}) \qquad \forall \boldsymbol{u} \in \mathscr{U} \tag{6}$$

So $\Psi_{\mathscr{C}}(\boldsymbol{u})$ is l.s.c. if and only if $\mathscr{C}$ is closed. A function F is said to be *concave, upper semicontinuous* (u.s.c.) if $-F$ is convex, l.s.c. The theory of concave functions thus parallels the theory of convex functions, with only the obvious and dual changes.

If F is finite on $\mathscr{C}$, the *Gâteaux variation* of F at $\boldsymbol{u} \in \mathscr{C}$ in the direction $\boldsymbol{v}$ is defined as

$$\delta F(\boldsymbol{u}; \boldsymbol{v}) = \lim_{\theta \to 0^+} \frac{F(\boldsymbol{u} + \theta \boldsymbol{v}) - F(\boldsymbol{u})}{\theta} \tag{7}$$

$F(\boldsymbol{u})$ is said to be *Gâteaux* (*or G-*) *differentiable* at $\boldsymbol{u}$ if $\delta F(\boldsymbol{u}; \boldsymbol{v}) = (DF(\boldsymbol{u}), \boldsymbol{v})$, where $DF: \mathscr{C} \subset \mathscr{U} \to \mathscr{U}^*$ is called the *Gâteaux derivative* of F. In finite-dimensional space, the Gâteaux variation is simply the directional derivative, and $DF = \nabla F$.

Let F and V be two real-valued functions. Throughout this article we assume that F and V are (a) convex or concave and (b) G-differentiable on the convex sets $\mathscr{C} \subset \mathscr{U}$ and $\mathscr{D} \subset \mathscr{V}$, respectively. Then the two *duality equations* between the paired spaces $\mathscr{U}$, $\mathscr{U}^*$ and $\mathscr{V}$, $\mathscr{V}^*$ can be given by

$$\boldsymbol{u}^* = DF(\boldsymbol{u}), \qquad \boldsymbol{v}^* = DV(\boldsymbol{v}) \tag{8}$$

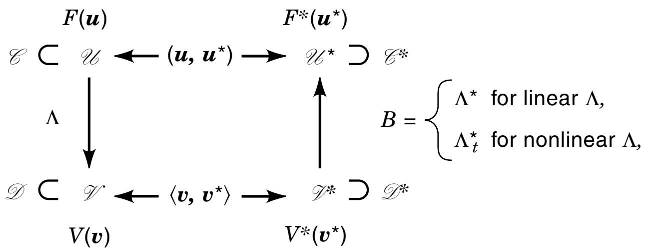

Figure 1. Framework in fully nonlinear systems.

In mathematical physics, the duality equation $\boldsymbol{v}^* = DV(\boldsymbol{v})$ is known as the *constitutive equation.* However, the duality equation $\boldsymbol{u}^* = DF(\boldsymbol{u})$ usually gives *natural boundary conditions* in variational boundary value problems.

Let $\mathscr{U}_a = \{\boldsymbol{u} \in \mathscr{U} |\ \boldsymbol{u} \in \mathscr{C},\ \Lambda\boldsymbol{u} \in \mathscr{D}\}$ be a so-called *feasible set.* On $\mathscr{U}_a$, the three types of canonical equations, Eqs. (2), (3) and (8), can be written in a so-called *fundamental equation:*

$$\Lambda^* DV(\Lambda\boldsymbol{u}) = DF(\boldsymbol{u}) \tag{9}$$

The system is called *physically linear* if both duality equations are linear. The system is called *geometrically linear* if the geometric operator $\Lambda : \mathscr{U} \to \mathscr{V}$ is linear. By the term *linear system* we mean that it is both geometrically and physically linear. In this case, if, for a given $\bar{\boldsymbol{u}}^* \in \mathscr{U}^*$, $F(\boldsymbol{u}) = (\boldsymbol{u}, \bar{\boldsymbol{u}}^*)$ is linear and $V(\boldsymbol{v}) = \frac{1}{2}\langle C\boldsymbol{v}, \boldsymbol{v}\rangle$ is quadratic, where $C : \mathscr{V} \to \mathscr{V}^*$ is a linear operator, then the fundamental equation can be written as

$$\Lambda^* C \Lambda \boldsymbol{u} = \bar{\boldsymbol{u}}^* \tag{10}$$

If C is symmetric, then the operator $K = \Lambda^* C \Lambda : \mathscr{U} \to \mathscr{U}^*$ is *self-adjoint* $K = K^*$. In *partial differential systems,* K is an *elliptic operator* if C is either *positive or negative definite,* whereas K is hyperbolic if C is *nonsingular and indefinite.*

The common mathematical structure in geometrically linear systems is shown in Fig. 1. In the textbook by Strang (1), this nice symmetrical structure can be seen from continuous theories to discrete systems. However, the symmetry in this structure is broken in geometrically nonlinear systems where Λ is a nonlinear operator. If we assume that $\boldsymbol{v} = \Lambda(\boldsymbol{u})$ is Gâteaux differentiable, then it can be split as

$$\Lambda = \Lambda_t + \Lambda_n \tag{11}$$

where Λ_t is the G-derivative of Λ and $\Lambda_n = \Lambda - \Lambda_t$, both of them depending on $\boldsymbol{u}$ (see Ref. 11). For a given $\bar{\boldsymbol{u}}^* \in \mathscr{U}^*$, the *virtual work principle* gives

$$\langle \delta\boldsymbol{v}(\bar{\boldsymbol{u}}; \boldsymbol{u}), \boldsymbol{v}^*\rangle = (\boldsymbol{u}, \Lambda_t^*(\bar{\boldsymbol{u}})\boldsymbol{v}^*) = (\boldsymbol{u}, \bar{\boldsymbol{u}}^*) \qquad \forall \boldsymbol{u} \in \mathscr{U} \tag{12}$$

In this case, the equilibrium operator $B = \Lambda_t^* : \mathscr{V}^* \to \mathscr{U}^*$ is the adjoint of Λ_t, which depends on the configuration variable. Then the equilibrium equation in geometrically nonlinear systems should be

$$\Lambda_t^*(\bar{\boldsymbol{u}})\boldsymbol{v}^* = \bar{\boldsymbol{u}}^* \tag{13}$$

The relation between the two bilinear forms is then

$$\langle \Lambda(\boldsymbol{u}), \boldsymbol{v}^*\rangle = (\boldsymbol{u}, \Lambda_t^*(\boldsymbol{u})\boldsymbol{v}^*) - G(\boldsymbol{u}, \boldsymbol{v}^*) \tag{14}$$

where $G(\boldsymbol{u}, \boldsymbol{v}^*)$ is the so-called *complementary gap function* introduced in Ref. 11:

$$G(\boldsymbol{u}, \boldsymbol{v}^*) = \langle -\Lambda_n(\boldsymbol{u})\boldsymbol{u}, \boldsymbol{v}^*\rangle \tag{15}$$

In geometrically nonlinear systems, this gap function recovers the duality theorems in convex optimization and plays an important role in nonconvex problems.

Example 1. Let us consider a mixed boundary value problem in electrostatics:

$$\operatorname{div}[\epsilon \operatorname{grad} \phi(x)] + \rho(x) = 0 \qquad \forall x \in \Omega \subset \mathbb{R}^n \tag{16}$$

$$\begin{aligned} \phi(x) &= 0 \qquad \forall x \in \Gamma_1 \\ \boldsymbol{n} \cdot \operatorname{grad} \phi(x) &= D_n \qquad \forall x \in \Gamma_2 \\ \Gamma_1 \cup \Gamma_2 &= \partial\Omega \end{aligned} \tag{17}$$

The configuration $\boldsymbol{u}$ is the *electrostatic potential* $\phi(x)$. The source variable is the *charge density* $\bar{\boldsymbol{u}}^* = \rho(x)$ in Ω and electric flux $\bar{\boldsymbol{u}}^* = D_n$ on Γ_2. ϵ is the *dielectric constant.* $\boldsymbol{n} \in \mathbb{R}^n$ is a unit vector normal to the boundary. Let $\Lambda = -\text{grad}$, and thus $\boldsymbol{v} = -\text{grad}\ \phi$ is the *electric field intensity,* denoted by $\boldsymbol{E}$. Let $\mathscr{D} = \mathscr{H}(\Omega; \mathbb{R}^n)$ be a *Hilbert space* with domain Ω and range $\mathbb{R}^n$, $\mathscr{C} = \{\phi \in \mathscr{H}(\Omega; \mathbb{R}) |\ \phi(x) = 0\ \forall x \in \Gamma_1\}$, and

$$F(\phi) = \int_\Omega \rho\phi \, d\Omega + \int_{\Gamma_2} \phi D_n \, d\Gamma - \Psi_{\mathscr{C}}(\phi) \tag{18}$$

$$V(\boldsymbol{E}) = \int_\Omega \frac{1}{2} \epsilon \boldsymbol{E}^T \boldsymbol{E} \, d\Omega + \Psi_{\mathscr{D}}(\boldsymbol{E}) \tag{19}$$

So on $\mathscr{C}$ and $\mathscr{D}$, F and V are finite, G-differentiable. Thus $\boldsymbol{D} = DV(\boldsymbol{E}) = \epsilon\boldsymbol{E}$ is the *electric flux density,* $\boldsymbol{u}^*(x) = DF(\phi) = \{\rho(x)\ \forall x \in \Omega,\ D_n\ \forall x \in \Gamma_2\}$. By the Gauss–Green theorem,

$$\begin{aligned} \langle \Lambda\phi, \boldsymbol{D}\rangle &= \int_\Omega (-\operatorname{grad}\phi) \cdot \boldsymbol{D} \, d\Omega \\ &= \int_\Omega \phi(\nabla \cdot \boldsymbol{D}) \, d\Omega - \oint_{\partial\Omega} \phi \boldsymbol{n} \cdot \boldsymbol{D} \, d\Gamma = (\phi, \Lambda^*\boldsymbol{D}) \end{aligned}$$

Hence, the adjoint operator Λ^* and the abstract equilibrium equation (3) are

$$\boldsymbol{u}^* = \Lambda^*\boldsymbol{D} = \begin{cases} \operatorname{div}\boldsymbol{D} = \rho & \text{in } \Omega \\ -\boldsymbol{n} \cdot \boldsymbol{D} = D_n & \text{on } \Gamma_2 \end{cases} \tag{20}$$

The fundamental equation, Eq. (10), in this problem is a Poisson equation and $K = \Lambda^* C \Lambda = -\epsilon\Delta$ is a Laplace operator for constant $\epsilon \in \mathbb{R}$.

FENCHEL–ROCKAFELLAR DUALITY

For a given function $V : \mathscr{V} \to \overleftrightarrow{\mathbb{R}}$, its *conjugate function* is defined by the following *Fenchel transformation:*

$$V^*(\boldsymbol{v}^*) = \sup_{\boldsymbol{v} \in \mathscr{V}} \{\langle \boldsymbol{v}, \boldsymbol{v}^*\rangle - V(\boldsymbol{v})\} \tag{21}$$

which is always l.s.c. and convex on $\mathscr{V}^*$. The following *Fenchel–Young inequality* holds:

$$V(\boldsymbol{v}) \geq \langle \boldsymbol{v}, \boldsymbol{v}^*\rangle - V^*(\boldsymbol{v}^*) \qquad \forall \boldsymbol{v} \in \mathscr{V}, \boldsymbol{v}^* \in \mathscr{V}^* \tag{22}$$

If V is strictly convex, G-differentiable on a convex set $\mathscr{D} \subset \mathscr{V}$, then Eq. (21) is the classical *Legendre transformation,* and the following relations are equivalent to each other:

$$\boldsymbol{v}^* = DV(\boldsymbol{v}) \Leftrightarrow \boldsymbol{v} = DV^*(\boldsymbol{v}^*) \Leftrightarrow V(\boldsymbol{v}) + V^*(\boldsymbol{v}^*) = (\boldsymbol{v}, \boldsymbol{v}^*) \quad (23)$$

In this section we assume that

$$(A1) \quad \begin{cases} V : \mathscr{V} \to \vec{\mathbb{R}} := (-\infty, +\infty] \text{ is proper, convex and l.s.c.} \\ F : \mathscr{U} \to \overleftarrow{\mathbb{R}} := [-\infty, +\infty) \text{ is proper, concave and u.s.c.} \end{cases} \quad (24)$$

The conjugate function of a concave function F is defined by

$$F^*(\boldsymbol{u}^*) = \inf_{\boldsymbol{u} \in \mathscr{U}} \{(\boldsymbol{u}, \boldsymbol{u}^*) - F(\boldsymbol{u})\} \quad (25)$$

Let $\mathscr{C} \subset \mathscr{U}$ be a nonempty convex set on which F is finite, G-differentiable, and define $\mathscr{C}^*$, $\mathscr{D}$, and $\mathscr{D}^*$ similarly for F^*, V, and V^*. Then on $\mathscr{C}^*$ and $\mathscr{D}^*$, the duality equations are invertible and

$$\boldsymbol{u} = DF^*(\boldsymbol{u}^*), \qquad \boldsymbol{v} = DV^*(\boldsymbol{v}^*) \quad (26)$$

Two extremum problems associated with the fundamental equation, Eq. (9), are

$$(\mathscr{P}_{\text{inf}}) \quad \text{minimize} \quad P(\boldsymbol{u}) = V(\Lambda \boldsymbol{u}) - F(\boldsymbol{u}) \quad \forall \boldsymbol{u} \in \mathscr{U} \quad (27)$$

$$(\mathscr{P}^*_{\text{sup}}) \quad \text{maximize} \quad P^*(\boldsymbol{v}^*) = F^*(\Lambda^* \boldsymbol{v}^*) - V^*(\boldsymbol{v}^*) \quad \forall \boldsymbol{v}^* \in \mathscr{V}^* \quad (28)$$

Note that $P : \mathscr{U} \to \vec{\mathbb{R}}$ is l.s.c., convex. It is finite at $\boldsymbol{u}$ if and only if the following *implicit constraint* of $(\mathscr{P}_{\text{inf}})$ is satisfied:

$$\boldsymbol{u} \in \mathscr{U}_a := \{\boldsymbol{u} \in \mathscr{U} | \boldsymbol{u} \in \mathscr{C}, \quad \Lambda \boldsymbol{u} \in \mathscr{D}\} \quad (29)$$

A vector $\bar{\boldsymbol{u}} \in \mathscr{U}_a$ is called an *optimal solution* (*or minimizer*) to $(\mathscr{P}_{\text{inf}})$ if the infimum is achieved at $\bar{\boldsymbol{u}}$ and is not $+\infty$. We write $P(\bar{\boldsymbol{u}}) = \min_{\boldsymbol{u}} P(\bar{\boldsymbol{u}})$. Similarly, the condition

$$\boldsymbol{v}^* \in \mathscr{V}^*_a := \{\boldsymbol{v}^* \in \mathscr{V}^* | \boldsymbol{v}^* \in \mathscr{D}, \quad \Lambda^* \boldsymbol{v}^* \in \mathscr{C}^*\} \quad (30)$$

is called the implicit constraint of $(\mathscr{P}^*_{\text{sup}})$. A vector $\bar{\boldsymbol{v}}^* \in \mathscr{V}^*_a$ is a *dual optimal solution* (*or maximizer*) to $(\mathscr{P}^*_{\text{sup}})$ if the supremum in $(\mathscr{P}^*_{\text{sup}})$ is achieved at $\bar{\boldsymbol{v}}^*$ and is not $-\infty$. We write $P^*(\bar{\boldsymbol{v}}^*) = \max_{\boldsymbol{v}^*} P^*(\boldsymbol{v}^*)$.

For any given F and V, we always have

$$\inf P(\boldsymbol{u}) \geq \sup P^*(\boldsymbol{v}^*) \qquad \forall \boldsymbol{u} \in \mathscr{U}, \boldsymbol{v}^* \in \mathscr{V}^* \quad (31)$$

The difference $\inf P - \sup P^*$ is the so-called *duality gap.* The duality gap is zero if P is convex.

A vector $\bar{\boldsymbol{u}} \in \mathscr{U}_a$ is called a *critical point* of P if P is G-differentiable at $\bar{\boldsymbol{u}}$ and $DP(\bar{\boldsymbol{u}}) = 0$, which gives the *Euler–Lagrange equation* of $(\mathscr{P}_{\text{inf}})$:

$$\Lambda^* DV(\Lambda \bar{\boldsymbol{u}}) - DF(\bar{\boldsymbol{u}}) = 0 \quad (32)$$

If $\mathscr{U}_a$ is an *open set,* the critical point $\bar{\boldsymbol{u}}$ should be a *global minimizer* of the convex function P on $\mathscr{U}_a$. Similarly, the critical condition $DP^*(\bar{\boldsymbol{v}}^*) = 0$ gives the *dual Euler–Lagrange equation* of $(\mathscr{P}^*_{\text{sup}})$:

$$\Lambda DF^*(\Lambda^* \bar{\boldsymbol{v}}^*) - DV^*(\bar{\boldsymbol{v}}^*) = 0 \quad (33)$$

If $\mathscr{V}^*_a$ is an open set, $\bar{\boldsymbol{v}}^*$ should be a *global maximizer* of the concave function P^* on $\mathscr{V}^*_a$.

We say that $(\mathscr{P}_{\text{inf}})$ is *stable* if there exists at least one vector $\boldsymbol{u}_0 \in \mathscr{U}$ such that F is finite at $\boldsymbol{u}_0$ and $V(\boldsymbol{v})$ is finite and continuous at $\boldsymbol{v} = \Lambda \boldsymbol{u}_0$.

Strong Duality Theorem 1. $(\mathscr{P}_{\text{inf}})$ is stable if and only if $(\mathscr{P}^*_{\text{sup}})$ has at least one solution and

$$\inf P = \max P^* \quad (34)$$

Dually, $(\mathscr{P}^*_{\text{sup}})$ is stable if and only if $(\mathscr{P}_{\text{inf}})$ has at least one solution and

$$\min P = \sup P^* \quad (35)$$

If $(\mathscr{P}_{\text{inf}})$ and $(\mathscr{P}^*_{\text{sup}})$ are both stable, then both have solutions and

$$+\infty > \min P = \max P^* > -\infty \quad (36)$$

This theorem shows that as long as the primal problem is stable, the dual problem is sure to have at least one solution. However, the existence conditions for the primal solution are stronger.

Existence and Uniqueness Theorem. Let $\mathscr{U}$ be a reflexive (i.e., $\mathscr{U} = \mathscr{U}^{**}$) Banach space with norm $\| \; \|$. We assume that the feasible set $\mathscr{U}_a \subset \mathscr{U}$ is a nonempty closed convex subset and conditions in (A1) hold. If $\mathscr{C}$ is bounded, or if P is coercive over $\mathscr{C}$, i.e. if

$$\lim P(\boldsymbol{u}) = +\infty \qquad \forall \boldsymbol{u} \in \mathscr{C}, \|\boldsymbol{u}\| \to \infty$$

Then the problem $(\mathscr{P}_{\text{inf}})$ has at least one minimizer. The minimizer is unique if P is strictly convex over $\mathscr{C}$.

All finite-dimensional spaces are reflexive. But some infinite-dimensional vector spaces are not reflexive. So the primal solution in infinite-dimensional systems may or may not exist. If the primal solution does not exist, the dual problem can provide a generalized solution of the problem.

Dual Equivalence Theorem. The following statements are equivalent to each others:

1. $(\mathscr{P}_{\text{inf}})$ is stable and has a solution $\bar{\boldsymbol{u}}$.
2. $(\mathscr{P}^*_{\text{sup}})$ is stable and has a solution $\bar{\boldsymbol{v}}^*$.
3. The extremality relation $P(\bar{\boldsymbol{u}}) = P^*(\bar{\boldsymbol{v}}^*)$ is satisfied.

On $\mathscr{U}_a$ and $\mathscr{V}^*_a$, the extremality condition $P(\bar{\boldsymbol{u}}) = P^*(\bar{\boldsymbol{v}}^*)$ and the Euler–Lagrange equations, Eqs. (32) and (33), are equivalent to each other. On the convex sets $\mathscr{U}_a$ and $\mathscr{V}^*_a$, the extremum problems $(\mathscr{P}_{\text{inf}})$ and $(\mathscr{P}_{\text{sup}})$ and the following variational inequalities are equivalent to each other in the sense that they have the same solution set.

$$\text{(PVI)} \qquad \langle DP(\bar{\boldsymbol{u}}), \boldsymbol{u} - \bar{\boldsymbol{u}} \rangle \geq 0 \qquad \forall \boldsymbol{u} \in \mathscr{U}_a \tag{35}$$

$$\text{(DVI)} \qquad \langle DP^*(\bar{\boldsymbol{v}}^*), \boldsymbol{v}^* - \bar{\boldsymbol{v}}^* \rangle \leq 0 \qquad \forall \boldsymbol{v}^* \in \mathscr{V}_a^* \tag{36}$$

Furthermore, if $\mathscr{C} = \{\boldsymbol{u} \in \mathscr{U} | \ \boldsymbol{u} \geq 0\}$ is a *convex cone,* $\mathscr{C}^* = \{\boldsymbol{u}^* \in \mathscr{U}^* | \ \boldsymbol{u}^* \geq 0\}$ is its *polar cone,* then these problems are equivalent to the following *nonlinear complementarity problem* (NCP):

$$\text{(NCP)} \qquad \boldsymbol{s} = DP(\boldsymbol{u}), \qquad \boldsymbol{u} \in \mathscr{C}, \boldsymbol{s} \in \mathscr{C}^*, \boldsymbol{u} \perp \boldsymbol{s} \tag{37}$$

where $\boldsymbol{s} \in \mathscr{C}^*$ is the so-called *vector of dual slacks.* The *complementarity condition* $\boldsymbol{u} \perp \boldsymbol{s}$ means that $\boldsymbol{u}$ and $\boldsymbol{s}$ are perpendicular to each other. Conditions in Eq. (39) are called the Karush–Kuhn–Tucker (KKT) *constraint qualification* in convex programming. To construct the dual complementarity problem, we need the inverse operator Λ^{-1} (see Ref. 12). In infinite-dimensional systems, to find Λ^{-1} is usually very difficult.

LAGRANGE DUALITY AND HAMILTONIAN

In order to study duality theory in nonconvex problems, we need the so-called Lagrangian form. Let $L: \mathscr{U} \times \mathscr{V}^* \to \overleftrightarrow{\mathbb{R}}$ be an arbitrarily given real-valued function. The following inequality is always true:

$$\sup_{\boldsymbol{v}^* \in \mathscr{V}^*} \inf_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) \leq \inf_{\boldsymbol{u} \in \mathscr{U}} \sup_{\boldsymbol{v}^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{40}$$

A point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is said to be a *minimax point* of L if

$$\sup_{\boldsymbol{v}^* \in \mathscr{V}^*} \inf_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) = L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = \inf_{\boldsymbol{u} \in \mathscr{U}} \sup_{\boldsymbol{v}^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{41}$$

A point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is said to be a *saddle point* of L if

$$L(\bar{\boldsymbol{u}}, \boldsymbol{v}^*) \leq L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \leq L(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) \qquad \forall (\boldsymbol{u}, \boldsymbol{v}^*) \in \mathscr{U} \times \mathscr{V}^* \tag{42}$$

A point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is said to be a *subcritical* (*or* ∂^-*-critical*) *point* of L if

$$L(\bar{\boldsymbol{u}}, \boldsymbol{v}^*) \geq L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \leq L(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) \qquad \forall (u, \boldsymbol{v}^*) \in \mathscr{U} \times \mathscr{V}^* \tag{43}$$

A point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is said to be a *supercritical* (*or* ∂^+*-critical*) *point* of L if

$$L(\bar{\boldsymbol{u}}, \boldsymbol{v}^*) \leq L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \geq L(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) \qquad \forall (\boldsymbol{u}, \boldsymbol{v}^*) \in \mathscr{U} \times \mathscr{V}^* \tag{44}$$

Obviously, the function L possesses a saddle point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ on $\mathscr{U} \times \mathscr{V}^*$ if and only if

$$\max_{\boldsymbol{v}^* \in \mathscr{V}^*} \inf_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) = \min_{\boldsymbol{u} \in \mathscr{U}} \sup_{\boldsymbol{v}^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \tag{45}$$

In general, we have the following connection between the minimax theorem and the saddle point theorem:

Minimax Theorem. If there exists a minimax point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \in \mathscr{U} \times \mathscr{V}^*$ such that

$$L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = \min_{\boldsymbol{u} \in \mathscr{U}} \max_{\boldsymbol{v}^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \max_{\boldsymbol{v}^* \in \mathscr{V}^*} \min_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{46}$$

then $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is a saddle point. Conversely, if $L(\boldsymbol{u}, \boldsymbol{v}^*)$ possesses a saddle point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \in \mathscr{U} \times \mathscr{V}^*$, then the following minimax theorem holds:

$$L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = \min_{\boldsymbol{u} \in \mathscr{C}} \max_{\boldsymbol{v}^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \max_{\boldsymbol{v}^* \in \mathscr{D}^*} \min_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{47}$$

This theorem shows that the existence of a saddle point implies the existence of a minimax point. However, the inverse result holds only on $\mathscr{C} \times \mathscr{D}^*$. This is because $\max_{\mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*)$ may not necessarily exist for all $\boldsymbol{u} \in \mathscr{U}$ and also $\min_{\mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*)$ may not necessarily exist for all $\boldsymbol{v}^* \in \mathscr{V}^*$.

The function $L(\boldsymbol{u}, \boldsymbol{v}^*)$ is said to be a *Lagrangian form* of problem $(\mathscr{P}^*_{\sup})$ if

$$L(\boldsymbol{u}, \boldsymbol{v}^*) = \langle \Lambda \boldsymbol{u}, \boldsymbol{v}^* \rangle - V^*(\boldsymbol{v}^*) - F(\boldsymbol{u}) \tag{48}$$

A vector $\bar{\boldsymbol{u}} \in \mathscr{U}$ is said to be a *Lagrange multiplier* for $(\mathscr{P}^*_{\sup})$ if $\bar{\boldsymbol{u}}$ is an optimal solution to $(\mathscr{P}_{\inf})$.

Dually, the Lagrangian form of problem $(\mathscr{P}_{\inf})$ is defined by

$$L^*(\boldsymbol{u}, \boldsymbol{v}^*) = -\langle \Lambda \boldsymbol{u}, \boldsymbol{v}^* \rangle + V(\Lambda \boldsymbol{u}) + F^*(\Lambda^* \boldsymbol{v}^*) \tag{49}$$

which is also called the *conjugate Lagrangian form.* A vector $\bar{\boldsymbol{v}}^* \in \mathscr{V}^*$ is said to be a *Lagrange multiplier* for $(\mathscr{P}_{\inf})$ if $\bar{\boldsymbol{v}}^*$ is an optimal solution to $(\mathscr{P}^*_{\sup})$. Obviously, we have $L + L^* = P + P^*$. If $\Lambda: \mathscr{U} \to \mathscr{V}$ and $\Lambda^*: \mathscr{V}^* \to \mathscr{U}^*$ are one-to-one and surjective, then the duals of the following results about L also hold for L^*.

A point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \in \mathscr{C} \times \mathscr{D}^*$ is said to be a critical point of L if L is G-differentiable at $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ with respect to both $\boldsymbol{u}$ and $\boldsymbol{v}^*$ separately and

$$D_{\boldsymbol{u}} L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = 0, \qquad \Rightarrow \Lambda^* \bar{\boldsymbol{v}}^* = DF(\bar{\boldsymbol{u}}) \tag{50}$$

$$D_{\boldsymbol{v}^*} L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = 0, \qquad \Rightarrow \Lambda \bar{\boldsymbol{u}} = DV^*(\bar{\boldsymbol{v}}^*) \tag{51}$$

It is easy to establish the following result:

Critical Points Theorem. If $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \in \mathscr{C} \times \mathscr{D}^*$ is either a saddle point or a super (or sub) critical point of L, then $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is a critical point of L, $DP(\bar{\boldsymbol{u}}) = 0$, $DP^*(\bar{\boldsymbol{v}}^*) = 0$ and

$$P(\bar{\boldsymbol{u}}) = L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = P^*(\bar{\boldsymbol{v}}^*) \tag{52}$$

If $F: \mathscr{U} \to \overleftarrow{\mathbb{R}}$ is u.s.c., concave, and $V: \mathscr{V} \to \overrightarrow{\mathbb{R}}$ is l.s.c., convex, then L is a saddle function, and

$$P(\boldsymbol{u}) = \sup_{\boldsymbol{v}^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*), \qquad P^*(\boldsymbol{v}^*) = \inf_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{53}$$

In this case, $P(\boldsymbol{u}) \geq L(\boldsymbol{u}, \boldsymbol{v}^*) \geq P^*(\boldsymbol{v}^*)$ $\forall (\boldsymbol{u}, \boldsymbol{v}^*) \in \mathscr{U} \times \mathscr{V}^*$, and we have

Saddle Point Theorem. $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is a saddle point of L if and only if $\bar{\boldsymbol{u}}$ is a primal solution of $(\mathscr{P}_{\inf})$, $\bar{\boldsymbol{v}}^*$ is a dual solution of $(\mathscr{P}^*_{\sup})$, and $\inf P = \sup P^*$.

If both $F: \mathscr{U} \to \overrightarrow{\mathbb{R}}$ and $V: \mathscr{V} \to \overrightarrow{\mathbb{R}}$ are convex, l.s.c., then $L: \mathscr{U} \times \mathscr{V}^* \to \overleftarrow{\mathbb{R}}$ is a supercritical function and

$$P(\boldsymbol{u}) = \sup_{\mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*), \qquad P^*(\boldsymbol{v}^*) = \sup_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{54}$$

In this case, both P and P^* are nonconvex and $P(\boldsymbol{u}) \geq L(\boldsymbol{u}, \boldsymbol{v}^*) \leq P^*(\boldsymbol{v}^*)$ $\forall (\boldsymbol{u}, \boldsymbol{v}^*) \in \mathscr{U} \times \mathscr{V}^*$.

Dual Max–Min Theorem. If $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \in \mathscr{C} \times \mathscr{D}^*$ is a supercritical point of L, then either

$$P(\bar{\boldsymbol{u}}) = \sup_{\boldsymbol{u}\in\mathscr{U}} P(\boldsymbol{u}) = \sup_{\boldsymbol{v}^*\in\mathscr{V}^*} P^*(\boldsymbol{v}^*) = P^*(\bar{\boldsymbol{v}}^*) \tag{55}$$

or

$$P(\bar{\boldsymbol{u}}) = \inf_{\boldsymbol{u}\in\mathscr{U}} P(\boldsymbol{u}) = \inf_{\boldsymbol{v}^*\in\mathscr{V}^*} P^*(\boldsymbol{v}^*) = P^*(\bar{\boldsymbol{v}}^*) \tag{56}$$

Proof. Since $P(\bar{\boldsymbol{u}}) = L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) = P^*(\bar{\boldsymbol{v}}^*)$, if $\bar{\boldsymbol{u}}$ maximizes P, then

$$\begin{aligned} P(\bar{\boldsymbol{u}}) &= \sup P(\boldsymbol{u}) = \sup_{\boldsymbol{u}} \sup_{\boldsymbol{v}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \sup_{\boldsymbol{v}^*} \sup_{\boldsymbol{u}} L(\boldsymbol{u}, \boldsymbol{v}^*) \\ &= \sup_{\boldsymbol{v}^*} P^*(\boldsymbol{v}^*) = P^*(\bar{\boldsymbol{v}}^*) \end{aligned} \tag{57}$$

as we can take the suprema in either order. If $\bar{\boldsymbol{u}}$ minimizes P, then

$$P(\bar{\boldsymbol{u}}) = \inf P(\boldsymbol{u}) = \inf_{\boldsymbol{u}} \sup_{\boldsymbol{v}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = L(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$$

Since $\bar{\boldsymbol{v}}^*$ is a critical point of P^*, it could be either a local extremum point or a saddle point of P^*. If $\bar{\boldsymbol{v}}^*$ is a saddle point of P^* and it maximizes P^* in the direction $\boldsymbol{v}_o^*$, then we have

$$P^*(\bar{\boldsymbol{v}}^*) = \sup_{\theta\geq 0} P^*(\bar{\boldsymbol{v}}^* + \theta \boldsymbol{v}_o^*) = \sup_{\boldsymbol{u}} \sup_{\theta\geq 0} L(\boldsymbol{u}, \bar{\boldsymbol{v}}^* + \theta\boldsymbol{v}_o^*) = \sup_{\boldsymbol{u}} P(\boldsymbol{u})$$

But $\bar{\boldsymbol{u}}$ is a minimizer of P. This contradiction shows that $\bar{\boldsymbol{v}}^*$ must be a minimizer of P^*.

In geometrically linear systems, the Lagrangian L is usually a saddle function for static problems. But in dynamic problems, L is usually a supercritical function. If V is a *kinetic energy* and F is a *potential energy,* then P is called the *total action* and P^* is called the *dual action.*

By using the Legendre transformation, the *Hamiltonian* $H: \mathscr{U} \times \mathscr{V}^* \to \overleftrightarrow{\mathbb{R}}$ can then be obtained from the Lagrangian as

$$H(\boldsymbol{u}, \boldsymbol{v}^*) = \langle \Lambda\boldsymbol{u}, \boldsymbol{v}^*\rangle - L(\boldsymbol{u}, \boldsymbol{v}^*) \tag{58}$$

If H is G-differentiable on $\mathscr{C} \times \mathscr{D}^*$, we have the following *Hamiltonian canonical equations:*

$$\Lambda\boldsymbol{u} = D_{\boldsymbol{v}^*} H(\boldsymbol{u}, \boldsymbol{v}^*), \qquad \Lambda^*\boldsymbol{v}^* = D_{\boldsymbol{u}} H(\boldsymbol{u}, \boldsymbol{v}^*) \tag{59}$$

If $\Lambda = d/dt$, its adjoint should be $\Lambda^* = -d/dt$. If $V(\Lambda\boldsymbol{u}) = \frac{1}{2}\langle\Lambda\boldsymbol{u}, C\Lambda\boldsymbol{u}\rangle$ is quadratic and the operator $K = \Lambda^*C\Lambda = K^*$ is self-adjoint, then the total action can be written as

$$I(\boldsymbol{u}) = \tfrac{1}{2}\langle \boldsymbol{u}, K\boldsymbol{u}\rangle - F(\boldsymbol{u}) \tag{60}$$

Let $I^c(\boldsymbol{u}) = -P^*(C\Lambda\boldsymbol{u})$, and thus the function $I^c: \mathscr{U} \to \overleftrightarrow{\mathbb{R}}$

$$I^c(\boldsymbol{u}) = \tfrac{1}{2}\langle \boldsymbol{u}, K\boldsymbol{u}\rangle - F^*(K\boldsymbol{u}) \tag{61}$$

is the so-called Clarke dual action (see Ref. 14). Let $K: \mathscr{C} \subset \mathscr{U} \to \mathscr{U}^*$ be a *closed, self-adjoint operator,* and let Ker $K = \{\boldsymbol{u} \in \mathscr{U} \mid K\boldsymbol{u} = 0 \in \mathscr{U}^*\}$ be the *null space* of K, then we have

Clarke Duality Theorem. If $\bar{\boldsymbol{u}} \in \mathscr{C}$ is a critical point of I, then any vector $\boldsymbol{u} \in$ Ker $K + \bar{\boldsymbol{u}}$ is a critical point of I^c. Conversely, if there exists a $\boldsymbol{u}_o \in \mathscr{C}$ such that $K\boldsymbol{u}_o \in \mathscr{C}^*$, then for a given critical point $\bar{\boldsymbol{u}}$ of I^c, any vector $\boldsymbol{u} \in$ Ker $K + \bar{\boldsymbol{u}}$ is a critical point of I.

A comprehensive study of duality theory in linear dynamics is given in Ref. 15.

PRIMAL–DUAL SOLUTIONS AND CENTRAL PATH

Let us now demonstrate how the above scheme fits in with finite-dimensional linear programming. Let $\mathscr{U} = \mathscr{U}^* = \mathbb{R}^n$, $\mathscr{V} = \mathscr{V}^* = \mathbb{R}^m$, with the standard inner products $(\boldsymbol{u}, \boldsymbol{u}^*) = \boldsymbol{u}^T\boldsymbol{u}^*$ in $\mathbb{R}^n$, and $\langle \boldsymbol{v}, \boldsymbol{v}^*\rangle = \boldsymbol{v}^T\boldsymbol{v}^*$ in $\mathbb{R}^m$. For fixed $\bar{\boldsymbol{u}}^* = \boldsymbol{c} \in \mathbb{R}^n$ and $\bar{\boldsymbol{v}} = \boldsymbol{b} \in \mathbb{R}^m$, the primal problem is a constrained linear optimization problem:

$$(\mathscr{P}_{\text{lin}}) \qquad \min_{\boldsymbol{u}\in\mathbb{R}^n} (\boldsymbol{c}, \boldsymbol{u}) \quad s.t. \quad A\boldsymbol{u} = \boldsymbol{b}, \boldsymbol{u} \geq 0 \tag{62}$$

where $A \in \mathbb{R}^{m\times n}$ is a matrix. To reformulate this linear constrained optimization problem in the model form $(\mathscr{P}_{\text{inf}})$, we need to set $\Lambda = A$, $\mathscr{C} = \{\boldsymbol{u} \in \mathbb{R}^n \mid \boldsymbol{u} \geq 0\}$, and $\mathscr{D} = \{\boldsymbol{v} \in \mathbb{R}^m \mid \boldsymbol{v} = \boldsymbol{b}\}$, and let

$$F(\boldsymbol{u}) = -(\boldsymbol{c}, \boldsymbol{u}) - \Psi_{\mathscr{C}}(\boldsymbol{u}), \qquad V(\boldsymbol{v}) = \Psi_{\mathscr{D}}(\boldsymbol{v})$$

The conjugate functions in this elementary case may be calculated at once as

$$\begin{aligned} V^*(\boldsymbol{v}^*) &= \sup_{\boldsymbol{v}\in\mathscr{D}} \langle \boldsymbol{v}, \boldsymbol{v}^*\rangle = \langle \boldsymbol{b}, \boldsymbol{v}^*\rangle \qquad \forall \boldsymbol{v}^* \in \mathscr{D}^* = \mathbb{R}^m \\ F^*(\boldsymbol{u}^*) &= \inf_{\boldsymbol{u}\in\mathscr{C}} (\boldsymbol{u}, \boldsymbol{u}^* + \boldsymbol{c}) = -\Psi_{\mathscr{C}^*}(\boldsymbol{u}^* + \boldsymbol{c}) \end{aligned}$$

where $\mathscr{C}^*$ is a polar cone of $\mathscr{C}$. Let $\boldsymbol{p} = -\boldsymbol{v}^* \in \mathbb{R}^m$, the dual problem $(\mathscr{P}^*_{\text{sup}})$ can be written as

$$(\mathscr{P}^*_{\text{lin}}) \qquad \max_{\boldsymbol{p}\in\mathbb{R}^m} \{P^*(\boldsymbol{p}) = \langle \boldsymbol{b}, \boldsymbol{p}\rangle - \Psi_{\mathscr{C}^*}(\boldsymbol{c} - A^*\boldsymbol{p})\} \tag{63}$$

The implicit constraint in this problem is

$$\boldsymbol{p} \in \mathscr{V}_a^* = \{\boldsymbol{p} \in \mathbb{R}^m \,|\, \boldsymbol{c} - A^*\boldsymbol{p} \geq 0\}$$

For a given $\alpha \in \mathbb{R}^+ := \{\alpha \in \mathbb{R} | \alpha \geq 0\}$, let

$$\Psi_\alpha(\boldsymbol{p}) = \tfrac{1}{2}\alpha \|(A^*\boldsymbol{p} - \boldsymbol{c})_+\|^2$$

where $(x)_+ = \max\{0, x\}$. We have

$$\lim_{\alpha\to\infty} \Psi_\alpha(\boldsymbol{p}) = \Psi_{\mathscr{C}^*}(\boldsymbol{c} - A^*\boldsymbol{p})$$

So the inequality constraint in $(\mathscr{P}^*_{\text{lin}})$ can be relaxed by the following so-called *external penalty method:*

$$(\mathscr{P}^*_p) \qquad \lim_{\alpha\to\infty} \max_{\boldsymbol{p}\in\mathbb{R}^m} \{P^*_p(\boldsymbol{p}; \alpha) = \langle \boldsymbol{b}, \boldsymbol{p}\rangle - \Psi_\alpha(\boldsymbol{p})\} \tag{64}$$

For any given sequence $\{\alpha_k\} \to +\infty$, $P^*_p: \mathbb{R}^m \to \mathbb{R}$ is always concave, and the solution of $(\mathscr{P}^*_p)$ should be also a solution of $(\mathscr{P}^*_{\text{lin}})$. The main disadvantage of the penalty method is that the problem $(\mathscr{P}^*_p)$ will become unstable when the penalty parameter α_k increases.

The Lagrangian $L(\boldsymbol{u}, \boldsymbol{v}^*)$ of $(\mathscr{P}^*_{\text{lin}})$ is

$$L(\boldsymbol{u},\boldsymbol{p}) = \langle A\boldsymbol{u}, -\boldsymbol{p}\rangle - \langle \boldsymbol{b}, -\boldsymbol{p}\rangle + (\boldsymbol{c},\boldsymbol{u}) = (\boldsymbol{b},\boldsymbol{p}) - (\boldsymbol{u}, A^*\boldsymbol{p} - \boldsymbol{c}) \quad (65)$$

But for the inequality constraint in $\mathscr{V}^*_a$, the Lagrange multiplier $\boldsymbol{u} \in \mathbb{R}^n$ has to satisfy the following KKT optimality conditions:

$$\begin{aligned} A\boldsymbol{u} = \boldsymbol{b}, \qquad & \boldsymbol{s} = \boldsymbol{c} - A^*\boldsymbol{p} \\ \boldsymbol{u} \geq 0, \qquad & \boldsymbol{s} \geq 0, \qquad \boldsymbol{s}^T\boldsymbol{u} = 0 \end{aligned} \quad (66)$$

The problem to find $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{s}})$ satisfying Eq. (66) is also known as the *mixed linear complementarity problem* (see Ref. 16)

Combining both the penalty method and the Lagrange method, we have

$$L_{pd}(\boldsymbol{u},\boldsymbol{p};\alpha) = L(\boldsymbol{u},\boldsymbol{p}) - \Psi_\alpha(\boldsymbol{p}) \quad (67)$$

The so-called *augmented Lagrangian method* for solving constrained problem $(\mathscr{P}^*_{\text{lin}})$ is then

$$(\mathscr{P}^*_{pd}) \qquad \min_{(\alpha,\boldsymbol{u})\in\mathbb{R}^+\times\mathbb{R}^n} \max_{\boldsymbol{p}\in\mathbb{R}^m} L_{pd}(\boldsymbol{u},\boldsymbol{p};\alpha) \quad (68)$$

Penalty–Duality Theorem. There exists a finite $\alpha_o > 0$ such that for any given $\alpha \in [\alpha_o, +\infty)$, the solution of the following saddle point problem:

$$\min_{\boldsymbol{u}\in\mathbb{R}^n} \max_{\boldsymbol{p}\in\mathbb{R}^m} L_{pd}(\boldsymbol{u},\boldsymbol{p};\alpha) \quad (69)$$

is also a solution of $(\mathscr{P}^*_{\text{lin}})$. Moreover, for a given penalty-duality sequence $(\alpha_k, \boldsymbol{u}_k) \in \mathbb{R}^+ \times \mathbb{R}^n$, the optimal solution $\overline{\boldsymbol{p}}_k$ of the following unconstrained problem

$$\max_{\boldsymbol{p}\in\mathbb{R}^m} L_{pd}(\boldsymbol{u}_k,\boldsymbol{p};\alpha_k) \quad (70)$$

is an optimal solution of $(\mathscr{P}^*_{\text{lin}})$ if and only if $\boldsymbol{p}_k \in \mathscr{V}^*_a$.

This theorem shows that by constructing a penalty-duality sequence $(\alpha_k, \boldsymbol{u}_k) \in [\alpha_o, +\infty) \times \mathbb{R}^n$, the constrained problem $(\mathscr{P}^*_{\text{lin}})$ can be relaxed by an unconstrained problem [Eq. (70)]. This method is much better than the pure penalty method. Detailed study of the augmented Lagrange methods and applications are given in Ref. 17.

By using the vector of dual slacks $\boldsymbol{s} \in \mathbb{R}^n$, the dual problem $(\mathscr{P}^*_{\text{lin}})$ can be rewritten as

$$(\mathscr{P}^*_{\text{lin}}) \qquad \max_{\boldsymbol{p}\in\mathbb{R}^m} \langle \boldsymbol{b},\boldsymbol{p}\rangle \quad s.t. \quad A^*\boldsymbol{p} + \boldsymbol{s} = \boldsymbol{c}, \qquad \boldsymbol{s} \geq 0 \quad (71)$$

We can see that the primal variable $\boldsymbol{u}$ is the Lagrange multiplier for the constraint $A^*\boldsymbol{p} - \boldsymbol{c} \leq 0$ in the dual problem. However, the dual variables $\boldsymbol{p}$ and $\boldsymbol{s}$ are, respectively, Lagrange multipliers for the constraints $A\boldsymbol{u} = \boldsymbol{b}$ and $\boldsymbol{u} \geq 0$ in the primal problem. These choices are not accidents.

Strong Duality Theorem 2. The vector $\overline{\boldsymbol{u}} \in \mathbb{R}^n$ is a solution of $(\mathscr{P}_{\text{lin}})$ if and only if there exists Lagrange multiplier $(\overline{\boldsymbol{p}}, \overline{\boldsymbol{s}}) \in \mathbb{R}^m \times \mathbb{R}^n$ for which the KKT optimality conditions [Eq. (66)] hold for $(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{s}) = (\overline{\boldsymbol{u}}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{s}})$. Dually, the vector $(\overline{\boldsymbol{p}}, \overline{\boldsymbol{s}}) \in \mathbb{R}^m \times \mathbb{R}^n$ is a solution of $(\mathscr{P}^*_{\text{lin}})$ if and only if there exists a Lagrange multiplier $\overline{\boldsymbol{u}} \in \mathbb{R}^n$ such that the KKT conditions [Eq. (66)] hold for $(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{s}) = (\overline{\boldsymbol{u}}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{s}})$.

The vector $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{s}})$ is called a *primal–dual solution* of $(\mathscr{P}_{\text{lin}})$. The so-called *primal–dual methods* in mathematical programming are those methods to find primal–dual solutions $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{s}})$ by applying variants of Newton's method to the three equations in Eq. (66) and modifying the search directions and step lengths so that the inequalities in Eq. (66) are satisfied at every iteration. If the inequalities are strictly satisfied, the methods are called *primal–dual interior-point methods.* In these methods, the so-called *central path* $\mathscr{C}_{\text{path}}$ plays a vital role in the theory of primal–dual algorithms. It is a parametrical curve of strictly feasible points defined by

$$\mathscr{C}_{\text{path}} = \{(\boldsymbol{u}_\tau, \boldsymbol{p}_\tau, \boldsymbol{s}_\tau)^T \in \mathbb{R}^{2n+m} | \tau > 0\} \quad (72)$$

where each point $(\boldsymbol{u}_\tau, \boldsymbol{p}_\tau, \boldsymbol{s}_\tau)$ solves the following system:

$$\begin{aligned} A\boldsymbol{u} = \boldsymbol{b}, \qquad & A^*\boldsymbol{p} + \boldsymbol{s} = \boldsymbol{c} \\ \boldsymbol{u} > 0, \qquad \boldsymbol{s} > 0, \qquad & u_i s_i = \tau, \quad i = 1, 2, \ldots, n \end{aligned} \quad (73)$$

This problem has a unique solution $(\boldsymbol{u}_\tau, \boldsymbol{p}_\tau, \boldsymbol{s}_\tau)$ for each $\tau > 0$ if and only if the *strictly feasible set*

$$\mathscr{F}_o = \{(\boldsymbol{u},\boldsymbol{p},\boldsymbol{s}) | \, A\boldsymbol{u} = \boldsymbol{b}, A^T\boldsymbol{p} + \boldsymbol{s} = \boldsymbol{c}, \boldsymbol{u} > 0, \boldsymbol{s} > 0\} \quad (74)$$

is nonempty. A comprehensive study of the primal–dual interior-point methods in mathematical programming has been given in Ref. 4

DUALITY IN FULLY NONLINEAR OPTIMIZATION

In fully nonlinear systems, $\Lambda(\boldsymbol{u})$ is a nonlinear operator. The nonlinear Lagrangian form is (see Ref. 11)

$$L(\boldsymbol{u},\boldsymbol{v}^*) = \langle \Lambda(\boldsymbol{u}), \boldsymbol{v}^*\rangle - V^*(\boldsymbol{v}^*) - F(\boldsymbol{u}) \quad (75)$$

The critical condition $\delta L(\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}^*; \boldsymbol{u}, \boldsymbol{v}^*) = 0 \; \forall(\boldsymbol{u}, \boldsymbol{v}^*) \in \mathscr{C} \times \mathscr{D}^*$ gives the canonic equations

$$D_{\boldsymbol{u}} L(\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}^*) = 0 \Rightarrow \Lambda^*_t(\overline{\boldsymbol{u}})\overline{\boldsymbol{v}}^* = DF(\overline{\boldsymbol{u}}) \quad (76)$$

$$D_v L(\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}^*) = 0 \Rightarrow \Lambda(\overline{\boldsymbol{u}}) = DV^*(\overline{\boldsymbol{v}}^*) \quad (77)$$

Since V is either convex or concave on $\mathscr{D}$, the inverse constitutive equation is equivalent to $\overline{\boldsymbol{v}}^* = DV(\Lambda(\overline{\boldsymbol{u}}))$. Then the fundamental equation in fully nonlinear systems should be

$$\Lambda^*_t(\overline{\boldsymbol{u}}) DV(\Lambda(\overline{\boldsymbol{u}})) = DF(\overline{\boldsymbol{u}}) \quad (78)$$

We can see that the symmetry is broken in geometrically nonlinear systems.

If Λ is a quadratic operator, the Taylor expansion of Λ at $\overline{\boldsymbol{u}}$ should be $\Lambda(\overline{\boldsymbol{u}} + \delta\boldsymbol{u}) = \Lambda(\overline{\boldsymbol{u}}) + \Lambda_t(\overline{\boldsymbol{u}})\delta\boldsymbol{u} + \frac{1}{2}\delta^2\Lambda(\overline{\boldsymbol{u}}; \delta\boldsymbol{u})$. We now assume that

$(A2)$ $\quad F : \mathscr{C} \to \mathbb{R}$ is linear and Λ is a quadratic operator such that $\delta^2\Lambda(\overline{\boldsymbol{u}}; \delta\boldsymbol{u}) = -2\Lambda_n(\delta\boldsymbol{u})$

Under this assumption, if $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}^*)$ is a critical point of L, we have $L(\boldsymbol{u}, \boldsymbol{v}^*) - L(\overline{\boldsymbol{u}}, \boldsymbol{v}^*) = G(\boldsymbol{u} - \overline{\boldsymbol{u}}, \boldsymbol{v}^*)$. If $V : \mathscr{D} \to \overline{\mathbb{R}}$ is convex, then (see Ref. 12)

$$(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \text{ is a saddle point of } L \text{ if and only if } \quad G(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) \geq 0 \quad \forall \boldsymbol{u} \in \mathscr{C} \tag{79}$$

$$(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \text{ is a supercritical point of } L \text{ if and only if } \quad G(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) < 0 \quad \forall \boldsymbol{u} \in \mathscr{C} \tag{80}$$

In this case, $P(\boldsymbol{u}) = \sup_{v^* \in \mathscr{V}^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = V(\Lambda(\boldsymbol{u})) - F(\boldsymbol{u})$. But its conjugate function will depend on the sign of G (see Ref. 12):

$$P^*(\boldsymbol{v}^*) = \begin{cases} \inf_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) & \text{if } G(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) \geq 0 \quad \forall \boldsymbol{u} \in \mathscr{U} \\ \sup_{\boldsymbol{u} \in \mathscr{U}} L(\boldsymbol{u}, \boldsymbol{v}^*) & \text{if } G(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) < 0 \quad \forall \boldsymbol{u} \in \mathscr{U} \end{cases} \tag{81}$$

We have the following interesting result:

Triality Theorem. Suppose that the assumption (A2) holds and $V : \mathscr{V} \to \bar{\mathbb{R}}$ is convex, proper and l.s.c. Let $\mathscr{C}_b \times \mathscr{D}_b^*$ be a neighborhood of a critical point $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ of L such that on $\mathscr{C}_b \times \mathscr{D}_b^*$, $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*)$ is the only critical point. Then if $G(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \geq 0$, we obtain

$$P(\bar{\boldsymbol{u}}) = \inf_{\boldsymbol{u} \in \mathscr{C}_b} \sup_{\boldsymbol{v}^* \in \mathscr{D}_b^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \sup_{\boldsymbol{v}^* \in \mathscr{D}_b^*} \inf_{\boldsymbol{u} \in \mathscr{C}_b} L(\boldsymbol{u}, \boldsymbol{v}^*) = P^*(\bar{\boldsymbol{v}}^*) \tag{82}$$

If $G(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) < 0$, we have either

$$P(\bar{\boldsymbol{u}}) = \inf_{\boldsymbol{u} \in \mathscr{C}_b} \sup_{\boldsymbol{v}^* \in \mathscr{D}_b^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \inf_{\boldsymbol{v}^* \in \mathscr{D}_b^*} \sup_{\boldsymbol{u} \in \mathscr{C}_b} L(\boldsymbol{u}, \boldsymbol{v}^*) = P^*(\bar{\boldsymbol{v}}^*) \tag{83}$$

or

$$P(\bar{\boldsymbol{u}}) = \sup_{\boldsymbol{u} \in \mathscr{C}_b} \sup_{\boldsymbol{v}^* \in \mathscr{D}_b^*} L(\boldsymbol{u}, \boldsymbol{v}^*) = \sup_{\boldsymbol{v}^* \in \mathscr{D}_b^*} \sup_{\boldsymbol{u} \in \mathscr{C}_b} L(\boldsymbol{u}, \boldsymbol{v}^*) = P^*(\bar{\boldsymbol{v}}^*) \tag{84}$$

The proof of this theorem was given in Ref. 18. This theorem can be used to solve some nonconvex variational problems (see Refs. 3, 18, and 19).

If $V : \mathscr{D} \to \bar{\mathbb{R}}$ is concave, then

$$(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \text{ is a saddle point of } -L \text{ if and only if } G(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) \leq 0 \quad \forall \boldsymbol{u} \in \mathscr{C} \tag{85}$$

$$(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}^*) \text{ is a subcritical point of } L \text{ if and only if } G(\boldsymbol{u}, \bar{\boldsymbol{v}}^*) > 0 \quad \forall \boldsymbol{u} \in \mathscr{C} \tag{86}$$

In this case, $P(\boldsymbol{u}) = \sup_{v^*} L(\boldsymbol{u}, \boldsymbol{v}^*)$. The dual problem depends also on the sign of the gap function and we have a similar triality theorem (see Ref. 18).

Example 2. Let us consider the minimization of the following nonconvex variational problem:

$$(\mathscr{P}_u) \qquad P(u) = \int_0^1 w(\Lambda(u))\, dt - \int_0^1 fu\, dt \to \min \qquad \forall u \in \mathscr{U}_a \tag{87}$$

where the source variable f is a given function and we let $f(1) = 0$; $w(v)$ could be either a convex or concave function of $v = \Lambda(u)$. As an example, we simply let $w(v) = \frac{1}{2}C(v - \lambda)^2$, with a given parameter $\lambda > 0$ and a material constant $C > 0$ and let Λ be a quadratic operator $\Lambda u = \frac{1}{2}[(d/dt)u]^2 = \frac{1}{2}(u_{,t})^2$. Then $w(\Lambda(u))$ is a double-well function of $\epsilon = u_t$, and

$$\Lambda_t(\bar{u})u = \bar{u}_{,t} u_{,t}, \qquad \Lambda_n(\bar{u})u = -\tfrac{1}{2}\bar{u}_{,t} u_{,t} \tag{88}$$

For a mixed boundary value problem, the convex set $\mathscr{C}$ is a hyperplane

$$\mathscr{C} = \{u \in \mathscr{H}[0,1] |\, u(0) = 0\}$$

and $\mathscr{D} = \{v \in \mathscr{H}[0, 1] |\, v(t) \geq 0 \ \forall t \in [0, 1]\}$ is a convex cone. Then on the feasible set $\mathscr{U}_a$, $P(u)$ is nonconvex, Gâteaux differentiable.

The *direct methods* for solving nonconvex variational problems are difficult. However, by the triality theorem, a closed-form solution of this problem can be easily obtained (see Ref. 12). To do so, we need first to find the conjugate functions. We let $F(u) = \int_0^1 uf\, dt \ \forall u \in \mathscr{C}$. On $\mathscr{D}$, $V(v) = \frac{1}{2}\int_0^1 C(v - \lambda)^2\, dt$ is quadratic. Then the constitutive equation $v^* = \sigma = DV(v) = C(v - \lambda)$ is linear.

$$F^*(u^*) = \inf_{u \in \mathscr{C}} \left\{ \int_0^1 uu^*\, dt + u(1)u^*(1) - \int_0^1 uf\, dt \right\} = -\Psi_{\mathscr{C}^*}(u^*) \tag{89}$$

$$\begin{aligned} V^*(\sigma) &= \sup_{v \in \mathscr{D}} \left\{ \int_0^1 \sigma v\, dt - V(v) \right\} \\ &= \int_0^1 \left(\frac{1}{2C}\sigma^2 + \lambda\sigma \right) dt + \Psi_{\mathscr{D}^*}(\sigma) \end{aligned} \tag{90}$$

where $\mathscr{C}^* = \{u^* \in \mathscr{H}[0, 1] |\, u^*(t) = f(t) \ \forall t \in (0, 1), u^*(1) = 0\}$ is a hyperplane, and $\mathscr{D}^* = \{\sigma \in \mathscr{H}[0, 1] |\, \sigma \neq 0\}$ ($\sigma = 0$ implies that $v = \lambda$.) Since $v = \frac{1}{2}u_{,t}^2 \geq 0 \ \forall u \in \mathscr{C}$ the range of $\sigma = DV(v)$ should be $\mathscr{D}_r^* = \{\sigma \in \mathscr{H}[0, 1] | -\lambda C \leq \sigma < +\infty\}$.

The Lagrangian $L : \mathscr{C} \times \mathscr{D}^* \to \mathbb{R}$ for this problem should be

$$L(u, \sigma) = \int_0^1 \left[\frac{1}{2}(u_{,t})^2\sigma - \left(\frac{1}{2C}\sigma^2 + \lambda\sigma \right) - fu \right] dt \tag{91}$$

The optimality conditions for this problem are

$$\frac{1}{2}u_{,t}^2 = \frac{1}{E}\sigma + \lambda \qquad \forall t \in (0, 1), u(0) = 0 \tag{92}$$

$$-[u_t\sigma]_{,t} = f(t) \qquad \forall t \in (0, 1), \sigma(1) = 0 \tag{93}$$

Let $\tau(t) = u_{,t}\sigma$. It is easy to find that

$$\tau(t) = \int_0^t -f(s)\, ds + \int_0^1 f(s)\, ds \tag{94}$$

The gap function in this problem is a quadratic function of u:

$$G(u, \sigma) = \langle \sigma, -\Lambda_n(u)u \rangle = \frac{1}{2}\int_0^1 \sigma u_{,t}^2\, dt$$

If $\lambda < 0$, then the gap function is positive on $\mathscr{D}_r^*$. In this case, $P(u)$ is convex, and the problem has a unique solution. If $\lambda >$

0, the gap function could be negative on $\mathscr{D}_r^*$. In this case, $P(u)$ is nonconvex and the primal problem may have more than one solution.

On $\mathscr{D}^*$, the conjugate function P^* obtained by Eq. (81) is well-defined:

$$P^*(\sigma) = -\int_0^1 \left[\frac{1}{2C}\sigma^2 + \lambda\sigma + \frac{1}{2}\tau^2/\sigma \right] dt \qquad (95)$$

The dual Euler–Lagrange equation in this example is a cubic algebraic equation:

$$2\sigma^2 \left(\frac{1}{C}\sigma + \lambda \right) = \tau^2 \qquad \forall t \in (0, 1) \qquad (96)$$

For a given $f(t)$ such that $\tau(t)$ is obtained by Eq. (94), this equation has at most three solutions σ_i ($i = 1, 2, 3$). Since $\tau = u_{,t}\sigma$, $u(0) = 0$, the analytic solution for this nonconvex variational problem is

$$u_i(t) = \int_0^t \frac{\tau(s)}{\sigma_i(s)} ds, \qquad i = 1, 2, 3 \qquad (97)$$

By the Triality Theorem we know that $P(u_i) = P^*(\sigma_i)$. The properties of u_i are given by the triality theorem. For certain given f and λ such that $\sigma_1 > 0 > \sigma_2 > \sigma_3$, u_1 is a global minimizer of P, u_2 is a local minimizer, and u_3 is a local maximizer of P. To see this, let $C = 1$; the conjugate function of

$$W^*(\sigma) = -\tfrac{1}{2}(\sigma^2 + 2\lambda\sigma + \tau/\sigma) \qquad (98)$$

is the well-known van der Waals *double-well function:*

$$W(u) = \tfrac{1}{2}(\tfrac{1}{2}u^2 - \lambda)^2 - \tau u \qquad (99)$$

Figure 2 shows the graphs of W (solid line) and W^* (dashed line). The Lagrangian associated with the problem min $W(u)$ is simply given as

$$L(u, \sigma) = \tfrac{1}{2}u^2\sigma - (\tfrac{1}{2}\sigma^2 + \lambda\sigma) - \tau u \qquad (100)$$

Figure 3 shows that L is a saddle function when $\sigma \geq 0$. L is concave if $\sigma < 0$.

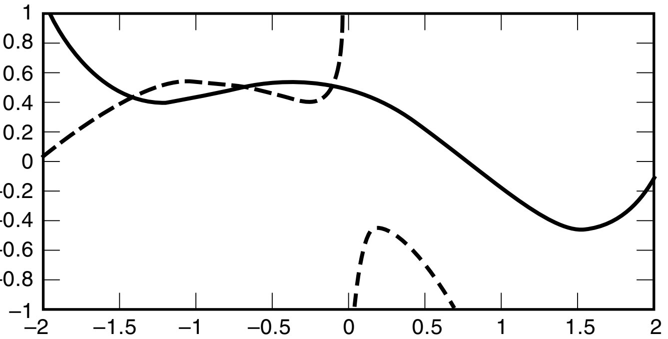

Figure 2. Graphs of $F(u)$ (solid) and $F^*(\sigma)$ (dashed).

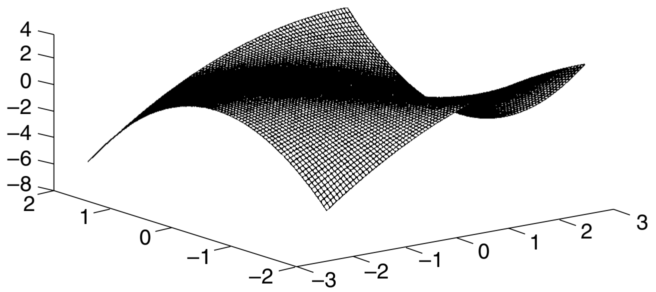

Figure 3. Lagrangian $L(u, \sigma)$.

CONCLUSIONS

Duality theory plays a crucial role in many natural phenomena. It can be used to study wider classes of problems in engineering and science. For geometrically linear systems, duality theory and methods are quite well understood. The excellent textbooks by Strang (1) and Wright (4) are highly recommended. An informal general result was proposed in Ref. 20.

General Duality Principle. For a given system $\mathscr{S}$, if there exists a geometrically linear operator $\Lambda : \mathscr{U} \to \mathscr{V}$ such that the primal system $\mathscr{S}_p = \{\mathscr{U}, \mathscr{V}; \Lambda\}$ and the dual system $\mathscr{S}_d = \{\mathscr{U}^*, \mathscr{V}^*; \Lambda^*\}$ are isomorphic, then

1. For each statement in the primal system $\mathscr{S}_p$, there exists a complementary statement, which is obtained by applying this statement to the dual system $\mathscr{S}_d$; and
2. For each valid theorem defined on the whole system $\mathscr{S} = \mathscr{S}_p \cup \mathscr{S}_d$, the dual theorem, which is obtained by changing all the concepts in the original theorem to their duals, is also valid on $\mathscr{S}$.

From the point of view of the *category theory* (see Ref. 21), the primal system $\mathscr{S}_p$ and the dual system $\mathscr{S}_d$ are said to be *isomorphic* if there exists a so-called *contravariant factor* $\mathscr{F}$ such that the map $\mathscr{F} : \mathscr{S}_p \to \mathscr{S}_d$ is *one-to-one and surjective.* The dual concepts include the paired variables ($\mathbf{u}$, $\mathbf{u}^*$), ($\mathbf{v}$, $\mathbf{v}^*$), conjugate functionals, as well as the dual operations (Λ, Λ^*), ($\geq$, $\leq$), (inf, sup), and so on.

In fully nonlinear systems, the one-to-one symmetrical relations between the primal and dual systems do not usually exist. The duality theory depends on the choice of the nonlinear operator Λ and the associated gap function. The triality theory reveals an intrinsic symmetry in fully nonlinear systems. For a given nonlinear system, the choice of Λ may not be unique, but a quadratic operator will make problems much easier. As long as the paired intermediate variables are defined correctly, the duality theory presented in this article can be used to develop both new theoretical results and powerful numerical methods. A comprehensive study and applications of the duality principle in nonconvex systems are given in Ref. 3. Primal–dual algorithms have been developed for both linear programming (see Ref. 4) and nonconvex problems (see Ref. 22). Triality theory can be used to develop algorithms for robust numerical solutions in fully nonlinear, nonconvex problems.

BIBLIOGRAPHY

1. G. Strang, *Introduction to Applied Mathematics,* Cambridge: Wellesley–Cambridge Univ. Press, 1986.
2. E. Tonti, A mathematical model for physical theories, *Rend. Accad. Lincei,* **LII**, I, 133–139; II, 350–356, 1972.
3. D. Y. Gao, *Duality Principles in Nonlinear Systems: Theory, Methods and Applications,* Dordrecht: Kluwer, 1999.
4. S. J. Wright, *Primal–Dual Interior-Point Methods,* Philadelphia: SIAM, 1996.
5. I. Ekeland and R. Temam, *Convex Analysis and Variational Problems,* Amsterdam: North-Holland, 1976.
6. R. T. Rockafellar, *Conjugate Duality and Optimization,* Philadelphia: SIAM, 1974.
7. M. J. Sewell, *Maximum and Minimum Principles,* Cambridge: Cambridge Univ. Press, 1987.
8. M. Walk, *Theory of Duality in Mathematical Programming,* Berlin: Springer-Verlag, 1989.
9. G. Auchmuty, Duality for non-convex variational principles, *J. Differ. Equ.,* **50**: 80–145, 1983.
10. J. F. Toland, A duality principle for non-convex optimization and the calculus of variations, *Arch. Rational Mech. Anal.,* **71**: 41–61, 1979.
11. D. Y. Gao and Strang G., Geometric nonlinearity: Potential energy, complementary energy, and the gap function, *Q. Appl. Math.,* **XLVII** (3): 487–504, 1989.
12. D. Y. Gao, Dual extremum principles in finite deformation theory with applications in post-buckling analysis of nonlinear beam model, *Appl. Mech. Rev., ASME,* **50** (11): S64–S71, 1997.
13. D. Y. Gao, Minimax and Triality Theory in Nonsmooth Variational Problems, in M. Fukushima and L. Q. Qi (eds.), *Reformulation—Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods,* Dordrecht: Kluwer, 1998, pp. 161–180.
14. I. Ekeland, *Convexity Methods in Hamiltonian Mechanics,* Berlin: Springer-Verlag, 1990.
15. B. Tabarrok and F. P. J. Rimrott, *Variational Methods and Complementary Formulations in Dynamics,* Dordrecht: Kluwer, 1994, p. 366.
16. R. W. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problems,* New York: Academic Press, 1992.
17. M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems,* Amsterdam: North-Holland, 1983.
18. D. Y. Gao, Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications, *IMA J. Appl. Math.,* 1998, in press.
19. D. Y. Gao, Duality in Nonconvex Finite Deformation Theory: A Survey and Unified Approach, in R. Gilbert, P. D. Panagiotopoulos, and P. Pardalos (eds.), *From Convexity to Nonconvexity, A Volume Dedicated to the Memory of Professor Gaetano Fichera,* Dordrecht: Kluwer, 1998, in press.
20. D. Y. Gao, *Complementary-duality theory in elastoplastic systems and pan-penalty finite element methods,* Ph.D. Thesis, Tsinghua University, Beijing, 1986.
21. R. Geroch, *Mathematical Physics,* Chicago: University of Chicago Press, 1985.
22. G. Auchmuty, Duality algorithms for nonconvex variational principles, *Numer. Funct. Anal. Optim.,* **10**: 211–264, 1989.

DAVID YANG GAO
Virginia Polytechnic Institute and State University

DUALITY OF MAGNETIC AND ELECTRIC CIRCUITS. See MAGNETIC CIRCUITS.

DUCTILE ALLOY SUPERCONDUCTORS. See SUPERCONDUCTORS, METALLURGY OF DUCTILE ALLOYS.

DUCTING. See REFRACTION AND ATTENUATION IN THE TROPOSPHERE.

DURATION MEASUREMENT. See TIME MEASUREMENT.

DVD-ROMS. See CD-ROMS, DVD-ROMS, AND COMPUTER SYSTEMS.

DYADIC GREEN'S FUNCTION. See GREEN'S FUNCTION METHODS.