INTEGRO-DIFFERENTIAL EQUATIONS

This article will focus on methods of solution. The aim is to show how a student or engineer can manipulate an integro-differential problem into a form that is simple to calculate. Few of these equations yield analytical solutions. The direct numerical approach of using finite differences for derivatives and sums for integrals relies on the capability of the computer and on the stability of the numerical algorithm. The methods described in this article aim to improve the stability of the eventual calculation by removing derivatives, and to minimize repetitive calculations (nested loops). These techniques make it possible to solve realistic problems with modest personal computers.

An integro-differential equation describes the influence of an accumulation of points upon the value and dynamics of each individual member of the collection. These equations are a balance between a quantity, its derivatives, and its integrals. The most significant applications of integro-differential equations are in modeling the impact of heredity and the dynamics of systems out of equilibrium. Heredity problems in engineering include analyzing fluid and heat flow, mechanical stress, and the accumulation of residual charge for materials with memory. The study of nonequilibrium systems is based on kinetic theory, where the properties of a gas are calculated as the average of individual molecular collisions. Integro-differential equations are applied in biology and economics as well as in physics and engineering.

A differential equation describes the dynamics of a quantity. It is a balance between the values of the quantity and its various rates of change *at a given moment.* One such balance is between the acceleration of a particle and the action of external forces, classical Newtonian mechanics. Basic examples are the one-dimensional mass-spring-damper equation and its electrical analog, the resistor–inductor–capacitor (*RLC*) circuit equation. These specific differential models each conserve a global quantity; for the mechanical example it is momentum, for the electrical example it is current. The implicit assumption in the differential description is that the future state of a system does not depend on its history. Erosion, fatigue, wear, failure, experience, heredity, evolution, karma—all these words express some observation about the impact of past dynamics on future dynamics. One example is the failure of mechanical components subjected to repetitive stress. Engineers routinely calculate the amount of twist of a metal bar subjected to a specific torque. If we assume this phenomenon to be purely differential, then the same amount of torque will always produce the same amount of twist. In reality, metal subjected to repeated strain experiences fatigue. Eventually the application of the same torque produces a different twist, perhaps a catastrophic event. The metal inherits a degradation of its elasticity, an integral of the history of deflections. This effect caused the breakup of two de Havilland Comet jet airliners during flight in 1954. The engineers of the day were unaware that the aluminum fuselage would experience metal fatigue as a result of the frequent cycles of cabin pressurization.

An integral equation describes the influence of all points in a field upon the value of any particular point. Integral equations express an equilibrium. The points can be spatial for a field of stress in a surface or rod, or they can be instants of time in an orbital trajectory, or they can be individual molecules in a gas in which the field is a statistical distribution of molecular velocities. When external conditions change suddenly so that the system is out of balance, energy or information must flow within the system to rearrange it into a new equilibrium. Describing this nonequilibrium process requires differential terms in addition to the original integral equation. Instantaneous equilibration is often assumed in engineering applications, for instance in thermodynamics, and integro-differential equations are avoided. However, the sharpening of technology into much smaller space and time scales has required more exacting physical models that account for nonequilibrium dynamics. This technological trend drives the continuing interest in solving integro-differential equations.

A method is described for transforming integro-differential equations with linear derivatives into purely integral forms, which are then solved by iteration. Knowledge of an approximate solution speeds the convergence of iteration. One method of developing such approximate solutions is described in the section that follows. By the very nature of approximation, such methods depend on the specifics of the particular integro-differential equation. It is best to view the development of the approximate solutions in the sample analysis as an example of the attitude and reasoning that may prove useful in other problems. After the discussion of linear equations a nonlinear system with hereditary effects is described. This nonlinear system describes the conflict between populations of predators and prey. This system is reduced to a single nonlinear integral equation for which an iterated solution is found. However, a much easier calculation of the solution is possible with an approximate equation developed from the nonlinear integral. This approximation recognizes the effect of the hereditary integral and casts the problem as a type of recursion formula $\{x(t) = f[x(t - \Delta t), t]\}$, eliminating the need for any iteration. A discussion of physical applications of integro-differential equations concludes the article.

A SAMPLE ANALYSIS

Consider the following linear, first-order, homogeneous integro-differential equation for unknown function $y(x)$:

$$\frac{dy(x)}{dx} = -\int_{x-\alpha}^{x+\beta} \frac{p(\xi)}{a(x)}\, y(\xi)\, d\xi \tag{1}$$

We will use this equation to demonstrate how a solution may be attempted. Both α and β are positive constants. This particular equation has a separable kernel

$$K(\xi, x) = p(\xi)/a(x) \tag{2}$$

and we suppose that $a(x)$ is not zero in the domain of interest. Can Eq. (1) be cast as a purely differential, or purely integral equation? If so, it may be possible to transform it to a standard form and solve it by established techniques. In this section we will look first at the effect of differentiating Eq. (1), then we will seek an approximate solution directly from the integro-differential form, then we will transform Eq. (1) into a purely integral form, and finally we will show specific examples.

Differentiating Eq. (1) results in

$$a(x)y''(x) + a'(x)y'(x) = p(x-\alpha)y(x-\alpha) - p(x+\beta)y(x+\beta) \tag{3}$$

where primes refer to differentiation with respect to x. In Eq. (3) derivatives of y at x depend on values of y at positions to either side of x. This form is a differential-difference equation. References 1 and 2 describe this type of equation. From Eq. (3), $y(x - \alpha)$ can be cast as depending on itself at higher x

$$y(x-\alpha) = \frac{a(x)y''(x) + a'(x)y'(x) + p(x+\beta)y(x+\beta)}{p(x-\alpha)} \tag{4}$$

This form of the equation is the basis of a numerical solution in cases where it is known that $y(x)$ decays exponentially with respect to positive x. Above a given coordinate, say x_2, y is assumed to be small and its derivatives are assumed to be zero. A solution is constructed for $x < x_2$ using Eq. (4). In a range $x < x_1$, where $x_1 < x_2$, the function $y(x)$ is exponentially larger than the starting value assumed, and $y(x)$ is considered an accurate solution. Care must be taken in the numerical treatment of the derivatives, and this is most directly accomplished by using closely spaced points and higher-order differences. If the kernel is not separable, then differentiation will not remove the integral. It will now contain the derivative of the kernel with respect to x, $K'(\xi, x)$.

Let us assume that $y(x)$ is a positive function that decays exponentially with respect to positive x. In this case both $a(x)$ and $p(x)$ are positive over the range of interest, $x_0 \leq x <$

∞. Let us seek a solution in the form

$$y(x) = \exp\left[-\int_{x_0}^{x} B(\eta)\,d\eta\right] \tag{5}$$

where x_0 is a reference coordinate where $y = 1$, an initial condition. Notice that $-y'/y = B$. Divide Eq. (1) by $y(x)$ and use Eq. (5),

$$a(x)B(x) = \int_{x-\alpha}^{x+\beta} p(\xi)\exp\left[-\int_{x}^{\xi} B(\eta)\,d\eta\right] d\xi \tag{6}$$

Notice that in Eq. (6) it is the ratio $y(\xi)/y(x)$ that appears in the integral with $p(\xi)$, and this ratio is given by the exponential involving $B(\eta)$. Given an estimate of function B, call it B_0, Eq. (6) can be used to find a possibly more accurate estimate B_1 by the method of successive approximations. This method, also known as the method of Picard, uses a prior iterant within the integral (B_0) to find a next iterant (B_1) from the equation. References 3 and 4 describe the validity and use of this method.

If the method of successive approximations converges to a solution, then the exact nature of the initial iterant B_0 is unimportant. However, the more accurately B_0 portrays the actual solution $B(x)$, the fewer iterants need to be calculated. We now seek an initial iterant from Eq. (6) by making whatever assumptions simplify this problem, while at the same time being mindful to avoid a trivial result by being too hasty. For the moment we will assume that $p(x)$ is weakly dependent on x within any band $(x - \alpha, x + \beta)$, and that $B(\xi)$ remains of the same order of magnitude for $(x - \alpha \leq \xi \leq x + \beta)$. The following approximations cascade from Eq. (6) by using these assumptions:

$$\begin{aligned}
a(x)B_0(x) &\approx \int_{x-\alpha}^{x+\beta} p(\xi)e^{-B_0(x)(\xi-x)}\,d\xi \approx p(x)\int_{x-\alpha}^{x+\beta} e^{-B_0(x)(\xi-x)}\,d\xi \\
&\approx \frac{p(x)}{-B_0(x)}\int_{x-\alpha}^{x+\beta} e^{-B_0(x)(\xi-x)}[-B_0(x)]\,d\xi \qquad (7)\\
&\approx \frac{p(x)}{B_0(x)}[e^{\alpha B_0(x)} - e^{-\beta B_0(x)}]
\end{aligned}$$

$$B_0(x) = \sqrt{\frac{p(x)}{a(x)}e^{\alpha B_0(x)}[1 - e^{-(\alpha+\beta)B_0(x)}]} \tag{8}$$

For a small $B_0(x)$ such that both $B_0(x)(\alpha + \beta) < 1$ and $B_0(x)\alpha < 1$, then

$$B_0(x) = \frac{p(x)}{a(x)}(\alpha + \beta) \tag{8a}$$

which is found by expanding the exponentials in Eq. (8). Notice that to be consistent, $p(x)/a(x)$ must be less than $(\alpha + \beta)^{-2}$. For $B_0(x)$ such that $B_0(x)(\alpha + \beta) > 1$ while $B_0(x)\alpha < 1$, which implies $\beta \gg \alpha$, then

$$B_0(x) = \sqrt{\frac{p(x)}{a(x)}} \tag{8b}$$

This case is consistent with $1/(\alpha + \beta) < \sqrt{p(x)/a(x)} < 1/\alpha$. Finally, for $B_0(x)$ such that both $B_0(x)(\alpha + \beta) > 1$ and $B_0(x)\alpha > 1$, then $B_0(x)$ is the root of a transcendental equation,

$$B_0(x) = \sqrt{\frac{p(x)}{a(x)}e^{\alpha B_0(x)}} \tag{8c}$$

If none of Eqs. (8a)–(8c) are applicable, then $B_0(x)$ must be found as a root of Eq. (8). The corresponding initial iterant $y_0(x)$ for Eq. (1) is given by using the appropriate result from Eqs. (8) or (8a)–(8c) in definition (5). The example shown below uses the simplest case, Eq. (8a),

$$y_0(x) = \exp\left[-(\alpha + \beta)\int_{x_0}^{x} \frac{p(\eta)}{a(\eta)}\,d\eta\right] \tag{9}$$

It is very important to capture the functional nature of $p(x)$ within the integral of Eq. (7). In the preceding, $p(x)$ was assumed to be very weakly dependent on x over the interval $(x - \alpha, x + \beta)$ in a manner similar to a constant or $\log(x)$. If instead, $p(x) = p_0(x)x$, where $p_0(x)$ is a weak function of x as used here, then the results in place of Eq. (8) are as follows:

$$B_0(x) = \sqrt[3]{\begin{aligned}&\frac{p_0(x)}{a(x)}e^{\alpha B_0(x)}\{1 + B_0(x)(x - \alpha) \\ &\qquad - [1 + B_0(x)(x + \beta)]e^{-(\alpha+\beta)B_0(x)}\}\end{aligned}} \tag{10}$$

The case of small $B_0(x)$ corresponding to Eq. (8a) is now

$$B_0(x) = \frac{p_0(x)}{a(x)}(\alpha + \beta)\left(x + \frac{\beta^2 - \alpha^2}{(\alpha + \beta)}\right) \tag{10a}$$

Note the additional linear factor in comparison to Eq. (8a). It is essential to retain that factor of $p(x)$ with significant variation within the integral of Eq. (7). We will only use the simplest B_0 and y_0, derived as Eqs. (8a) and (9), respectively, to illustrate a first iterant with Eq. (11):

$$a(x)B_1(x) = \int_{x-\alpha}^{x+\beta} p(\xi)\exp\left[-(\alpha + \beta)\int_{x}^{\xi} \frac{p(\eta)}{a(\eta)}\,d\eta\right] d\xi \tag{11}$$

and a $y_1(x)$ can be constructed from the $B_1(x)$ of Eq. (11). A $y_1(x)$ can also be written explicitly from the integral of Eq. (1) by using $y_0(x)$

$$y_1(x) = y(x_0) - \int_{x_0}^{x}\int_{\xi-\alpha}^{\xi+\beta} \frac{p(\eta)}{a(\xi)}y_0(\eta)\,d\eta\,d\xi \tag{12}$$

Recall that $y(x_0) = 1$ in this particular case. Whether the first iterant sought is B_1 from Eq. (11) or y_1 from Eq. (12), a double integration is required after the zeroth iterants B_0 and y_0 are calculated. It would be very discouraging to do all this work and then find that our iteration was diverging. An effort to reduce repetitive integration follows.

Equation (12) is a purely integral form of Eq. (1) when the subscripts on y are removed and $y(x_0)$ is arbitrary. By reversing the order of integration it is possible to reformulate this equation as a single integration over the unknown $y(\eta)$

with a new kernel

$$\begin{aligned} y(x) - y(x_0) &= -\int_{x_0}^{x}\int_{\xi-\alpha}^{\xi+\beta} \frac{p(\eta)}{a(\xi)} y(\eta)\,d\eta\,d\xi \\ &= -\int_{x_0-\alpha}^{x+\beta} p(\eta) M(\eta, x) y(\eta)\,d\eta \end{aligned} \tag{13}$$

The new kernel factor $M(\eta, x)$ is given below for this example. The method of reversing the order of integration and generating kernels of this type will be described in the section titled Linear Equations:

$$\begin{aligned} M(\eta, x) = &\left\{\int_{x_0}^{\eta+\alpha} \frac{d\xi}{a(\xi)};\, [x_0 - \alpha] \le \eta \right. \\ &\left. < [\min(x_0 + \beta, \min(x_0 + \alpha + \beta, x) - \alpha)]\right\} \\ &+ \left\{\int_{x_0}^{x} \frac{d\xi}{a(\xi)};\, \min[x_0 + \beta, \min(x_0 + \alpha + \beta, x) - \alpha] \right. \\ &\left. \le \eta < [x_0 + \beta]\right\} \\ &+ \left\{\int_{\eta-\beta}^{\min(\eta+\alpha, x)} \frac{d\xi}{a(\xi)};\, [x_0 + \beta] \le \eta \le [x + \beta]\right\} \end{aligned} \tag{14}$$

The function min(a, b, . . .), used in $M(\eta\, x)$, selects the minimum of its arguments. $M(\eta\, x)$ is the sum of three terms, each defined over a different range of η, and these ranges are functions of x. The new kernel $K(\eta, x) = p(\eta)M(\eta, x)$ can be calculated once from known functions $a(x)$ and $p(x)$, and by the explicit operations of Eq. (14). The derivation of Eqs. (13) and (14) proceeds directly from Eq. (1), without requiring any specialized assumptions, as were used in the development of B_0. Now the original integro-differential equation has been transformed into a purely integral form, a Volterra equation (variable upper limit) of the second kind [inhomogeneous if $y(x_0) \neq 0$]. The method of successive approximations applied to Eq. (13) proceeds more quickly because each iterant of $y(x)$ is now the result of a single integration.

Two specific numerical examples follow. In both cases $a(x) = a_0x$, and $p(x) = p_0x$, where a_0 and p_0 are constants. Solutions are sought in the range $[(x_0 = 1) \le x \le (x_1 = 3)]$, though calculations must consider the wider range $(1 - \alpha, 3 + \beta)$. In these cases $y(1) = 1$. $B_0(x)$ is found as the root of Eq. (10), and a $y_0(x)$ is calculated from Eq. (5). The kernel $K(\eta, x) = p(\eta)M(\eta, x)$ is calculated on the basis of Eq. (14). Two iterants, y_1 and y_2, are then found by the method of successive approximations from Eq. (13). Figure 1 shows y_0, y_1, and y_2 for $a_0 = 1$, $p_0 = 8$, $\alpha = 0$, and $\beta = 1.45$. Iterants y_0 and y_1 are quite smooth; with y_2 the point-to-point numerical noise becomes noticeable (point locations are shown for y_2). This noise diminishes as more closely spaced points are used. In this case $y(x)$ has a rapid exponential decay. A second case has $a_0 = 1$, $p_0 = 0.08$, $\alpha = 0.55$, and $\beta = 1.45$. Figure 2 shows the three iterants of y, which decay gently with x. In both cases y_0 and y_1 bracket y_2. Figures 3 and 4 show the kernel $K(\eta, x)$ for the first case (both appear similar). Two views are given to help visualize this surface over the full range of the calculation.

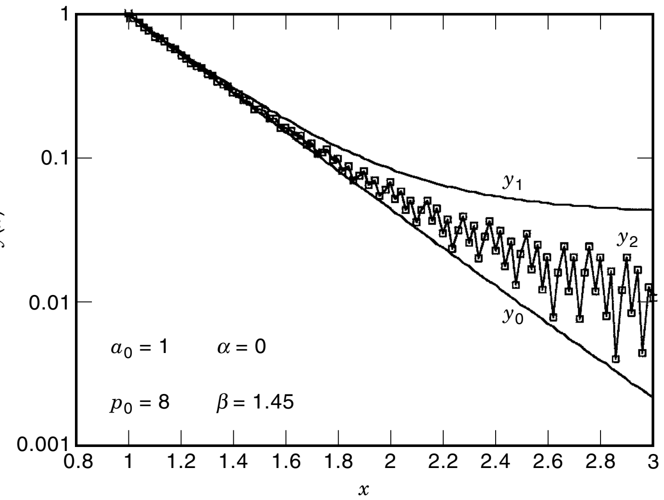

Figure 1. Three iterants for the $y(x)$ of Eq. (1) when $a(x) = x$, $p(x) = 8x$, $\alpha = 0$, and $\beta = 1.45$. The zeroth iterant $y_0(x)$ is found by an approximation to its logarithmic derivative $B_0 = -\{d\ln[y_0(x)]/dx\}$ that is given by Eq. (10). The iteration is applied to Eq. (13), which is a single integral form of Eq. (1) with a new kernel $p(\eta)M(\eta, x)$ that is described by Eq. (14). Convergence is rapid. This function decays by two orders of magnitude for $1 \le x \le 3$. The relative error is comparable to $y(x)$ at low amplitude. This error diminishes as more points are used (point locations shown for y_2).

LINEAR EQUATIONS

Casting a linear, first-order integro-differential equation into a simple integral form is very useful because then it can be solved by the method of successive approximations. This transformation involves switching the order of integration of a double integral, an operation mentioned without explanation in the section titled A Sample Analysis. This transforma-

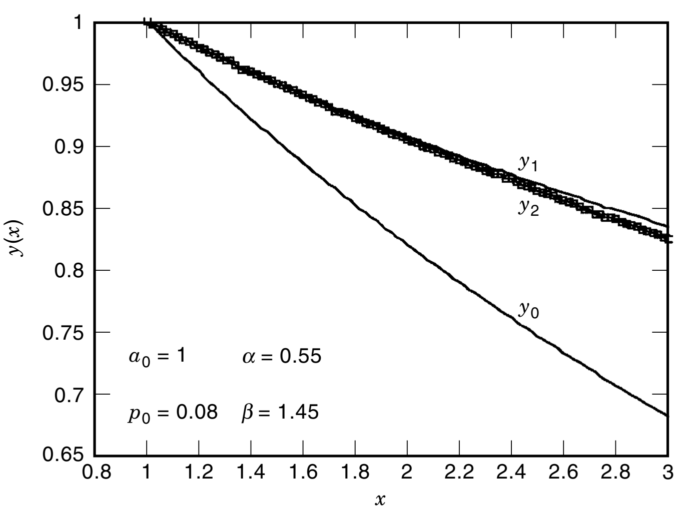

Figure 2. Three iterants for the $y(x)$ of Eq. (1) when $a(x) = x$, $p(x) = 0.08x$, $\alpha = 0.55$, and $\beta = 1.45$. Another case similar to that of Fig. 1. Here the function decays very gently, and the relative error is small.

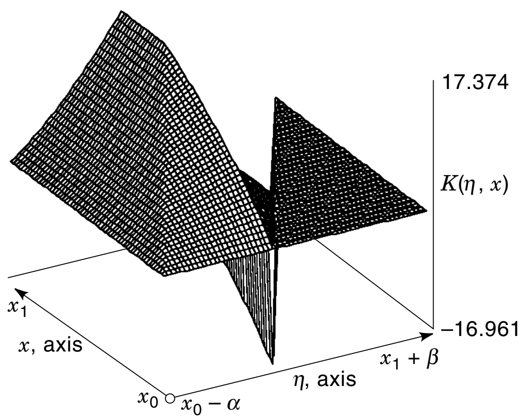

Figure 3. A surface plot of $K(\eta, x) = p(\eta)M(\eta, x)$, the kernel used in the example of Fig. 1. The kernel for the second example has the same shape but is of different magnitude. This view extends over the full area of the calculation in the (η, x) plane. The problem has been converted to an integral equation with single integration and a known kernel.

tion will be illustrated for the following equation:

$$\frac{dy(x)}{dx} + b(x)y(x) + c(x) = \int_{x-\alpha}^{x+\beta} K_1(\xi, x)y(\xi)\,d\xi + \int_{x_0}^{x} K_2(\xi, x)y(\xi)\,d\xi \tag{15}$$

We assume that over the domain of interest, $x_0 \le x \le x_1$, none of b, c, K_1 and K_2 become infinite. Also, α and β are positive constants. The labels $V_1(x)$ and $V_2(x)$ will be used to represent the integrals over K_1 and K_2, respectively. Now Eq. (15) is seen as a linear, first-order differential equation with an inhomogeneous term $V_1(x) + V_2(x) - c(x)$. This is formally integrated to

$$y(x) = e^{-\int_{x_0}^{x} b(\gamma)\,d\gamma} \left\{ y(x_0) + \int_{x_0}^{x} [V_1(\xi) + V_2(\xi) - c(\xi)] e^{\int_{x_0}^{\xi} b(\gamma)\,d\gamma}\,d\xi \right\} \tag{16}$$

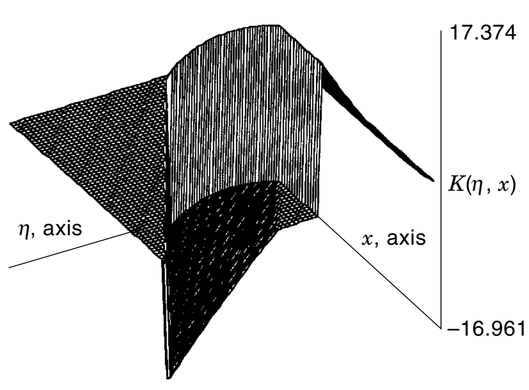

Figure 4. A surface plot of $K(\eta, x) = p(\eta)M(\eta, x)$ seen from a different orientation. This view shows features of the surface that are hidden in Fig. 3. This example shows that integral equations can have smooth solutions even with discontinuous kernels.

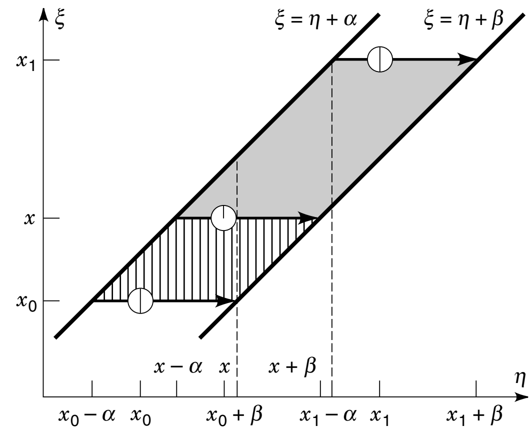

Figure 5. The area where double integration with the K_1 kernel is reversed. Horizontal arrows show the original direction of integration across the area. Reversed integration is shown by vertical hatching. As x moves up the ξ axis from x_0 to x_1, the horizontal arrow from $\eta = x - \alpha$ to $\eta = x + \beta$ moves vertically through the integration area. Reversed integration is done below this rising arrow. Here vertical integration proceeds in sections as η moves from limits $x_0 - \alpha$ to $x_1 + \beta$ (the new outer integral). ξ is integrated successively from: x_0 to the $\eta + \alpha$ boundary line, x_0 to x, the $\eta - \beta$ boundary line to the $\eta + \alpha$ boundary line, and the $\eta - \beta$ boundary line to x (the new inner integral). The limits are conditional statements because the transitions between vertical sections depend on the slant and width of the area.

$$\begin{aligned} y(x) = {} & y(x_0)e^{-\int_{x_0}^{x} b(\gamma)\,d\gamma} - \int_{x_0}^{x} c(\xi)e^{-\int_{\xi}^{x} b(\gamma)\,d\gamma}\,d\xi \\ & + \int_{x_0}^{x}\int_{\xi-\alpha}^{\xi+\beta} e^{-\int_{\xi}^{x} b(\gamma)\,d\gamma} K_1(\eta, \xi)y(\eta)\,d\eta\,d\xi \\ & + \int_{x_0}^{x}\int_{x_0}^{\xi} e^{-\int_{\xi}^{x} b(\gamma)\,d\gamma} K_2(\eta, \xi)y(\eta)\,d\eta\,d\xi \end{aligned} \tag{17}$$

The order of double integration will now be reversed. This is done to achieve single integral forms $\int M(\eta, x)y(\eta)d\eta$ with kernels $M(\eta, x)$ that are integrals of known functions. The K_1 and K_2 integrations of Eq. (17) occur over specific areas of the (η, ξ) plane determined by the limits. Figure 5 is a schematic of the area of integration for K_1. Figure 6 is a similar sche-

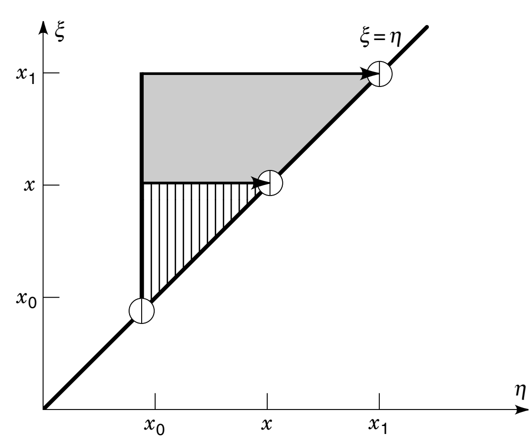

Figure 6. The area where double integration with the K_2 kernel is reversed. The original integration of $x_0 \le \xi \le x$ and $x_0 \le \eta \le \xi$ is reversed to $x_0 \le \eta \le x$ and $\eta \le \xi \le x$.

matic for K_2. In Figs. 5 and 6 these integrations would be visualized as progressing horizontally through the respective areas (see arrows). To reverse the order of integration is to progress vertically through the integral areas (see vertical hatching). The original double integrals could each become sums of several "reversed" terms. Each of the new, reversed double integrals would account for a portion of the original (η, ξ) area. The limits of the reversed integrals could be conditional statements that depend on the shape of the area boundary. The result here is shown as Eqs. (18) through (22):

$$y(x) = y(x_0)e^{-\int_{x_0}^{x} b(\gamma)\,d\gamma} - \int_{x_0}^{x} c(\xi)e^{-\int_{\xi}^{x} b(\gamma)\,d\gamma}\,d\xi + I_1(x) + I_2(x) \tag{18}$$

$$\begin{aligned} I_1(x) = &\int_{x_0-\alpha}^{\min[x^*-\alpha, x_0+\beta]} \int_{x_0}^{\eta+\alpha} \ldots d\xi\,d\eta \\ &+ \int_{\min[x^*-\alpha, x_0+\beta]}^{x_0+\beta} \int_{x_0}^{x} \ldots d\xi\,d\eta + \int_{x_0+\beta}^{x+\beta} \int_{\eta-\beta}^{\min[\eta+\alpha, x]} \ldots d\xi\,d\eta \end{aligned} \tag{19}$$

where the integrands are

$$e^{-\int_{\xi}^{x^*} b(\gamma)\,d\gamma} K_1(\eta, \xi) y(\eta)$$

for the first term of I_1, and

$$e^{-\int_{\xi}^{x} b(\gamma)\,d\gamma} K_1(\eta, \xi) y(\eta) \tag{20}$$

for the last two terms of I_1. The function x^* is defined as

$$x^* = \min[x, (x_0 + \alpha + \beta)] \tag{21}$$

Finally, for I_2,

$$I_2(x) = \int_{x_0}^{x} y(\eta) \left\{ \int_{\eta}^{x} e^{-\int_{\xi}^{x} b(\gamma)\,d\gamma} K_2(\eta, \xi)\,d\xi \right\} d\eta \tag{22}$$

The original equation is now in a purely integral form with single integrations. New kernels, $M_1(\eta, x)$ (three terms) and $M_2(\eta, x)$, are defined as integrals of the products of an integrating factor and original kernels $K_1(\eta, \xi)$ and $K_2(\eta, \xi)$, respectively. The sample analysis describes solving a particular equation from this point.

Linear integro-differential equations with second-order derivatives can be transformed into a Volterra form in a manner similar to first-order equations. Consider the following second-order equation with the same integrals V_1 and V_2 as in Eq. (16):

$$\frac{d^2y(x)}{dx^2} + b(x)\frac{dy(x)}{dx} + c(x)y(x) + d(x) = V_1(x) + V_2(x) \tag{23}$$

Define the function $p(x) = dy(x)/dx$. Now Eq. (23) becomes a linear, first-order equation for $p(x)$ with inhomogeneous term $V_1(x) + V_2(x) - d(x) - c(x)y(x)$. This is integrated once for $p(x)$ by using an integrating factor $\exp[\int b(x)dx]$, and specifying an initial condition $p(x_0)$. The result for $p(x)$ is integrated from x_0 to x yielding $y(x) - y(x_0)$

$$\begin{aligned} y(x) = &y(x_0) + p(x_0)\int_{x_0}^{x} e^{-\int_{x_0}^{x} b(\gamma)\,d\gamma}\,d\xi \\ &- \int_{x_0}^{x}\int_{x_0}^{\xi} d(\eta)e^{-\int_{\eta}^{\xi} b(\gamma)\,d\gamma}\,d\eta\,d\xi \\ &+ \int_{x_0}^{x}\int_{x_0}^{\xi} [V_1(\eta) + V_2(\eta) - c(\eta)y(\eta)]e^{-\int_{\eta}^{\xi} b(\gamma)\,d\gamma}\,d\eta\,d\xi \end{aligned} \tag{24}$$

The first three terms after the equal sign in Eq. (24) are all known, the fourth term contains $y(x)$ within double and triple integrals. Let $f(x, x_0)$ represent the sum of the three known terms in Eq. (24) and $H(x, \eta)$ represent the integral factor

$$H(x, \eta) = \int_{\eta}^{x} e^{-\int_{\eta}^{\xi} b(\gamma)\,d\gamma}\,d\xi \tag{25}$$

Using these definitions, and Eq. (21) for x^*, the form of Eq. (24) with only single integrals is

$$\begin{aligned} y(x) = &f(x, x_0) + \int_{x_0-\alpha}^{\min[(x^*-\alpha), (x_0+\beta)]} \\ &y(\zeta)\left\{\int_{x_0}^{\zeta+\alpha} H(x^*, \eta)K_1(\zeta, \eta)\,d\eta\right\} d\zeta \\ &+ \int_{\min[(x^*-\alpha), (x_0+\beta)]}^{x_0+\beta} y(\zeta)\left\{\int_{x_0}^{x} H(x, \eta)K_1(\zeta, \eta)\,d\eta\right\} d\zeta \\ &+ \int_{x_0+\beta}^{x+\beta} y(\zeta)\left\{\int_{\zeta-\beta}^{\min[(\zeta+\alpha), x]} H(x, \eta)K_1(\zeta, \eta)\,d\eta\right\} d\zeta \\ &+ \int_{x_0}^{x} y(\zeta)\left\{\int_{\zeta}^{x} H(x, \eta)K_2(\zeta, \eta)\,d\eta\right\} d\zeta \\ &- \int_{x_0}^{x} y(\zeta)c(\zeta)H(x, \zeta)\,d\zeta \end{aligned} \tag{26}$$

The five integrals shown for Eq. (26) can be combined into a single integration from $(x_0 - \alpha)$ to $(x + \beta)$ by multiplying each kernel with a difference of Heaviside unit step functions to define limits. This was done to calculate examples from Eq. (14) in the sample analysis.

The linear equations discussed in the preceding have all described initial value problems. If boundary values are placed on $y(x)$ or its first derivative at two points (x_0, x_1), then the solution of a second-order equation is based on the characteristic functions, or eigenfunctions, of the differential part of the equation. A solution of the form $y(x) = C_0y_0(x) + C_1y_1(x) + C_2y_2(x) + \ldots$ is assumed, where the y_i are eigenfunctions corresponding to the eigenvalues λ_i. The coefficients C_i are found in exactly the same way as in boundary value problems involving nonhomogeneous differential equations. The three steps to the solution are: substitute the eigenfunction expansion into Eq. (23), multiply by a particular eigenfunction y_i to solve for its coefficient C_i, and integrate over the interval (x_0, x_1). In purely differential problems the result is a series of equations, one for each of the coefficients C_i. For integro-differential equations the result is a series of equations linking each coefficient to a weighted sum of coefficients, $C_i = \Sigma w_nC_n$. The weights w_n result from integration and may be difficult to calculate. This matrix relationship among the coefficients reflects the nature of the original equation. The magnitude C_i of each mode $y_i(x)$ is linked to the magnitudes of the other modes in solution $y(x)$ by the integrals involving K_1 and K_2.

VOLTERRA ANIMALS

Volterra introduced the following system of coupled, nonlinear, first-order, integro-differential equations to describe the dynamics of survival for a population of predators $y(t)$ and a population of prey $x(t)$:

$$\begin{aligned}\left[\frac{1}{x(t)}\right]\frac{dx(t)}{dt} &= a(t) - b(t)y(t) - \int_c^t K_y(t,s)y(s)\,ds \\ \left[\frac{1}{y(t)}\right]\frac{dy(t)}{dt} &= -\alpha(t) + \beta(t)x(t) + \int_c^t K_x(t,s)x(s)\,ds\end{aligned} \tag{27}$$

These equations show rates of population growth that are dependent on three factors: herd size or predator density, encounters between species, and hereditary influences. Prey $x(t)$ are adversely affected by encounters with predators, $-b(t)x(t)y(t)$, and by evolutionary improvements in these predators, $-x(t)\int K_y(t, s)y(s)\,ds$. Predators are adversely affected by too high a population of their own kind, $-\alpha(t)y(t)$. Reference 4 discusses this system in detail. The hereditary integral is described for heredity coefficients (K_x and K_y) of the form $K(t - s)$ under various names: it is the "renewal equation" in Ref. 2, "convolution" in Ref. 3, the "superposition integral" in Ref. 5, and an integral with a "displacement kernel" in Ref. 6. We will describe an approximate method of solution for Eq. (27) that makes few assumptions about the coefficients a, b, α, and β, or the kernels K_x and K_y.

Figure 7 is a particular example of Eqs. (27) for $0 \le t \le 20$, $c = 0$ ("the creation"), $a = b = 2$, $\alpha = \beta = 1$, $K_x = K_y = -0.05$, and initial conditions $x(0) = 1$ and $y(0) = 2$. Figure 8 is a phase diagram for this case where the initial conditions and the direction of time's arrow are indicated. Without hereditary influences ($K_x = K_y = 0$), the nonlinear, purely differential system has a periodic trajectory that is a noncircular closed path on the xy phase plane (a "vortex cycle"). In general, neither $x(t)$ nor $y(t)$ can be expressed in terms of elementary functions. The effect of the hereditary integrals is to cause a "drift" in the solutions, seen as a rising trend for this example. Additional long-term effects for this case are a diminishing impact of predators (y) on prey (x) and a shortening of the time between cycles. Converging or diverging populations that either grow or diminish can clearly be simulated by changing the magnitudes and the signs of the coefficients and kernels. More interesting effects arise when these factors are time dependent.

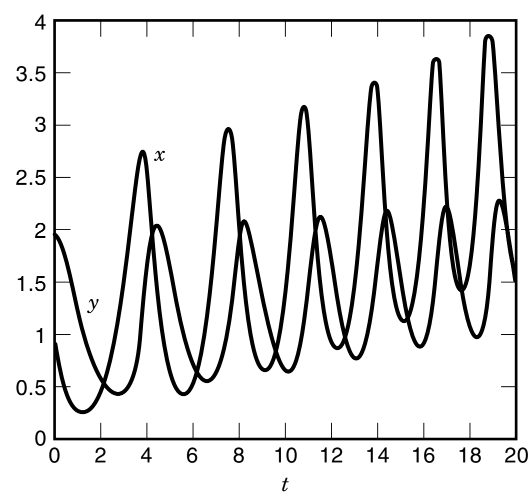

Figure 7. Population histories of predators $y(t)$ and prey $x(t)$ from the Volterra model of Eq. (27) with $a = b = 2$, $\alpha = \beta = 1$, $K_x = K_y = -0.05$, $c = 0$, $y(c) = 2$, and $x(c) = 1$. In this example heredity causes the populations to increase, diverge, and cycle more often.

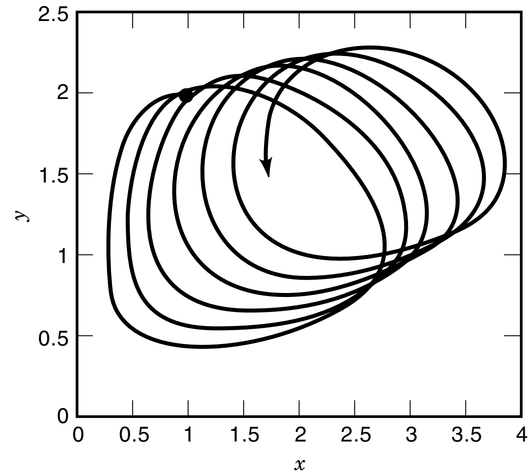

Figure 8. Phase diagram for the Volterra model example of Fig. 7. If heredity coefficients K_x and K_y are zero, then this curve is a closed noncircular path called a vortex cycle. The initial point and the direction of time's arrow are shown. Heredity causes a drift in the cyclic action.

Equations (27) are integrated once

$$\begin{aligned}\ln\left[\frac{x(t)}{x_c}\right] &= \int_c^t a(s)\,ds \\ &\quad - \int_c^t b(s)y(s)\,ds - \int_c^t\int_c^s K_y(s,u)y(u)\,du\,ds \\ \ln\left[\frac{y(t)}{y_c}\right] &= -\int_c^t \alpha(s)\,ds + \int_c^t \beta(s)x(s)\,ds \\ &\quad + \int_c^t\int_c^s K_x(s,u)x(u)\,du\,ds\end{aligned} \tag{28}$$

The order of double integration is reversed, and then the following functions are defined:

$$\begin{aligned}A(t) &= \int_c^t a(s)\,ds \\ \Lambda(t) &= \int_c^t \alpha(s)\,ds \\ M_x(t,u) &= \int_u^t K_x(s,u)\,ds \\ M_y(t,u) &= \int_u^t K_y(s,u)\,ds\end{aligned} \tag{29}$$

Now the equations are

$$\begin{aligned}\ln\left[\frac{x(t)}{x_c}\right] &= A(t) - \int_c^t [b(u)+M_y(t,u)]y(u)\,du \\ \ln\left[\frac{y(t)}{y_c}\right] &= -\Lambda(t) + \int_c^t [\beta(u)+M_x(t,u)]x(u)\,du\end{aligned} \tag{30}$$

The equation for $y(t)$ is substituted into the equation for $x(t)$, yielding

$$\ln\left[\frac{x(t)}{x_c}\right] = A(t) - y_c\int_c^t [b(u)+M_y(t,u)]e^{-\Lambda(u)+\int_c^u[\beta(w)+M_x(u,w)]x(w)\,dw}\,du \tag{31}$$

This nonlinear equation for $x(t)$ would appear to be an excellent form on which to apply the method of successive approximations. Figure 9 is a display of twenty-three successive approximations to Eq. (31) for the specific example described in Figs. 7 and 8. Forty-one points are used in this calculation, and the range is restricted to $0 \le t \le 10$. The zeroth iterant is $x_c = 1$ for the entire range, and calculated values of $x(t)$ larger than $6x_c$ are reset to x_c. The solution is seen to chip its way into the unknown like a pickax repeatedly driven into concrete. This is because the derivative of the solution at its leading edge depends on the integral of its history, so each iterant only advances the solution a small amount in time. It would be more efficient to calculate a solution by advancing forward in time rather than iterating over the entire time domain. The calculation of $x(t)$ requires iteration because $x(t)$ appears on both sides of Eq. (31). If $x(t)$ could be shown to depend only on its history, and not also on its present value, then the solution would be a recursion formula and calculations would be speedier. A solution of this type is achieved by assuming that $x(t) \approx x(t - \Delta t)$ for Δt sufficiently small. The integral in Eq. (31) is split into two terms, the first with limits $c \le u \le t - \Delta t$, and the second with limits $t - \Delta t \le u \le t$. The second integral is now approximated by a two-point trapezoid rule [$2 \times \text{integral}/\Delta t = \text{integrand}(t) + \text{integrand}(t - \Delta t)$]. The trapezoid rule integrand at t has the form

$$L_y(t,t)e^{-\Lambda(t)+\int_c^t L_x(t,w)x(w)\,dw} \tag{32}$$

where $L_y(t, u) = b(u) + M_y(t, u)$ and $L_x(t, u) = \beta(u) + M_x(t, u)$. The two-point trapezoid approximation is used again for the integral in Eq. (32), which now has the form

$$\begin{aligned}&\{L_y(t,t)e^{-\Lambda(t)+\int_c^{t-\Delta t} L_x(t,w)x(w)\,dw}\}e^{\int_{t-\Delta t}^t L_x(t,w)x(w)\,dw} \\ &\quad\approx \{\ldots\}e^{(\Delta t/2)[L_x(t,t)x(t)+L_x(t,t-\Delta t)x(t-\Delta t)]} \\ &\quad\approx \{\ldots\}e^{(\Delta t/2)[L_x(t,t)+L_x(t,t-\Delta t)]x(t-\Delta t)}\end{aligned} \tag{33}$$

The resulting approximation in place of Eq. (31) is

$$\begin{aligned}\ln\left[\frac{x(t)}{x_c}\right] =& A(t) - y_c\int_c^{t-\Delta t} L_y(t,u)e^{-\Lambda(u)+\int_c^u L_x(u,w)x(w)\,dw}\,du \\ &- \frac{y_c\Delta t}{2}L_y(t,t-\Delta t)e^{-\Lambda(t-\Delta t)+\int_c^{t-\Delta t} L_x(t-\Delta t,w)x(w)\,dw} \\ &- \left\{\frac{y_c\Delta t}{2}L_y(t,t)e^{-\Lambda(t)+\int_c^{t-\Delta t} L_x(t,w)x(w)\,dw}\right\} \\ &e^{(\Delta t/2)[L_x(t,t)+L_x(t,t-\Delta t)]x(t-\Delta t)}\end{aligned} \tag{34}$$

Notice that for $\Delta t = 0$ we recover the original equation. Here the calculation is an explicit operation moving forward in time. This result makes it much easier to calculate the example shown in Figs. 7 and 8 than by iteration (161 points span the range $0 \le t \le 20$).

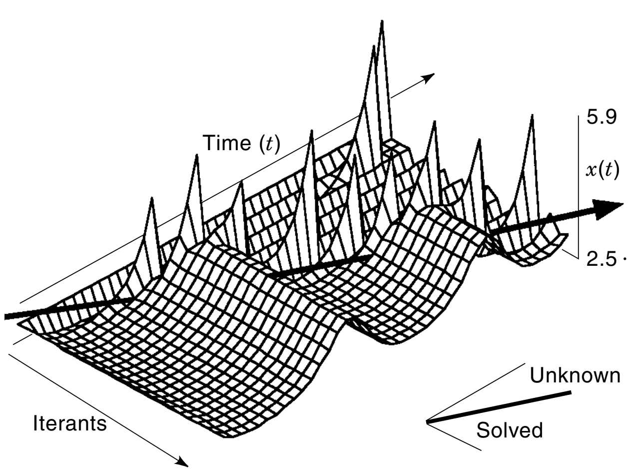

Figure 9. A sequence of successive approximations to Eq. (31) for the case shown in Figs. 7 and 8. Equation (31) is a nonlinear integral form of the Volterra predator and prey model. Each iterant builds up the solution sequentially, even though iteration occurs over the full time interval ($0 \le t \le 10$ in these calculations of 23 iterants). This is because the derivative of the solution at the present moment depends on the integral of its history. The iteration had an upper limit of six times the initial condition; any point calculated above this limit was reset to the initial condition. A better method of calculation is based only on prior events.

APPLICATIONS

Equation (1) in the sample analysis section is a form of the Boltzmann equation for the drift of a cloud of electrons along a constant electric field through a uniform molecular gas. Physical quantities are as follows: x is electron kinetic energy in units of eV, $y(x)$ is the distribution function of electron kinetic energy in units of $\text{eV}^{-3/2}$, $a(x) = a_0 x = (1/3)(E/N)^2(x/Q)$, E is the electric field in V/cm, N is the particle density of the gas in cm^{-3}, Q is the electron-molecule elastic collision cross section in cm^2, $p(x) = p_0 x = Sx$, S is the electron–molecule inelastic collision cross section in cm^2, $\alpha = 0$, $B_0(x)$ is an approximation for the logarithmic derivative of $y(x)$, β is large so $\beta B_0(x) > 1$, and $xB_0(x) > 1$ (this model of electron kinetics is for energies x above the range of thermal motion, $x \gg 0.03$ eV). Both Q and S are assumed to be only mildly dependent on x. $B_0(x)$ is given by Eq. (10) and then $y_0(x)$ is

$$y_0(x) = \exp\left\{-\int_0^x \frac{\sqrt{3Q(\xi)S(\xi)}}{E/N}\,d\xi\right\} \tag{35}$$

Typical parameters in experiments might be $Q = 10^{-15}$, $S = 3 \times 10^{-16}$, $E = 1000$, and $N = 10^{18}$. Good approximations for distribution functions of electron energy in nitrogen mixtures have been calculated from this result by using cross section data. Literature on the Boltzmann equation is vast. The pervasive approximation is that the system is never far from thermal equilibrium $f_0(x)$ ("Maxwellian" distribution), so that the nonequilibrium solution $f(x)$ is a perturbation given by an expansion $f(x) = f_0(x)(1 + \phi_1(x) + \phi_2(x) + \ldots)$, where succeeding terms are of smaller magnitude. The full development of this Chapman–Enskog method is quite involved (see Ref. 7). The electron energy distribution may be far from thermal equilibrium with the gas molecules in an electric discharge because of the high electric fields. A Chapman–Enskog expansion for electrons might require the calculation of many terms. The alternative is to expand the electron distribution function in a series of spherical harmonics defined by an axis aligned with the electric field. This creates a sequence of linked equations. Once the zeroth-order equation is solved for the leading term of the expansion, then the first-order term can be solved, and so on. The zeroth term describes the average energy of an isotropic cloud of electrons, and the first term describes the drift of this cloud along the field, a current. The result given by Eq. (35) is an approximation to the isotropic part of the electron distribution. References 8, 9, and 10 describe kinetic theory and the mathematics of the Boltzmann equation. References 11, 12, and 13 describe the theory for electrons in a gas.

The mechanical constitutive equation of a material relates the stress tensor to the deformation tensor for a solid, and to the rate of strain tensor for a fluid. Many engineering materials are characterized by linear isotropic constitutive relations: the generalized Hooke's law for solids; and the Newtonian fluid. Technology is rapidly increasing the application of nonlinear "engineered" solids and plastic "rheological" fluids. These materials can have stress dependent on the deformation and rate of strain in a nonlinear way, on higher velocity derivatives, on anisotropies of their internal structure, and on the history of their deformation and motion. In the most general case the constitutive relation is an integro-differential equation that relates stress to the entire history of the material. One example follows:

$$m\frac{d^2u(t)}{dt^2} + au(t) + \int_0^t K(t-s)\frac{du(s)}{ds}\,ds = q(t) \tag{36}$$

This is a mass-spring-damper equation with heredity in the damping term. Here u is distance, t is time, m is mass, a is the spring constant, K is a renewal kernel, and q is a forcing term. References 14 and 15 describe this equation. Reference 2 shows how to solve linear, constant-coefficient renewal equations with Laplace transforms. The solution of Eq. (36) by the methods of this article is

$$\begin{aligned} mu(t) = {} & u(0)\left[m + \int_0^t \int_0^s K(x)\,dx\,ds\right] \\ & + mp(0)t + \int_0^t (t-s)q(s)\,ds \\ & - \int_0^t u(s)\left\{a(t-s) + \int_0^{t-s} K(x)\,dx\right\} ds \end{aligned} \tag{37}$$

where $u(0)$ and $p(0)$ are the initial conditions of the displacement u and its first derivative $p = du/dt$. Notice that the kernel is a function of one variable. For $K = q = 0$ the problem collapses to a harmonic oscillator, and it is easy to show that $\sin(\sqrt{a/m}\,t)$ is a solution of the reduced form of Eq. (37). Volterra had shown that kernels of the type $K(t - s)$ produce periodic solutions (see Refs. 4, 16, and 17). In general we can expect heredity to alter the frequency of oscillations, to introduce a damping, and to shift the mean position, all during the course of time.

An integro-differential equation for heat transfer occurs when the constitutive relation between heat flux and temperature gradient in the material is a hereditary integral. In a similar way, an integro-differential equation describes the evolution of an electric field in the vicinity of a nonconducting material dielectric with a memory of its charging history (a "Maxwell–Hopkinson dielectric"). These and other applications are described in Ref. 18, a mathematician's treatise on integro-differential equations.

The technological application of mechanics with heredity and of nonequilibrium kinetics is likely to drive future efforts to improve the solution of integro-differential problems. These problems arise in the development of nonequilibrium processes, such as plasma-chemical reactors for modifying material surfaces, and in the development of synthetic materials with engineered physical properties. Another thrust to solving these equations is the desire to improve our understanding of natural phenomena and materials. It is not hard to imagine that natural flows like lava or glaciers, and natural cycles like climate and weather, can have a hereditary factor. It would be interesting to have a method for easily estimating a distant cycle-time average of a quantity influenced by heredity, be it metal fatigue or species extinction.

BIBLIOGRAPHY

1. B. Sherman, The difference-differential equation of electron energy distribution in a gas, *J. Math. Anal. Appl.,* **1**: 342–354, 1960.
2. R. Bellman and K. L. Cooke, *Differential-Difference Equations,* New York: Academic Press, 1963.
3. F. B. Hildebrand, *Methods of Applied Mathematics,* 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1965.
4. H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations,* Washington, DC: United States Atomic Energy Commission, USGPO, 1960; New York: Dover, 1962.
5. F. B. Hildebrand, *Advanced Calculus for Applications,* Englewood Cliffs, NJ: Prentice-Hall, 1962.
6. J. Matthews and R. L. Walker, *Mathematical Methods of Physics,* 2nd ed., Menlo Park, CA: W. A. Benjamin, 1970.
7. S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases,* 3rd ed., London: Cambridge University Press, 1970.
8. T. I. Gombosi, *Gaskinetic Theory,* New York: Cambridge University Press, 1994.
9. C. Cercignani, *Mathematical Methods in Kinetic Theory,* 2nd ed., New York: Plenum Press, 1990.
10. C. Cercignani, *Theory and Application of the Boltzmann Equation,* New York: Elsevier, 1975.
11. T. Holstein, Energy distribution of electrons in high frequency gas discharges, *Phys. Rev.,* **70** (5, 6): 1946.
12. W. L. Nighan, Electron energy distributions and collision rates in electrically excited N_2, CO, CO_2, *Phys. Rev. A,* **2** (5): 1989–2000, 1970.
13. B. E. Cherrington, *Gaseous Electronics and Gas Lasers,* New York: Pergamon Press, 1979.
14. E. Volterra, On elastic continua with hereditary characteristics. *J. Appl. Mech.,* **18**: 273–279, 1951.

15. V. Volterra, Sur la théorie mathématique des phénomenès héréditaires. *J. Math. Pures Appl.,* **7**: 249–298, 1928.
16. V. Volterra, *Leçons sur les équations intégrales et les équations integro-differentielles,* Paris: Gauthier-Villars, 1913.
17. V. Volterra, *Leçons sur la théorie mathématique de la lutte pour la vie,* Paris: Gauthier-Villars, 1931.
18. F. Bloom, *Ill-Posed Problems for Integrodifferential Equations in Mechanics and Electromagnetic Theory,* Philadelphia: SIAM, 1981.

MANUEL GARCIA
Lightning Publications