The McKendrick Partial Differential Equation and its Uses in Epidemiology and Population Study

Barbara Lee Keyfitz
Department of Mathematics
University of Houston
Houston, Texas 77204-3476

Nathan Keyfitz
American Academy of Arts and Sciences
Initiatives on Children
136 Irving Street
Cambridge, Massachusetts 02318

Abstract

One way of modeling the evolution in time of an age-structured population is to set up the evolution process as a partial differential equation in which time and age are the independent variables. The resulting equation, known as the McKendrick equation, has received attention recently from mathematicians. Some advantages of the PDE are that it can easily be adapted to include more detail in the model, including explicit time-dependence in the coefficients and even some nonlinear effects. The initial-boundary conditions for the McKendrick equation, imposed by the population model, are not the standard side conditions one sees in PDE theory for an evolution equation. In the simplest case, the problem reduces to a well-known model in demography, the Lotka integral equation.

In this paper we explain the solution of the McKendrick model and compare the McKendrick equation with other common models for age-structured populations (the Leslie matrix and the difference equation,

as well as the integral equation) in several ways. The approaches differ in their suitability for computation, their ease of generalization, and their adaptability to different demographic objectives and other biological applications. With small intervals of age and time all forms are identical, but if the intervals are finite, differences will appear in the numerical results. The structure of solutions of the partial differential equation contributes to better understanding and computation of population models.

1 Introduction

From any age a person can go on to the next age or die, and at certain ages a woman can give birth to a child aged zero. With nothing beyond these considerations as postulates, the essential features of deterministic population mathematics can be derived. Indeed models can be derived in many ways, of which the literature contains at least four. We will speak of the integral, difference, and partial differential equations, and of the projection matrix. In appearance these are grossly dissimilar, yet since they proceed from the same starting point and end at the same conclusion, they must somehow be equivalent.

In the first part of the paper, we reproduce and compare these formulations. To keep the exposition as simple as possible, we consider the single-sex, female population model. The two-sex, female-dominated model can be derived from this in a straightforward way, which we omit. There are, of course, more complex reproductive models which make assumptions beyond those of our introductory sentence. Similar comparisons could be made for these.

The two postulates can be expressed as 'life-table data', in the form of quantities $\ell(x)$, the probability of surviving from birth to age x, and $m(x)$, the probability of a woman of age x giving birth in a unit of time.

The simplest formulation of most models assumes that these quantities depend only on age and are invariant in time. In the concluding section, we discuss dropping this assumption.

The history of the matrix formulation starts with Edwin Cannan [2] who traced lines of survivorship and birth on a Lexis diagram, the two-dimensional representation of lifelines in a space with coordinate axes of age and time, Figure 1. His arithmetical formulation was used (apparently independently) by Whelpton, [22]. Later, Bernardelli, [1], Lewis [12], and especially Leslie

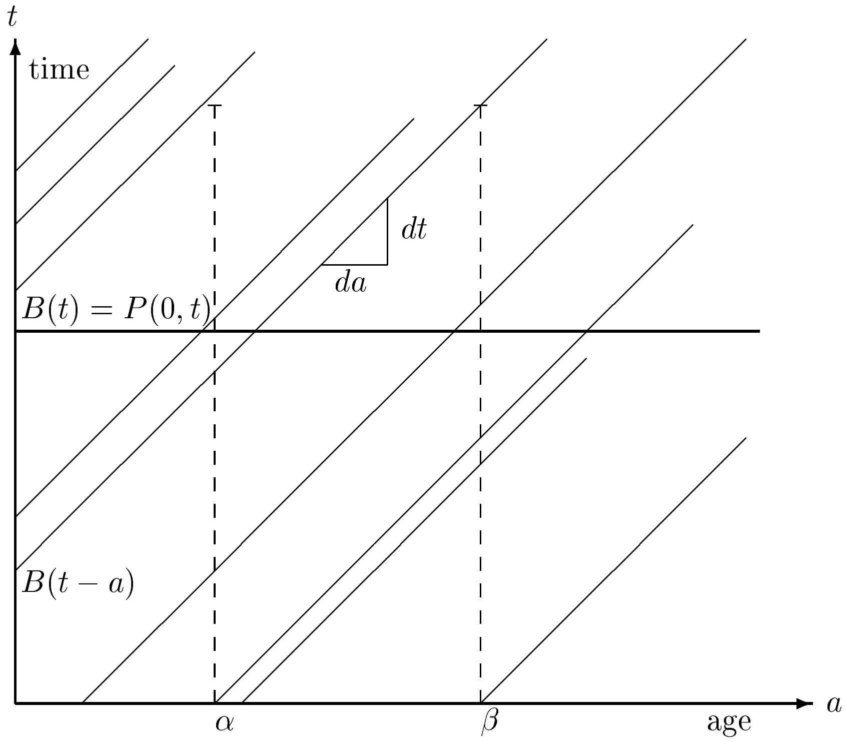

Figure 1: Lexis diagram representing mortality and reproduction in a population.

[10] showed that it was linear and readily translatable into matrix terms. Their formulation permitted ready application of the theory of matrices with non-negative elements, especially certain ergodic theorems, which had been developed early in the century.

The matrix formulation concentrates on the relation between one horizontal line and the next one in time. The matrix operator pre-multiplies the age distribution at time t, to produce the age distribution at time $t+1$. We write

$$\bar{K}^{(t+1)} = M\bar{K}^{(t)}, \tag{1}$$

representing the population as a (column) vector $\bar{K}^{(t)}$ whose j-th entry is the number of women in the j-th age group alive at time t. The entries of the projection matrix, M, are determined from the life-table $\ell(x)$ and $m(x)$, [7]. For example, when the age interval is the same as the time interval, then the $j+1$-st entry of $\bar{K}^{(t+1)}$ consists precisely of the survivors of the j-th age group in $\bar{K}^{(t)}$, so each row of M, below the first, consists of a single nonzero entry to the left of the diagonal. Thus, defining

$$m_j = \int_0^1 m(j+s)\,ds \quad \text{and} \quad L_j = \int_0^1 \ell(j+s)\,ds$$

as discrete approximations to the continuous quantities, one can write $M = S + B$, where S, the survivorship matrix, has nonzero entries

$$S_{j+1,j} = \frac{L_j}{L_{j-1}}\,. \tag{2}$$

The first row of M is formed by considering births to women alive between time t and time $t+1$. It is standard practice to approximate the number of births credited to the j-th age group in $\bar{K}^{(t)}$ by supposing the women in that group have birth rate m_j for the first half of the time interval, and the survivors to the next age have birth rate m_{j+1} for the second half. (We present a justification and some comments in the concluding section.) We thus have

$$M_{1,j} = \frac{L_0}{2}\left(m_{j-1} + \frac{L_j}{L_{j-1}} m_j\right) \tag{3}$$

as the non-zero entries in the birth matrix. Equation (1) gives an explicit formula for advancing the population through a single time-step. Using matrix analysis, one solves (1) by expressing $\bar{K}^(t)$ in terms of invariants (eigenvalues and eigenvectors) of the matrix.

The difference equation method, first published by Thompson [19] and applied by Cole [3] starts with the same part of the Lexis diagram — the relation of births at time t to those of the preceding generation. Now, iterating this relation to express all quantities in terms of the age zero cohort at earlier times, one can derive a difference equation whose order is the number of ages of mother recognized. This equation was originally solved by generating functions. In a sense the method goes back to Leonardo of Pisa, known as Fibonacci, who in the year 1202 discussed the birth of rabbits in an enclosure, starting with a single pair that produces a pair after a month and a second pair after a further month, with all generations having the same law of reproduction and no mortality for the first two months of life. The difference equation can be thought of as a method for solving the matrix equation. Fibonacci's equation, apparently the earliest formulation of modern population dynamics, had to wait several centuries for a general solution in terms of generating functions.

The first appearance of the integral equation in its demographic application seems to have been due to Sharpe and Lotka, [17]. The Lotka formulation focuses on a special part of the plane of the Lexis diagram: since we may assume that $m(x)$ is zero except for x in an interval between α and β, say, we can connect the births at a time t with the births of the mothers of those children α to β years earlier, doing so via the interval from age α to age β on the horizontal line for time t. Thus, if $B(t)$ is the number of births at time t, then

$$B(t) = \int_{\alpha}^{\beta} B(t-a)\ell(a)m(a)\,da + G(t). \tag{4}$$

Here (in Lotka's formulation), $G(t)$ simply aggregates the births at time t due to women who were in the population at time zero.

[We do not here attempt to sort out the priorities in the discovery; the interested reader may refer to a paper by Samuelson, [16].]

Lotka's expression of the relation between two successive generations of births is the homogeneous integral equation, with G absent; in our notation

$$B(0) = \int_{\alpha}^{\beta} B(t-a)\ell(a)m(a)\,da.$$

Sharpe and Lotka solved it by seeking roots of the characteristic equation,

$$\int_{\alpha}^{\beta} e^{-ra}\ell(a)m(a)\,da = 1,$$

let us say $r_1, r_2, \ldots$. They show that only one of these roots can be real, all others being complex with negative real parts, so that their disappearance in the course of time is assured. They refer to the initial population, necessary if the process is to be completely specified, but in effect they dismiss it by pointing out that in the long run only the real root counts. This solution method is not determinate, since it does not provide for a start to the process. The function $G(t)$ is one way of starting off the process; $G(t)$ disappears in the course of the 50 or so years following, but its effect lingers on through later generations, only during the course of centuries of the given fixed rates descending close to its limiting value of zero. Ultimately Lotka [13] evolved a way of solving the non-homogeneous equation, (4), starting with the homogeneous form, finding an answer that includes several arbitrary constants, and then showing how to assign the constants in such fashion that the result fits the non-homogeneous form with $G(t)$ added on the right.

If one assumes G to be known, then Feller [4] showed that (4) could be solved with Laplace transforms. Once $B(t)$ is known for all t, then the complete population distribution at any time can be found by using the survival rates.

Historically, these competing methods enjoyed a measure of success; they are compared in [9, 7]. Until recently, no one paid much attention to a fourth formulation of the problem, via a partial differential equation. In fact, however, the partial differential equation is the most detailed and explicit of the four ways of drawing out the consequences of the two postulates. It forms a good basis, in fact, for seeing why the different approaches give almost the same answers, and how they differ. We introduce the equation in the next section, and compare it with the first three approaches in the last section.

2 A Partial Differential Equation for Age-Structured Populations

The partial differential equation, which we will call the McKendrick form, is written

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial a} + \mu(a)P = 0, \tag{5}$$

where $P = P(a,t)$, for a and t nonnegative, represents the density of the population of age a at time t. This equation can be seen as a representation of the *local* effects of population dynamics; it is derived from the fact that, since age and time are measured in the same units, the rate of change with respect to time of the size of the population of age a would exactly balance the derivative with respect to age, were it not for the *force of mortality* or or instantaneous death rate, $\mu(a) = -\ell'(a)/\ell(a)$. Equation (5) is an exact representation of the dynamics in the interior of the Lexis diagram, if one assumes that the postulates are exact. The behavior along the coordinate axes is described by the boundary conditions,

$$P(a,0) = Q(a), \qquad \text{at } t=0;\ a \geq 0, \tag{6}$$

representing the initial population, and

$$P(0,t) = \int_\alpha^\beta P(x,t)m(x)\,dx = \int_0^\infty P(x,t)m(x)\,dx\,, \tag{7}$$

where m is zero outside $[\alpha, \beta]$, representing the newborns at each t. The standard method for solving an equation of the type (5), a first-order, linear partial differential equation, is to note the existence of distinguished curves in the a-t plane, called characteristics, along which (5) reduces to an ordinary differential equation. That is, representing the curve parametrically as $t = t(s)$, $a = a(s)$, and representing P evaluated on the curve by

$$\widetilde{P}(s) = P(a(s), t(s))$$

then

$$\frac{d\widetilde{P}}{ds} = \frac{d}{ds}P(a(s), t(s)) = \frac{\partial P}{\partial a}\frac{da}{ds} + \frac{\partial P}{\partial t}\frac{dt}{ds}\,. \tag{8}$$

Now, if we choose the curves $t(s)$ and $a(s)$ by the conditions

$$\frac{dt}{ds} = 1, \quad \frac{da}{ds} = 1, \tag{9}$$

then the right-hand side of (8) becomes the differential operator in (5), and so

$$\frac{d\widetilde{P}}{ds} + \mu(a(s))\widetilde{P} = 0 \tag{10}$$

along the curves in (9).

This general method works even if the coefficients in (5) are variable and even (with some modifications) if they depend on P. In the present case it is clear that the significance of the solution curves of (9) is that they are the diagonal lines of the Lexis diagram. In fact, equation (5) was formulated by noting how the derivative of P in the direction of (9) contributes to the population balance.

Now the solution of (5) is found using the method of characteristics as follows.

If (a_0, t_0) is any point in the first quadrant, and P_0 the value of P there, then P is determined along the entire solution to (9),

$$t = t_0 + s; \quad a = a_0 + s, \tag{11}$$

by

$$\begin{aligned}\widetilde{P}(s) &= P_0 \exp\left(-\int_0^s \mu(a_0+\sigma)\,d\sigma\right) = P_0 \exp\left(-\int_{a_0}^{a_0+s} \mu(\rho)\,d\rho\right) \\ &= P_0 \frac{\ell(a_0+s)}{\ell(a_0)} \end{aligned} \tag{12}$$

since

$$\ell(x) = \exp\left(-\int_0^x \mu(\alpha)\,d\alpha\right).$$

Thus we obtain a unique solution to (5) for all $a \geq 0$ and $t \geq 0$ if we specify the value of P along a *non-characteristic curve* — a curve which intersects each diagonal line non-tangentially exactly once. The positive a- and t-axes, taken together, constitute such a curve.

To find $P(a,t)$ below the diagonal $t = a$, we take $t_0 = 0$, $a_0 \geq 0$, $P_0 = Q(a_0)$ in (11) and (12), and obtain

$$\widetilde{P}(s) = P(a_0+s, s) = Q(a_0)\frac{\ell(a_0+s)}{\ell(a_0)},$$

and so, since $t = s$ and $a = a_0 + s$ in (11), we have

$$P(a,t) = Q(a-t)\frac{\ell(a)}{\ell(a-t)}, \qquad \text{if } t \leq a\,. \tag{13}$$

Above the diagonal we take $t = t_0$, $a_0 = 0$, $P_0 = P(0,t_0)$, and so (refchr2) and (12) become

$$\widetilde{P}(s) = P(s, t_0+s) = P(0,t_0)\ell(s)$$

or

$$P(a,t) = P(0,t-a)\ell(a)\,, \qquad \text{if } t > a\,. \tag{14}$$

We have not yet used the boundary condition (7). When we do so, we obtain

$$\begin{aligned} P(0,t) &= \int_0^t P(x,t)m(x)\,dx + \int_t^\infty P(x,t)m(x)\,dx \\ &= \int_0^t P(0,t-x)\ell(x)m(x)\,dx \\ &\quad + \int_t^\infty Q(x-t)\frac{\ell(x)}{\ell(x-t)}m(x)\,dx, \end{aligned} \tag{15}$$

substituting (14) and (13) for the two integrands. Now $P(0,t) = B(t)$, the births at time t, and the second integral represents the births at time t to women who were in the population at time zero. This quantity is exactly what we called $G(t)$ in (4). Hence (15) becomes (4), one of the forms of the Lotka integral equation.

Thus we see that using the method of characteristics to solve the partial differential equation leads directly to the integral equation (4). Formulating the problem via the McKendrick equation is a conceptual, rather than a practical, advance. It contributes to the theory a more thoroughgoing rationale. The partial differential equation is more transparent in its derivation than the renewal equation, and lends itself easily to variations. For example, one could make μ both time- and age-dependent, and one could consider the effect of migration with an equation similar to (5):

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial a} = -\mu(a,t)P + f(a,t), \tag{16}$$

and the same solution procedure as before would lead to an equation analogous to (4). As a consequence of this clarity and flexibility, the McKendrick equation is now becoming a basic tool of mathematical modeling of nonlinear phenomena in demography and epidemiology. The reader is referred to the monographs of Hoppensteadt, [6], and Webb, [21] for some of the realizations of this approach.

3 Comparison of the Approaches

The partial differential equation (5) is often referenced as von Foerster [20]. Yet in fact it is given by McKendrick [14] fully 30 years earlier. In the pioneering and little-known article McKendrick considers various transitions important in epidemology, and the simple case, a homogeneous epidemic with non-recovery, is one of several. Yet his equation [14, p. 122] is our (5) virtually in the same notation. Unwitting rediscovery by those who have not read the literature is common in science, especially in fields where the literature is scattered through many unlikely sources. Few demographers regularly follow the *Proceedings of the Edinburgh Mathematical Society*. Joel E. Cohen is the scholar who first drew our attention to the McKendrick article. As mentioned in the previous section, mathematicians working on nonlinear population models have also rediscovered the McKendrick formulation and are using it as a springboard for generalizations to nonlinear, interacting or nonequilibrium populations, [6, 21].

It is possible to unify the different approaches along the following lines. We claim that, in the context of arguing from the initial two postulates, the partial differential equation approach is fundamental. Conceptually, it expresses exactly the *local* consequences of the first postulate and the *global* consequences of the second, while distinguishing between the nature of the two rather dramatically. The local postulate of population evolution at positive ages is expressed by (5), a standard 'balance of mass' or 'kinematic' equation, familiar to all mathematicians who work in continuum modeling. Integration by the method of characteristics is equally familiar.

On the other hand, the boundary condition (7) is unusual for first-order hyperbolic equations, and that it would lead to (4) is somewhat surprising. (By contrast, the reduction of a partial differential equation to an integral equation on the boundary is a common solution feature for elliptic equations such as the potential equation.)

In particular, the integral equation (4) involves a *compact operator*; as a consequence, it is characterized by a discrete spectrum of eigenvalues, and the solution can be expressed as a superposition of eigenfunctions. In fact, this is a principal tool used in solving it. The fact that the problem given by (5), (6) and (7) is equivalent to (4) shows that it, too, involves a compact operator. Since this need not be the case for a partial differential equation on an unbounded region, the boundary condition (7) has the remarkable property of making the problem compact. This is not true of standard boundary conditions for hyperbolic partial differential equations. (One interesting consequence is that the inverse problem is not well-posed, [15].)

The projection and matrix problems also have discrete spectra; however, they are finite-dimensional, and all finite-dimensional problems have this property. The fact that they are approximations to a compact operator helps justify the observation that the spectra are related to each other.

Aside from straightening out the fact that McKendrick's work was rediscovered unknowingly by von Foerster and cited as his, we draw attention to the different formulations. If the two discrete forms are taken to limiting values as the interval of time and age shrinks to zero, the four methods are identical. Suppose one thinks of the discrete models as approximations (numerical approximations, for example) to the McKendrick model. In this light, there is no reason to couple the 'mesh spacings' of age and time, or even to discretize both simultaneously. Although it is not in common use for age-structured single-sex models, a 'semi-discrete' model

$$\frac{dK(t)}{dt} = A\bar{K}(t), \tag{17}$$

can be derived from (5) by simply aggregating the age dependence of P into intervals while retaining the time continuum. The population vector $\bar{K}$ will have the same meaning as before but is now augmented continuously, with A a matrix expressing the rates of change. Ordinary differential equation models, in the manner of the celebrated logistic equation, are particularly amenable to incorporating nonlinear effects, and semi-discrete systems which generalize (17) are ubiquitous as models for interacting species and qualitative models for nonlinear processes in both population study and epidemiology, [6, 21].

Standard finite-difference methods for approximating partial differential equations, especially characteristic methods, suggest that the Leslie matrix, in (1), with the coefficients determined by (2) and (3), is but one example of an approximation procedure. A simple generalization, in which the increments for age and time need not be identical, is well-known, [7] (in fact, the natural scales are different, since population characteristics are usually reported by individual years of age in five- or ten-year units of time between censuses). One obtains (1), which is an explicit formula for advancing to the next unit of time, by differencing both (5) and the boundary condition (7). A standard (from the numerical analysis viewpoint) discretization of (7) would result in a fully explicit (past) or fully implicit (present) equation, relating births to population at a single instant of time. In a population which is changing, these estimates are biased. The approximation (3), justified in Section 1 by heuristic reasoning, is now seen to correspond to splitting the difference in the approximation to (7), while the contribution to the survivor part of M was computed using an explicit (or forward-time) difference scheme in approximating (5). The effects of some of these numerical variants are discussed in [9, 7].

From the numerical point of view, the forward approximation is only first-order accurate, and a higher-order approximation could be devised by differencing (5) in a more sophisticated way. We suggest a procedure for doing this in the Appendix. However, differences in the results are probably not great enough to justify the more complicated coding effort. As an alternative way of deriving the matrix method, the Lotka integral can be converted to finite form by supposing the population to be concentrated at spikes at discrete ages on the horizontal line at time t. When used in place of M in (1), this gives a lower future population than does the Leslie matrix. This is consistent with the choice of an explicit scheme. The true projected value will in fact depend on the way the population is distributed within the age intervals.

Under our assumption that the McKendrick model is exact, one could classify these and other matrix and approximate integral solutions by their order of accuracy and direction of bias.

Do the four different approaches serve different purposes in the empirical study of population, or do they merely exemplify different mathematical methods? On the evidence of current research, it rather appears that they do serve diverse purposes. The Thompson-Cole difference equation, which is the least flexible of the models presented here, is mostly applied to animal populations. The Lotka equation and the McKendrick approach differ mainly in their treatment of the initial data, which is not expressed conveniently in the Lotka model. However, for many kinds of theoretical analysis, concerning ergodicity for example, the exact form of the initial population is immaterial, and the Lotka equation is exact. It is not surprising that the Lotka equation is widely used for this sort of analysis. That the Leslie matrix is used for routine population forecasting is surely a consequence of the fact that it utilizes data in the form available from standard census figures. It offers some possibility of generalization to nonstationary rates because of the existing highly developed theory of positive matrices, [5], but does not adapt well to replacing linear by nonlinear models.

The McKendrick approach was virtually not used at all, except by von Foerster in connection with cell division, until quite recently, when it has become the model of choice for mathematicians working on population dynamics and epidemiology models. Its virtues in this regard are twofold: on the one hand, its accurate distinction between local (in age and time) and nonlocal (in age, though not in time) implications of survivorship and fertility make it a very flexible tool for exploring more general assumptions of time-changing rates or nonlinear influences. In addition, it has the attribute of being exact with regard to discretization, and hence can be used to deduce a variety of approximation schemes, tailored to fit the form of the data or the computational power of the engine. Fortunately for those interested in applications, the amount of theory required to understand the solution of the basic partial differential equation model is not great. One looks forward to seeing the McKendrick equation become part of standard undergraduate courses in applied mathematics and a standard tool of mathematical demography.

4 Appendix: Use of the McKendrick Equation to Derive Numerical Approximations

If one identifies the Leslie matrix as resulting from a 'numerical approximation' to the McKendrick equation, then other approximations can be devised. Several principles govern this procedure. The main questions are how to approximate the spatial derivative and how to replace the time derivative. The standard spatial discretization in the Leslie method is a *forward* difference: $\bar{K}_j$ is the population between ages j and $j+1$. Suppose we retain this, for convenience, and then look at the McKendrick equation along the characteristics, (10). (We take advantage of the fact that we are really dealing with an ordinary differential equation here.) A 'forward Euler' approximation to (10) is

$$\widetilde{P}(s+1) - \widetilde{P}(s) = -\mu(a(s))\widetilde{P}(s)$$

and since $\widetilde{P}(s+1) = P(a(s+1), t(s+1)) = \bar{K}_{j+1}^{t+1}$, we get a survivorship matrix S with subdiagonal elements (b_j is the element in the j-th column and $j+1$-st row)

$$b_j = 1 - \mu_j\,.$$

The Leslie matrix has L_j/L_{j-1} here. In fact, μ_j would have to be approximated by numerical differentiation, [8, page 227], so one does not have exactly the same expression.

On the other hand, the improved Euler approximation is

$$\widetilde{P}(s+1) - \widetilde{P}(s) = -\frac{1}{2}\Big(\mu(a(s))\widetilde{P}(s) + \mu(a(s+1))\widetilde{P}(s+1)\Big) \qquad (18)$$

and this leads to

$$b_j = \frac{1 - \frac{\mu_j}{2}}{1 + \frac{\mu_{j+1}}{2}}\,.$$

In principle, this is second order accurate, while the original approximation is only first order. However, it may be awkward to have an expression involving derivatives of $\ell(x)$, which are harder to calculate and more sensitive to inaccuracies in the data than the integrals L_j.

Rather than replace (18) by yet higher order approximations which will suffer from the same instability, an alternative approach is to use methods such as cell averaging which look at (5) as an equation evolving the 'cell' of population between mesh points (endpoints of age cohorts). Such methods were originally devised for conservation laws; the McKendrick equation is an example of a conservation law with a source term. Current research in scientific computing is directed at devising higher-order methods of this type, see [11, 18].

References

[1] H. Bernardelli. Population waves. *Journal of Burma Research Society*, 31:1–18, 1941.

[2] E. Cannan. The probability of cessation of the growth of population in England and Wales during the next century. *Economic Journal*, 5:505–515, 1895.

[3] L. Cole. The population consequences of life history phenomena. *Quarterly Review of Biology*, 19:103–137, 1954.

[4] W. Feller. On the integral equation of renewal theory. *Annals of Mathematical Statistics*, 12:243–267, 1941.

[5] M. Golubitsky, E. Keeler, and M. Rothschild. Convergence of the age structure: Applications of the projective metric. *Theoretical Population Biology*, 7:84–93, 1975.

[6] F. Hoppensteadt. *Mathematical Theories of Populations: Demographics, Genetics and Epidemics*. Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[7] N. Keyfitz. Reconciliation of population models: Matrix, integral equation and partial fraction. *Journal of the Royal Statistical Society, Series A*, 130:61–83, 1967.

[8] N. Keyfitz. *Introduction to the Mathematics of Population*. Addison-Wesley, Reading, 1968.

[9] N. Keyfitz. *Applied Mathematical Demography*. Wiley-Interscience, New York, 1977.

[10] P. H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 33:183–212, 1945.

[11] R. J. LeVeque. *Numerical Methods for Conservation Laws.* Birkhäuser Verlag, Basel, 1992.

[12] E. G. Lewis. On the generation and growth of a population. *Sankhya, Indian Journal of Statistics*, 6:93–96, 1942.

[13] A. J. Lotka. Théorie analytique des associations biologiques. Part II. Analyse démographique avec application particulière à l'espèce humaine, 1939.

[14] A. G. McKendrick. Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–130, 1926.

[15] W. Rundell. Determining the birth function for an age structure population. *Mathematical Population Studies*, 1:377–395, 1989.

[16] P. A. Samuelson. Resolving a historical confusion in population analysis. *Human Biology*, 48:559–580, 1976.

[17] F. R. Sharpe and A. J. Lotka. A problem in age-distribution. *Philosophical Magazine*, 21:435–438, 1911.

[18] J. C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations.* Wadsworth & Brooks/Cole, Pacific Grove, 1989.

[19] W. R. Thompson. On the reproduction of organisms with overlapping generations. *Bulletin of Entomological Research*, 22:147–172, 1931.

[20] H. Von Foerster. Some remarks on changing populations. *The Kinetics of Cell Proliferation*, pages 382–407, 1959.

[21] G. F. Webb. *Theory of Nonlinear Age-Dependent Population Dynamics.* Marcel Dekker, New York, 1985.

[22] P. K. Whelpton. An empirical method of calculating future population. *Journal of the American Statistical Association*, 31:457–473, 1936.