PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The Numerical Solution of Non-Singular Linear Integral Equations

L. Fox and E. T. Goodwin

Phil. Trans. R. Soc. Lond. A 1953 **245**, 501-534
doi: 10.1098/rsta.1953.0005

THE NUMERICAL SOLUTION OF NON-SINGULAR LINEAR INTEGRAL EQUATIONS

By L. FOX, D.Phil., and E. T. GOODWIN, Ph.D.

The National Physical Laboratory, Teddington, Middlesex

(*Communicated by E. C. Bullard, F.R.S.—Received* 4 *September* 1952—*Revised* 7 *November* 1952)

The integral equations discussed and illustrated are those of Fredholm, with fixed limits in the integral and including the eigenvalue problem, and of Volterra, with a variable upper limit in the integral. The methods are mostly based on finite-difference theory, the integrals being replaced by formulae for numerical quadrature. Computational details are given for several methods, and there is a discussion of error analysis for Volterra's equation. Some methods are given for accelerating the convergence of classical iterative processes.

1. Introduction

We shall follow the continental practice of referring to integral equations with fixed limits in the integral as equations of Fredholm type, and to those with a variable limit as equations of Volterra type. The equations of Fredholm type of the first and second kinds are respectively given by

$$\int_a^b k(x,y)\,f(y)\,\mathrm{d}y = g(x), \tag{1}$$

$$\int_a^b k(x,y)\,f(y)\,\mathrm{d}y = g(x) + f(x). \tag{2}$$

The corresponding equations of Volterra can be written as

$$\int_a^x k(x,y)\,f(y)\,\mathrm{d}y = g(x), \tag{3}$$

$$\int_a^x k(x,y)\,f(y)\,\mathrm{d}y = g(x) + f(x). \tag{4}$$

Another type of equation of frequent occurrence is given by

$$\lambda \int_a^b k(x,y)\,f(y)\,\mathrm{d}y = f(x). \tag{5}$$

This is an eigenvalue problem, invariably associated with fixed limits in the integral, and we shall therefore class it with equations of the Fredholm type.

In equations (1) to (4) f is the wanted function, all other functions being known either analytically, graphically or numerically. In equation (5) both the eigenfunction f and the eigenvalue λ have to be calculated.

In this paper we restrict ourselves to non-singular integral equations, by which we mean here that the limits of integration, whether fixed or variable, shall remain finite, and that both $k(x,y)$, the 'kernel', and the required f should be well behaved throughout the range of integration. More specifically, our methods require that the definite integrals should be calculable by numerical quadrature, using known formulae in the theory of finite differences.

Most of the methods presented are indeed based on the theory of finite differences, and have analogy with those used for the solution of differential equations. We have demonstrated in previous papers (Fox 1949; Fox & Goodwin 1949) methods of solving ordinary

differential equations with various boundary conditions. When the boundary conditions were specified at a single point we used step-by-step methods, and when boundary conditions were given at distinct points relaxation methods were found convenient. In the numerical solution of integral equations we find a similar difference in technique. *Fredholm's equations are conveniently treated by solving simultaneous equations, Volterra's equations by solving recurrence relations.*

In part A of the paper we discuss, with illustrative examples, the solution of Fredholm's equations (1), (2) and (5). In every case the integral is represented by a quadrature formula involving values of the wanted function f at pivotal points in the range of integration. We shall usually use a simple formula with a difference correction or truncation error of known form. For equations (1) and (2) the simple formula leads to a set of algebraic linear simultaneous equations whose solution gives a first approximation to the required pivotal values of f. This approximation can be improved by use of the difference-correction technique (Fox 1949).

The same process applied to equation (5) involves the determination of one or more latent roots and vectors of a matrix, of order equal to the chosen number of pivotal points. Standard methods are available, and relaxation is again often convenient.

For all these problems we have an alternative to the use of a difference correction. This consists in computing approximate solutions for different interval lengths, and calculating a more accurate result by the method of the 'deferred approach to the limit' (Richardson & Gaunt 1927).

We also make some observations on means of accelerating the convergence of some classical iterative methods.

Part B is concerned with the solution of Volterra's equations. The suggested procedure here is to calculate successive pivotal values by recurrence. Simple quadrature formulae could be used, a first approximation calculated, and the solution improved by the use of the difference correction (Fox & Goodwin 1949). It is often less laborious, however, to use at the outset quadrature formulae of high accuracy; it is not difficult to decide, after a few steps, the number of significant differences which must be retained in the quadrature.

Several examples, some with known solutions, are solved in illustration of the method, and there is a discussion of building-up error and its possible suppression.

It is probable that singular integral equations are of more practical importance and mathematical interest. Non-singular equations, however, are of quite frequent occurrence, and the methods given here enable accurate solutions to be obtained without a prohibitive expenditure of time and energy.

A. EQUATIONS OF FREDHOLM TYPE

In this part we shall be concerned with the solution of integral equations of the following forms:

$$\int_a^b k(x,y)\,f(y)\,\mathrm{d}y = g(x) + f(x), \tag{2}$$

$$\lambda \int_a^b k(x,y)\,f(y)\,\mathrm{d}y = f(x), \tag{5}$$

$$\int_a^b k(x,y)\,f(y)\,\mathrm{d}y = g(x). \tag{1}$$

In each case we shall express the integral as a finite sum of terms of the form $a_r k(x_s, y_r) f(y_r)$, and for this purpose we need to examine some formulae for numerical integration.

2. Numerical integration

In general, formulae using differences are preferred to those of the Lagrangian type. We can truncate the former when contributions become negligible, whereas without an examination of the differences the degree of an adequate Lagrangian polynomial is not known. The most useful finite-difference integration formulae can be written in the form

$$\frac{1}{h}\int_a^{a+nh} f(x)\,dx = \tfrac{1}{2}f_0 + f_1 + \ldots + f_{n-1} + \tfrac{1}{2}f_n + \Delta, \tag{6}$$

where $f_0 = f(a)$, $f_r = f(a+rh)$, and Δ, the difference correction, is a function of the differences of f. The form of the difference correction depends, as in the corresponding formulae of differentiation, on whether we wish to use central differences or forward and backward differences.

Using central differences we have, for the simplest case when $n = 1$ and the integral is taken between adjacent pivotal points, the formula

$$\frac{1}{h}\int_a^{a+h} f(x)\,dx = \tfrac{1}{2}(f_0 + f_1) - \tfrac{1}{12}\mu\delta^2 f_{\frac{1}{2}} + \tfrac{11}{720}\mu\delta^4 f_{\frac{1}{2}} \ldots, \tag{7}$$

and, in the general case,

$$\left.\begin{aligned} \frac{1}{h}\int_a^{a+nh} f(x)\,dx &= \tfrac{1}{2}f_0 + f_1 + \ldots + f_{n-1} + \tfrac{1}{2}f_n + \Delta, \\ \Delta &= \left(-\tfrac{1}{12}\mu\delta + \tfrac{11}{720}\mu\delta^3 \ldots\right)(f_n - f_0). \end{aligned}\right\} \tag{8}$$

Using forward differences we have, corresponding to (7), the equation

$$\frac{1}{h}\int_a^{a+h} f(x)\,dx = \tfrac{1}{2}(f_0 + f_1) - \tfrac{1}{12}\Delta^2 f_0 + \tfrac{1}{24}\Delta^3 f_0 - \tfrac{19}{720}\Delta^4 f_0 \ldots, \tag{9}$$

and, in the general case,

$$\left.\begin{aligned} \frac{1}{h}\int_a^{a+nh} f(x)\,dx &= \tfrac{1}{2}f_0 + f_1 + \ldots + f_{n-1} + \tfrac{1}{2}f_n + \Delta, \\ \Delta &= \left(-\tfrac{1}{12}\Delta^1 + \tfrac{1}{24}\Delta^2 - \tfrac{19}{720}\Delta^3 \ldots\right)(f_n - f_0). \end{aligned}\right\} \tag{10}$$

All these formulae use differences obtained from pivotal points outside the range of integration. A formula using only pivotal points inside the range can be obtained from equation (10) by recasting the difference correction into the form

$$\Delta = -\left(\tfrac{1}{12}\nabla^1 + \tfrac{1}{24}\nabla^2 + \tfrac{19}{720}\nabla^3 + \ldots\right) f_n + \left(\tfrac{1}{12}\Delta^1 - \tfrac{1}{24}\Delta^2 + \tfrac{19}{720}\Delta^3 \ldots\right) f_0. \tag{11}$$

This is the difference correction of Gregory's integration formula.

The Gregory formula is our main ally in the solution of Fredholm integral equations. We are concerned with their solution only at pivotal points inside the specified range of integration; the kernel may not even be defined outside the range, so that the use of formulae like (8) may become impossible. Central-difference formulae should not, however, be overlooked. The correction involves fewer terms, and the smaller coefficients give rise to slightly better accuracy. They are particularly useful when the kernel has different forms in different parts of the range of integration, and examples of the use of the various formulae, given in the following sections, will illustrate some possibilities.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

3. Solution of Fredholm's equations of the second kind

We treat first Fredholm's equation of the second kind, given by equation (2).

Using equation (6) for the representation of the integral, we can write equation (2) in the form

$$h[\tfrac{1}{2}k(x,0)\,f_0+k(x,1)\,f_1+\ldots+k(x,n-1)\,f_{n-1}+\tfrac{1}{2}k(x,n)\,f_n+\Delta(x)] = g(x)+f(x),$$

where $k(x,s)$ denotes the value of $k(x,y)$ at the point (x,sh). In this equation we can give any value to x for which k is defined, but the same $n+1$ values of f, at the equidistant pivotal points of the range (a,b), are present for any x. Since $f(x)$ also appears on the right, however, it is desirable to choose the x's to be pivotal points of (a,b). Taking all pivotal points into consideration we can then replace the integral equation by a set of $(n+1)$ linear simultanous equations typified by

$$h[\tfrac{1}{2}k(r,0)\,f_0+k(r,1)\,f_1+\ldots+k(r,n-1)\,f_{n-1}+\tfrac{1}{2}k(r,n)\,f_n+\Delta_r] = g_r+f_r,$$

where now $k(r,s)$ denotes the value of $k(x,y)$ at the point (rh,sh), and r takes the values $0,1,\ldots,n$. Combining the f_r on the left with those on the right, we can write the equations in the form

$$\left.\begin{aligned}
&\{1-\tfrac{1}{2}hk(0,0)\}f_0-hk(0,1)\,f_1\ldots-hk(0,n-1)\,f_{n-1}-\tfrac{1}{2}hk(0,n)\,f_n = -g_0+h\Delta_0,\\
&-\tfrac{1}{2}hk(1,0)\,f_0+\{1-hk(1,1)\}f_1\ldots-hk(1,n-1)\,f_{n-1}-\tfrac{1}{2}hk(1,n)\,f_n = -g_1+h\Delta_1,\\
&\qquad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&-\tfrac{1}{2}hk(n-1,0)\,f_0-hk(n-1,1)\,f_1\ldots\\
&\qquad\qquad+\{1-hk(n-1,n-1)\}f_{n-1}-\tfrac{1}{2}hk(n-1,n)\,f_n = -g_{n-1}+h\Delta_{n-1},\\
&-\tfrac{1}{2}hk(n,0)\,f_0-hk(n,1)\,f_1\ldots-hk(n,n-1)\,f_{n-1}+\{1-\tfrac{1}{2}hk(n,n)\}f_n = -g_n+h\Delta_n.
\end{aligned}\right\}\quad(12)$$

We will briefly consider some possibilities in the solution of these equations.

If we use the Gregory formula the Δ_r are linear functions of the f_r. If, further, we knew *a priori* the order of the last significant difference in (11), these linear functions would be known. The Δ_r in (12) could then be taken over to the left, and a set of equations produced for the f_r which would have different coefficients, and only the g_r on the right. This is effectively equivalent to the use of a Lagrangian formula for the representation of the integral. Such a method will be found convenient for Volterra's equation, where pivotal values of f are calculated *successively*, and it is relatively easy to determine the adequate Lagrangian polynomial. In Fredholm's equation most of the work, involving the approximate calculation of the *whole solution*, must be done before this information becomes available.

We might, alternatively, choose a very small interval h, with the intention of making the Δ_r negligible. This may involve the solution of a large number of linear equations, with great labour and the danger of errors through ill-conditioning. The number of equations could be reduced by using a formula like Simpson's rule, but again we cannot, without further work, say much about the accuracy of the results. Finally, we might solve the equations keeping up to fourth differences in Δ, and again keeping up to sixth differences, comparing the results. None of these methods is quite satisfactory. They involve a great deal of labour, or are uncertain as regards accuracy, or both.

An iterative approach seems to be more attractive. In this method we obtain a first approximation to the required solution by neglecting the Δ_r in equations (12), using, that

is, the simplest of all quadrature formulae. For each x we then calculate and difference the quantities $k(x, 0)f_0$, $k(x, 1)f_1$, etc., using the approximate f_r just calculated. All the Δ_r can then be calculated and inserted as corrections to the right-hand side of equations (12). Corrections to f are then obtained trivially, and the process repeated until there is no further change. In practice, convergence is extremely rapid, more than two cycles rarely being required.

We can express this technique symbolically as follows. If A denotes the square matrix of the coefficients of f_r, g the vector with coefficients g_r, and Δ the vector with components Δ_r, we solve in succession the equations

$$\begin{aligned} Af^{(0)} &= -g, \\ Af^{(1)} &= h\Delta(f^{(0)}), \\ Af^{(2)} &= h\Delta(f^{(1)}), \\ &\cdots\cdots\cdots\cdots, \end{aligned}$$

the final solution being given by

$$f = f^{(0)} + f^{(1)} + f^{(2)} + \ldots.$$

This approach is identical with that of Fox (1949) for problems involving differential equations of boundary-value type.

The choice of the interval h is of course rather arbitrary. We want to keep to a minimum the number of linear equations, but the interval must not be so large that the finite-difference equations are meaningless. Since the differences are examined, the method guards against the possibility of obtaining wrong results from this cause. It also ensures that neither too few nor too many differences are retained in the quadrature formulae, the point of truncation not being necessarily the same for all x.

4. Example 1

We now illustrate the proposed method by means of a simple example with a known solution. The equation

$$\frac{4}{\pi}\int_0^{\frac{1}{2}\pi} \cos(x-y)\, f(y)\, dy = \frac{2}{\pi}\cos x + f(x) \tag{13}$$

has the solution

$$f(x) = \sin x.$$

We will try to solve equation (13) by the method of the previous section, using the Gregory integration formula, to obtain four-decimal values of $f(x)$ at $x = 0°\,(10°)\,90°$. To nullify the effects of rounding errors in the computation, especially in the solution of the linear equations, five decimals are retained throughout.

There are here ten linear equations of type (12), but their coefficients are such that their solution is not difficult. We notice first that the matrix of coefficients, after dividing the first and last rows by 2, is symmetric about the centre of its array of elements, a property called centro-symmetric by Aitken (1939), and characterized by the relation

$$a_{r,s} = a_{n-r+1,\, n-s+1},$$

where n is the order of the matrix.

If n is even we can split the n equations into two sets of $\frac{1}{2}n$ equations, thus considerably simplifying the work of solution. If n is odd we have two sets of orders $\frac{1}{2}(n-1)$ and $\frac{1}{2}(n+1)$ respectively.

Consider, for example, the four equations of centro-symmetric form given by

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3+a_{14}x_4 &= b_1,\\ a_{21}x_1+a_{22}x_2+a_{23}x_3+a_{24}x_4 &= b_2,\\ a_{24}x_1+a_{23}x_2+a_{22}x_3+a_{21}x_4 &= b_3,\\ a_{14}x_1+a_{13}x_2+a_{12}x_3+a_{11}x_4 &= b_4. \end{aligned}\right\} \tag{14}$$

Adding respectively equations (1) and (4), (2) and (3) of the set (14) we produce two equations for the variables (x_1+x_4), (x_2+x_3) in the form

$$\left.\begin{aligned} (a_{11}+a_{14})\,(x_1+x_4)+(a_{12}+a_{13})\,(x_2+x_3) &= b_1+b_4,\\ (a_{21}+a_{24})\,(x_1+x_4)+(a_{22}+a_{23})\,(x_2+x_3) &= b_2+b_3. \end{aligned}\right\} \tag{15}$$

By subtraction of the same pairs of equations we produce two equations for the variables (x_1-x_4), (x_2-x_3) in the form

$$\left.\begin{aligned} (a_{11}-a_{14})\,(x_1-x_4)+(a_{12}-a_{13})\,(x_2-x_3) &= b_1-b_4,\\ (a_{21}-a_{24})\,(x_1-x_4)+(a_{22}-a_{23})\,(x_2-x_3) &= b_2-b_3. \end{aligned}\right\} \tag{16}$$

The required solution to (14) can be obtained trivially from the solutions of (15) and (16), and this process is quicker by a factor of four than the direct solution of (14).

The matrix will always be centro-symmetric if the kernel is an even function of $x-y$, and in this case it will also be symmetric in the usual sense, that is,

$$a_{r,s}=a_{s,r} \quad \text{or} \quad k(x,y)=k(y,x).$$

We see further that the matrices of equations (15) and (16) are then also symmetric, and the equations can be solved compactly by the method of Cholesky as illustrated by Fox (1950*b*).

Simplifications of this kind are quite common in the solution of integral equations. In particular, a symmetric kernel will always produce a symmetric matrix for the coefficients of the linear equations, a property of some importance in the eigenvalue problem of equation (5).

Returning to our integral equation (13), we first set up equations of type (12), divide the first and last rows by two, carry out the procedure resulting in (15) and (16), and obtain the following equations, in which $h=\frac{1}{18}\pi$ and $f_r=f(rh)$.

f_0+f_9	f_1+f_8	f_2+f_7	f_3+f_6	f_4+f_5	
+0·44444	−0·12872	−0·14241	−0·15178	−0·15654	$=-0{\cdot}31831+\frac{1}{2}h(\Delta_0+\Delta_9)$,
−0·12872	+0·70178	−0·32995	−0·35166	−0·36268	$=-0{\cdot}73750+h(\Delta_1+\Delta_8)$,
−0·14241	−0·32995	+0·63494	−0·38907	−0·40129	$=-0{\cdot}81597+h(\Delta_2+\Delta_7)$,
−0·15178	−0·35166	−0·38907	+0·58533	−0·42766	$=-0{\cdot}86964+h(\Delta_3+\Delta_6)$,
−0·15654	−0·36268	−0·40129	−0·42766	+0·55894	$=-0{\cdot}89689+h(\Delta_4+\Delta_5)$;

f_0-f_9	f_1-f_8	f_2-f_7	f_3-f_6	f_4-f_5	
+0·44444	−0·09012	−0·06641	−0·04067	−0·01370	$=-0{\cdot}31831+\frac{1}{2}h(\Delta_0-\Delta_9)$,
−0·09012	+0·85378	−0·10773	−0·06598	−0·02222	$=-0{\cdot}51640+h(\Delta_1-\Delta_8)$,
−0·06641	−0·10773	+0·92062	−0·04861	−0·01637	$=-0{\cdot}38049+h(\Delta_2-\Delta_7)$,
−0·04067	−0·06598	−0·04861	+0·97023	−0·01002	$=-0{\cdot}23302+h(\Delta_3-\Delta_6)$,
−0·01370	−0·02222	−0·01637	−0·01002	+0·99662	$=-0{\cdot}07847+h(\Delta_4-\Delta_5)$.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The first step is to solve these equations with the Δ_r neglected, thus obtaining first approximations $f_0^{(0)}, f_1^{(0)}, \ldots, f_9^{(0)}$. For each x we then calculate and difference the quantities $hk(x,0)f_0^{(0)}$, $hk(x,1)f_1^{(0)}$, etc., and obtain an approximation to $h\Delta_x$ from equation (11). The latter are then inserted on the right-hand side of the linear equations, and with a trivial amount of calculation corrections $f_0^{(1)}, f_1^{(1)}, \ldots, f_9^{(1)}$ are obtained. The process should be repeated until further corrections are negligible.

In this case the maximum discrepancy between $f_x^{(0)}+f_x^{(1)}$ and $\sin x$ is only one unit in the fourth decimal place. We record in table 1 the first approximation $f_x^{(0)}$, a typical table (for $x = 20°$) from which Δ_2 is calculated, the corrections $f_x^{(1)}$, and a comparison of our solution with the analytical result $\sin x$. Note that in the calculation of Δ_2 seventh differences are effectively zero and can safely be ignored.

TABLE 1

			computation of Δ_2 $[\frac{a}{9}\cos(20-y)^0 f_{(y)}^{(0)}]$									
x	$f^{(0)}$	y								$f^{(1)}$	$f^{(0)}+f^{(1)}$	$\sin x$
0°	0·00004	0°	0·00001							−0·00001	0·0000	0·0000
				+3839								
10°	0·17546	10°	0·03840		0					−0·00180	0·1737	0·1736
				3839		−463						
20°	0·34556	20°	0·07679		−463		+ 56			−0·00354	0·3420	0·3420
				3376		407		+48				
30°	0·50516	30°	0·11055		870		104		− 9	−0·00518	0·5000	0·5000
				2506		303		39		−13		
40°	0·64941	40°	0·13561		1173		143		−22	−0·00668	0·6427	0·6428
				+1333		−160		+17		+ 5		
50°	0·77394	50°	0·14894		1333		160		−17	−0·00796	0·7660	0·7660
				0		0		0		− 1		
60°	0·87494	60°	0·14894		1333		160		−18	−0·00900	0·8659	0·8660
				−1333		+160		−18				
70°	0·94937	70°	0·13561		1173		+142			−0·00976	0·9396	0·9397
				−2506		+302						
80°	0·99494	80°	0·11055		−871					−0·01022	0·9847	0·9848
				−3377								
90°	1·01028	90°	0·07678							−0·01038	0·9999	1·0000

$h\Delta_2 = 0{\cdot}00615$

5. EXAMPLE 2

We now consider a more practical example taken from a paper by E. R. Love (1949), in which he shows that the field due to two equal circular coaxial conducting disks at equal or at equal and opposite potentials, with zero potential at infinity, can be obtained from a function $f(x)$ defined by the integral equation

$$f(x) \pm \int_{-1}^{1} \frac{a}{\pi\{a^2+(x-y)^2\}} f(y)\,dy = 1, \tag{17}$$

in which the disks are of radius r and separated by a distance ar. Love shows further, by analytical methods, that there exists a unique, continuous, real and even solution, and that it can be expressed as a convergent series of the form

$$f(x) = 1 + \sum_{n=1}^{\infty} (\mp 1)^n \int_{-1}^{1} k_n(x,y)\,dy, \tag{18}$$

where the iterated kernels $k_n(x,y)$ are given by

$$k_1(x,y) = \frac{a}{\pi\{a^2+(x-y)^2\}}, \quad k_n(x,y) = \int_{-1}^{1} k_{n-1}(x,t)\,k_1(t,y)\,dt. \tag{19}$$

This method of solution is somewhat laborious, whereas the suggested finite-difference approach is relatively trivial.

Consider first the positive sign in equation (17), with $a = 1$. If we take an interval $h = \frac{1}{4}$, there are nine pivotal points and nine linear equations for solution. We find, however, that the matrix is again of centro-symmetric type, so that we can solve the equations in two lots of five and four respectively. In the four equations analogous to (16) we see, moreover, that all the constant terms are zero, so that, for the first approximation, $f_1^{(0)} = f_{-1}^{(0)}, f_{\frac{3}{4}}^{(0)} = f_{-\frac{3}{4}}^{(0)}$, etc., and the function is even. The corresponding Δ are therefore also equal, and subsequent corrections $f^{(1)}$ have the same property. The final solution is also, therefore, an even function.

The equations corresponding to (15) can be written down, and with the usual adjustments to produce symmetry and allowance for the equality of f_r and f_{-r}, we find the following five equations:

f_{-1}	$f_{-\frac{3}{4}}$	$f_{-\frac{1}{2}}$	$f_{-\frac{1}{4}}$	f_0	
0·52387	0·04724	0·04407	0·04100	0·01989	$= \frac{1}{2}(1-h\Delta_{-1}),$
0·04724	1·10407	0·10596	0·10345	0·05093	$= 1-h\Delta_{-\frac{3}{4}},$
0·04407	0·10596	1·11937	0·12583	0·06366	$= 1-h\Delta_{-\frac{1}{2}},$
0·04100	0·10345	0·12583	1·14324	0·07490	$= 1-h\Delta_{-\frac{1}{4}},$
0·01989	0·05093	0·06366	0·07490	0·53979	$= \frac{1}{2}(1-h\Delta_0).$

(20)

To four decimals we obtain the first approximate solution $f^{(0)}$, the difference corrections, for which differences up to the sixth were significant, and a more accurate solution f as shown in table 2.

TABLE 2

x	-1	$-\frac{3}{4}$	$-\frac{1}{2}$	$-\frac{1}{4}$	0
$f^{(0)}$	0·7554	0·7152	0·6838	0·6645	0·6581
$h\Delta$	−0·0001	+0·0005	+0·0008	+0·0010	+0·0010
f	0·7557	0·7149	0·6832	0·6638	0·6574

Subsequent corrections could hardly do more than change the rounding off in the fourth decimal.

If we take the negative sign in the integral equation the linear equations corresponding to (20) are only slightly affected. The coefficients of Δ change sign, the off-diagonal coefficients change sign, and the diagonal terms have the following values:

0·47612, 0·89593, 0·88063, 0·85676, 0·46021.

The equations are rather ill-conditioned compared with (20), and both solutions and difference corrections have larger values. The results are shown in table 3.

TABLE 3

x	-1	$-\frac{3}{4}$	$-\frac{1}{2}$	$-\frac{1}{4}$	0
$f^{(0)}$	1·6364	1·7470	1·8365	1·8934	1·9127
$h\Delta$	0·0014	0·0027	0·0034	0·0035	0·0035
f	1·6397	1·7520	1·8424	1·8997	1·9191

6. Other methods and comments

(i) *The deferred approach to the limit*

For casual computers unused to manipulating differences and finite-difference equations, the method used by Richardson & Gaunt (1927) for the solution of ordinary differential equations, and applied by Hartree & Womersley (1937) and Fox (1950*a*) for partial differential equations, can be applied also to the solution of integral equations.

The theory of the method depends on the fact that, under certain conditions and with the use of certain formulae, the difference between the true solution of the integral equation and the approximation obtained by using simple quadrature formulae at an interval h can be written in the form

$$E(h) = Bh^2 + Ch^4 + \ldots. \tag{21}$$

Thus if we compute approximate solutions f_r for intervals h_r, then the accurate solution f is related to the f_r by the equations

$$\left.\begin{aligned} f - f_1 &= Bh_1^2 + Ch_1^4 + \ldots, \\ f - f_2 &= Bh_2^2 + Ch_2^4 + \ldots, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{22}$$

Taking only two approximations, we ignore the term in h^4, eliminate B and find

$$f = f_2 + \frac{h_2^2}{h_1^2 - h_2^2}(f_2 - f_1). \tag{23}$$

If three approximations are available we can eliminate both B and C and find a more accurate expression for f.

The simplest approximate quadrature formula is that given by equation (6) with Δ ignored, and this is satisfactory as regards the theory underlying equation (21). The usual technique is to obtain approximate solutions for intervals h, $\frac{1}{2}h$, ... If two such solutions are known, equation (23) reduces to the simple form

$$f = f(\tfrac{1}{2}h) + \tfrac{1}{3}\{f(\tfrac{1}{2}h) - f(h)\}, \tag{24}$$

while for three approximations we have

$$f = f(\tfrac{1}{4}h) + \tfrac{19}{45}\{f(\tfrac{1}{4}h) - f(\tfrac{1}{2}h)\} - \tfrac{1}{45}\{f(\tfrac{1}{2}h) - f(h)\}. \tag{25}$$

For the second example of the last section we find the following approximations:

x	± 1	$\pm\frac{1}{2}$	0
$h = 1$	1·5980		1·8400
$h = \frac{1}{2}$	1·6268	1·8188	1·8940
$h = \frac{1}{4}$	1·6364	1·8365	1·9127

If we now extrapolate, using equation (24) for $h = \frac{1}{2}$ and $\frac{1}{4}$, we obtain the values 1·6396, 1·8424 and 1·9189, which agree very closely with those of table 3. Extrapolation at 1 and 0, using equation (25) for $h = 1$, $\frac{1}{2}$ and $\frac{1}{4}$, produces the results 1·6398 and 1·9194.

This method, though of considerable simplicity and usefulness, has the drawback that the error of the final extrapolated value is uncertain, though some idea of its magnitude can often be obtained by extrapolating in various ways. In the present problem, for example,

two kinds of extrapolation at $x = 0$ and 1 produced values differing at most by 5 units in the last figure, so that results could be given to three decimals with some confidence.

(ii) *Relaxation methods*

In the previous examples we solved the linear equations by direct methods. It may happen that the diagonal coefficients of the matrix are large compared with the off-diagonal terms. In this case iteration or relaxation methods can be used with convenience. The set of equations (20), for example, can be solved very rapidly by relaxation. After the first approximation has been computed in this way the Δ are calculated as before and inserted as 'still unliquidated residuals'.

Alternatively, relaxation can be used in conjunction with the method of the deferred approach to the limit. In this connexion it is particularly useful, for solutions are relatively easily calculated at a large interval, and these serve as useful 'initial guesses' for the solution at a smaller interval.

(iii) *The use of central differences*

When the kernel exists outside the range of integration it is possible to calculate approximate values of the wanted function at points outside the range, and hence to use central-difference formulae for the calculation of the corrections Δ.

The procedure here is to calculate the first approximation, for points internal to the range, in exactly the same way as before. When these values are available we can compute approximations at external points directly from the integral equation. For equation (13), for example, we could compute results for $x = -z$ direct from the equation

$$f(-z) = \frac{4}{\pi}\int_0^{\frac{1}{2}\pi} \cos(z+y)\, f(y)\, \mathrm{d}y - \frac{2}{\pi}\cos z.$$

In the integration only values at internal points are used. We again ignore difference corrections, since at this stage we want the external values to fit on smoothly with the internal ones, that is, to have an error arising from the same source.

When the external approximations have been produced, more differences can be added at each end of the range in the quadratures. This has two advantages. First, we can use the more accurate central-difference integration formulae, and, more important, differences of higher order are available. The latter are clearly welcome whenever the contribution from the last available difference used in the ordinary Gregory formula is significant.

(iv) *A classical iterative method*

The solution to Love's problem given by equation (17) can also be computed iteratively from the formula

$$f^{(n)}(x) = 1 \mp \int_{-1}^{1} \frac{a}{\pi\{a^2+(x-y)^2\}} f^{(n-1)}(y)\, \mathrm{d}y, \tag{26}$$

where $f^{(n)}(x)$ is the nth approximation or iterate to the required $f(x)$. If $f^{(0)} = 1$ this method is identical with that presented by equations (18) and (19), but we could use any $f^{(0)}$ as a start, the rate of convergence depending on the difference between $f^{(0)}$ and the true $f(x)$.

If we take $f^{(0)} = 1$ in equation (26) we can calculate, by numerical quadrature using the Gregory formula, values of $f^{(1)}$ at every pivotal value of x. Repeating the process we can

next evaluate $f^{(2)}(x)$, and if the iterative cycle is convergent, the steady solution $f^{(n)}(x)$ is the required solution of our problem.

There is a considerable amount of work in this unrefined process, but we can shorten it by using methods for the acceleration of convergence of sequences. A useful device is that of Aitken (1926), who showed that, in certain circumstances, a second sequence $g^{(n)}$, obtained from a given sequence $f^{(n)}$ by the relation

$$g^{(n)} = \frac{f^{(n+1)}f^{(n-1)} - f^{(n)2}}{f^{(n+1)} + f^{(n-1)} - 2f^{(n)}}, \tag{27}$$

converges to the same value as $f^{(n)}$, and more rapidly.

In applying this to Love's problem we used an interval $h = \frac{1}{8}$ in the quadrature, to ensure the accuracy of the Gregory formula. Table 4 shows successive values of the sequences $f^{(n)}$ and $g^{(n)}$ for the three points $x = 0$, $\frac{1}{2}$ and 1. There is seen to be striking unanimity in the $g^{(n)}$, which also bear comparison with the results obtained by the previous methods. A final check on the last $g^{(n)}$, which at this stage may have a small accumulation of rounding errors, can easily be obtained by substitution in equation (26).

Table 4

case 1						case 2					
$x=0$		$x=\frac{1}{2}$		$x=1$		$x=0$		$x=\frac{1}{2}$		$x=1$	
$f^{(n)}$	$g^{(n)}$	$f^{(n)}$	$g^{(n)}$	$f^{(n)}$	$g^{(n)}$	$f^{(n)}$	$g^{(n)}$	$f^{(n)}$	$g^{(n)}$	$f^{(n)}$	$g^{(n)}$
1·0000	—	1·0000	—	1·0000	—	1·0000	—	1·0000	—	1·0000	—
0·5000	—	0·5396	—	0·6476	—	1·5000	—	1·4604	—	1·3524	—
0·7288	0·6573	0·7483	0·6832	0·8047	0·7558	1·7288	1·9191	1·6691	1·8425	1·5095	1·6394
0·6249	0·6573	0·6536	0·6831	0·7336	0·7558	1·8327	1·9189	1·7638	1·8422	1·5806	1·6398
0·6720	0·6573	0·6965	0·6831	0·7659	0·7558	1·8798	1·9190	1·8067	1·8424	1·6129	1·6395
0·6506	—	0·6770	—	0·7513	—	1·9012	—	1·8262	—	1·6275	—

Other methods, such as Euler's transformation for the acceleration of convergence of alternating series, can also be used with advantage.

7. Eigenvalue problems

We now turn to the next equation of Fredholm type, the eigenvalue problem, given by the equation

$$\lambda \int_a^b k(x,y)\, f(y)\, dy = f(x). \tag{5}$$

Solutions now exist only for discrete values of λ, and the problem is to find one or more of such λ values (eigenvalues) and the associated function $f(x)$ (eigenfunction).

Following the technique of § 3 we replace the integral equation by a set of simultaneous equations, given by

$$\left.\begin{aligned}
\{\tfrac{1}{2}k(0,0)-\mu\}f_0+k(0,1)\,f_1+\ldots+k(0,n-1)\,f_{n-1}+\tfrac{1}{2}k(0,n)\,f_n+\Delta_0 &= 0,\\
\tfrac{1}{2}k(1,0)\,f_0+\{k(1,1)-\mu\}f_1+\ldots+k(1,n-1)\,f_{n-1}+\tfrac{1}{2}k(1,n)\,f_n+\Delta_1 &= 0,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
\tfrac{1}{2}k(n-1,0)\,f_0+k(n-1,1)\,f_1+\ldots+\{k(n-1,n-1)-\mu\}f_{n-1}+\tfrac{1}{2}k(n-1,n)\,f_n+\Delta_{n-1} &= 0,\\
\tfrac{1}{2}k(n,0)\,f_0+k(n,1)\,f_1+\ldots+k(n,n-1)\,f_{n-1}+\{\tfrac{1}{2}k(n,n)-\mu\}f_n+\Delta_n &= 0,
\end{aligned}\right\} \tag{28}$$

where $\mu = 1/h\lambda$, $f_r = f(rh)$.

As before, the Δ_r are linear functions of f_r, and could be incorporated in the coefficients of f_r if we knew in advance the order of the last significant difference in the chosen quadrature formula. Again we prefer to treat the Δ_r as correction terms, whose inclusion in (28) can be used to improve an approximate solution.

The set of equations (28) can be written in the matrix form

$$(A-\mu)\,f+\Delta f=0, \tag{29}$$

and we solve this, as before, by an iterative procedure. The first step is to neglect Δ, obtaining first approximations $\mu^{(0)}$ and $f^{(0)}$ from the equation

$$(A-\mu^{(0)})\,f^{(0)}=0. \tag{30}$$

This is a standard 'latent-root' problem, and many methods are available for its treatment. We note that the equation is homogeneous in f, so that $f=0$ everywhere is a trivial solution, non-trivial solutions existing only for discrete values of μ, those for which the determinant $|A-\mu|=0$. To each such μ there corresponds a vector f, unique save for an arbitrary constant of multiplication.

If A is symmetric all the roots μ are real. In our case A can be made symmetric, *provided the kernel is symmetric*, by dividing its first and last rows by 2. This is equivalent to writing equations (29) and (30) in the forms

$$(B-D\mu)\,f+D\Delta f=0, \tag{31}$$

$$(B-D\mu^{(0)})\,f^{(0)}=0, \tag{32}$$

where B is the symmetric matrix, D a diagonal matrix with $\frac{1}{2}$ in its first and last rows, unity elsewhere.

We shall use the method of relaxation for the solution of (32). This method is particularly useful when we want to pick out a particular solution, for which we may have a rough approximation to f. The procedure, given in more detail by Fox (1948, 1949), is as follows.

(i) Make a guess at numerical values of f at all pivotal points.

(ii) Calculate an approximation to the corresponding value of μ by the use of Rayleigh's principle, giving

$$\mu=\frac{f'Bf}{f'Df}, \tag{33}$$

where a dash denotes matrix transposition.

(iii) Calculate the residuals $(B-D\mu)\,f$, and try to liquidate them by making changes in the guessed values of f. When no further improvement seems possible take the new values of f, insert in (33) and calculate a new μ, repeating the cycle until the residuals are negligible and there is no further change in μ.

(iv) To get a more accurate solution we now compute the $\Delta f^{(0)}$, insert them in (31), and obtain a better approximation to μ from the equation

$$\mu=\frac{f'(B+D\Delta)\,f}{f'Df}, \tag{34}$$

recalculate the residuals, now given by

$$R=(B-D\mu+D\Delta)\,f, \tag{35}$$

and repeat the relaxation cycle until μ settles down and residuals are again negligible.

Note that in the liquidation of the residuals, especially those of equation (35), it is essential not to vary unduly the *magnitude* of the vector f. This is most easily effected by keeping constant the largest component of f.

Again the process may have to be repeated if the Δ_r have changed further, but the number of repetitions is usually small.

8. Example 3

As a simple example, consider the equation

$$\lambda \int_0^{\frac{1}{2}\pi} \sin(x+y)\, f(y)\, dy = f(x). \tag{36}$$

It is easy to show that there are only two non-trivial solutions, given by

$$\lambda^{-1} = \tfrac{1}{2} \pm \tfrac{1}{4}\pi, \quad f(x) = \cos x \pm \sin x. \tag{37}$$

Taking an interval $h = \frac{1}{12}\pi$, there are seven pivotal points in the range, and seven equations of the form (31), given here by

$$\left.\begin{array}{ccccccc|l}
f_0 & f_1 & f_2 & f_3 & f_4 & f_5 & f_6 & \\
-\frac{1}{2}\mu & 0{\cdot}1294 & 0{\cdot}2500 & 0{\cdot}3536 & 0{\cdot}4330 & 0{\cdot}4830 & 0{\cdot}2500 & +\frac{1}{2}\Delta_0 = 0, \\
0{\cdot}1294 & 0{\cdot}5000-\mu & 0{\cdot}7071 & 0{\cdot}8660 & 0{\cdot}9659 & 1{\cdot}0000 & 0{\cdot}4830 & +\ \Delta_1 = 0, \\
0{\cdot}2500 & 0{\cdot}7071 & 0{\cdot}8660-\mu & 0{\cdot}9659 & 1{\cdot}0000 & 0{\cdot}9659 & 0{\cdot}4330 & +\ \Delta_2 = 0, \\
0{\cdot}3536 & 0{\cdot}8660 & 0{\cdot}9659 & 1{\cdot}0000-\mu & 0{\cdot}9659 & 0{\cdot}8660 & 0{\cdot}3536 & +\ \Delta_3 = 0, \\
0{\cdot}4330 & 0{\cdot}9659 & 1{\cdot}0000 & 0{\cdot}9659 & 0{\cdot}8660-\mu & 0{\cdot}7071 & 0{\cdot}2500 & +\ \Delta_4 = 0, \\
0{\cdot}4830 & 1{\cdot}0000 & 0{\cdot}9659 & 0{\cdot}8660 & 0{\cdot}7071 & 0{\cdot}5000-\mu & 0{\cdot}1294 & +\ \Delta_5 = 0, \\
0{\cdot}2500 & 0{\cdot}4830 & 0{\cdot}4330 & 0{\cdot}3536 & 0{\cdot}2500 & 0{\cdot}1294 & -\frac{1}{2}\mu & +\frac{1}{2}\Delta_6 = 0,
\end{array}\right\} \tag{38}$$

where $f_r = f(\frac{1}{12}r\pi)$, $\Delta_r = \Delta(f_r)$.

Equations (38) are of centro-symmetric type, and, as in § 3, can be split into two sets of four and three equations respectively. We have

$$\left.\begin{array}{cccc|l}
f_0+f_6 & f_1+f_5 & f_2+f_4 & f_3 & \\
0{\cdot}2500-\frac{1}{2}\mu & 0{\cdot}6124 & 0{\cdot}6830 & 0{\cdot}7071 & +\frac{1}{2}(\Delta_0+\Delta_6) = 0, \\
0{\cdot}6124 & 1{\cdot}5000-\mu & 1{\cdot}6730 & 1{\cdot}7320 & +\ (\Delta_1+\Delta_5) = 0, \\
0{\cdot}6830 & 1{\cdot}6730 & 1{\cdot}8660-\mu & 1{\cdot}9318 & +\ (\Delta_2+\Delta_4) = 0, \\
0{\cdot}3536 & 0{\cdot}8660 & 0{\cdot}9659 & 1{\cdot}0000-\mu & +\ \Delta_3 = 0,
\end{array}\right\} \tag{39}$$

and

$$\left.\begin{array}{ccc|l}
f_6-f_0 & f_5-f_1 & f_4-f_2 & \\
0{\cdot}2500+\frac{1}{2}\mu & 0{\cdot}3536 & 0{\cdot}1830 & +\frac{1}{2}(\Delta_6-\Delta_0) = 0, \\
0{\cdot}3536 & 0{\cdot}5000+\mu & 0{\cdot}2588 & +\ (\Delta_5-\Delta_1) = 0, \\
0{\cdot}1830 & 0{\cdot}2588 & 0{\cdot}1340+\mu & +\ (\Delta_4-\Delta_2) = 0.
\end{array}\right\} \tag{40}$$

Now if Δ_r is neglected, there are non-trivial solutions to (38) for seven values of μ. Equation (38) has been replaced by (39) and (40), and four roots now come from the vanishing of the determinant corresponding to (39), three from that corresponding to (40). It is easy to show that, to four decimals, all but one root of each of these determinants vanish. There are therefore only two solutions for which the eigenvalue is finite.

The non-zero roots μ are not identical, so that the root of (39) must give a trivial solution to (40), that is, for this root, $f_0 = f_6, f_1 = f_5, f_2 = f_4$, and also $\Delta_0 = \Delta_6$, etc. We can then write equations (39) in the form

$$\left.\begin{array}{llll} f_0 & f_1 & f_2 & f_3 \\ 0{\cdot}2500-\frac{1}{2}\mu & 0{\cdot}6124 & 0{\cdot}6830 & 0{\cdot}3536 \quad +\frac{1}{2}\Delta_0 = 0, \\ 0{\cdot}6124 & 1{\cdot}5000-\mu & 1{\cdot}6730 & 0{\cdot}8660 \quad +\Delta_1 = 0, \\ 0{\cdot}6830 & 1{\cdot}6730 & 1{\cdot}8660-\mu & 0{\cdot}9659 \quad +\Delta_2 = 0, \\ 0{\cdot}3536 & 0{\cdot}8660 & 0{\cdot}9659 & 0{\cdot}5000-\frac{1}{2}\mu \quad +\frac{1}{2}\Delta_3 = 0, \end{array}\right\} \tag{41}$$

and the matrix is still symmetric.

The solution of (41) corresponds to the first of the known analytical solutions. A similar argument shows that the odd solution and its corresponding μ are given by equation (40), with $f_0 - f_6$ replaced by f_0, etc.

The qualitative nature of the solutions has thus been confirmed by numerical arguments, and we now proceed to the solution of (41).

Neglecting Δ, and taking the initial guess

$$f_0 = f_1 = f_2 = f_3 = 1{\cdot}0000,$$

we find the corresponding value of μ, calculated by equation (33), to be 4·8079. After calculation of residuals and a little relaxation, we obtain

$$f = 0{\cdot}8500,\ 1{\cdot}0500,\ 1{\cdot}1500,\ 1{\cdot}2000,$$

and a new calculation of μ gives a value 4·8657. Further relaxation gives a final result

$$f^{(0)} = 0{\cdot}8492,\ 1{\cdot}0400,\ 1{\cdot}1600,\ 1{\cdot}2009, \quad \mu^{(0)} = 4{\cdot}8660.$$

We now calculate the Δ_r, for which the computation corresponding to Δ_2, using the Gregory formula, is shown in table 5. The computed values of Δ are +0·0372, 0·0456, 0·0507 and 0·0526. When these values are inserted into equation (34) we obtain the new value $\mu = 4{\cdot}9097$. The new residuals given by (35) are small and necessitate no alteration in f, the final result being

$$f = 0{\cdot}8492,\ 1{\cdot}0400,\ 1{\cdot}1600,\ 1{\cdot}2009; \quad \mu = 4{\cdot}9097, \quad \lambda^{-1} = 1{\cdot}2854.$$

This solution agrees within one unit in the last figure with the analytical solution (37).

TABLE 5

$12y/\pi$	$[\sin(\frac{1}{6}\pi+y)f(y)]$					
0	0·4246					
		+3108				
1	0·7354		− 416			
		2692		−723		
2	1·0046		1139		+310	
		1553		−413		+99
3	1·1599		1552		+409		−84
		+ 1		− 4		+15
4	1·1600		1556		+424	
		−1555		+420		
5	1·0045		−1136			
		−2691				
6	0·7354					

$\Delta_2 = +0{\cdot}0507$

The fact that the vector suffers less change than the eigenvalue when the difference correction is applied is no unusual feature of eigenvalue problems, and was noticed by Fox (1947, 1949) in connexion with the solution of differential equations. This phenomenon is not confined to oscillatory functions. Consider, for example, the integral equation

$$\lambda \int_0^1 (x^2+y^2)\, f(y)\, \mathrm{d}y = f(x). \tag{42}$$

It is easy to show that this has two eigenvalues, given by the roots of the equation

$$4\lambda^2+30\lambda-45 = 0,$$

and the eigenfunctions have the form $x^2-\frac{1}{3}+\lambda^{-1}$.

In this case, using an interval $h = \frac{1}{4}$, two applications of the difference correction are needed to produce the final λ, but the vector is correct to within one unit in the fourth decimal after the first application. Results are shown in table 6.

TABLE 6

x	first approx.	second approx.	third approx.	analytical solution
0	0·3197	0·3091	0·3090	0·3090
$\frac{1}{4}$	0·3622	0·3523	0·3522	0·3522
$\frac{1}{2}$	0·4897	0·4818	0·4818	0·4818
$\frac{3}{4}$	0·7024	0·6977	0·6977	0·6977
1	1·0000	1·0000	1·0000	1·0000
	$\lambda=1{\cdot}2292$	$\lambda=1{\cdot}2807$	$\lambda=1{\cdot}2812$	$\lambda=1{\cdot}2812$

9. EXAMPLE 4

Integral equations of eigenvalue type can be constructed, with the use of the Green's function, from a corresponding differential equation (see, for example, Lovitt 1950). Such problems are of interest since the kernel, incorporating the Green's function, has a discontinuous derivative. For example, the differential equation

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}+\frac{\lambda f}{x} = 0, \quad f = 0 \text{ for } x = 0, 4 \tag{43}$$

is equivalent to the integral equation

$$\lambda \int_0^4 \frac{k(x,y)}{y} f(y)\, \mathrm{d}y = f(x), \tag{44}$$

where

$$\left.\begin{aligned} k(x,y) &= x(1-\tfrac{1}{4}y) \quad \text{for} \quad 0\leqslant x\leqslant y,\\ &= y(1-\tfrac{1}{4}x) \quad \text{for} \quad y\leqslant x\leqslant 4. \end{aligned}\right\} \tag{45}$$

The integral must now be taken in two parts, and we write equation (44) in the form

$$\lambda\left[(1-\tfrac{1}{4}x)\int_0^x f(y)\, \mathrm{d}y + x\int_x^4 \frac{1}{y}(1-\tfrac{1}{4}y)\, f(y)\, \mathrm{d}y\right] = f(x). \tag{46}$$

One limit of integration is now variable, in distinction to all our previous examples. Equation (6) can still be used for the representation of the integral, but the Gregory formula for the difference corrections is not now convenient, especially when x in (46) is near an end of the range. There are two difference corrections $\Delta^{(1)}(x)$ and $\Delta^{(2)}(x)$, corresponding to each

of the integrals in equation (46), and we shall be able to evaluate them by using the central-difference formulae of equations (7) and (8).

We first observe, directly from the integral equation, that $f(0) = f(4) = 0$. Then, taking an interval $h = \frac{1}{2}$, we set up equations corresponding to (28), and, neglecting difference corrections, produce the following first approximation in the usual way:

x	0	$\frac{1}{2}$	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4
f	0·000	0·580	0·893	1·000	0·953	0·797	0·566	0·293	0·000

$$\lambda = 0{\cdot}9221.$$

We note in passing that this first approximation, both to f and λ, agrees within a unit in the last figure given with that obtained by Fox (1949) from the differential equation (43), using the same interval and the approximation

$$h^2 \mathrm{d}^2 f/\mathrm{d}x^2 = \delta^2 f.$$

The next step is the calculation of the difference corrections, which we can write in the form

$$\Delta(x) = (1 - \tfrac{1}{4}x)\,\Delta^{(1)}(x) + x\Delta^{(2)}(x),$$

where $\Delta^{(1)}(x)$, $\Delta^{(2)}(x)$ are respectively the corrections associated with

$$\int_0^x f(y)\,\mathrm{d}y \quad \text{and} \quad \int_x^4 \frac{1}{y}(1 - \tfrac{1}{4}y)\,f(y)\,\mathrm{d}y.$$

When the integrands are tabulated we find that either fourth or fifth differences are negligible, and other differences can be extrapolated with confidence, to enable central-difference formulae to be used for the computation of the difference corrections. Self-explanatory results are given in table 7, extrapolated differences appearing in brackets. The new application of Rayleigh's principle gives a value $\lambda = 0{\cdot}9175$, agreeing within one unit in the last figure with the analytical solution. The final residuals are positive for $x > 2$, negative for $x < 2$, suggesting that values of $f(x)$ for $x > 2$ should be increased (by about 1 unit), while those for $x < 2$ should be decreased (by about half a unit). With these adjustments the values agree with the analytical solution within about half a unit in the last figure.

TABLE 7

y	$f(y)$					$\Delta^{(1)}$	$\left[\frac{1}{y}(1-\frac{1}{4}y)f(y)\right]$				$\Delta^{(2)}$	Δ
0	0·000		(−337)		(− 9)							
		+580		(+70)		(0)						
$\frac{1}{2}$	0·580		−267		(− 9)		+0·0249	1·015		(+110)		
		+313		+61		(0)			−345		(−18)	
1	0·893		−206		− 9		0·0444	0·670		+ 92		(0)
		+107		+52		+ 2			253		−18	
$1\frac{1}{2}$	1·000		−154		− 7		0·0593	0·417		74		+5
		− 47		+45		− 4			179		13	
2	0·953		−109		−11		0·0701	0·238		61		−3
		−156		+34		+10			118		16	
$2\frac{1}{2}$	0·797		− 75		− 1		0·0777	0·120		45		+7
		−231		+33		(0)			73		9	
3	0·566		− 42		(− 1)		0·0826	0·047		36		0
		−273		(+32)		(0)			37		− 9	
$3\frac{1}{2}$	0·293		(− 10)		(− 1)		+0·0847	0·010		+ 27		(0)
									− 10		(− 9)	
4								0·000		(+ 18)		(0)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

10. The deferred approach to the limit

The method of the deferred approach to the limit can also be used for eigenvalue problems. For example, if we take the same differential equation (43), but with $f = 0$ at $x = 0$ and 1, the integral equation is

$$\lambda \int_0^1 \frac{k(x,y)}{y} f(y)\, dy = f(x),$$

where

$$k(x,y) = x(1-y) \quad \text{for} \quad 0 \leqslant x \leqslant y,$$
$$= y(1-x) \quad \text{for} \quad y \leqslant x \leqslant 1.$$

Table 8 shows the values of f and λ, at intervals of $h = \frac{1}{4}$ and $\frac{1}{8}$, obtained by using the simplest quadrature formula, with Δ neglected. The results of the extrapolation process are also given, and these values agree with the analytical solution within less than a unit in the last figure.

Table 8

$8x$	$h=\frac{1}{4}$	$h=\frac{1}{8}$	extrapolated values
0	0·000	0·000	0·000
1	—	0·609	—
2	0·940	0·937	0·936
3	—	1·049	—
4	1·000	1·000	1·000
5	—	0·835	—
6	0·593	0·594	0·594
7	—	0·307	—
8	0·000	0·000	0·000
	$\lambda=3{\cdot}743$	$\lambda=3{\cdot}688$	$\lambda=3{\cdot}670$

11. Fredholm's equation of the first kind

Treatment, both theoretical and practical, of Fredholm's first equation

$$\int_0^1 k(x,y)\, f(y)\, dy = g(x) \tag{1}$$

is difficult and complicated. It is easy to see, for some types of kernel, that solutions are possible only for a restricted class of functions $g(x)$. For example, we observe, by successive differentiation of (1), that if $k(x,y)$ satisfies a linear differential equation of the form

$$\left\{\frac{\partial^n}{\partial x^n} + p_1(x)\frac{\partial^{n-1}}{\partial x^{n-1}} + \ldots + p_n(x)\right\} k(x,y) = 0,$$

then solutions of (1) exist only if $g(x)$ satisfies the same differential equation. Thus for kernels of this kind there is no well-behaved solution for arbitrary well-behaved functions $k(x,y)$ and $g(x)$, and it is known theoretically that this is true for all regular kernels. Fredholm's equation of the first kind is then almost always a singular integral equation, in the sense of this paper.

The problem is further complicated by the fact that, given the necessary relation between k and g, there may be an infinity of solutions. For a kernel of the form $k(x,y) = \sum_1^n a_r(x)\, b_r(y)$, for example, it is clear that no solution exists unless $g(x)$ has the form

$$g(x) = \sum_1^n c_r a_r(x),$$

and then any solution is possible for which

$$\int_0^1 b_r(y)\, f(y)\, \mathrm{d}y = c_r \quad (r = 1, \ldots, n).$$

We need only consider a form for $f(y)$ which contains more than n arbitrary constants to confirm the existence of an infinity of solutions.

These phenomena must have some counterpart in any attempt at evaluating solutions by numerical methods. Our standard technique of replacing the integral equation by a set of linear algebraic equations leads in this case to the following equations:

$$\left.\begin{array}{ccccccc} f_0 & f_1 & f_2 & \ldots & f_n & & \\ \frac{1}{2}k(0,0) & k(0,1) & k(0,2) & \ldots & \frac{1}{2}k(0,n) & = \frac{1}{h}g_0 - \Delta_0, \\ \frac{1}{2}k(1,0) & k(1,1) & k(1,2) & \ldots & \frac{1}{2}k(1,n) & = \frac{1}{h}g_1 - \Delta_1, \\ \ldots & \ldots & \ldots & \ldots & \ldots & \\ \frac{1}{2}k(n,0) & k(n,1) & k(n,2) & \ldots & \frac{1}{2}k(n,n) & = \frac{1}{h}g_n - \Delta_n. \end{array}\right\} \tag{47}$$

Now suppose the kernel is a polynomial in y of degree p. We first try to solve the equations with Δ neglected. If the interval h is such that $n \geqslant p+1$, there is a linear relation between the columns (and therefore also between the rows) of the matrix of coefficients, and the matrix is singular. If g is arbitrary we cannot then find the first approximate solution; if g is such that there is a corresponding linear relation between its pivotal values an infinity of such solutions can be found.

It will be noticed that nothing is gained by including the Δ in the left of equations (47), modifying the coefficients accordingly (see §§ 3 and 14). The effect of this is to multiply each column by a constant, which will not affect the singularity of the matrix.

A solution can be obtained if the matrix is non-singular, and therefore the number of pivotal points must be at most equal to $p+1$, or $n \leqslant p$. Any function $g(x)$ will now yield a solution to the algebraic equations, and we have to consider whether it is possible to improve this solution, regarded as an approximation to the solution of the integral equation. In the first place we know that the integral equation has no solution unless $g(x)$ is of special form, so that the approximation obtained for arbitrary $g(x)$ is not likely to be 'smooth'; it will not difference, and no Δ can be calculated. If $g(x)$ has the required form then the approximation may or may not be 'smooth'. Even in the latter case, however, we are unlikely to be able to calculate an accurate Δ, since k is already a polynomial of degree p, the integrand is kf, and we have only $p+1$ pivotal points. We may, of course, improve the first approximation to some extent by using as many differences as are available with the Gregory formula, or, equivalently, by allowing for Δ from the beginning by using a formula such as Simpson's rule or one of the 'strip' formulae of integration (see, for example, Bickley 1939). The accuracy cannot, however, be specified or estimated, and only in the special case when the solution f is a constant can we calculate accurately the difference corrections and proceed to obtain an accurate solution of the integral equation.

Some alleviation of this difficulty is made possible by the choice of a different kind of quadrature formula. Typical of the proposed class is the formula of Gauss, for which

$$\int_0^1 k(x,y)\,f(y)\,dy = \sum_0^n A_r k(x,y_r)\,f(y_r), \tag{48}$$

where the A_r are known constants and the y_r are selected non-equidistant points in the range of integration, both A_r and y_r depending only on n. This formula, involving $n+1$ pivotal points, is exact if the integrand is a polynomial of degree $2n$, and therefore provides a better chance of using an exact quadrature formula while still retaining a non-singular matrix.

A further difficulty arises, for any quadrature formula, when we consider the behaviour of the kernel in the x-direction. If the kernel is a polynomial of degree m in x, there will be a linear relation between the rows if there are more than $m+1$ pivotal values of x, and the matrix is then singular.

Consideration of kernels of polynomial form is of course academic, since such equations can be solved quite simply by analysis. They do, however, indicate the kind of difficulty likely to be experienced when dealing with integral equations with regular kernels, for all finite-difference formulae of integration, including those of the Gauss type, depend for their reliability on the accuracy with which the functions involved can be represented by polynomials.

In illustration of the arguments presented above we proceed to consider some simple integral equations.

(1)
$$\int_0^1 (x+y)\,f(y)\,dy = g(x). \tag{49}$$

If we take more than two algebraic equations the matrix is singular because the kernel is linear in y, and we find numerically, what is obvious analytically, that g must also be linear. Using the simple quadrature formula, we take pivotal points at 0 and 1 and obtain the equations

$$\tfrac{1}{2}f_1 = g_0 - \Delta_0,$$
$$\tfrac{1}{2}f_0 + f_1 = g_1 - \Delta_1.$$

The first approximation is $f_1 = 2g_0$, $f_0 = 2g_1 - 4g_0$, and this cannot be improved. It is a correct solution of the integral equation only if f is a constant, say unity, for which

$$g(x) = x + \tfrac{1}{2}.$$

If $g(x) = \frac{1}{2}x + \frac{1}{3}$, a correct solution is given by $f(x) = x$. This cannot be found by using the simple quadrature formula, but is obtained from the Gauss formula for $n = 1$, for which the equations are given by

$$(x=0) \quad 0{\cdot}5y_1 f(y_1) + 0{\cdot}5y_2 f(y_2) = \tfrac{1}{3},$$
$$(x=1) \quad 0{\cdot}5(1+y_1)\,f(y_1) + 0{\cdot}5(1+y_2)\,f(y_2) = \tfrac{5}{6},$$

where
$$y_1 = \tfrac{1}{2} - \frac{1}{2\sqrt{3}}, \quad y_2 = \tfrac{1}{2} + \frac{1}{2\sqrt{3}}.$$

Solving these equations we find

$$f(y_1) = \tfrac{1}{2} - \frac{1}{2\sqrt{3}}, \quad f(y_2) = \tfrac{1}{2} + \frac{1}{2\sqrt{3}},$$

which is correct.

(2)
$$\int_0^1 (x+y^5)\,f(y)\,dy = g(x). \tag{50}$$

Here neither the simple quadrature formula nor any Gauss formula could produce an accurate solution, since the kernel is linear in the x-direction, and no more than two pivotal points can be used.

(3) $$\int_0^1 \sqrt{(x^2+y^2)}\, f(y)\, dy = g(x). \tag{51}$$

The solution $f(x) = x$ corresponds to $g(x) = \frac{1}{3}\{(1+x^2)^{\frac{3}{2}} - x^3\}$.

Consider first the approximate solution obtained by using the simplest quadrature formula, with Δ neglected, at an interval $h = \frac{1}{4}$. We obtain the following results:

$f(0)$	$f(\frac{1}{4})$	$f(\frac{1}{2})$	$f(\frac{3}{4})$	$f(1)$
0·0299	0·2832	0·3995	0·9580	0·6886

These values are not very close to the true solution, and in particular are not smooth, so that Δ cannot be calculated.

Using a more accurate integration formula of the same type, that of Simpson, for which, at the same interval,

$$\int_0^1 k(x,y)\, f(y)\, dy = \tfrac{1}{12}[k(x,0)\, f(0) + 4k(x,\tfrac{1}{4})\, f(\tfrac{1}{4}) + 2k(x,\tfrac{1}{2})\, f(\tfrac{1}{2}) + 4k(x,\tfrac{3}{4})\, f(\tfrac{3}{4}) + k(x,1)\, f(1)], \tag{52}$$

we obtain the following results:

$f(0)$	$f(\frac{1}{4})$	$f(\frac{1}{2})$	$f(\frac{3}{4})$	$f(1)$
0·0448	0·2124	0·5992	0·7185	1·0329

These values are in general more accurate than the former, but further correction is impossible.

Repeating the calculations at an interval $h = \frac{1}{8}$, we obtain the following results:

	$f(0)$	$f(\frac{1}{8})$	$f(\frac{1}{4})$	$f(\frac{3}{8})$	$f(\frac{1}{2})$	$f(\frac{5}{8})$	$f(\frac{3}{4})$	$f(\frac{7}{8})$	$f(1)$
Simple quadrature	0·0144	0·1464	0·1573	0·6822	−0·2338	1·8157	−0·5076	1·7323	0·4003
Simpson's rule	0·0216	0·1098	0·2360	0·5117	−0·3507	1·3618	−0·7614	1·2992	0·6004

These solutions, obtained at a smaller interval than the former, are much less accurate, bearing little resemblance to the true solution.

Consider now the solutions obtained by the more powerful Gaussian formulae. We tabulate below the results for various values of the x of equation (48). In each case the values chosen for x subdivide the range of integration into equal intervals:

n	y_r	$f(y_r)$			
1	0·2113	0·7887			
	0·2092	0·7893			
2	0·1127	0·5000	0·8873		
	0·1113	0·5021	0·8856		
3	0·0694	0·3300	0·6700	0·9306	
	0·0698	0·3295	0·6705	0·9302	
4	0·0469	0·2308	0·5000	0·7692	0·9531
	0·0479	0·2288	0·5033	0·7645	0·9575

We find, as with the more simple quadrature formula, that, as the interval decreases, solutions at first become more accurate (and in this respect the Gauss formula is much the better of the two), but after a time begin to get worse again.

This phenomenon can be regarded in the following way. If we write the set of equations in the matrix form

$$Af = g + \Delta, \tag{53}$$

let us now consider the Δ as the residual vector $Af - g$ obtained by substituting the true solution f into equation (53). There is no doubt that, as the interval decreases, Δ also decreases, but the solution is given by

$$f = A^{-1}(g + \Delta),$$

and the matrix A^{-1} has coefficients which rapidly become larger, so that there is no reason to expect the 'correction vector' $A^{-1}\Delta$ to get smaller.

This clearly corresponds to the case of a polynomial-type kernel in which, for a small enough interval, the matrix A becomes singular. Any finite-difference method in which solutions do not steadily improve in accuracy cannot be regarded as satisfactory, since we have in general no means of determining the 'best' interval to use, and in any case we cannot estimate the accuracy of our 'best' solution.

We note also that in all our approximate solutions the error is an oscillatory function of x. Inspection of the integral equation gives the reason for this, in that $f(y)$ could have oscillatory variation without causing great changes in the $g(x)$. Accurate estimation of f would thus be expected to be difficult, and this fact has led other workers—for example, Bullard & Cooper (1948) and Kreisel (1949)—to consider the evaluation of some suitably smoothed function of f. We do not pursue these methods here, since we are concerned throughout this paper with methods of finding accurate solutions or, equivalently, of determining the error of approximate solutions.

B. EQUATIONS OF VOLTERRA TYPE

We now turn to equations of the Volterra type, in which one of the limits in the integral is variable. Here there is no eigenvalue problem, and a further simplification permits attention to be focused on the integral equation (4) of the second kind.

The simplification relates to the connexion between Volterra's equations of the first and second kinds. For non-singular kernels, the only type considered in this paper, the first equation can always be transformed into the second. For example, taking

$$\int_0^x k(x, y)\, f(y)\, \mathrm{d}y = g(x), \tag{3}$$

we obtain by differentiation the equation

$$k(x, x)\, f(x) + \int_0^x \frac{\partial}{\partial x} k(x, y)\, f(y)\, \mathrm{d}y = \frac{\mathrm{d}}{\mathrm{d}x} g(x),$$

an integral equation of the second kind unless $k(x, x) \equiv 0$. In the latter case we repeat the differentiation, producing finally the equation

$$k^{(n)}(x, x)\, f(x) + \int_0^x \frac{\partial^{n+1}}{\partial x^{n+1}} k(x, y)\, f(y)\, \mathrm{d}y = \frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} g(x), \tag{54}$$

where
$$k^{(n)}(x, x) = \left[\frac{\partial^n}{\partial x^n} k(x, y)\right]_{y=x}$$

is the first non-zero derivative of $k(x, y)$ with respect to x when y is equated with x.

In practical examples it is possible that the functions $k(x, y), g(x)$ may be given as functions of x only by graphical or numerical results, and it may be objected that it is then not possible to obtain accurate high-order derivatives. Unfortunately, it seems that in such a case a solution can be computed accurately only to the extent to which such derivatives can be obtained. As an example, consider the very simple case when $k(x, y) = (x-y)^2$ and the function $g(x)$ is only known in graphical form. The analytical solution is then $f(x) = \frac{1}{2}g'''(x)$, and no matter what method is adopted it will not be possible to obtain the solution to better accuracy than that to which the third derivation of $g(x)$ can be computed.

Part B of this paper is therefore concerned solely with methods of obtaining accurate solutions of the Volterra equation of the second kind, given by

$$\int_0^x k(x, y) f(y) \, dy = f(x) + g(x). \tag{4}$$

12. Possible methods of solution

We first consider methods of solution analogous to those used for Fredholm's equation. Using finite-difference methods, the initial step is the replacement of the integral equation by a set of algebraic linear simultaneous equations. In equation (4) we let x take the set of pivotal values $0, h, 2h, \ldots$, use an integration formula of type (6), and produce, in analogy with (12), the following set of equations:

$$\left.\begin{aligned} &f_0 &&= -g_0,\\ &-\tfrac{1}{2}hk(1, 0) f_0 + \{1-\tfrac{1}{2}hk(1, 1)\} f_1 &&= -g_1 + h\Delta_1,\\ &-\tfrac{1}{2}hk(2, 0) f_0 - hk(2, 1) f_1 + \{1-\tfrac{1}{2}hk(2, 2)\} f_2 &&= -g_2 + h\Delta_2,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{55}$$

It is not here necessary to solve the equations simultaneously; the matrix of coefficients is already in triangular form, and successive values of f are calculable immediately from successive equations. As in the case of Fredholm's equation, all finite-difference methods differ only in the treatment of the difference corrections Δ, and there are two main possibilities. First, we could get approximate solutions, at two or more different intervals and with Δ ignored throughout, producing a more accurate answer by the method of the deferred approach to the limit. Alternatively, from the approximate solution at one interval we could calculate approximate values of Δ, replace them in the linear equations, and improve the first approximation. These methods might be called indirect. Secondly, Δ could be expressed in Lagrangian form and incorporated in equations (55), an accurate solution then being obtained in one application of the solving process. This might be called the direct method, and, as already mentioned in § 3, has advantages for Volterra's equation. Before going on to discuss the direct method we shall, in the next section, treat an example by indirect methods.

13. Example 5. Indirect methods

We take an example from a paper by Friedlander (1941), in which it is required to solve the integral equation

$$\int_0^x \frac{-2}{(x-y+2)^2} f(y)\,dy = f(x) - \frac{1}{(x+2)^2}. \qquad (56)$$

(*a*) Using first the method of the deferred approach to the limit, we solve equations of type (55), with Δ neglected, for each of the intervals $h = \frac{1}{2}$, $h = \frac{1}{4}$ as far as $x = 5{\cdot}0$. The left half of table 9 shows these results, and on the right are shown the extrapolated values given by equation (24). Values at intermediate points at the smaller interval are omitted.

(*b*) The alternative indirect method is to correct a first approximation, obtained as in (*a*), by computing approximate Δ_r and including them in equations (55). Difficulties are at once apparent. The computation of Δ_1, for example, with $h = \frac{1}{2}$, necessitates accurate integration between the limits 0 and $\frac{1}{2}$, when no values of the integrand within this range are available. For some kernels integration could be effected, as in § 9, by considering points outside the range, for which approximate f are already available and k is calculable. In the present example the kernel $2(2\frac{1}{2}-y)^{-2}$ approaches its singularity too quickly to permit accurate determination of Δ_1 at this interval, and the same applies to the Δ corresponding to the first few pivotal values of x.

With $h = \frac{1}{4}$, however, the singularity is too far away to cause real difficulty, and all Δ can be calculated. Moreover, as x increases, and more tabular points are included in the interval 0 to x, it becomes possible to increase the interval to $h = \frac{1}{2}$, and later to use the more convenient Gregory formula. For this example it was possible to change the interval at $x = 2$ from $h = \frac{1}{4}$ to $h = \frac{1}{2}$, and the results after one application of the difference correction, again with intermediate points omitted, is shown in the last column of table 9.

Table 9

	approximate solutions		final values	
x	$f(x)$ $(h=\frac{1}{2})$	$f(x)$ $(h=\frac{1}{4})$	method (*a*) $f(x)$	method (*b*) $f(x)$
0	0·25000	0·25000	0·25000	0·25000
$\frac{1}{2}$	0·12444	0·12497	0·12515	0·12515
1	0·06872	0·06924	0·06941	0·06942
$1\frac{1}{2}$	0·04143	0·04184	0·04198	0·04197
2	0·02690	0·02722	0·02733	0·02733
$2\frac{1}{2}$	0·01859	0·01883	0·01891	0·01891
3	0·01352	0·01372	0·01379	0·01378
$3\frac{1}{2}$	0·01026	0·01041	0·01046	0·01046
4	0·00805	0·00818	0·00822	0·00821
$4\frac{1}{2}$	0·00649	0·00659	0·00662	0·00662
5	0·00536	0·00545	0·00548	0·00547

The discrepancy between the results of the two methods is at most one unit in the last figure. This could be due to real error, since neither method has been pursued to its final conclusion. It is worth noting, however, that it would be difficult to obtain closer agreement without retaining another figure in the calculations, since a combination of the extrapolation process with unfavourable roundings could produce an error of nearly one unit.

14. The direct method

The method we have found most effective for Volterra's equation again uses the basic equations (55), with Δ given by the Gregory formula (11). The differences in this formula are now expressed as functions of the pivotal values, so that, retaining throughout differences up to a predetermined order, a particular equation of set (55) can now be written in the form

$$\{1-ha_0k(n,n)\}f_n - h[a_1k(n,n-1)f_{n-1}+a_2k(n,n-2)f_{n-2}+\ldots+a_1k(n,1)f_1+a_0k(n,0)f_0] = -g_n, \quad (57)$$

from which f_n can be calculated directly. The a_r are numerical constants depending only on the order of differences retained in the quadrature formula.

We notice first that this formula cannot be used for the calculation of the first few values of f, since not enough differences exist for the application of Gregory's formula. Initial values can be calculated without difficulty, however, in the case of Volterra's equation, from the Taylor series, for which the derivatives are obtainable by repeated differentiation of the given integral equation. Differentiating n times, we obtain the relation

$$f^{(n)}(x)+g^{(n)}(x) = \sum_{s=0}^{n-1}\frac{\mathrm{d}^{n-1-s}}{\mathrm{d}x^{n-1-s}}\{k^{(s)}(x,x)f(x)\}+\int_0^x\frac{\partial^n}{\partial x^n}\{k(x,y)\}f(y)\,\mathrm{d}y.$$

Putting $x=0$, we find

$$f^{(n)}(0) = \sum_{s=0}^{n-1}\left[\frac{\mathrm{d}^{n-1-s}}{\mathrm{d}x^{n-1-s}}\{k^{(s)}(x,x)f(x)\}\right]_{x=0}-g^{(n)}(0), \quad (58)$$

a recurrence relation from which successive values of $f^{(n)}(0)$ can be evaluated. It should be noted that in the important case for which $k(x,y)$ is a function only of $(x-y)$, equation (58) reduces to the simpler form

$$f^{(n)}(0) = \sum_{s=0}^{n-1}k^{(s)}(0)f^{(n-1-s)}(0)-g^{(n)}(0). \quad (59)$$

When sufficient starting values are available from the Taylor series to make possible the use of Gregory's quadrature formula, successive values of $f(x)$ are calculated from equation (57). To determine the order of differences which must be retained it is necessary, for the first step, to record and difference the individual terms of the integrand $k(x,y)f(y)$, and it may be desirable to check, at a later stage, that the order retained is still adequate.

It is found that the same coefficient a_r is associated with the terms of the integrand r steps removed from either end of the range of integration. If the last difference retained, moreover, is of order p, then in general $a_s = 1$ for $s>p$. We give in table 10 some values of a_s corresponding to values of p from 1 to 6.

Table 10

a \ p	0	1	2	3	4	5	6
a_0	0·5	0·416̇	0·375	0·3486̇	0·32986̇	0·31559 19312	0·30422 45370
a_1	1·0	1·083̇	1·16̇	1·24583̇	1·32083̇	1·39217 92328	1·46038 35979
a_2	1·0	1·0	0·9583̇	0·87916̇	0·76̇	0·62397 48677	0·45346 39550
a_3	1·0	1·0	1·0	1·02638̇	1·10138̇	1·24408 06878	1·47142 85714
a_4	1·0	1·0	1·0	1·0	0·98125	0·90990 41005	0·73939 31878
a_5	1·0	1·0	1·0	1·0	1·0	1·01426 91799	1·08247 35450
a_6	1·0	1·0	1·0	1·0	1·0	1·0	0·98863 26058
a_7	1·0	1·0	1·0	1·0	1·0	1·0	1·0

These coefficients a_s require modification if in equations (11) and (57) $n \leqslant 2p$, that is, when in equation (11) any forward and any backward difference involve the same pivotal value or values. This change is necessary only for values of n lying between p and $2p$, for when n is less than p the difference of order p does not exist and the Gregory formula is not accurate. In this case either more initial values must be obtained from the Taylor series, or the interval of tabulation reduced.

The required modification is simple, the coefficient a_s being replaced by $a_s + a_{n-s} - 1$. This can be computed most effectively by tabulating a_s, with argument s, and sliding alongside a table written in reverse order, as shown in table 11. The modified coefficient is then obtained by adding adjacent entries and subtracting unity.

A similar device may be used, when $k(x, y) = k(x-y)$, to assist the calculation of the summation in equation (57). In this case $k(n, s)$ is a function of a single variable and may be written k_{n-s}. Then $a_t k_t$ is tabulated for argument t, in reverse order, and recorded on a movable strip alongside a table of $a_s f_s$ tabulated with argument s, as shown in table 12. Since $a_s = 1$ for $s > p$, the values of f_n are written down without modification as they are calculated, and the strip moved down one place to make $a_1 k_1$ adjacent to the value f_{n-1} just calculated.

TABLE 11

s	a_s	a_{n-s}
0	a_0	1
1	a_1	1
⋮	⋮	⋮
$n-p-1$	a_{n-p-1}	1
$n-p$	a_{n-p}	a_p
⋮	⋮	⋮
p	a_p	a_{n-p}
$p+1$	1	a_{n-p-1}
⋮	⋮	⋮
$n-1$	1	a_1
n	1	a_0

TABLE 12

s	$a_s f_s$	$a_t k_t$	t
0	$a_0 f_0$	k_n	n
1	$a_1 f_1$	k_{n-1}	$n-1$
⋮	⋮	⋮	⋮
p	$a_p f_p$	k_{n-p}	$n-p$
$p+1$	f_{p+1}	k_{n-p-1}	$n-p-1$
⋮	⋮	⋮	⋮
$n-p-1$	f_{n-p-1}	k_{p+1}	$p+1$
$n-p$	f_{n-p}	$a_p k_p$	p
⋮	⋮	⋮	⋮
$n-1$	f_{n-1}	$a_1 k_1$	1

15. Example 5. Direct method with $k(x, y) \equiv k(x-y)$

We now repeat the previous example, equation (56), by the direct method.

The kernel is a function of $(x-y)$, so that the recurrence relation (58) can be used for the calculation of derivatives in the Taylor expansion. The latter can be written in the form

$$f(x) = \tau_0 - \tfrac{1}{2}x\tau_1 + (\tfrac{1}{2}x)^2\tau_2 - \ldots, \tag{60}$$

where

$$\tau_n = (-2)^n f^{(n)}(0)/n!,$$

and satisfies the recurrence relation

$$\tau_n = \tfrac{1}{4}(n+1) + \tau_0 + \frac{\tau_1}{\binom{n}{1}} + \frac{\tau_2}{\binom{n}{2}} + \ldots + \frac{\tau_{n-1}}{\binom{n}{n-1}}. \tag{61}$$

The direct use of equation (60) would involve the calculation of many values of τ, and the work can be reduced considerably by means of the Euler transformation given by

$$z = \frac{\frac{1}{2}x}{1+\frac{1}{2}x}, \quad f(x) = (1-z)\sum_{s=0}^{\infty}(-z)^s \Delta^s \tau_0. \tag{62}$$

The series in equation (62) converges more rapidly than that of (60), and reduces by a factor of two the necessary number of values of τ.

At an interval $h = \frac{1}{4}$ we find that in the quadrature only differences up to order four need be retained, so that it is sufficient to calculate by Taylor series the first four values f_0, f_1, f_2 and f_3, corresponding to $x = 0$, $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ respectively. We now calculate f_n for $n = 4(1)8$, using equation (57) with modified coefficients $a_s + a_{n-s} - 1$. These coefficients, based on table 10, with $p = 4$ and using five decimals, are given in table 13.

TABLE 13

s \ n	4	5	6	7	8
0	0·31111	0·32986	0·32986	0·32986	0·32986
1	1·42222	1·30208	1·32083	1·32083	1·32083
2	0·53333	0·86806	0·74792	0·76667	0·76667
3	1·42222	0·86806	1·20278	1·08264	1·10139
4	0·31111	1·30208	0·74792	1·08264	0·96250
5	—	0·32986	1·32083	0·76667	1·10139
6	—	—	0·32986	1·32083	0·76667
7	—	—	—	0·32986	1·32083
8	—	—	—	—	0·32986

We next prepare a table of $hk_{n-s}f_s$, given by

$$hk_{n-s}f_s = -\frac{f_s}{2[2+\frac{1}{4}(n-s)]^2},$$

and record the values to an extra decimal. Then, since $k = 0{\cdot}5$, f_4 $(x = 1)$ is obtained from equation (57) in the form

$$1{\cdot}03889 f_4 = \frac{1}{3^2} - \sum_{s=0}^{3} \frac{a'_s f_s}{2(3-\frac{1}{4}s)^2}, \tag{63}$$

the modified coefficients a'_s being taken from the first column in table 13. This value should be checked against that calculable from the Taylor series, and when agreement is obtained we can proceed with confidence to calculate the remaining f_n of this group from the equations

$$1{\cdot}04123 f_n = \frac{1}{(2+\frac{1}{4}n)^2} - \sum_{s=0}^{n-1} \frac{a'_s f_s}{2[2+\frac{1}{4}(n-s)]^2}, \tag{64}$$

the a'_s being taken in succession from columns 5, 6, 7 and 8 of table 13.

The tables of $hk_{n-s}f_s$ are collected together in table 14. The first column $(n = 4)$ can be recorded immediately from the values obtained from the Taylor series, and in each successive column one more value can be added when the corresponding f has been calculated. The summations of equations (63) and (64) are effected by placing side-by-side corresponding columns of tables 13 and 14 and accumulating the products of adjacent terms.

TABLE 14

s \ n	4	5	6	7	8
0	0·013889	0·011834	0·010204	0·008889	0·007812
1	0·011533	0·009691	0·008257	0·007120	0·006202
2	0·010012	0·008274	0·006953	0·005924	0·005108
3	0·009099	0·007370	0·006091	0·005118	0·004361
4	—	0·006855	0·005553	0·004589	0·003856
5	—	—	0·005278	0·004275	0·003533
6	—	—	—	0·004146	0·003358
7	—	—	—	—	0·003318

For $n>8$, f_n can now be calculated, using equation (57) with the unmodified coefficients a_s given in table 10. The function ha_tk_t is tabulated with argument t in reverse order, and recorded on a movable strip. This is placed adjacent to the column a_sf_s so that ha_1k_1 is opposite the value of f_{n-1} just determined. The summation of products can now be effected conveniently, and in this way the solution completed up to $x = 5{\cdot}0$. The final results are shown in table 15.

TABLE 15

movable strip t	ha_tk_t	n	x	$f(x)$	a_nf_n
20	0·01020	0	0·00	0·25000	0·08247
19	0·01097	1	0·25	0·17443	0·23039
18	0·01183	2	0·50	0·12515	0·09595
17	0·01280	3	0·75	0·09212	0·10146
16	0·01389	4	1·00	0·06941	0·06811
15	0·01512	5	1·25	0·05344	0·05344
14	0·01653	6	1·50	0·04198	⋮
13	0·01814	7	1·75	0·03359	⋮
12	0·02000	8	2·00	0·02733	⋮
11	0·02216	9	2·25	0·02257	⋮
10	0·02469	10	2·50	0·01891	⋮
9	0·02768	11	2·75	0·01605	⋮
8	0·03125	12	3·00	0·01378	⋮
7	0·03556	13	3·25	0·01195	⋮
6	0·04082	14	3·50	0·01046	⋮
5	0·04734	15	3·75	0·00923	⋮
4	0·05452	16	4·00	0·00822	⋮
3	0·07282	17	4·25	0·00736	⋮
2	0·06133	18	4·50	0·00662	⋮
1	0·13045	19	4·75	0·00600	⋮
		20	5·00	0·00547	0·00547

In this example we retained throughout the same interval $h = \frac{1}{4}$. This is not always essential, and the chance should not be overlooked of increasing the interval, to reduce the number of terms in the summation and therefore both the labour and rounding error. Here it is, in fact, possible to use $h = \frac{1}{2}$ after $x = 2$, as shown in the indirect method (b).

16. DIRECT METHOD WITH $k(x, y) \not\equiv k(x-y)$

As an example of an integral equation for which the kernel is not a function of $(x-y)$ only, we transform the equation (56) of the previous example by writing

$$\phi(x) = (x+2)^2 f(x),$$

a substitution suggested by the form of the integral equation and the rate of decrease of $f(x)$ with increasing x. Equation (56) becomes

$$\int_0^x \frac{-2(x+2)^2}{(x-y+2)^2(y+2)^2}\phi(y)\,dy = \phi(x) - 1. \qquad (65)$$

For the Taylor expansion we must here use the general formula (58) to produce successive derivatives at $x = 0$. Writing

$$\tau_n = (-2)^n \phi^{(n)}(0)/n!,$$

giving the Taylor expansion

$$\phi(x) = \tau_0 - (\tfrac{1}{2}x)\,\tau_1 + (\tfrac{1}{2}x)^2\tau_2 - \ldots,$$

we obtain after some manipulation the recurrence relation

$$\tau_n = \frac{1}{n}\tau_{n-1} + \sum_{s=1}^{n-2} \frac{(n-1-s)!\,s!}{n!} \sum_{m=0}^{n-2-s} \{(s-1)(n-s-m) - 2s\}\tau_m. \tag{66}$$

In this way we calculate, at an interval $h = \frac{1}{4}$, the first four values of $\phi(x)$. Successive values can then be obtained from the equation corresponding to (57), which can here be written in the form

$$(1+\tfrac{1}{8}a_0)\,\phi_n = 1 - \tfrac{1}{8}(x+2)^2\,[a_0 k_n k_0 \phi_0 + a_1 k_{n-1} k_1 (\phi_1 + \phi_{n-1}) + \ldots], \tag{67}$$

where, as before,

$$k_s = \frac{-2}{(2+\frac{1}{4}s)^2},$$

and where the coefficients a_s are modified or not according to the value of n.

The values of $\phi(x)$ so obtained, having five significant figures everywhere, provided an excellent check on the $f(x)$ calculated by the previous method, no $f(x)$ being in error by much more than half a unit in the fifth decimal.

17. Checking. Laplace transforms

It would not be standard practice to solve an integral equation in two different forms, as in the last two sections, and other checks are necessary. Apart from the routine checks such as differencing, a useful auxiliary check can often be provided by means of the Laplace transformation, when the kernel is of the form $k(x-y)$. This method provides a formal solution, but the computation may be laborious and only practicable for some values of the argument.

The Laplace transformation of the integral equation

$$\int_0^x k(x-y)\,f(y)\,\mathrm{d}y = f(x) + g(x)$$

is given by

$$\mathscr{L}(k)\,\mathscr{L}(f) = \mathscr{L}(f) + \mathscr{L}(g),$$

so that

$$\mathscr{L}(f) = -\frac{\mathscr{L}(g)}{1-\mathscr{L}(k)}.$$

Applying the inverse transformation, when this is applicable, we obtain

$$f(x) = -\frac{1}{2\pi\mathrm{i}}\int_{\gamma-\mathrm{i}\infty}^{\gamma+\mathrm{i}\infty} \frac{\mathrm{e}^{px}\,\mathscr{L}(g)}{1-\mathscr{L}(k)}\,\mathrm{d}p, \tag{68}$$

where the path of integration lies to the right of all the poles of the integrand. It is often possible to obtain from (68) an analytical expression for $f(x)$ suitable for computational purposes.

For the previous example given by equation (56) we have

$$g = -\frac{1}{(x+2)^2}, \quad \mathscr{L}(g) = -\{\tfrac{1}{2} + p\,\mathrm{e}^{2p}\,\mathrm{Ei}\,(-2p)\},$$

where $\mathrm{Ei}\,(x)$ denotes the exponential integral and

$$-\mathrm{Ei}\,(-2p) = \int_{2p}^{\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t.$$

The kernel has the same form as $2g(x)$, and we therefore obtain

$$\mathscr{L}(f) = \frac{\frac{1}{2}+p\,\mathrm{e}^{2p}\,\mathrm{Ei}\,(-2p)}{2\{1+p\,\mathrm{e}^{2p}\,\mathrm{Ei}\,(-2p)\}}. \tag{69}$$

This expression does not tend to zero as $p\to\infty$, so that the inverse transformation cannot be applied directly. If $F(x)$ denotes the indefinite integral of $f(x)$, however, then

$$\mathscr{L}(F) = \frac{1}{p}\mathscr{L}(f),$$

and the inverse transformation can be applied directly to $\mathscr{L}(F)$ to give

$$F(t) = \frac{1}{2} - \frac{1}{2\pi\mathrm{i}}\int_{\gamma-\mathrm{i}\infty}^{\gamma+\mathrm{i}\infty} \frac{\mathrm{e}^{pt}}{4p\{1+p\,\mathrm{e}^{2p}\,\mathrm{Ei}\,(-2p)\}}\,\mathrm{d}p. \tag{70}$$

It is easily shown that the expression in braces in equation (70) has no zeros in the complex plane, excluding the negative real axis along which there is a cut on account of the singular behaviour of $\mathrm{Ei}\,(-2p)$. We can therefore deform the contour of integration and write

$$F(t) = \frac{1}{2} - \frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)} \frac{\mathrm{e}^{pt}}{4p\{1+p\,\mathrm{e}^{2p}\,\mathrm{Ei}\,(-2p)\}}\,\mathrm{d}p. \tag{71}$$

The contribution from the circle round the origin is $\frac{1}{2}\pi\mathrm{i}$. Along the negative real axis $\mathrm{Ei}\,(x\,\mathrm{e}^{\pm\pi\mathrm{i}}) = \mathrm{Ei}\,(x) \pm \pi\mathrm{i}$, so that

$$F(t) = \frac{1}{4} - \frac{1}{4}\int_0^\infty \frac{\mathrm{e}^{-x(t+2)}}{[1-x\,\mathrm{e}^{-2x}\,\mathrm{Ei}\,(2x)]^2+\pi^2x^2\,\mathrm{e}^{-4x}}\,\mathrm{d}x,$$

or

$$f(t) = \frac{1}{4}\int_0^\infty \frac{x\,\mathrm{e}^{-x(t+2)}}{[1-x\,\mathrm{e}^{-2x}\,\mathrm{Ei}\,(2x)]^2+\pi^2x^2\,\mathrm{e}^{-4x}}\,\mathrm{d}x. \tag{72}$$

This expression is not very convenient for the routine computation of $f(t)$. For $t = 5$, however, $f(t)$ can be evaluated by numerical quadrature from (72), yielding the values $f(5) = 0{\cdot}0054700$, $\phi(5) = 0{\cdot}26803$, which agree exactly with the values computed by the direct numerical method.

Though this provides a useful check on the complete calculation, it cannot be generally inferred that the numerical values for all $x<5$ are certain to be correct. Much will depend on the form of the kernel and the behaviour of the solution. In the computation of $f(x)$, for example, any error introduced at an early stage will quickly fade out, since the coefficient by which it is subsequently multiplied is small. This point receives attention in a later section, on error accumulation. Before going on to discuss this important aspect we shall make brief comments on an iterative method, for use with Volterra's equation, analogous to that described in § 6 and employed on Fredholm's equation.

18. A classical iterative method

For the solution of the integral equation (56) Friedlander used an iterative method, given by the relations

$$f^{(0)}(x) = \frac{1}{(x+2)^2}\,, \quad f^{(n)}(x) = \frac{1}{(x+2)^2} - \int_0^x \frac{2}{(x-y+2)^2} f^{(n-1)}(y)\,\mathrm{d}y. \tag{73}$$

He evaluated $f^{(1)}(x)$ analytically, and computed $f^{(2)}$, $f^{(3)}$ and $f^{(4)}$ by numerical integration, taking $f^{(4)}(x)$ as an acceptable solution of the integral equation. In fact, his errors in the

fourth decimal, at $x = 1$, 2, 3 and 4, are 4, 5, 13 and 25 units respectively. The labour of this method is somewhat excessive, and it is clear that the order of iterate necessary to give accurate results is likely to increase with x. This process is therefore less satisfactory than in the case of Fredholm's equation, though again the Aitken extrapolation method and the Euler transformation can be used to accelerate convergence. For example, using an interval $h = \frac{1}{4}$, we find at $x = 1{\cdot}0$ the following values of the first four iterates: 0·11111, 0·05849, 0·07126, 0·06916. Aitken's process applied to the last three values produces the result 0·06946, in error by only five units in the fifth place. Again at $x = 4{\cdot}0$, we have the following series of terms, taken from Friedlander's paper, tabulated and differenced for use with Euler's transformation:

(+)	0·0278			
		+102		
(−)	0·0380		−190	
		− 88		+154
(+)	0·0292		− 36	
		−124		+ 72
(−)	0·0168		+ 36	
		− 88		
(+)	0·0080			

The Euler transformation produces a value 0·0078, considerably closer to the true value 0·0082 than the 0·0057 given by Friedlander. (This value 0·0078 is unreliable in the last figure, since Friedlander tabulated only to four decimals the terms u_n of a series $f = \Sigma 2^n u_n$. The above table contains values of $2^n u_n$.)

19. Error analysis

In the evaluation of recurrence relations, or in the computation involved in any step-by-step process, it is desirable to know the effect of introducing an error at any stage. An error of this kind is not a blunder, for whose detection the differencing check is usually sufficient, but is introduced whenever a number is rounded off to a prescribed number of decimals or significant figures. This error is propagated throughout the rest of the work, and its effect may die away, remain approximately constant, or increase indefinitely. In the latter case we have a *building-up error* which, unless detected and removed, will severely limit the accuracy of the results.

The analysis of error in integral equations is similar to that of first order differential equations. If we want a particular integral of the equation

$$\frac{dy}{dx} + p(x)\, y = q(x), \tag{74}$$

the numerical process will introduce, through rounding errors, multiples of the complementary function, the solution of

$$\frac{dy}{dx} + p(x)\, y = 0. \tag{75}$$

If the complementary function dies away with increasing x the effect is not serious; if it increases at the same rate as the particular integral we shall lose accuracy in decimals but not in significant figures; if it increases faster than the particular integral its effect is serious.

It is not possible here to give a general discussion on the conditions governing the incidence of building-up error in the numerical solution of Volterra's equation (4). We shall, in this section, give some simple examples of building-up error and of methods of mitigating its effect which are representative of some broad classes of integral equations, providing by analogy appropriate results for more general cases.

From equation (57) it is apparent that building-up error is likely to be associated with positive values in the kernel, corresponding in the Fredholm case with ill-conditioning in the simultaneous equations. It would also appear that when $k(x, y)$ varies, the effect is more or less severe according as it increases or decreases with decreasing y. As typical examples we consider separately the cases in which the kernel is of exponential or oscillatory type.

If $k(x, y) = \lambda\, \mathrm{e}^{\mu(x-y)}$ the equation governing the building-up error, corresponding to (75) for the complementary function of the first order differential equation (74), is given by

$$f(x) = \lambda \int_0^x \mathrm{e}^{\mu(x-y)} f(y)\, \mathrm{d}y. \tag{76}$$

On differentiation this reduces to the form

$$f' = (\lambda+\mu) f,$$

and there will be an exponential build-up of error if $\lambda+\mu>0$.

If $k(x, y) = \lambda \cos \mu(x-y)$ the kernel has its maximum value at the upper limit in the integral. The 'error equation' can be reduced to the form

$$f''-\lambda f'+\mu^2 f = 0$$

with exponentially increasing solutions if $\lambda>0$, $|2\mu|<|\lambda|$.

If $k(x, y) = \lambda \sin \mu(x-y)$ the kernel has its maximum at the lower limit, and we find

$$f'' = \mu(\lambda-\mu) f$$

with exponentially increasing solutions if $\mu(\lambda-\mu)>0$.

Other simple equations can be constructed which are amenable to this type of analysis, and it is often possible to obtain an idea of the behaviour of more complicated equations by analogy with one of these cases. The integral equation (56) has a kernel which is negative everywhere, with a modulus which steadily decreases. It can therefore be compared with (76), with both λ and μ negative, giving no error build-up. Even if the sign of the kernel were changed, giving a positive λ, the decrease in modulus is so rapid that there is still no substantial error accumulation.

Another approach to error analysis, for a kernel of the form $k(x-y)$, is provided by the Laplace transformation. The previous analysis is qualitative, in the sense that it examines the behaviour of the complementary function, introduced by a single rounding error. In fact, such a rounding error is made at every point, and it is more accurate to take for the total error propagation the solution of the integral equation obtained from (4) by replacing $g(x)$ by a step-function $s(x)$, for which $\frac{1}{2} \geqslant s(x) \geqslant -\frac{1}{2}$ for all x. The solution of this equation is given formally by

$$f(x) = -\frac{1}{2\pi \mathrm{i}} \int_{\gamma-\mathrm{i}\infty}^{\gamma+\mathrm{i}\infty} \frac{\mathrm{e}^{px}\mathscr{L}(s)}{1-\mathscr{L}(k)}\, \mathrm{d}p, \tag{77}$$

which is equation (68) with g replaced by s.

The presence of exponentially increasing factors in the solution of (77) is in general governed by the form of the denominator of the integrand, and will exist if the denominator has a zero in the right half plane. By this theorem we can again confirm the lack of building-up error in the solution of equation (56), or even of the same equation with the sign of the kernel reversed.

We note also that errors build up more rapidly than the wanted solution, the only really serious condition, when the zero of the function $1-\mathscr{L}(k)$ having the largest positive real part is also a zero of $\mathscr{L}(g)$. Serious building-up error is not therefore a very common occurrence, and in this respect integral equations are again more similar to first order differential equations than to those of higher order.

We cannot here give a general discussion of methods of eliminating error build-up in the cases where this occurs, but some possibilities might be mentioned. The most obvious and simple method is to retain extra 'guarding' figures throughout, a method whose success depends solely on the rate of error increase. Alternatively, it is sometimes possible to transform the integral equation to another form, in the solution of which errors are less likely to accumulate.

Examination of equation (57) and the simple examples given earlier in this section shows that severe error build-up occurs when $k(x, x)$ is large and positive. If $g(x)$ is such that the wanted solution does not itself increase in the same way, then retention of guarding figures may lead to a prohibitive amount of computation. The alternative procedure, however, is often possible.

On differentiating equation (4) with respect to x we obtain an equation which can be written in the form

$$f(x)+\frac{1}{k(x,x)}\int_0^x \frac{\partial}{\partial x}k(x,y)\,f(y)\,\mathrm{d}y=\frac{g'(x)}{k(x,x)}+\frac{f'(x)}{k(x,x)}. \tag{78}$$

Now $k(x, x)$ is large and positive, but $f(x)$ does not vary rapidly, so that $f'(x)$ is small. We can therefore solve (78) by a method of successive approximation, solving first for $f(x)$ by neglecting the last term in (78), calculating the latter approximately from the approximate $f(x)$ so obtained, and continuing the iterative procedure until there is no further change in f. It might here be remarked that the direct solution of (78), obtained by using finite-difference formulae for the derivative as well as integral involved, is not satisfactory, since a backward-difference formula must inevitably be used for the derivative, causing a partial cancellation in the coefficient of $f(x)$ on the two sides of the equation.

As an example we consider the equation

$$\int_0^x 10\cos(x-y)\,f(y)\,\mathrm{d}y=f(x)-(1-5x)\sin x, \tag{79}$$

which is known to have the analytical solution $\sin x$. If this equation is solved by the direct method the error builds up alarmingly, a solution to four decimals at an interval $h=0{\cdot}1$ having an error of 638 units at $x=1{\cdot}5$. The alternative form of this equation is given by

$$f(x)-\int_0^x \sin(x-y)\,f(y)\,\mathrm{d}y=\tfrac{1}{2}(x\cos x+\sin x)-\tfrac{1}{10}\cos x+\tfrac{1}{10}f'(x). \tag{80}$$

At an interval of $h=0{\cdot}2$, this equation was solved, neglecting its last term, as far as $x=4{\cdot}0$. From this approximate $f(x)$ an approximation to $f'(x)$ was calculated at every pivotal point

by the central-difference formula. These values were then used in equation (80) and a second approximation calculated for the wanted function $f(x)$. This second approximation has a maximum error of two units in the fourth decimal. Theoretically this approximation should here agree with the analytical solution, since the solution of equation (80) with the last term omitted is $\sin x - \frac{1}{10}$, differing from the solution of (79) only by a constant.

Finally, we mention an alternative treatment of equations like (79), for which the 'particular integral' remains of fairly constant size while the 'complementary function' increases very rapidly. For this method we need to know, by the Laplace transform method or otherwise, the required solution for some fairly large value of x.

Consider, for example, the problem of equation (79), for which we have the additional information that $f(1{\cdot}5) = 0{\cdot}9975$. The direct method gives a solution f_1 which can be regarded as a combination of the true solution f and a contaminating building-up error E, so that

$$f_1 = f + E.$$

Now any different $g(x)$ for which the true particular integral builds up will produce a solution f_2 which, after a few steps, is some multiple of E, so that

$$f_2 = kE.$$

If we know the error $f_1 - f$ for some value of x, we can therefore correct everywhere our approximation f_1 and produce an accurate f from the formula

$$f = f_1 - \frac{1}{k} f_2.$$

Table 16 shows a portion of the solution f_1 obtained by the direct method applied to equation (79), and the discrepancy between f_1 and the true solution $\sin x$. The table also gives a second solution f_2 obtained by replacing $g(x)$ by unity. We find that to make our solution correct at $x = 1{\cdot}5$ we have to take $k = 0{\cdot}1925$, and the final column then gives the correction to f_1 to be applied at other values of x. The final error has a maximum of one unit.

TABLE 16

x	$f_1(x)$	$E(x)$	$f_2(x)$ $(g=1)$	$0{\cdot}1925 f_2(x)$
0·7	0·6443	1	1	0
0·8	0·7174	0	3	1
0·9	0·7835	2	9	2
1·0	0·8419	4	23	4
1·1	0·8924	12	63	12
1·2	0·9353	33	169	33
1·3	0·9723	87	455	88
1·4	1·0091	236	1228	236
1·5	1·0613	638	3314	638

20. SUMMARY AND CONCLUSION

In this paper we have discussed, and illustrated in some detail, methods for solving linear integral equations of both Fredholm and Volterra types. Most of the methods are based on the theory of finite differences, so that only non-singular integral equations received attention. For singular integral equations numerical methods would be much more difficult to apply. Such equations are probably best treated analytically, either to

obtain a mathematical solution or to transform them to non-singular type; if numerical methods are used they could be expected to produce only approximate results, with errors not easily calculable.

The general methods given here are applicable to certain kinds of non-linear integral equations, though they would generally involve the solution of non-linear simultaneous algebraic equations, for which no very satisfactory general methods are yet available. Finally, they can also be used for some integro-differential equations, either by methods of successive approximation or by direct methods in which both integrals and derivatives are replaced by finite-difference expressions.

The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

References

Aitken, A. C. 1926 *Proc. Roy. Soc. Edinb.* **46**, 289.
Aitken, A. C. 1939 *Determinants and matrices.* Edinburgh: Oliver and Boyd.
Bickley, W. G. 1939 *Math. Gaz.* **23**, 352.
Bullard, E. C. & Cooper, R. I. B. 1948 *Proc. Roy. Soc.* A, **194**, 332.
Fox, L. 1947 *Proc. Roy. Soc.* A, **190**, 31.
Fox. L. 1948 *Quart. J. Mech. Appl. Math.* **1**, 253.
Fox, L. 1949 *Proc. Camb. Phil. Soc.* **45**, 50.
Fox, L. 1950*a* *Phil. Trans.* A, **242**, 345.
Fox, L. 1950*b* *J. R. Statist. Soc.* B, **12**, 120.
Fox, L. & Goodwin, E. T. 1949 *Proc. Camb. Phil. Soc.* **49**, 373.
Friedlander, F. G. 1941 *Proc. Camb. Phil. Soc.* **37**, 134.
Hartree, D. R. & Womersley, J. R. 1937 *Proc. Roy. Soc.* A, **161**, 353.
Kreisel, G. 1949 *Proc. Roy. Soc.* A, **197**, 160.
Love, E. R. 1949 *Quart. J. Mech. Appl. Math.* **2**, 428.
Lovitt, W. V. 1950 *Linear integral equations.* New York: Dover Publications Inc.
Richardson, L. F. & Gaunt, J. A. 1927 *Phil. Trans.* A, **226**, 299.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES