NUMERICAL SOLUTIONS OF DIFFERENTIAL EQUATIONS

H. Levy and E. A. Baggott

NUMERICAL SOLUTIONS OF DIFFERENTIAL EQUATIONS

H. LEVY

M.A., D.SC., F.R.S.E.

PROFESSOR OF MATHEMATICS AT THE IMPERIAL COLLEGE OF SCIENCE, UNIVERSITY OF LONDON

AND

E. A. BAGGOTT

M.SC., A.R.C.S., D.I.C.

LECTURER IN MATHEMATICS, THE POLYTECHNIC, REGENT STREET, LONDON

DOVER PUBLICATIONS, INC.

PREFACE

THE methods described in this volume have been developed and tried out in practice during more than ten years in the Mathematical Laboratory at the Imperial College of Science and Technology. Although a great deal of it is here published for the first time, much of it has formed part of the systematic instruction of the many hundreds of students who have passed through the Department of Mathematics of that college during these years.

This, the first volume, concerns itself only with the actual solution of ordinary differential equations and the numerical examination of many of their properties. The determination of Characteristic Numbers (*Eigenwerte*) and the investigation of Orthogonal Properties in general are, however, omitted. These will be included in Vol. II, since such properties are primarily of importance in connection with the practical solution of partial differential equations. It is for this reason also that no attempt has been made to examine in detail the special properties of well-known equations (Legendre, Emden, Mathieu, etc.), except where these have illustrated general methods applicable to classes of equations of similar types.

Although frequent use has been made of Finite Difference methods, little knowledge of that subject is here required, and as far as possible such use has been accompanied by full explanations of the meaning, and sometimes of the derivation of the formulæ. The authors desire to express their appreciation of the typing assistance so generously given by Miss R. E. Taylor and the help in sketching some of the integral curves by Mr. A. W. King.

August, 1934.

FIRST AMERICAN EDITION, 1950

DOVER PUBLICATIONS, INC.
1780 BROADWAY, NEW YORK 19, N. Y.

Published in England under the title
"Numerical Studies in Differential Equations"

PRINTED AND BOUND IN THE UNITED STATES OF AMERICA

CONTENTS

CHAPTER I

PAGE

GRAPHICAL INTEGRATION OF DIFFERENTIAL EQUATIONS 1

1. General remarks.
2. Propositions relating to the integration of $\frac{dy}{dx} = f(x, y)$.
3. The isoclinals of a differential equation of the first order.
   3.1. Singular points on the curve $f(x, y) = 0$.
   3.2. The envelope locus of $f(x, y, c) = 0$.
   3.3. General discussion of the integral curves and isoclinals of a differential equation of the first order.
   3.4. Application to the practical problem of integration.
   3.5. Approximations at points to the integral curves and isoclinals.

CHAPTER II

GRAPHICAL METHODS FOR DETAILED SOLUTION . . 35

3.6. Graphical integration of $\frac{dy}{dx} = f(x, y)$.
3.61. First approximation to the solution of $\frac{dy}{dx} = f(x, y)$, subject to $x = x_0$, $y = y_0$.
3.62. Sequence method of refining the first approximate solution.
3.7. Convergence of the sequence method of successive approximations.
3.8. Extension of the range of integration.
4. Systems of simultaneous equations.
4.1. Graphical integration of a system of two simultaneous equations of the first order.
4.2. Proof of convergence of the sequence method of successive approximations applied to a system of two first order equations.
5. Extension of the methods of integration to equations of higher order.
5.1. Integration of second order equations.
5.11. The first approximation. Two methods.
5.12. Improving the accuracy of the first approximation.
5.13. Linear differential equations of the second order,
$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = R.$$
5.14. Linear differential equations of the second order of the type
$$\frac{d^2y}{dx^2} + Qy = R.$$
5.15. Integration of second order differential equations when the boundary conditions are at both ends of the range.
5.16. Convergence of the sequence method applied to $\frac{d^2y}{dx^2} = f(x, y)$ subject to $x = a, y = b$.
5.17. Convergence of the sequence method applied to $\frac{d^2y}{dx^2} = f(x, y)$ subject to $x = a, y = A;\ x = b, y = B$.

CHAPTER III

NUMERICAL SOLUTION OVER A LIMITED RANGE . . 63

6. Solution in series.
6.1. Picard's method applied to first order equations.
6.11. Convergence of the process in a special case.
6.12. Application to simultaneous equations and to equations of the second and higher orders.
6.13. Picard's method when an approximate solution is available in tabular form.
6.2. Solution by expansion as a Taylor's series. Equations of any order.
6.21. Application to simultaneous equations.
6.3. The method of Frobenius.
7. The methods of Euler, Runge, and Kutta.
7.1. Method of Euler.
7.2. Modified method of Euler.

PAGE

7.3. Method of Runge.
7.31. Piaggio's modification of Runge's method.
7.4. Kutta's modification of Runge's method.
7.41. Kutta's third-order approximations.
7.42. Kutta's fourth-order approximations.
8. Comparison of accuracy of methods.
9. Collected formulæ.
9.1. Application to simultaneous equations.
10. Method of correcting initial values.

CHAPTER IV

FORWARD INTEGRATION OF FIRST ORDER EQUATIONS . 111

11. Forward integration.
11.1 Accuracy of tabulated values.
11.2. Magnitude of increment in x.
11.3. Checking processes.
11.4. Estimation of total error over range of integration.
11.5. Changing the increment in x.
11.6. Starting the integration.
12.0. METHOD I. The Adams–Bashforth method.
12.01. METHOD II. A method of successive approximation.
12.02. METHOD III. A new method of forward integration.
12.03. METHOD IV.
12.1. METHOD V. Alternative methods not requiring a difference table.
12.11. METHOD VI.
12.12. METHOD VII.
12.2 Comparison of the methods.

CHAPTER V

SIMULTANEOUS EQUATIONS AND EQUATIONS OF THE SECOND AND HIGHER ORDERS 141

13. Simultaneous equations.
14. Equations of higher order than the first.
15. A check for second-order equations solved as two simultaneous equations.
16. Systematic methods of integrating and correcting equations of the types

$$(a)\ \frac{d^2y}{dx^2} = f(x, y);\quad (b)\ \frac{d^2y}{dx^2} = f\left(x, y, \frac{dy}{dx}\right);\quad (c)\ \left.\begin{aligned}\frac{dy}{dx} &= f(x, y, z)\\ \frac{dz}{dx} &= g(x, y, z)\end{aligned}\right\}$$

16.1. A special method of integrating equations of the type $\frac{d^2y}{dx^2} = f(x, y)$ subject to $x = x_0,\ y = y_0,\ \frac{dy}{dx} = \left(\frac{dy}{dx}\right)_0$.
17. Integration of equations of the types $\frac{d^2y}{dx^2} + \phi(x, y) = 0$ and $\frac{d^2y}{dx^2} = f\left(x, y, \frac{dy}{dx}\right)$ subject to $x = x_0,\ y = y_0;\ x = x_1,\ y = y_1$.
17.1. To obtain a more accurate solution of the equations in 17.
17.2. A sequence method applicable over a wide range.

CHAPTER VI

SPECIAL METHODS APPLICABLE TO LINEAR DIFFERENTIAL EQUATIONS OF THE SECOND ORDER . . . 182

18. Use of some properties of second order linear equations to derive any numerical solution from a particular tabulated solution.
18.1. Numerical reduction to Normal Form; Method I.
18.2. Numerical reduction to Normal Form; Method II.
19. General theorems on upper and lower bounds to the solution of $y'' + \rho(x)\, y = 0$.
20. Approximate solutions for the cases where $\rho(x)$ is expressed as (a) series in ascending powers, (b) series in descending powers, (c) Fourier series, (d) series of exponentials.
21. Differential equations involving an unspecified small constant.
22. Differential equations involving an unspecified large constant.

CHAPTER I

GRAPHICAL INTEGRATION OF DIFFERENTIAL EQUATIONS

1. General remarks.
2. Propositions relating to the integration of $\frac{dy}{dx} = f(x, y)$.
3. The isoclinals of a differential equation of the first order.
   3.1. Singular points on the curve $f(x, y) = 0$.
   3.2. The envelope locus of $f(x, y, c) = 0$.
   3.3. General discussion of the integral curves and isoclinals of a differential equation of the first order.
   3.4. Application to the practical problem of integration.
   3.5. Approximations at points to the integral curves and isoclinals.

Descriptive Process for First Order Equations

1. General remarks.

If x is an independent variable, y a dependent variable, $\frac{dy}{dx}, \frac{d^2y}{dx^2}, \ldots \frac{d^ny}{dx^n}$, the first n differential coefficients of y with respect to x, then a differential equation is a relation between all or some of the numbers

$$x, y, \frac{dy}{dx}, \ldots \frac{d^ny}{dx^n},$$

in which one differential coefficient at least occurs.

Thus
$$\frac{dy}{dx} = x^2 + y^2$$
$$(1 - x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + y + 1 = 0$$

are examples of differential equations.

The order of such an equation is defined as the order of the highest differential coefficient present in it. The two cases cited above, for example, are of orders *one* and *two* respectively.

By a solution of the differential equation is here to be understood, a function of x denoted y, such that if the differential coefficients $\frac{dy}{dx}, \frac{d^2y}{dx^2}$, etc., are found and the values inserted in the differential equation, the latter (now a function of x alone) is identically satisfied.

For example, in the second equation written above

$$(1 - x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + y + 1 = 0,$$

a solution is $y = x - 1$, since

$$\frac{dy}{dx} = 1 \quad \text{and} \quad \frac{d^2y}{dx^2} = 0,$$

and the equation becomes, when these values are inserted,

$$0 - x \times 1 + x - 1 + 1 \equiv 0.$$

In the present chapter it is proposed to regard the function y, given, not as an explicit expression written in terms of powers of x, or literal functions of x such as $\sin x$ or $\log x$, but by a graph in the plane of the two rectangular axes OX and OY. Remembering that $\frac{dy}{dx}$ from this point of view is the slope of the tangent to the curve at any point (x, y), we may suppose that from such a curve of y against x, another curve of $\frac{dy}{dx}$ against x may be drawn.

What is the most accurate and convenient method of determining such a derived graph need not for the moment be considered. It will be dealt with in a later chapter. It suffices merely here to assert that if the original curve is everywhere continuous, then the curve for $\frac{dy}{dx}$ can certainly be found. Similarly, the curves representing $\frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}$, etc., may also be derived and plotted to the same base x.

On this view of the functional form of y, it can now be stated that a solution of the differential equation is found for a range of values of x when a graph has been obtained for that range, such that when its ordinate is taken to represent y, and all the necessary derived curves for $\frac{dy}{dx}, \frac{d^2y}{dx^2} \ldots$ are obtained, and these values are inserted in the differential equation, the latter is zero at each value of x in the range.

A few matters immediately call for comment. In the first place, be it noted that the final verification of the solution is not, as in the previous case, that terms in the final form of the differential equation are explicit expressions in x, which reduce to zero in the aggregate, but are mere numbers which when collected together sum up to zero. The verification, therefore, in this case is of an arithmetic rather than of an algebraic nature.

In the second place, and as a consequence, another distinction is apparent between the two types of solution in their verification. In the case of the graphical form of the solution it is evident that a graph can at best represent a function to a restricted degree of accuracy. The limitation arises in the last resort from the severely practical difficulty of placing two points on a chart, closer together than a certain minimum distance. This implies a degree of indefiniteness in the graphical solution which may conveniently be represented as a margin of error in that function. At least a corresponding margin of error will be present in each of the graphs for $\frac{dy}{dx}, \frac{d^2y}{dx^2}$ derived from this necessarily approximate form for y. In point of fact the errors in the derived curves may be much greater, the magnitude depending especially on the particular method that is adopted in estimating the values of the differential coefficients. This will be discussed later. For the moment it suffices to remark that when these values of y and its differential coefficients are inserted in the equation the terms may not sum up to zero at each value of x, but reduce to a number which is small in comparison with the individual terms that go to make up the sum.

These considerations suggest that, for precision, a modification in our definition of a solution of the differential equation must be made.

Suppose y is graphed as a function of x and suppose, at each value of x, a small possible margin of error $\epsilon(x)$ is attached. This will specify in the $x - y$ plane not really an individual curve but a region within which the curve lies. If then within this narrow region there exists a curve such that when the values of y and $\frac{dy}{dx}$, $\frac{d^2y}{dx^2}$, etc. derived from it are inserted in the equation, the latter is accurately satisfied, we shall say that the original graph about which the region was defined is a solution of the differential equation, with a margin of error $\epsilon(x)$.

On this definition it is evident that in practice three things are required :

(i) A method of determining the approximate solution as a graph of y.

(ii) An estimate of the margin of error $\epsilon(x)$ at each value of x.

(iii) An assurance that somewhere within the band defined by y and $\epsilon(x)$ an accurate solution of the differential equation exists.

The first step in the determination of the solution of the differential equation is, if possible, to find a rough approximation to its solution and some idea of the accuracy of that approximation. This will in general be carried out by finding upper and lower limits within which the solution must lie. Whether any further examination is necessary will, of course, depend on the accuracy to which the solution is desired. The second step, therefore, consists in refining this approximate estimate, and for this purpose, as for the initial step itself, many methods are available, but the particular method that should be chosen will depend on certain factors. In the first place, the range of the independent variable for which the solution is required will affect the selection; but even more than this, the degree of accuracy with which the solution is desired over that range will exercise a predominating influence in the choice. In addition to this there are such factors as the labour involved in the actual computation by any method, whether a computing machine is or is not available, and whether the method is suitable for mechanical computation; and finally, the number of intermediate positions along the range of the independent variable for which the solution is desired. All these factors, and others more intimately connected with the exact nature of the differential equation itself, enter into the choice.

As the subject develops and alternative methods are offered for the determination of a solution these points will require to be specially noted. We proceed, therefore, to the first step, to determine a rough approximation to the solution and upper and lower bounds to its accuracy. For this purpose it is simpler to bear in mind a geometrical or a graphical interpretation of the variables and of the relationship involved in the equation.

2. Propositions relating to the integration of $\frac{dy}{dx} = f(x, y)$.

If (x, y) be the co-ordinates of a point lying on a curve in a plane, then $p \equiv \frac{dy}{dx}$ represents the slope of the tangent at (x, y) to that curve.

A differential equation of the first order

$$p = f(x, y) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

attaches a certain direction to every point in the plane.

If a family of curves can be found such that at every point of every member, condition (1) is satisfied, then that family is called the integral family of curves of the differential equation.

Since a relation of the form

$$\phi(x, y, p) = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

can be reduced to a set of equations of type (1) merely by solving algebraically for p in (2), we may refer to the integral curves of (2) in the same sense.

If equation (2) is algebraic in p of order n, then at any point (x, y) there are n values of the slope, and n branches of the integral curves pass through that point. No two integral curves formed by pursuing corresponding branches across the field can meet. For such branches correspond because they are solutions of the same differential equation (1), and p is then uniquely determined at each point. It follows that through a given point there can be one and only one solution of the equation (1).

Some useful consequences follow immediately from these considerations.

(i) In the two equations

$$\frac{dy}{dx} = F(x, y) \quad \text{and} \quad \frac{dz}{dx} = F(x, z)$$

if $y_0 = z_0$ at $x = x_0$, then $y \equiv z$, since the solutions are unique.

Hence if at $x = x_0$, $y_0 > z_0$, then everywhere $y > z$, since they cannot cross.

(ii) In the equations

$$\frac{dy}{dx} = F(x, y) \quad \text{and} \quad \frac{dz}{dx} = \lambda F(x, z)$$

if $\lambda > 1$ and $y_0 = z_0$ at $x = x_0$, then $z > y$ from there onwards.

The two solutions have only one point in common, viz. the starting-point $x = x_0$, $y = y_0$, $z = z_0$.

(iii) In the equations

$$\frac{dy}{dx} = F(x, y) \quad \text{and} \quad \frac{dz}{dx} = \phi(x) \,.\, F(x, z)$$

where

$$\phi(x) > \lambda > 1$$

then if y and z start at a common point, from there onwards $z > y$ and they never meet again.

(iv) Throughout a region of the (x, y) plane which includes a point through which the solution of

$$\frac{dy}{dx} = f(x, y)$$

is required, if

$$M(x, y) > f(x, y) > m(x, y),$$

then the required solution lies intermediately between that of

$$\frac{dy}{dx} = M(x, y)$$

and

$$\frac{dy}{dx} = m(x, y).$$

Example 1.—The solution of

$$\frac{dy}{dx} = (x + ye^{-y})$$

which passes through (0, 0) and lies in the first quadrant, between $y = 0$ and $y = 0{\cdot}5$ is intermediate between that of

$$\frac{dy}{dx} = x + y$$

and

$$\frac{dy}{dx} = x + ye^{-0\cdot5}.$$

The solutions of these linear equations are respectively

$$y = -1 - x + e^x$$

and

$$y = e[e^{xe^{-0\cdot5}} - 1] - xe^{0\cdot5}$$

At the limits of the range, viz. at $x = 0{\cdot}5$, these, therefore, provide as upper and lower limits to the value of y, the two numbers 0·1487 and 0·1390, whose mean, 0·1438 is certainly less than 3·5 per cent. in error.

Example 2.—The solution of

$$\frac{dy}{dx} = \frac{x^2 + y^2}{x^2 + a^2} \equiv \frac{x^2}{x^2 + a^2}\left(1 + \frac{y^2}{x^2}\right)$$

for the range a to λa lies between

$$\frac{dy}{dx} = M\left(1 + \frac{y^2}{x^2}\right) \quad \text{and} \quad \frac{dy}{dx} = m\left(1 + \frac{y^2}{x^2}\right),$$

which are both soluble equations, since

$$\tfrac{1}{2} = m < \frac{x^2}{x^2 + a^2} < M = \frac{\lambda^2}{1 + \lambda^2}.$$

(v) In the two equations

$$\frac{dy}{dx} = F(x, y) \quad \text{and} \quad \frac{dz}{dx} = F(x, z) + \lambda$$

where $\lambda > 0$, if at $x = x_0$, $y = y_0$, $z = z_0$, then $z > y$ and always remains so, since where they would meet again $\frac{dz}{dx}$ would be again greater than $\frac{dy}{dx}$.

(vi) In the equation

$$\frac{dy}{dx} = F(x, y) + \psi(x)$$

if

$$M > \psi(x) > m > 0,$$

then the solution of the equation lies between that of

$$\frac{dy}{dx} = F(x, y) + m \quad \text{and} \quad \frac{dy}{dx} = F(x, y) + M.$$

Example 3.—In the equation

$$\frac{dy}{dx} = \frac{x^2 + y^2}{x^2 + a^2} = \frac{x^2}{x^2 + a^2} + \frac{y^2}{x^2 + a^2}$$

the solution lies between those of

$$\frac{dy}{dx} = \frac{y^2}{x^2 + a^2} \quad \text{and} \quad \frac{dy}{dx} = 1 + \frac{y^2}{x^2 + a^2}$$

taking the range of x to be 0 to ∞.

The simple methods just outlined may suffice to indicate roughly the general trend of any particular solution, although they are not necessarily very useful in practice. This arises from the fact that in effect we have made the determination of the upper and lower bounds of the solution dependent on the solution of two other equations generally of simpler type. Where these are themselves not directly capable of solution the process does not assist us much, if at all. We turn, therefore, to a consideration of yet another method, in this case of a graphical nature, which will provide us not only with a knowledge, detailed up to a certain degree, of any one particular solution of the equation, but with a picture of the general trend of all the solutions of the equation. For many purposes it is precisely this general knowledge that is required rather than detailed information about any one special solution passing through a particular point.

3. Isoclinals of a differential equation of the first order.

It is necessary that the conclusions, which will presently be drawn in the general discussion of the equation $\phi(x, y, p) = 0$ from a geometrical standpoint, should be expressed in concise mathematical form, and for this purpose the two following paragraphs are required.

3.1. *Singular points on the curve $f(x, y) = 0$.*

Consider the equation $f(x, y) = 0$.

If this be solved for y in terms of x and this value of y be substituted in the equation

$$u = f(x, y),$$

u will vanish identically. It follows, therefore, that the derivatives of u with regard to x will also vanish.

Thus

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

or

$$f_x + f_y \cdot y' = 0,$$

i.e.

$$y' = -f_x/f_y \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.1)$$

For any curve $f(x, y) = 0$, (3.1) in general determines the value of the gradient y'.

Differentiating (3) again,

$$f_{xx} + 2f_{xy} \cdot y' + f_{yy} \cdot y'^2 + f_y \cdot y'' = 0 \quad . \quad . \quad . \quad (4)$$

Now (3.1) fails to determine y' when $f_x = f_y = 0$; in which case

$$f_{xx} + 2f_{xy}\, y' + f_{yy} \cdot y'^2 = 0 \quad . \quad . \quad . \quad (4.1)$$

and this equation in general determines the two possible slopes at a double point.

In these circumstances there are three equations

$$\left.\begin{aligned} f_x &= 0 \\ f_y &= 0 \\ f(x, y) &= 0 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

to be satisfied, and since two equations only are necessary to determine x and y, it is clear that any points which satisfy equations (5) are necessarily singular to the curve.

3.2. *The Envelope locus of $f(x, y, c) = 0$.*

If $f(x, y, c) = 0$ represent a system of curves obtained by giving real values to c, then

$$\left.\begin{aligned} f(x, y, c) &= 0 \\ \frac{\partial f}{\partial c} &= 0 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

represents the envelope of the system $f(x, y, c) = 0$.

Consider the intersections of the curves

$$f(x, y, c) = 0 \text{ and } f(x, y, c + \delta c) = 0.$$

$$0 = f(x, y, c + \delta c) = f(x, y, c) + \frac{\partial}{\partial c}[f(x, y, c)] \,.\, \delta c + \lambda \delta c^2.$$

where λ is finite in general.

Thus when δc tends to zero

$$\frac{\partial}{\partial c} \,.\, f(x, y, c) = 0.$$

It follows that equations (6) determine the locus of the ultimate intersections of each member of the system with the neighbouring member.

Every member of the system $f(x, y, c) = 0$ touches this locus.

For the slope at a point (x_1, y_1) on c_1 is given by

$$f_x(x_1, y_1, c_1) + f_y(x_1, y_1, c_1)\frac{dy}{dx} = 0 \quad . \quad . \quad . \quad (7)$$

The slope at the same point on the locus (6) is determined from

$$f_x(x_1, y_1, c_1) + f_y(x_1, y_1, c_1)\frac{dy}{dx} + f_c(x_1, y_1, c_1)\frac{dc}{dx_1} = 0 \,. \quad (8)$$

But from equations (6), $f_c = 0$.

Hence (7) and (8) determine identical values for $\frac{dy}{dx}$.

3.3. General discussion of the integral curves and isoclinals of a differential equation of the first order.

A first order differential equation

$$\phi(x, y, p) = 0 \quad . \quad . \quad . \quad . \quad (9)$$

besides defining an integral system also determines a system of curves

$$\phi(x, y, c) = 0 \quad . \quad . \quad . \quad . \quad . \quad (9.1)$$

This may be regarded as a system which includes the loci of successive points on the integral curves for which $\frac{dy}{dx}$