Garrett Birkhoff and Applied Mathematics

David M. Young

Introduction

Garrett Birkhoff contributed to many areas of mathematics during his long and distinguished career. He is, of course, very well known for his work in algebra and in lattice theory. However, in this article we will focus on his work in applied mathematics, including the numerical solution of elliptic partial differential equations, reactor calculations and nuclear power, and spline approximations. We will also give a very brief discussion of his work on fluid dynamics. Additional information on Birkhoff's work in applied mathematics can be found in many of the publications listed below; see especially [11].

The author gratefully acknowledges the contributions of Richard Varga and Carl de Boor. Varga contributed the section entitled "Reactor Calculations and Nuclear Power", and de Boor contributed the section entitled "Spline Approximations".

The Numerical Solution of Elliptic Partial Differential Equations

In this section we describe two aspects of Birkhoff's work on the numerical solution of elliptic partial differential equations (PDE), his role in the automation of "relaxation methods", and his work on the dissemination of information on the numerical selection of elliptic PDE. Additional work of Birkhoff in this area is described in the section entitled "Reactor Calculations and Nuclear Power".

David M. Young is the Ashbel Smith Professor of Mathematics and Computer Sciences at The University of Texas at Austin. His e-mail address is young@cs.utexas.edu.

The Automation of Relaxation Methods

With the advent of high-speed computers in the 1940s Birkhoff became very interested in their possible use for obtaining numerical solutions to problems involving elliptic PDE. Many such problems could be "reduced", by the use of finite difference methods, to the solution of a (usually) large system of linear algebraic equations where the matrix was very sparse. However, because of the relatively low speeds and the very limited memory sizes of computers which were then available, the direct solution of such systems was usually out of the question.

On the other hand, many such large linear systems were actually being solved by R. V. Southwell and his associates in England using relaxation methods and without using computers; see [39]. Relaxation methods involve first choosing an initial guess for the unknown solution, u, at each grid point and then computing at each point the "residual", i.e., a number which measures the amount by which the linear equation for that point fails to be satisfied. One can eliminate, or "relax", the residual at a given grid point by suitably modifying the value of u at that point. (If one "overcorrects" or "overrelaxes", then the sign of the residual is changed.) Of course the residuals at nearby grid points are also changed when the value of u at a particular grid point is changed. By repeated use of relaxation a skilled person could soon achieve a situation where all of the residuals were very small and where the values of u at the grid points provided a satisfactory solution to the problem.

In the late 1940s when I asked Birkhoff for a thesis topic, he suggested that I work on the "automation" of relaxation methods. Actually there was already a systematic iteration procedure available, namely, the "Liebmann method" [34] (which is a special case of the Gauss-Seidel method). However, the Liebmann method is often exceedingly slow. Another method that was available at the time was Richardson's method [36]. This method involves the use of a number of parameters. However, at the time it was not obvious how the parameters should be chosen. (It was discovered later that by a suitable choice of the parameters, which could be found using Chebyshev polynomials, one could obtain very rapid convergence; see, e.g., [38] and [43].)

Largely as a result of the stimulus, encouragement, and many useful suggestions provided by Birkhoff, I was able to develop a method which is now called the "successive overrelaxation" (SOR) method and which is described in [41, 42]. (The SOR method was developed independently by Frankel [31], who called it the "extrapolated Liebmann" method.) The SOR method provides an order-of-magnitude improvement in convergence as compared to the Gauss-Seidel method for many linear systems corresponding to the numerical solution of elliptic PDE. Thus, for a class of problems corresponding to the Dirichlet problem the number of iterations required for convergence with the SOR method is proportional to h^{-1}, where h is the grid size, as compared with h^{-2} as required with the Gauss-Seidel method.

The SOR method, with generalizations, modifications, and extensions (see, e.g., Varga [40]), was used extensively for engineering and scientific computations for many years. Eventually it was superseded by other methods, such as preconditioned conjugate gradient methods and methods based on the use of Chebyshev polynomials.

Further discussion of Birkhoff's role in the automation of relaxation methods can be found in [11].

Dissemination of Information on the Numerical Solution of Elliptic PDE

Birkhoff was very active in the dissemination of information on the numerical solution of elliptic PDE. This activity included the preparation of a book with Robert Lynch (see [16]) and playing a leading role in the arranging of two conferences on "Elliptic Problem Solvers". The first of these conferences was held in Santa Fe in 1980 and led to a publication; see [37]. The second conference was held in Monterey in 1982 and also led to a publication; see [20].

The book with Lynch provides an excellent survey of many topics, including formulations of typical elliptic problems and classical analysis, difference approximations, direct and iterative methods, variational methods, finite element methods, integral equation methods, and a description of the ELLPACK software package. The book contains a wealth of information and is recommended reading for anyone interested in working in this area.

The two conferences provided, among other things, forums for discussions about the ELLPACK software package that was being developed at Purdue University by John Rice and his associates. Contributions to ELLPACK were made by a number of other institutions. For example, several iterative programs were contributed by The University of Texas.

David Kincaid and David Young, who directed the development at The University of Texas of the ITPACK software package for solving large sparse linear systems by iterative methods, regard Birkhoff as the "godfather" of the project. For several years he had been patiently but seriously suggesting that such a package be developed. The implementation of his idea was delayed in part by uncertainty as to how to choose the iteration parameters, such as omega for the SOR method, and how to decide when to terminate the iteration process. Eventually, as described in the book by Hageman and Young [32] and in the paper by Kincaid et al. [33], these and other obstacles were largely overcome and the ITPACK software package was completed.

Reactor Calculations and Nuclear Power

Garrett Birkhoff was intimately associated with reactor computations which played an essential role in the design of nuclear power reactors. This arose primarily from his role as a consultant to the Bettis Atomic Power Laboratory from 1955 through the early 1960s.

As a brief background, analytical models of nuclear reactors were brand new in the early 1950s, unlike the case of analytical fluid dynamics, which had enjoyed two hundred years of development. Fortuitously, high-powered digital computers were also making their appearance in the early 1950s. Because building full-scale nuclear reactors was both expensive and very time consuming, it was prudent and farsighted then to look to digital computers to numerically solve the associated nuclear reactor models. Even more fortuitous was the simultaneous emergence in 1950 of David M. Young's thesis [41], which contained an analytic treatment of the SOR iterative method for numerically solving second-order elliptic boundary problems.

In that exciting period when nuclear reactors were first being considered for naval ships, Bettis hired in 1954 five new Ph.D.s—Harvey Amster, Elis Gelbard, and Stanley Stein in physics, and Jerome Spanier and Richard Varga in mathematics—all of whom made contributions to various aspects of nuclear reactor theory. There is no doubt that detailed discussions with the energetic consultant, Garrett Birkhoff, helped solidify many of their emerging ideas. Garrett loved the challenge of working in new research areas, and his enthusiasm was infectious!

But Garrett's contributions to reactor theory and reactor computations were much more than just the random discussions of a consultant with Bettis people. Three solid contributions of his stand out. Early on he saw the relevance of nonnegative matrices (or, more generally, operators which leave a cone invariant) to nuclear reactor theory, and this can be seen in his publications [3] and [21]. In the latter paper the now well-known terms *essentially nonnegative* and *essentially positive* matrices, as well as *supercritical*, *critical*, and *subcritical* multiplicative processes, were first introduced. Second, while SOR-type iterative methods were being used for solving reactor problems at Bettis, alternating-direction (implicit), or ADI, iterative methods were similarly used for solving reactor problems at the Knolls Atomic Power Laboratory. The superiority of ADI iterative methods over the SOR method had been shown by Peaceman and Rachford [35] and by Douglas and Rachford [30], both for special Laplace-type problems in a rectangle. Garrett observed, in a classroom lecture at Harvard University, that the *commuting nature* of certain matrices may not hold in regions other than a rectangle, a property implicitly used in [30] and in [35]. This observation was the impetus for two research papers, [22] and [28], where many positive and negative results for such ADI schemes were presented.

Garrett was also very much interested in *semi-discrete* approximations of time-dependent problems, such as the heat-conduction equation; here "semi-discrete" means that time remains a continuous variable while other variables, usually the space variables, are discretized. This was researched in his paper [28], where Padé approximations to the function $\exp(z)$ were connected with time-stepping schemes for parabolic-like partial differential equations.

In no uncertain terms, Garrett Birkhoff, through his own research and his collaboration with others, left an indelible mark on nuclear reactor theory.

Spline Approximation

Birkhoff materially influenced the early development of spline theory and practice through his consulting work for General Motors Research. This work started in 1959 when General Motors decided that perhaps widespread use of nuclear energy was not just around the corner and needed some other useful problems for some of the members of its Nuclear Engineering Department to work on. One of the problems posed was the mathematical representation of automobile surfaces in order to exploit the recently developed numerically controlled milling machines for the cutting of dies needed for the stamping of outer and inner panels. The idea was to determine the free parameters in a suitably flexible mathematical model so as to fit closely to measurements taken from the finished physical model of the car. There was also the hope that eventually the design process itself could be carried out entirely on computers.

Birkhoff was quick to recommend the use of cubic splines (i.e., piecewise cubic polynomial functions with two continuous derivatives) for the representation of smooth curves. He was familiar with their use in naval design through his contact with the David Taylor Model Basin, and he also knew of their use at Boeing through a report written by MacLaren. Furthermore, in joint work with Henry Garabedian (see [14]) he developed what we would now call a four-mode, twelve-parameter C^1 macro finite element consisting of eight harmonic polynomial pieces, as a bivariate generalization of cubic spline interpolation, capable of interpolating a C^1 surface to a given rectangular mesh of cubic splines. This method eventually led de Boor to the now standard method of bicubic spline interpolation.

Subsequently, W. J. Gordon of General Motors Research developed the technique of spline blending for fitting smooth surfaces to an arbitrary (rectangular) smooth mesh of curves. This method too has become standard. Some mathematical aspects of blending are taken up in [15].

Birkhoff observed that the cubic spline is a good approximation to the draftsman's (physical) spline only when the latter is nearly flat. He contributed to the mathematical understanding of a more accurate model of the latter; see [26]. His insight into mechanics also made it obvious to him that a cubic spline which vanishes at all its modes must necessarily have exponential growth in at least one direction. The resulting paper [12] on the error in cubic spline interpolation was the first one to demonstrate and make use of the exponential decay of the fundamental functions of spline interpolation for "reasonable" breakpoint sequences.

The survey paper [13] provides a very good record for the many and wide-ranging suggestions concerning interpolation and approximation to univariate and bivariate data which Birkhoff made in those early days.

Somewhat later, in [4], a paper on local spline approximation by moments, Birkhoff proposed what is probably the first spline quasi-interpolant, i.e., a method of approximation that is local, stable, and aims only at reproducing all polynomials of a certain degree (rather than at matching function values).

Birkhoff's method is now treated as a special case of the de Boor-Fix quasi-interpolant. Already the above-mentioned survey contains detailed ideas about the use of splines in the numerical solution of integral and differential equations. The case of eigenvalue calculations for second-order ordinary differential equations via the Rayleigh-Ritz method is worked out in detail in [27], while the use of tensor-product splines in the numerical solution of partial differential equations is examined in [29] and in other work by Schultz. Since rectangular meshes cannot handle all practically important situations, Birkhoff also investigated splines on triangular meshes in [8, 1, 17]. The theme of multivariate interpolation was taken up one more time, but this time by Birkhoff the algebraist in [9].

Numerical Fluid Dynamics

In this section a very brief discussion of Birkhoff's work in numerical fluid dynamics will be given. For additional information the reader should see his two books, which are cited below, as well as his survey article [10].

Birkhoff worked extensively in numerical fluid dynamics, especially from the middle 1940s to the late 1950s. He was greatly influenced by the work of John von Neumann in fluid dynamics and in the then-emerging field of high-speed computing.

In 1981 Birkhoff was invited to give the John von Neumann lecture at the SIAM meeting in Troy, New York. This lecture led to the publication of a very informative survey article in numerical fluid dynamics; see [10].

It seems truly unfortunate that Birkhoff will not be around to witness the many advances in numerical fluid dynamics which will undoubtedly take place in the next twenty-five to fifty years and which in many cases will benefit from his ideas.

References

[1] R. E. Barnhill, G. Birkhoff, and W. J. Gordon, *Smooth interpolation in triangles*, J. Approx. Theory **8** (1973), 114–128.

[2] G. Birkhoff, *Hydrodynamics: A study of logic, fact, and similitude*, Princeton University Press, Princeton, NJ, 1950.

[3] ____, *Reactor criticality in neutron transport theory*, Rend. Mat. Appl. **22** (1963), 1–25.

[4] ____, *Local spline approximation by moments*, J. Math. Mech. **16** (1967), 987–990.

[5] ____, *Piecewise bicubic interpolation and approximation in polygons*, Approximation with Special Emphasis on Spline Functions (I. J. Schoenberg, ed.), Academic Press, New York, 1969, pp. 185–221.

[6] ____, *Tricubic polynomial interpolation*, Proc. Nat. Acad. Sci. **68** (1971), 1162–1164.

[7] ____, *The role of modern algebra in computing*, Computers in Algebra and Number Theory (G. Birkhoff and Marshall Hall, eds.), Proc. SIAM-AMS Sympos. Appl. Math., New York, 1970, vol. IV, 1971, pp. 1–47.

[8] ____, *Piecewise analytic interpolation and approximation in triangulated domains*, Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (A. K. Aziz, ed.), Academic Press, New York, 1972, pp. 363–385.

[9] ____, *The algebra of multivariate interpolation*, Constructive Approaches to Mathematical Models (C. V. Coffman and G. J. Fix, eds.), Academic Press, New York, 1979, pp. 345–363.

[10] ____, *Numerical fluid dynamics*, SIAM Rev. **65** (1983), 1–34.

[11] ____, *Fluid dynamics, reactor computations, and surface representation*, A History of Scientific Computing (S. G. Nash, ed.), ACM Press, New York, 1990, pp. 63–87.

[12] G. Birkhoff and C. de Boor, *Error bounds for spline interpolation*, J. Math. Mech. **13** (1964), 827–835.

[13] ____, *Piecewise polynomial interpolation and approximation*, Approximation of Functions (H. L. Garabedian, ed.), Elsevier, New York, 1965, pp. 164–190.

[14] G. Birkhoff and H. L. Garabedian, *Smooth surface interpolation*, J. Math. Phys. **39** (1960), 258–268.

[15] G. Birkhoff and W. J. Gordon, *The draftsman's and related equations*, J. Approx. Theory **1** (1968), 199–208.

[16] G. Birkhoff and R. E. Lynch, *Numerical solution of elliptic problems*, SIAM Studies in Appl. Math., vol. 6, SIAM, Philadelphia, PA, 1984.

[17] G. Birkhoff and L. Mansfield, *Compatible triangular finite elements*, J. Math. Anal. Appl. **47** (1974), 531–553.

[18] G. Birkhoff and A. S. Priver, *Hermite interpolation errors for derivatives*, J. Math. Phys. **46** (1967), 440–447.

[19] ____, *Optimal smoothing of Gaussian periodic data*, Indiana Univ. Math. J. **21** (1967), 103–113.

[20] G. Birkhoff and A. Schoenstadt, eds., *Elliptic problem solvers*, II, Academic Press, Orlando, FL, 1984.

[21] G. Birkhoff and R. S. Varga, *Reactor criticality and non-negative matrices*, J. Soc. Indust. Appl. Math. **6** (1958), 354–377.

[22] ____, *Implicit alternating direction methods*, Trans. Amer. Math. Soc. **92** (1959), 13–24.

[23] ____, *Discretization errors for well-set Cauchy problems*, J. Math. Phys. **44** (1965), 1–23.

[24] ____, eds., *Numerical solutions of field problems in continuum physics*, SIAM-AMS Proceedings, Academic Press, New York, 1970.

[25] G. Birkhoff and E. H. Zarantonello, *Jets, wakes and cavities*, Academic Press, San Diego, 1957.

[26] G. Birkhoff, C. de Boor, B. Swartz, and B. Wendroff, *Rayleigh-Ritz approximation by piecewise cubic polynomials*, SIAM J. Numer. Anal. **3** (1966), 188–203.

[27] G. Birkhoff, H. Burchard, and D. Thomas, *Non-linear interpolation by splines, pseudo splines and elastica*, General Motors Res. Publ., 1965, p. 468.

[28] G. Birkhoff, M. H. Schultz, and R. S. Varga, *Piecewise Hermite interpolation in one and two variables with applications*, J. Numer. Math. **11** (1968), 232–256.

[29] G. Birkhoff, R. S. Varga, and D. M. Young, *Alternating direction implicit methods*, Advances in Computers, vol. 3 (F. Alt, ed.), Academic Press, New York, 1962, pp. 189–273.

[30] J. Douglas Jr. and H. H. Rachford Jr., *On the numerical solution of heat conduction problems in two or three space variables*, Trans. Amer. Math. Soc. **82** (1956), 421–439.

[31] S. P. Frankel, *Convergence rates of iterative treatments of partial differential equations*, MTAC **4** (1950), 65–75.

[32] L. A. Hageman and D. M. Young, *Applied iterative methods*, Academic Press, New York, 1981.

[33] D. R. Kincaid, J. R. Respess, D. M. Young, and R. G. Grimes, *ITPACK 2C: A Fortran package for solving large sparse linear systems by adaptive accelerated iterative methods*, ACM Trans. Math. Software **8** (1981), 302–322.

[34] H. Liebmann, *Die Angenaharte Ermittleung Harmonischer Functionen und Konformer Abbildungen*, Sitzungsber, Akad. Wiss. Math. Phys. K1. **47** (1918), 385–416.

[35] D. W. Peaceman and H. H. Rachford Jr., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Indust. Appl. Math. **3** (1955), 28–41, 525–547.

[36] L. F. Richardson, *The approximate arithmetical solution by finite differences of physical problems involving differential equations with an application to the stresses in a masonry dam*, Philos. Trans. Roy. Soc. London Ser. A210 (1910), 307–357.

[37] M. H. Schultz, ed., *Elliptic problem solvers*, Academic Press, New York, 1981.

[38] G. Shortley, *Use of Tschebyscheff polynomial operators in the numerical solution of boundary value problems*, J. Appl. Phys. **24** (1953), 392–397.

[39] R. V. Southwell, *Relaxation methods in theoretical physics*, Oxford University Press, 1946.

[40] R. S. Varga, *Matrix iterative analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

[41] D. M. Young, *Iterative methods for solving partial difference equations of elliptic type*, Doctoral Thesis, Harvard University, Cambridge, MA, 1950 .

[42] ____, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Amer. Math. Soc. **76** (1954), 92–111.

[43] ____, *On Richardson's method for solving linear systems with positive definite matrices*, J. Math. Phys. **32** (1954), 243–255.