Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: NUMERICAL SOLUTION OF LINEAR INTEGRAL EQUATIONS

To: Mathematics Group

From: Alan J. Perlis

Date: September 16, 1948

Introduction:

The purpose of this report is a survey of numerical methods for solving integral equations. Only linear integral equations of real variables will be considered. There are three important types of such equations:

(a) The Fredholm equations of the second kind

$$f(x) = g(x) + \lambda \int_a^b K(x,y) f(y) dy \quad \text{(non-homogeneous)}$$

$$f(x) = \lambda \int_a^b K(x,y) f(y) dy \quad \text{(homogeneous)}$$

(b) The Volterra equation

$$f(x) = g(x) + \int_a^x K(x,y) f(y) dy$$

and

(c) The equation of the first kind

$$g(x) = \int_a^b K(x,y) f(y) dy$$

In each of the above equations the function f(x) is to be obtained from the given functions g(x) and K(x,y). Under quite general conditions (a), and (b) possess unique solutions whereas (c) will, in general, have many solutions except under certain restrictive constraints.

Thus far, at least by comparison, say, with linear algebraic equations and differential equations, little has been accomplished in

developing numerical methods for solving integral equations. The superiority over differential equations in expressing many physical problems has only recently been exploited.

In general, it would be nicest to attack each equation by a method best suited to its own structure. This puts a premium on analysis. The attitude taken here will be to develop quite general numerical methods which can most easily be treated by a computer of the digital type.

On the same level of importance as the design of a solving method for an equation is an appreciation of the status of errors in each step of the problem. Besides the usual numerical round-off and truncation errors, all methods of solution introduce an important stability error caused by the lapse of a transcendental problem into a discrete one. In general, the numerical methods show excellent stability (at least where unique solutions to the problem exist) due chiefly to the smoothing properties of integration.

2. The Linear Integral Equation of the First Kind

The equation to be considered is

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy \qquad (1)$$

with the functions $f(x)$ and $K(x,y)$ given and $g(y)$ to be determined. Examples of this type of equation are:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iyx}\, g(y)\, dy \qquad \text{(The Fourier transform)}$$

and

$$f(x) = \int_0^{\infty} e^{-xy}\, g(y)\, dy \qquad \text{(The Laplace transform)}$$

Often the solution $g(y)$ will not be unique; for, to the solution of (1) must be added solutions, if they exist, of

$$0 = \int_a^b K(x,y)\, g_i(y)\, dy \qquad (i = 1, 2, \ldots)$$

In both of the above mentioned examples definite inversion formulae of the type:

$$g(y) = \int_L h(x,y)\, f(x)\, dx \qquad (L, \text{ a path in the } x \text{ plane})$$

yield anslytic expressions for $g(y)$. In general, such inversion formulae do not exist. The integral equation does not always have solutions, since for a given $K(x,y)$ $f(x)$ must possess certain properties dependent upon the kernel. If $K(x,y)$ is a polynomial in x, then $f(x)$ must be a polynomial in x. Likewise if $K(a,y) = 0$ for all y and if the equation is satisfied for all y, then $f(a) = 0$. The function $f(x)$ is not perfectly arbitrary, for if $K(x,y)$ is a polynomial in x, then $f(x)$ must be a polynomial. That is, we must restrict $f(x)$ to a class of functions associated with $K(x,y)$. A theorem to this effect will be discussed later. In certain special cases an equation of the first kind can be reduced to one of the second kind and here $f(x)$ can be highly arbitrary.

As regards the uniqueness of the solution $g(y)$, where $g(y)$ is constrained to be continuous, the following may be said. If the kernel, $K(x,y)$, is "definite", i.e.

$$\int_a^b \int_a^b K(x,y)\, w(x)\, w(y)\, dx\, dy > 0$$

for all continuous $w(x)$, then the equation

$$\int_a^b K(x,y)\, w(y)\, dy = 0$$

has no continuous solution $w(y)$ except $w(y) = 0$. In particular, if two continuous solutions $g_1(y)$ and $g_2(y)$ of the equation

exist, then, for K(x,y) definite, $g_1(y) = g_2(y)$; i.e. the solution is unique. In general a given K(x,y) will not be definite but may be made so by the following artifice. Form:

$$L(x,y) = \int_a^b K(x,t)\, K(y,t)\, dt .$$

Since L(x,y) is symmetric its characteric values are real. Furthermore, if $\phi_n(x)$ is a characteristic function,

$$\phi_n(x) = \lambda_n \int_a^b L(x,y)\, \phi_n(y)\, dy$$

From which it follows that $\lambda_n \geq 0$ since

$$\int_a^b \phi_n^2(x)\, dx = \lambda_n \int_a^b \phi_n(x)\, dx \int_a^b L(x,y)\, \phi_n(y)\, dy$$

and

$$= \lambda_n \int_a^b \phi_n(x)\, dx \int_a^b \phi_n(y)\, dy \int_a^b K(x,t) K(y,t)\, dt$$

$$= \lambda_n \int_a^b \left(\int_a^b K(x,t)\, \phi_n(t)\, dt \right)^2 dx$$

hence $\lambda_n \geq 0$. Indeed if the $\{\phi_n(x)\}$ form a complete set,

$$\int_a^b K(x,t)\, \phi_n(x)\, dx \neq 0 \quad (n = 1, 2, \ldots)$$

so that $\lambda_n > 0$. Hence L(x,y) is definite and the integral equation:

$$u(x) = \int_a^b L(x,y)\, g(y)\, dy$$

where

$$u(x) = \int_a^b f(t)\, K(x,t)\, dt$$

has a unique continuous solution. If the function g(y) is restricted to be continuous f(x) will usually be required to satisfy linear conditions of the form:

$$\int_a^b w(x)\, f(x)\, dx = 0 , \quad \int_a^b w(x)\, K(x,y)\, dx = 0$$

An analogy to linear algebraic equations proves fruitful. The equation is analogous to the system:

$$f_x = \sum_{y=1}^{n} K_{x,y}\, g_y \quad (x = 1, 2, 3, \ldots, m)$$

and can be obtained from the system in the limit as $m,n \to \infty$. The properties of the system depend on the relation between m and n. Hence three cases can occur:

(1) $m > n$ implies the existence of relations of the form

$$\sum_{x=1}^{x=m} a_x K_{x,y} = 0 \qquad \text{for all } y$$

and then consistency will demand that f_x satisfy:

$$\sum_{x=1}^{x=m} a_x f_x = 0$$

In the limit these relations may take several forms, an obvious one being:

$$\int_a^b a(x) f(x) dx = 0$$

(2) $m = n$. A unique set of g_ys exists and is given by:

$$g_y = \sum_{x=1}^{x=m} a_x K_{x,y} \qquad y = (1, 2, \ldots, m)$$

provided no relation

$$\sum_{x=1}^{x=m} a_x K_{x,y} = 0 \qquad \text{(all y)}$$

is satisfied.

(3) $m < n$ yields an infinite number of sets of solutions but particular sets may be obtained by imposing $(n - m)$ linear conditions on g_y.

Hence, in general, it is too much to expect that the solution will be unique. An important question, particularly in numerical calculations, is the range in x over which f(x) is given. Decreasing the range may well involve making g(y) zero over a portion of the range to maintain uniqueness. In the particular case of the moment problem of Stieltjes, f(x) is defined only over an enumerable set of values $\{x_i\}_{i=1}^{i=\infty}$ and the side conditions induce a unique g(y).

In general, then, if the conditions are satisfied which will make the solution unique we shall expect to obtain an expression of the form:

$$g(y) = H_K f(y)$$

where H_K may be built up of linear combinations of the type

$$p(y) \frac{d^n f(y)}{dy}, \quad f(y+\beta), \quad \int_a^b H(y,x) f(x) dx$$

In two cases mentioned the inversion formula is of the highly desirable form:

$$g(y) = \int_L H(y,x) f(x) dx \qquad \text{(L some path)}$$

Only rarely will this integral be of the same type as that defining f(x). For a g(y) may be discontinuous at some $y = y_o$, and still yield a continuous f(x), hence the integral for g(y) must be improper, that is H(y,x) must be discontinuous or the limits be infinite.

In those cases where the equations of condition are not satisfied it will then not be possible to obtain an exact representation of f(x) by an integral. But by analogy to the method of least squares, the problem of solving the integral equation is essentially equivalent to that of minimizing an integral, i.e. a problem in the calculus of variations. The point being to obtain the best possible approximation to f(x) in the least squares sense by an integral of the given type. As in the least squares proceedure we may introduce a positive weight function q(x) which weights f(x) according to some law. Then if

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy$$

$$\int_a^b q(x) K(x,s) f(x)\, dx = \int_a^b q(x) K(x,s)\, dx \int_a^b K(x,y)\, g(y)\, dy$$

or

$$\phi(s) = \int_a^b L(s,y)\, g(y)\, dy$$

where

$$L(s,y) = \int_a^b K(x,s) K(x,y)\, q(x)\, dx$$

and

$$\phi(s) = \int_a^b q(x) K(x,s) f(x)\, dx.$$

If we assume that k(x,y) is "perfect", i.e.

$$\int_a^b K(x,y)\, u(y)\, dy = 0$$

implies, if u(y) is continuous, that u(y) = 0. The L(s,y) is definite, and symmetric. Hence the solution is unique, but may not possess a continuous solution g(y). This can be answered by ascertaining the existence or non-existence of a continuous function g(y) which minimizes

$$W = \int_a^b q(x) \left(f(x) - \int_a^b K(x,y)\, g(y)\, dy \right)^2 dx$$

over a set of values in x. This minimum may not always exist but W may be made as small as desired, and thus an approximate representation for f(x) is obtained.

A rewarding way of examining the integral equation of the first kind is as an operator equation:

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy \quad \text{as} \quad f(s) = L_K\, g(s)$$

and then the problem is to find L^{-1}_k such that

$$g(s) = L_K^{-1} f(s)$$

and, in general, will not be unique. For numerical work the problem of uniqueness becomes largely an academic one, in any case. The interesting question here is under what general conditions can a solution be found; how accurate will it be; and is there a quite general, easily arithmetized algorithm which will yield a reasonably accurate result. A general theorem by Bateman proves useful.

The equation is:

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy.$$

Either K(x,y) is symmetric or it is symmetrized by the operation:

$$u(s) = \int_a^b f(x) K(s,x)\, dx = \int_a^b K(s,x)\, dx \int_a^b K(x,y) g(y)\, dy$$

$$u(s) = \int_a^b K(s,x) K(x,y)\, dx \int_a^b g(y)\, dy = \int_a^b L(s,y) g(y)\, dy$$

Henceforth K(x,y) will be assumed to be symmetric. The theorem then is:

(i) If K(x,y) has a complete set of characteristic functions, $\phi_n(x)$, satisfying

$$\phi_n(x) = \lambda_n \int_a^b K(x,y)\, \phi_n(y)\, dy$$

and

(ii) f(x) can be expanded in an absolutely and uniformly convergent "Fourier" series

$$f(x) = \sum_{n=0}^{\infty} C_n \phi_n(x)$$

then there exists a g(y) such that

$$\left| f(x) - \int_a^b K(x,y)\, g(y)\, dy \right|$$

may be less than any arbitrary $\epsilon > 0$. The proof of the theorem provides the algorithm for the solution.

Proof:

Define: $$F(z,x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} z^{2n+1} L_K^{2n+1} f(x) \quad ; \quad L_K f(x) = \int_a^b K(x,y) f(y)\, dy$$

and

$$H(z,x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} z^{2n} L_K^{2n} f(x)$$

The series are abs. convergent for finite x. For:

Let $|f(x)| < \tilde{f}$

$|K(x,y)| < \tilde{K}$

They imply

$$|L_K^n f(x)| < (b-a)^n \tilde{f} \tilde{K}^n$$

and so each series becomes less than the exponential series.

Pick an M and write:

$$g(y) = z \int_0^M F(y,z)\,dz$$

$$\int_a^b K(x,y)\,g(y)\,dy = z \int_a^b K(x,y)\,dy \int_0^M F(y,z)\,dz$$

$$= z \int_0^M dz \int_a^b F(y,z)\,K(x,y)\,dy$$

Integration term by term yields

$$2 \int_a^b K(x,y)\,F(y,z)\,dy = 2 \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{n!} L_K^{2n+2} f(x)$$

$$= -\frac{d}{dz} H(z,x)$$

Therefore:

$$\int_a^b K(x,y)\,g(y)\,dy = -\int_0^M \frac{d}{dz} H(z,x)\,dz$$

$$= f(x) - H(x,M)$$

An expression for large M, of H(x,M), is to be found.

Use: $f(x) = \sum_{n=1}^{\infty} C_n \phi_n(x)$

Form: $\int_a^b K(x,y)\,f(y)\,dy = \sum_{n=1}^{\infty} \frac{C_n}{\lambda_n} \phi_n(x)$ which converges.

Also:

$$L_K^r f(x) = \sum_{n=1}^{\infty} \frac{C_n}{\lambda_n^r} \phi_n(x) \qquad (r = 1, 2, \ldots)$$

So:
$$H(x,z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{n!} L_K^{2n} f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{n!} \left(\sum_{n=1}^{\infty} \frac{c_n}{\lambda_n^{2n}} \phi_n(x) \right)$$

Now: $\sum_{n=m}^{\infty} |c_n \phi_n(x)| < \epsilon$ implies, for m large enough,
$$\sum_{n=m}^{\infty} \left| \frac{c_n}{\lambda_m^{2n}} \phi_n(x) \right| < \left| \frac{\epsilon}{\lambda_m^{2n}} \right|$$

since $\lambda_{n+1} > \lambda_{n+2}$ (by formal ordering of the eigenfunctions)

Hence:
$$H(x,z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{n!} \left(\sum_{n=1}^{m} \frac{c_n \phi_n(x)}{\lambda_n^{2n}} \right) + \delta$$

where
$$|\delta| < \epsilon e^{-z^2/\lambda_m^2}$$

So rearranging:
$$H(x,z) = \sum_{n=1}^{m} c_n \phi_n(x) e^{-z^2/\lambda_n^2} + \delta.$$

Finally:
$$H(x,z) = \sum_{n=1}^{\infty} c_n \phi_n(x) e^{-z^2/\lambda_n^2}$$

so now
$$\left| \sum_{n=p}^{\infty} c_n e^{-z^2/\lambda_n^2} \phi_n(x) \right| < \left| \sum_{n=p}^{\infty} c_n \phi_n(x) \right| < \epsilon/2$$

and $\left| \sum_{n=1}^{p} c_n e^{-M^2/\lambda_n^2} \phi_n(x) \right| < \epsilon/2$, for M large enough.

Therefore $|H(x,M)| < \epsilon.$

Hence $\left| \int_a^b K(x,y)\, g(y)\, dy - f(x) \right| < \epsilon$

and so the solution is
$$g(y) \sim z \int_0^M F(y,z)\, dz$$

If, in addition, g(y) is constrained to remain finite over the entire range of integration, more can be said. It has been shown that
$$\left| \int_a^b K(x,y)\, g(y)\, dy - f(x) \right| < \epsilon$$

but not that
$$\lim_{\epsilon \to 0} g_\epsilon(y) = g(y)$$

If, in addition to the finiteness condition on $g(y)$ the "derived series" $\sum |\lambda_n s_n \tilde{\phi}_n|$ is convergent, $\tilde{\phi}_n = |\phi_n|_{\max}$ on (a,b) then there exists a $g(y)$ defined by $g(y) = ? \int_0^\infty F(y,z)\,dz$

such that

$$f(x) = \int_a^b K(x,y)\,g(y)\,dy$$

where $F(y,z)$ is defined by the series on page (7).

As an example of this method of solution, consider the equation:

$$e^{-s^2} = \int_0^\infty g(t) \cos st\, dt$$

The solution is known to be $g(t) = \frac{1}{\sqrt{\pi}} e^{-t^2/4}$

The function $F(t,x)$ is formed:

$$F(t,x) = x\int_0^\infty e^{-s^2}\cos st\, ds - x^3\int_0^\infty \cos t s_1\, ds_1 \int_0^\infty \cos s_1 s_2\, ds_2 \int_0^\infty e^{-s_3^2}\cos s_3 s_2\, ds_3 + \cdots\cdots$$

$$= e^{-t^2/4} \sum_{n=0}^{\infty} \frac{(-1)^n \left(\sqrt{\pi}\right)^{2n+1}}{n!\, 2^{n+1}} x^{2n+1}$$

$$g(t) = 2\int_0^\infty e^{-t^2/4}\left(\sum_{n=0}^{\infty} \frac{(-1)^n \left(\sqrt{\pi}\right)^{2n+1}}{n!\, 2^{n+1}} x^{2n+1}\right) dx$$

$$= \frac{1}{\sqrt{\pi}} e^{-t^2/4} \int_0^\infty d\left(e^{-\pi/2\, x^2}\right)$$

$$= \frac{1}{\sqrt{\pi}} e^{-t^2/4}$$

3. Solution in Terms of Orthogonal Functions:

The conditions for the existence of a solution of the equation of the first kind may be rather neatly stated in terms of orthogonal functions. The treatment will be restricted to equations and functions satisfying the following conditions:

(i) The variable y is real on (a,b); the variable x is real on the same finite interval of definition (a,b).

(ii) The functions f(x), k(x,y) are real.

(iii) The kernel k(x,y) is bounded and integrable squared (L^2) in x and y; the function f(x) is integrable squared (L^2) in (a,b). Hence g(y) is L^2 in (a,b).

It will be convenient to say that if

$$\int_a^b \left(f(y) - g(y) \right)^2 dy = 0$$

Then f(y) = g(y).

A set of real functions $\{\phi_n(x)\}$ are ortho-normal L^2 on (a,b) if they satisfy:

$$\int_a^b \phi_n(x)\, \phi_m(x)\, dx = \delta_{mn} \quad , \quad m,n = 1, 2, 3, \cdots$$

If the set is orthogonal and L^2 on (a,b) it may be made ortho-normal. The set is said to be complete if no function u(x) in L^2 exists for which

$$\int_a^b \phi_n(x)\, u(x)\, dx = 0 \quad \text{(all } n\text{)}$$

If the set $\{\phi_n(x)\}$ is not complete, it may be completed by adjoining to it a set $\{\phi_n'(x)\}$ (13) which is called the complementary set to $\{\phi_n(x)\}$.

Bessel's inequality is:

$$\sum_{n=1}^{\infty} \left(\int_a^b f(x)\, \phi_n(x)\, dx \right)^2 = \sum_{n=1}^{\infty} f_n^2 \leq \int_a^b f(x)^2\, dx$$

If the set $\{\phi_n(x)\}$ is complete the equality holds. In addition, for a complete set, the relation

$$\int_a^b f(x)\, g(x)\, dx = \sum_{n=1}^{\infty} f_n\, g_n$$

is valid. If not complete the equality holds if and only if

$$\int_a^b g(x)\,\phi_n(x)\,dx = 0 \qquad n = 1, 2, \ldots$$

implies

$$\int_a^b f(x)\,g(x)\,dx = 0$$

for all $g(x)$ on $L^2(a,b)$. This can be put another way. In order that $f(x)$ be orthogonal to all $g(x)$ which satisfy the above, it is sufficient that $f(x)$ be orthogonal to the set $\{\phi_n'(x)\}$ complementary to $\{\phi_n(x)\}$.

The fundamental Riesz-Fischer theorem states: if $\{f_n\}_{n=1}^{\infty}$ exists such that $\sum_1^{\infty} f_n^2 < \infty$ then there exists and $f(x)$ in $L^2(a,b)$ such that

$$f_n = \int_a^b f(x)\,\phi_n(x)\,dx \qquad n = 1, 2, \ldots$$

where

$$\int_a^b \phi_m(x)\,\phi_n(x)\,dx = 0 \qquad m \neq n = 1, 2, \ldots$$

$$\int_a^b \phi_n^2(x)\,dx < \infty \qquad n = 1, 2, \ldots$$

and if the set $\{\phi_n(x)\}$ is complete, $f(x)$ is unique except possibly for a set of measure zero.

One point which may now be made clear is that there functions for which the equation (1) has no solution. For if $f(x)$ L^2 on some (a,b), then for some orthogonal set on (a,b);

$$c_n = \int_a^b f(x)\,\Phi_n(x)\,dx = \int_a^b A_n(y)\,g(y)\,dy$$

where

$$A_n(y) = \int_a^b K(x,y)\,\Phi_n(x)\,dx$$

Then

$$c_n^2 \leq \int_a^b A_n^2(y)\,dy \int_a^b g(y)^2\,dy.$$

Let

$$\int_a^b A_n^2(y)\,dy = a_n^2$$

so that $\frac{c_n^2}{a_n^2}$ be bounded is a necessary condition if $f(x)$ is to be a solution of (1); however it is not sufficient. But:

$$\sum_{n=1}^{N} a_n^2 = \int_a^b \left[\sum_{n=1}^{N} A_n^2(y)\right] dy \leq \int_a^b \int_a^b K^2(x,y)\,dy\,dx.$$

so that $\sum_{n=1}^{\infty} G_n^2$ is convergent and furthermore depends only on the kernel. But there are series $\sum_{n=1}^{\infty} c_n^2$ and $\sum_{n=1}^{\infty} a_n^2$ which converge, and yet for which $\frac{c_n^2}{a_n^2}$ is unbounded. For example: $\frac{c_n^2}{a_n^2} = n$ and

$$\sum_{n=1}^{\infty} c_n^2 = \sum_{n=1}^{\infty} n a_n^2 = \sum_{n=1}^{\infty} \frac{a_n'^2}{n^3}$$ i.e. both series converge.

It is possible to reduce the solution of the given equation to that of the following problem: Being given a demunerable set of constants $\{c_n\}$ and a set of functions $\{A_n(y)\}$, to find a g(y) which satisfies:

$$c_n = \int_a^b A_n(y)\, g(y)\, dy \qquad n = 1, 2, \ldots$$

The following may be said at once: If the $A_n(y)$ form an ortho-normal set on (a,b) the necessary and sufficient condition for a solution g(y) is that

$$\sum_{n=1}^{\infty} c_n^2 < \infty$$

then the solution is the limit, in the mean, of

$$\sum_{n=1}^{N} c_n A_n(y).$$

Naturally, if $\{A_n(y)\}$ be not closed there must be added, to the above, solutions of:

$$\int_a^b A_n(y)\, g(y)\, dy = 0$$

If the $\{A_n(y)\}$ are not ortho-normal, they may be made linearly independent by suppressing those which obey relations of the form:

$$A_n(y) = \sum_{i=1}^{n-1} \gamma_i^{(n)} A_i(y)$$

This creates similar equations among the c's:

$$c_n = \sum_{i=1}^{n-1} \gamma_i^{(n)} c_i.$$

The remaining equations are ortho-normalized. Hence one gets series of the form

$$\sum_{n=1}^{\infty} b_n^2 = \sum_{n=1}^{\infty} (a_1^{(n)} c_1 + \cdots a_n^{(1)} c_n)$$

which must converge and the solution is the limit in the mean of

$$\sum_{n=1}^{N} b_n B_n(y).$$

where $\{B_n(y)\}$ are the derived ortho-normal set. Conversely if this problem is solved, where the C_n are the Fourier coefficients of f(x), the equation is solved.

The solutions previously given demand knowledge of the characteristic values and functions of the kernel or of the symmetrized definite kernel constructed from it.

It will now be show that solving the integral equation is equivalent to that of maximizing a functional equation whose maximum value, if attained, is the solution. The conditions that the maximum be attained, then, are the conditions previously obtained which insure the solutions.

The functional H(g) is defined as

$$H[g] = \frac{\left(\int_a^b f(x)\,\tilde{g}(x)\,dx\right)^2}{\int_a^b\int_a^b L(x,y)\,\tilde{g}(x)\,\tilde{g}(y)\,dx\,dy}$$

where L(x,y) is the previously defined symmetrized kernel:

$$L(x,y) = \int_a^b K(x,t)\,K(y,t)\,dt;$$

f(x) is defined by:

$$f(x) = \int_a^b K(x,y)\,g(y)\,dy.$$

and g(x) is any $L^2(a,b)$ function. The denominator of H[g] is bounded away from zero if L(x,y) is "closed" or if f(x) is orthogonal to each member of the complimentary set of eigen functions of L(x,y). This will be assumed. Define now the "Fourier" coefficients:

$$C_i = \int_a^b f(x)\,\phi_i(x)\,dx$$

$$\tilde{g}_i = \int_a^b \tilde{g}(x)\,\phi_i(x)\,dx$$

where the $\{\phi_i(x)\}$ are the characteristic functions of L(x,y). Then H[g] can be written as:

$$H[g] = \frac{\left(\sum_{n=1}^{\infty} \tilde{g}_n f_n\right)^2}{\sum_{n=1}^{\infty} \frac{\tilde{g}_n^2}{\lambda_n^2}}$$

Now introduce the identity:

$$\sum_{i=1}^{N} a_i^2 \sum_{i=1}^{N} b_i^2 = \left(\sum_{n=1}^{N} a_i b_i\right)^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(a_i b_j - a_j b_i)^2$$

Now as $N \to \infty$, if $\sum_{n=1}^{\infty} a_i^2$ and $\sum_{n=1}^{\infty} b_i^2$ converge then $\sum_{n=1}^{\infty} a_i b_i$ converges and hence the series

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(a_i b_j - a_j b_i)^2$$

converges. The substitution $a_i = \lambda_i c_i, b_i = \tilde{g}_i/\lambda_i$ yields, the inequality

$$\sum_{i=1}^{N} \lambda_i^2 c_i^2 \geq \frac{\left(\sum_{i=1}^{N} \tilde{g}_i c_i\right)^2}{\sum_{i=1}^{N} \tilde{g}_i^2/\lambda_i^2}$$

Now let $N \to \infty$; the right side approaches H(g) and if the series converges H(g) is bounded. Furthermore $\sum_{i=1}^{\infty} \tilde{g}_i^2/\lambda_i^2$ converges so that the identity can be written:

$$H(g) = \sum_{i=1}^{\infty} \lambda_i^2 c_i^2 - \frac{1}{2}\left(\sum_{i=1}^{\infty} \tilde{g}_i^2/\lambda_i^2\right)\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left(\tilde{g}_i c_j \frac{\lambda_j}{\lambda_i} - \tilde{g}_j c_i \frac{\lambda_i}{\lambda_j}\right)^2$$

Now put $g_i = \lambda_i^2 c_i$. Hence for the function g(x) possessing these Fourier coefficients H(g) attains the maximum. Hence if the series $\sum_{i=1}^{\infty} \lambda_i^2 c_i^2$ converges and $\sum_{i=1}^{\infty} \lambda_i^4 c_i^2$ converges, then there exists a $\tilde{g}(x)$ whose Fourier coefficients are $\lambda_i^2 c_i$ and which maximizes the functional H(g); furthermore, this $\tilde{g}(x)$ satisfies the integral equation (1).

For numerical work, one would define $\tilde{g}_i = \lambda_i^2 c_i$ for $i \leq N$ and $\tilde{g}_i = 0$ for $i > N$. The N being selected as large as necessary for a good approximation. The method used for the variational proceedure would be, perhaps the Trefftz variation of the Rayleigh Ritz method. But here again a knowledge of the characteristic values of L(x,y) are necessary.

4. The Method of Steepest Descents

The equation to be solved is:

$$f(x) + \int_a^b K(x,y)\,g(y)\,dy = 0$$

We form L (x,y):

$$\int_a^b f(x)\,K(x,y)\,dx = \int_a^b K(x,y)\,dx \int_a^b K(x,t)\,g(t)\,dt$$

$$L(y,t) = \int_a^b K(x,y)\,K(x,t)\,dx$$

$$u(y) = \int_a^b L(y,t)\,g(t)\,dt = -v(y).$$

So we wish to find a g (t) such that

$$\int_a^b L(y,t)\,g(t)\,dt + v(y) = 0$$

and the solution of this problem is the same as of (1).

The method of steepest descents can be used as an iterative method providing certain conditions are satisfied. It will be assumed that L(x,y) is such that, for all v(y) under consideration,:

(I) $$\int_a^b v(y)\,dy \int_a^b L(y,t)\,v(t)\,dt \geq m \int_a^b v^2(t)\,dt \quad ; m > 0$$

(II) $$\int_a^b \left\{ \int_a^b L(y,t)\,v(t)\,dt \right\}^2 dy \leq M \int_a^b v^2(y)\,dy \quad ; M > 0.$$

(I) can be written as:

$$\int_a^b dx \left(\int_a^b K(x,y)\,v(y)\,dy \right)^2 \geq m \int_a^b v^2(t)\,dt$$

Now if the characteristic functions of L (y,t) form a closed set, an m can be found such that (I) holds true for all v(y) which are $L^2(a,b)$. If the characteristic functions of L (y,t) do not form a closed set, a set $\{\phi'(y)\}$ is adjoined to yield the complete set $\{\Phi(y)\} = \{\phi(y)\} + \{\phi'(y)\}$
Then for there to exist an m, must all v(y) be orthogonal to the complementary set $\{\phi'(y)\}$.

Apply the Schwartz inequality, now, to (II) and obtain:

$$\int_a^b dt \left| \int_a^b K(y,t) v(y)\,dy \right|^2 \le \int_a^b dt \int_a^b |L(y,t)|^2 dy \int_a^b v(y)^2 dy$$

$$\le \int_a^b v^2(y)\,dy \int_a^b \int_a^b (L(y,t))^2 dy\,dt$$

$$\le M \int_a^b v^2(y)\,dy.$$

So (II) follows if the kernel L (y,t) is $L^2(a,b)$. As in the case of algebraic equations the functional W(g), defined by

$$W(g) = \tfrac{1}{2} \int_a^b g(y)\,dy \int_a^b L(y,t)\, g(t)\,dt + \int_a^b v(y)\, g(y)\,dy$$

is considered. Squaring

$$\int_a^b g(y)\,dy \int_a^b L(y,t)\, g(t)\,dt + \int_a^b v(y)\, g(y)\,dy.$$

yields the inequality:

$$\left\{ 2 \int_a^b g(y)\,dy \int_a^b L(y,t)\, g(t)\,dt \right\} W(g) + \left\{ \int_a^b v(y)\, g(y)\,dy \right\}^2 \ge 0.$$

Schwartz's inequality applied to the second integral and condition (I) applied to the first give:

$$2m \left(\int_a^b g^2(y)\,dy \right) W(g) + \int_a^b g^2(y)\,dy \int_a^b v^2(y)\,dy \ge 0$$

Hence: $$W(g) \ge -\frac{1}{m} \int_a^b v^2(y)\,dy$$

and W(g) has finite lower bound independant of g(y).

Now, to improve and approximation, $g^{(v)}(y)$, we consider the functional:

$$W(g^{(v+1)}) = W(g^{(v)} + Y w(y))$$

where Y, a real number, and w(y), a function, are to be determined.

Then: $$W(g^{(\nu+1)}) = \tfrac{1}{2}\int_a^b (g^{(\nu)}(y) + \gamma w^{(\nu)}(y))dy \int_a^b L(y,t)[g^{(\nu)}(t) + \gamma w^{(\nu)}(t)]dt + \int_a^b v(y)\{g^{(\nu)}(y) + \gamma w^{(\nu)}(y)\}dy.$$

Collecting terms in powers of γ:

$$W(g^{(\nu+1)}) = W(g^{(\nu)}) + \gamma\{\tfrac{1}{2}\int_a^b w^{(\nu)}(y)dy \int_a^b L(y,t)\, g^{(\nu)}(t)dt + \int_a^b v(y)\, w^{(\nu)}(y)dy\} + \gamma^2 \int_a^b w^{(\nu)}(y)dy \int_a^b L(y,t) w^{(\nu)}(t)dt;$$

or, using an obviously convenient notation:

$$W(g^{(\nu+1)}) = W(g^{(\nu)}) + \gamma W_1(g^{(\nu)}, w^{(\nu)}) + \gamma^2 W_2(w^{(\nu)}).$$

Then, completing the square in γ:

$$W(g^{(\nu+1)}) = W(g^{(\nu)}) + W_2(w^{(\nu)})\left\{\left(\gamma + \tfrac{1}{2}\frac{W_1(g^{(\nu)}, w^{(\nu)})}{W_2(w^{(\nu)})}\right)^2 - \tfrac{1}{4}\left(\frac{W_1(g^{(\nu)}, w^{(\nu)})}{W_2(w^{(\nu)})}\right)^2\right\}$$

The largest decrease in W(g) occurs when

$$\gamma = -\tfrac{1}{2}\frac{W_1(g^{(\nu)}, w^{(\nu)})}{W_2(w^{(\nu)})}.$$

In agreement with the result obtained in linear algebraic systems where the correction vector $w^{(\nu)}$ is taken along the gradient of a functional W and turns out to be the residue, the function $w^{(\nu)}(y)$ is taken as:

$$w^{(\nu)}(y) = \int_a^b L(y,t)\, g^{(\nu)}(t)dt + v(y).$$

Hence, the next approximation is:

$$g^{(\nu+1)}(y) = g^{(\nu)}(y) + \gamma w^{(\nu)}(y) = g^{(\nu)}(y) - w^{(\nu)}(y)\frac{\int_a^b (w^{(\nu)}(y))^2 dy}{\int_a^b\int_a^b L(y,t)\, w^{(\nu)}(y) w^{(\nu)}(t)dydt}.$$

Furthermore the decrease in the functional W is:

$$W(g^{(\nu)}) - W(g^{(\nu+1)}) = \frac{\left(\int_a^b w^{(\nu)}(y)\,dy\right)^2}{2\int_a^b w^{(\nu)}(y)\,dy \int_a^b L(y,t)\,w^{(\nu)}(t)\,dt}$$

Hence the sequence $\{W g^{(\nu)}\}$ is monotonic decreasing and it has a finite lower bound. Therefore the sequence converges to a unique limit. It follows that

$$\lim_{\nu\to\infty} \left(W(g^{(\nu)}) - W(g^{(\nu+1)}) \right) = 0$$

and the denominator satisfies (II) so that the equation yields the result

$$\lim_{\nu\to\infty} \int_a^b (w^{(\nu)}(y))^2 dy = 0$$

The sequence $\{w^{(\nu)}(y)\}$ converges to the null function.

Hence: $$\int_a^b L(y,t)\left[g^{(\nu)}(t) - g^{(\nu+1)}(t)\right] dt = 0$$

But, by (i) and by the Schwartz inequality:

$$\left(\int_a^b (g^{(\nu)}(y) - g^{(\nu+1)}(y))^2 dy\right)\left[\int_a^b \left(\int_a^b L(y,t)(g^{(\nu)}(t) - g^{(\nu+p)}(t))^2 dy\right]\right.$$

$$\geq \left(\int_a^b (g^{(\nu)}(y) - g^{(\nu+p)}(y))\,dy \int_a^b L(y,t)(g^{(\nu)}(t) - g^{(\nu+p)}(t))\,dt\right)^2$$

$$\geq m^2 \left(\int_a^b [g^{(\nu)}(y) - g^{(\nu+p)}(y)]^2 dy\right)^2$$

Hence:

$$\left|\int_a^b L(y,t)(g^{(\nu)}(t) - g^{(\nu+p)}(t))\,dt\right| \geq m \left|\int_a^b (g^{(\nu)}(y) - g^{(\nu+p)}(y))\,dy\right|$$

Therefore the sequence, $g^{(\nu)}(y)$, converges to a limit function $g(y)$.

Now if the transformation is assumed to be bounded i.e. satisfies:

$$\left| \int_a^b L(y,t)\, g(t)\, dt \right| \leq K |g(t)| ;$$

Then the "triangle" inequality gives:

$$\left| \int_a^b L(y,t)\, g(t)\, dt + v(y) \right| \leq \left| \int_a^b L(y,t) g^{(v)}(t) + v(y) \right| + \left| \int_a^b L(y,t) \left\{ g(t) dt - g^{(v)}(t) \right\} dt \right|$$

and, in the limit, as $v \to \infty$, it is seen that the limit function $g(t)$ satisfies the given equation. for

$$\left| \int_a^b L(y,t)\, g(t)\, dt + v(y) \right| \leq |w^{(v)}(y)| + K |g(t) - g^{(v)}(t)| .$$

5. A Method of Back Substitution:

The method of back substitution can be applied to the study of integral equations of the first kind:

$$f(x) = \int_a^b L(x,y)\, g(y)\, dy$$

where L (x,y) satisfies the condition of the previous section. An iteration algorithm,

$$g^{(\nu+1)}(y) = g^{(\nu)}(y) + \gamma \left(\int_a^b L(y,x)\, g^{(\nu)}(x)\, dx - f(y) \right),$$

can be easily obtained, for

$$\int_a^b L(x,y) \left[(g(y) - g^{(\nu)}(y)) + g^{(\nu)}(y) \right] dy = f(x).$$

Hence:

$$g(y) - g^{(\nu)}(y) = - \int P(y,x)\, dx$$

where the integral operator $P(y,x)\phi(x)\,dx$ is the operator (if it exists) that has the property

$$\int_a^b P(y,x)\, dx \left(\int_a^b L(x,y)\, p(y)\, dy \right) = p(y)$$

for a p(y) in $L^2(a,b)$. In general, such an operator does not exist; and in every case of interest is not known; for this is equivalent to having solved the problem. Then replace the operator by a real scalar operator, γ, and $g^{(\nu+1)}(y)$ is defined by

$$g^{(\nu+1)}(y) = g^{(\nu)}(y) + \gamma \left[\int_a^b L(y,x)\, g^{(\nu)}(x)\, dx - f(y) \right].$$

This equation is analogous to that obtained in the method of steepest descent; however, in this case γ, being a real constant, does not vary from one iteration to the next. This reduces the rate of convergence but lessens considerably the number of calculations involved. Furthermore, with the kernel satisfying conditions I and II a γ can always be chosen such that the iteration method converges, i.e.

$$\lim_{\nu \to \infty} \left(g^{(\nu+1)}(y) - g^{(\nu)}(y) \right) = 0$$

for almost all y on (a,b). To prove this select

$$g^{(0)}(y) = -\gamma f(y).$$

Then, using an obvious operator formalism,

$$g^{(1)}(y) = \gamma(E + \gamma L) f(y) - \gamma f(y).$$

Hence:

$$g^{(\nu)}(y) = \gamma(E+\gamma L)^{\nu} f(y) - (E+\gamma L)^{\nu-1} f(y) \ldots - \gamma f(y),$$

so that

$$g^{(\nu+1)}(y) - g^{(\nu)}(y) = \gamma(E+\gamma L)^{\nu+1} f(y).$$

Then it is necessary and sufficient for convergence of the iteration process that

$$\lim_{\nu \to \infty} (E+\gamma L)^{\nu+1} = [0]$$

where $[0]$ is the null operator.

Since all of the functions $g^{(\nu)}(y)$, $\nu = 0,1,2,\ldots$ are L^2 (a,b), it is sufficient to consider a function $\phi_i(y)$ satisfying

$$\phi_i(x) = \lambda_i \int_a^b L(x,y)\, \phi_i(y)\, dy.$$

Hence $\lim_{\nu \to \infty} (E + \gamma L)^{\nu} \phi_i(y) = \lim_{\nu \to \infty} \left(1 + \frac{\gamma}{\lambda_i}\right)^{\nu} \phi_i(y)$

Since both λ_i and γ are real it is necessary and sufficient that γ satisfy:

$$0 < -\gamma/\lambda_i < 1.$$

Furthermore from the nature of L(x,y) it follows that, since all λ_i are positive, γ must be negative. Let the λ_i be so ordered that

$$\lambda_1 \leq \lambda_2 \leq \lambda_3 \ldots \lambda_n \leq \lambda_{n+1}.$$

Then, convergence is insured if $0 < -\gamma < \lambda_1$. The conditions on L(x,y) insure that $\lambda_1 > 0$.

An obvious estimate for λ min. $= \lambda_1$ is obtained by considering the defining equation

$$\phi_1(x) = \lambda_1 \int_a^b L(x,y)\, \phi_1(y)\, dy.$$

Then:

$$\int_a^b \phi_i(x)^2 dx = \lambda_1 \int_a^b \phi_1(x) dx \int_a^b L(x,y)\, \phi_1(y) dy .$$

By Schwartz' inequality:

$$\int_a^b \phi_i(x)^2 dx \leq \lambda_1 \left(\int_a^b \int_a^b |L(x;y)|^2 dy\, dx \right)^{1/2} \int_a^b \phi_1(x) dx$$

or, using an accepted notation:

$$\lambda_1 \geq \frac{1}{\| L(x,y) \|} .$$

Hence a possible γ is

$$- \frac{1}{2 \| L(x,y) \|} .$$

Furthermore, even if L(x,y) is not "closed" but is "definite", the above estimate will insure convergence.

6. Solution by Quadrature:

A rather obvious approach is to reduce the integral equation to a set of linear algebraic equations, whose solution will presumably give an approximation to the function g(y). This may be done in any of several ways.

Firstly, the integral may be approximated over some set of values $\{y_i\}_{i=1}^{i=n}$ by a suitable quadrature formulated. As an example, use may be made of the Gauss quadrature formula. The range (a,b) is transformed to (-1, 1) and the integral is approximated by:

$$\int_{-1}^{1} K(x,y)\, g(y)\, dy = \sum_{i=1}^{n} \alpha_i^{(n)} K(x, y_i)\, g(y_i)$$

where the values $\alpha_i^{(n)}$ are real constants which depend on the number of points chosen. The set of values $\{y_i\}_{i=1}^{i=n}$ are the n solutions of $P_n(y) = 0$, where $P_n(y)$ is the n^{th} Legendre polynomial. These zeroes are real, distinct, and lie in the range (-1,1). They have been tabulated for fairly large values of n. By this double choice of weight factors and ordinates the integral, evaluated by n points, is exact if the functions integrated are polynomials of degree 2n - 1 or less. For small n this is a decided improvement over the Lagrangian interpolation polynomial. The given function f(x) is then evaluated at the set of points $\{x_i\}_{i=1}^{i=n}$ with each $x_i = y_i$ i=1,2,...n and the system of equations:

$$f(x_j) = \sum_{i=1}^{n} \alpha_i^{(n)} K(x_j, y_i)\, g(y_i) \qquad j = 1, 2, \ldots, n.$$

is solved for $g(y_i)$ (i = 1,2,...,n) the answer being obtained as a set of values $\{g(y_i)\}_{i=1}^{i=n}$. The integral equation may be used to afford some idea of the error. Once the approximations are introduced, there is little point in debating over uniqueness of solution. Obviously this method may well give a solution to the algebraic porblem where there is none to the integral equation. However, if no relation of the form

$$\sum_{i=1}^{n} P_i\, K(x_j, y_i) = 0 \quad (\text{all } x_j)$$

is satisfied, this method should give a reasonable approximation commensurate with the number of points used. In actuality if the ranges in x and y are different, a set of linear equations with a singular matrix will be obtained providing the intervals in x and y are different.

In addition to the method mentioned any of the iterative schemes for solving simultaneous equations may be used to advantage.

However the important factor of solution stability must be kept in mind. The reduction to a system of algebraic equations involves one approximate quadrature and the remaining calculations are algebraic. An iterative scheme like the steepest descent or back-substitution methods involves repeated integrations each of which is approximate but essentially no solution of a set of linear equations.

It is a matter of extreme complexity to make a precise "error" analysis when an integral equation is solved by numerical methods; however, the need for such an analysis is pressing. Of course once a solution is obtained the error is immediately calculable from the equation itself. An a priori estimation of error, i.e. of the needed fineness of the integration scheme, or of the degree of approximating polynomials, however, is usually a problem of the same order of magnitude as the solution of the given equation. Since iteration methods tend to forestall the unlimited accruing of round-off and truncation errors which occur in each calculation, it appears that an iterative method would show more stability.

The functions can be maintained as tables of functional values for given ordinates or as tables of coefficients of interpolation polynomials, in which case presumably the quadratures can be performed exactly except for round-off error. Certainly, the case of the kernel, a function two variables, if 100 points were required, 10^4 values need be stored, whereas an interpolation polynomial of 99 degree would require as much storage for its co-efficients.

7. The Method of Fourier Integrals

We wish to obtain a numerical solution to an integral equation of the type

$$f(x) = \int_{-\infty}^{\infty} K(x,y)\, g(y)\, dy.$$

This method will be limited to those particular cases where k(x,y) is of the type k(x-y), k(x+y) or k(xy). By a simple variable transformation each of these can be transformed into k(x-y). So we wish to solve

$$f(x) = \int_{-\infty}^{\infty} K(x-y)\, g(y)\, dy.$$

Formally we can obtain a solution in the following way.
Formally:

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{iux} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} dx \int_{-\infty}^{\infty} K(x-y)\, g(y) dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y) dy \int_{-\infty}^{\infty} K(x-y) e^{iux} dx.$$

let x = t + y and obtain

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y) dy \int_{-\infty}^{\infty} K(t) e^{iu(t+y)} dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y) dy\, e^{iyu} \int_{-\infty}^{\infty} K(t) e^{iut} dt.$$

Therefore:

$$F(u) = \sqrt{2\pi}\, G(u)\, K(u)$$

where G(u) and K(u) are the Fourier transforms respectively, of g(y) and k(t).

Finally we obtain:

$$G(u) = \frac{1}{\sqrt{2\pi}} \frac{F(u)}{K(u)}$$

and our desired result is

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du$$

The result is purely formal. The convergence of the improper integrals and the validity of interchange of integration orders must be investigated. The following theorem, however, will be of use in many numerical problems. It is:

Theorem: If: (i) $f(x) \in L^2(-\infty, \infty)$

(ii) $K(x) \in L^2(-\infty, \infty)$

Then if and only if

(iii) $\frac{F(u)}{K(u)} \in L^2(-\infty, \infty)$

is $g(y) \in L^2(-\infty, \infty)$.

We are specifically interested in a numerical solution when the given function f(x) has certain properties:

(i) It is defined, either through lack of complete information or for convenience, at but a finite number of points.

(ii) These points, except perhaps for an assumed constant value at ∞, are distributed over a finite portion of the real axis.

(iii) It will usually be the case that the values tabulated over the set of points $\{x_i\}$ will not be $f(x_i)$ but $f(x_i) + \epsilon(x_i)$ or even more likely $f(x_i) + \epsilon(x)$, where $\epsilon(x)$ is an error function. That is, the simplest case is where the error at any point is functionally defined by

$$\epsilon(x_j) = \epsilon(j)$$

Usually, through having interpolated by one means of another to obtain each $f(x_j)$, the error at any point x_j, will depend on all points, i.e. $\epsilon = \epsilon(x_j)$.

Naturally we are going to assume that a unique solution to the equation exists. And, due to the incompleteness with which f(x) is given, we must expect to get at best an approximation to g(y).

The Numerical Solution

(a) $$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{iux} dx$$

$$\sqrt{2\pi} F(u) = \int_{-x_0}^{x_0} f(x) e^{iux} dx + I(x_0, u)$$

The cut-off at $x = \pm x_0$ in the integration interval and the subsequent dropping of $I(x_0,u)$ is equivalent to folding the transform $F(u)$ against the transform of the unit step function:

$$m(x) = \begin{cases} 0 & , x > x_0 \\ 1 & -x_0 \le x \le x_0 \\ 0 & x < -x_0 \end{cases}$$

i.e., $\int_{-\infty}^{\infty} F(t) \sin \frac{x_0}{x_0}(u-t)dt$. The precise effect this has on the solution is difficult to estimate. Hence we have here a first limitation. It is necessary that a finite x_0 exist such that $I(x_0,u)$ will be sufficiently small. Since $F(u)$ exists this would appear trivial, but if $f(x)$ is a tabulated function or an experimentally measured quantity, knowledge out to x_0, at least, is required. It is possible then then that the range of definition of $f(x)$ may be so limited that $I(x_0,u)$ will be of such a size as to a render a good solution impossible. This difficulty will be discussed later. For the moment we will assume $I(x_0,u)$ to be sufficiently small.

The evaluation of the transforms requires that there be available tables of sines and cosines or that they may be conveniently manufactured in the machine during the process of solution. We assume sets of trigonometric functions will be available.

Further, on the x scale, we assume, since this will usually be most convenient, the values of $f(x)$ to be spaced equidistantly; then let

$$e^{ixu} \to e^{2\pi i h w} \quad ; \; h = 0, 1, 2, \ldots h_{max}$$

and let $x_0 \to h_{max}$. Then $h = \frac{h_{max}\, x}{x_0}$. We have:

$$\sqrt{2\pi}\, F(u) = \int_{-h_{max}}^{h_{max}} [f(x)]_{x \to h} e^{2\pi i h w} \frac{x_0}{h_{max}} dh + I(h_{max}, w)$$

We can always make $dh = 1$. Using, in effect the Euler-MacLaurin summation formula:

$$\sqrt{2\pi}\, F(u) \underset{u \to w}{\simeq} \frac{x_0}{h_{max}} \sum_{-h_{max}}^{h_{max}} [f(x)]_{x \to h} e^{2\pi i h w}.$$

Similarly, for the kernel,

$$\sqrt{2\pi}\, K(u)_{u \to w} \simeq \frac{x_0}{h_{max}} \sum_{-h_{max}}^{h_{max}} [K(x)]_{x \to h} e^{2\pi i h w}$$

Form the quotient:

$$\frac{\sum_{-h_{max}}^{h_{max}} [f(x)]_{x\to h}\, e^{2\pi i h w}}{\sum_{-h_{max}}^{h_{max}} [K(x)]_{x\to h}\, e^{2\pi i h w}} = B_j : \quad j = 0, 1, 2, \ldots \frac{W_{max}}{\Delta w}$$

which is to represent $\frac{F(u)}{K(u)}$ at $\frac{W_{max}}{\Delta w}$ values of w corresponding to u.

That means that for $0 \leq w \leq w_{max}$, u runs over $0 \leq u \leq \frac{2\pi h_{max}}{X_o} W_{max}$

Next we form:

$$g(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(u)\, e^{-iyu}\, du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)}\, e^{-iyu}\, du$$

$$2\pi\, g(y) = \int_{-u_{max}}^{u_{max}} \frac{F(u)}{K(u)}\, e^{-iyu}\, du + J(u_{max}, y).$$

The Euler-MacLaurin integration scheme gives:

$$2\pi\, g(y)_{y\to z} = \frac{u_{max}}{m_{max}} \sum_{m=-u_{max}}^{m=+u_{max}} \left[\frac{F(u)}{K(u)}\right]_{u\to m} e^{-2\pi i m z}\, \Delta m$$

and again let $\Delta m = 1$. The variable ranges are as follows:

(i) x: 0 to x_o

(ii) h: 0 to H(= h_{max}); $\Delta h = 1$

(iii) w: 0 to W(= w_{max}) (say 0·250 of a cycle)

hence from H values of the integrand, we get

$\frac{W}{\Delta w}$ values of F(u).

(iv) u: 0 to $\frac{2\pi HW}{X_o}$

of the $\frac{W}{\Delta w}$ values of $\left[\frac{F(u)}{K(u)}\right]_{u\to w}$ we may use all, or an equidistant subset in evaluating g(y).

Hence while (iv) holds,

(v) m: 0 to M in M jumps; $\Delta m = 1$.

and $\frac{W}{\Delta w} = Ml$; l an integer.

(vi) z: 0 to $Z(= z_{max})$

hence from M values of $\left[\frac{F(u)}{K(u)}\right]_{u \to w}$ we obtain

$Z/\Delta z$ values of $g(y)_{y \to z}$

for which:

(vii) y: 0 to $\frac{2\pi MZ}{4\pi HW} x_0 = \frac{MZ}{HW} x_0$.

Having obtained g(y), the problem is essentially solved.

More may be said about the actual evaluation of the transforms. We can always choose our zero on the x axis so as to maintain the following symmetry:

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_0^{x_0} \left\{ [f(x) + f(-x)] \cos ux + i [f(x) - f(-x)] \sin ux \right\} dx$$

and likewise for K(u). Both F(u) and K(u) are evaluated over the same sets of points. And

$$G(u) = \frac{1}{\sqrt{2\pi}} \frac{F_{cos}(u) + i F_{sin}(u)}{K_{cos}(u) + i K_{sin}(w)} = \frac{1}{\sqrt{2\pi}} \frac{(F_c + i F_s)(K_c - i K_s)}{K_c^2(u) + K_s^2(u)}.$$

Since we integrate over $(-\infty, \infty)$ all odd terms vanish. Hence we must furnish:

(i) $K_c^2(u) + K_s^2(u)$

(ii) $F_c K_c$

(iii) $-F_s K_s$

and

g(y) over the range 0 to y_{max} will be the (cosine + sine) transform of G(u), while being the (cosine - sine) transform of G(u) over $-y_{max}$ to 0.

Cut-off errors: In general, it is very difficult to assess the error in the solution of this problem. The way they arise is at once evident though their accumulative effect on the solution is difficult to determine. The errors are the usual ones in a numerical calculation. We may list them here. They are:

(i) The function f(x), either through expediency, or in fact, is incompletely defined an, as mentioned before, will never, except possibly for a constant value, be given over the entire range $(-\infty, \infty)$.

(ii) In addition, each value, $f(x_i)$ will be contaminated by errors. These will be, at least, truncation errors.

(iii) The infinite integrals must always, unless they can be evaluated exactly by some analytic procedure, be cut off at some finite value. This is equivalent, as mentioned before, to the folding of the precise transform against the step function $\frac{\sin ux_o}{ux_o}$ if x_o is the cut-off value. This means that the frequency spectrum of f(x) is warped by this factor. What is more important, though, the warping cannot be removed form g(y) by any unfolding.

However, we can make some assumptions about f(x) for large X and develop an error formula for the cut-off. We observe that we have two cut-off integrals to evaluate:

(a) $$\int_{x_o}^{\infty} f(-x)\, e^{-iux}\, dx$$

(b) $$\int_{x_o}^{\infty} f(x)\, e^{iux}\, dx$$

As $x_o \to \infty$, these errors $\to 0$; and the high frequency components of F(u) should be largely transformed at any reasonably large x_o and similarly for the cut-off error in the evaluation of K(u)

The integration scheme

The evaluation of F(u) and G(u) is to be done numerically and in those cases where it cannot easily be evaluated analytically K(u) will be also so treated.

Since we are evaluating special types of integrals, i.e. Fourier transforms, it is best that we use an integration scheme which allows us to make most use of the duplicative properties of sines and cosines.

To make best use of the computer's limited storage capacity we must use a process that forms the integration components from a minimum storage list of sines and cosines. For the moment we will stand pat on the Euler-MacLaurin integration scheme using equidistantly spaced points. It is:

$$\int_1^n f(t)dt = \frac{1}{2}[f(1)+f(n)] + \sum_{n=2}^{n-1} f(n) - \int_1^n f'(t)\,S_1(t)dt$$

where

$$S_1(t) = -\frac{1}{2}\sum_{n=1}^{\infty} \frac{\sin 2\pi nt}{2n\pi}$$

We are to evaluate an integral of the form:

$$\int_0^{X_0} f(x)\,e^{iux}dx = \frac{W_0}{H}\int_0^H f(h)\,e^{2\pi ihw}dh$$

where $H = X_0 = h_{max}$

The sum formula yields:

$$\int_0^H f(h)e^{2\pi ihw}dh = \frac{1}{2}[f(0)+f(H)e^{2\pi iHw}] + \sum_{n=1}^{H-1} f(n)\,e^{2\pi inw} - \int_0^H \frac{d}{dh}(f(h)e^{2\pi ihw})S_1(h)dh.$$

An analysis of the error term is desirable for it develops certain facts about the relation of the h and w intervals. Naturally as the number of intervals $\longrightarrow \infty$, the function $S(h) \longrightarrow 0$. The functional relationship between the error and the interval length resides, of course in S(h). We will assume that $|f(h)| < M$, M a finite constant, over(0,H). This is not unreasonable. We consider the error for W = 0:

$$E(0,H) = M\int_0^H \left(\sum_1^{\infty} \frac{1}{2n\pi}\sin 2n\pi h\right)dh$$

$$= M\int_0^H (n - \tfrac{1}{2} - [h])dh = 0$$

([h] is the least integer in (0,1), (1,2),...etc)
For W = M, M = integer

$$E(M,H) = \sum_{n=1}^{\infty}\left(\frac{1}{2n\pi}\right)\frac{M}{2\pi im}\int_0^H e^{2\pi imh}\sin 2\pi nh\,dh = \frac{MH}{4m^2\pi}$$

For the error to be zero, except at $W = e^{2\pi ihw}$, it is necessary that $e^{2\pi ihw}$ be orthogonal to

$$\sum_{n=1}^{\infty} \frac{\sin 2\pi n h}{2\pi n}$$ over (0,H). This it cannot be.

Hence it would appear that, since $W = 0$ is a value we use, we should though the peak about $W = 1$ is narrow, limit W to lie certainly in the range $0 \leq W \leq C < 1$, to minimize the error. This of course is a consequence of the fact that we are fortunate to be dealing with periodic functions, $e^{2\pi ihw}$, super-imposed upon, as we have assumed, a function containing practically no frequencies above some finite cut-off point. This concludes the analytic discussion of the errors exclusive of those occurring from round-off, truncation, or in the function f(x) itself. These latter can probably be treated statistically and this may actually be done later.

Convergence Factors

Before going on to the more numerical treatment we must mention some important considerations concerning convergence. We assume that F(u) and K(u) exist and are finite. Hence F(u) must have a higher order zero at infinity than K(u) to insure convergence of

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du$$

But it is precisely those features which we lack in any approximate solution which make this extremely difficult to achieve. Every error that occurs in the procedure has the effect of adding into F(u), Fourier components which extend the range in u on which F(u) maintains appreciable value. It is only through a sacrifice in "resolution" of g(y) that we will be able to obtain a convergent numerical solution. Nevertheless, admitting the necessity, and even justice, of such a resolution depletion we can obtain solutions which are quite good. The extent of this "smudging out" of g(y) will depend, then, on the nature of g(y), the extent of definition of f(x), and, of course, the accuracy we maintain in the numerical computations. A detailed example will be shown later where the function g(y) is a sum of δ - functions. Even though its transform does not converge we are able to obtain a partially resolved solution; indeed in this worst possible case we do obtain a resolution as fine as desirable by merely extracting an increasing number of values of f(x) and refining our integration scheme, accordingly. Hence to obtain a g(y), we must usually expect to be forced to introduce a convergence factor. A convenient one, since up to constant amplitude and half-width factors it transforms into itself, is e^{-ax^2}. Another is $\frac{\sin^2 ax}{(ax)^2}$ and still a third is $x^2 e^{-ax^2}$. As in the formation of

the first transforms F(u), K(u), this "folding function" or convergence factor gives us a solution at the expense of resolution. For, since we cannot calculate

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du.$$

We introduce the factor T(u) and get:

$$g^*(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} T(u) e^{-iyu} du$$

$$= \int_{-\infty}^{\infty} g(y-s)\, t(s)\, ds.$$

In short, our solution is "folded against" the transform of T(u). The step function yields a $g^*(y)$ which has oscillations of decreasing amplitudes and obscures much of the result. The function e^{-au^2} transforms into $e^{-\frac{1}{4a}s^2}$, up to an amplitude factor. This means that the greater the value is u at which $\frac{F(u)}{K(u)}$ remains well behaved, the narrower will be the spread by which each point will be weighted in the $g^*(y)$. This is almost poetic in its justice. For in just this way does the incomplete definition of f(x) affect the result.

A specific example

Consider the Laplace integral equation

$$f(\xi) = \int_0^{\infty} e^{-\xi\eta} g(\eta)\, d\eta$$

The transformation

$$\eta = e^{-y}, \quad \xi = e^{x}$$

yields

$$e^x f(e^x) = \int_{-\infty}^{\infty} e^{-e^{(x-y)}} e^{x-y} g(e^{-y})\, dy$$

or

$$e^x \tilde{f}(x) = \int_{-\infty}^{\infty} e^{-e^{(x-y)}} e^{x-y} \tilde{g}(y)\, dy.$$

We form:

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^x \tilde{f}(x) e^{ixu} dx$$

$$K(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-e^t} e^t e^{iut} dt$$

and their quotient to give:

$$\tilde{g}(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du.$$

With this kernel, K(u) can be evaluated analytically and is easily seen to be the Euler integral for $\frac{1}{\sqrt{\pi}} \Gamma(1 + iu)_0$. Likewise a necessary and sufficient condition for the behavior of f(x) for large x(i.e. x_0) can be stated. With respect to the two cut off integrals it is as follows: The cut off integrals are:

(a) $$\int_{x_0}^{\infty} e^{-x} \tilde{f}(-x) e^{-iux} dx$$

(b) $$\int_{x_0}^{\infty} e^{x} \tilde{f}(x) e^{iux} dx.$$

Then for convergence, it is both necessary and sufficient that, for (a), for large x, f(-x), be $O(e^{nx})$, $n < 1$ or f(x) be $O(e^{-nx})$, $n < 1$.; and that for (b), for large x, f(x) be $O(e^{nx})$; $n < -1$. For both these cases an error formula can be simply obtained by a single integration by parts. They are:

$$(a) \quad \left| \int_{x_0}^{\infty} e^{-x} O(e^{nx}) e^{-iux} dx \right| \leq \left| \frac{1}{n-1} O(e^{(n-1)x_0}) \right| + \frac{|u|}{|n-1|} \left| \int_{x_0}^{\infty} O(e^{(n-1)x_0}) e^{-iux} dx \right|$$

$$\left| \int_{x_0}^{\infty} e^{-x} O(e^{nx}) e^{-iux} dx \right| \leq \frac{O(e^{(n-1)x_0})}{|n-1| - |u|}.$$

and (b) yields similarly:

$$\left| \int_{x_0}^{\infty} e^{x} f(x) e^{iux} dx \right| \leq \frac{O e^{(n+1)x_0}}{|n+1| - |u|}$$

A more specific result holds for the cut-off in forming K(u). We observe that these two formulas yield an exact expression for the error in the two cut-off's. They are:

$$\int_{x_0}^{\infty} e^{-e^t} e^t e^{itu} dt = -\Gamma_{\ln x_0}(1+iu) + \Gamma(1+iu)$$

and $$\int_{-\infty}^{-x_0} e^{-e^t} e^t e^{itu} dt = -\Gamma_{\ln \frac{1}{x_0}}(1+iu)$$

where $\Gamma_j(x)$ referes to the incomplete gamma function of x with integration limit j.

Having formed the quotient $\frac{F(u)}{\Gamma(1+iu)}$ we observe that:

$$|\Gamma(1+iu)| = \frac{\sqrt{\pi}}{|u|}\frac{1}{|\sinh \pi u|}$$

so that K(u) diminishes quite rapidly. Assuming that a convergence factor is necessary let us introduce e^{-au^2} as one. Our choice of "a" is governed by the behaviour of $\frac{F(u)}{K(u)}$ as u increases. The necessary existence of undamped Fourier terms in F(u) implies that for some value of u, the quotient will begin to grow without bound. We select an "a" such that $\frac{F(u)}{K(u)}e^{-au^2}$ will be essentially zero for all u beyond this critical value. Hence we have:

$$\tilde{g}^*(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{F(u)}{\Gamma(1+iu)}e^{-au^2}e^{-iuy}\,du$$

$$= \frac{1}{\sqrt{2}a}\int_{-\infty}^{\infty}\tilde{g}(y-t)\,e^{-(t/2a)^2}\,dt$$

and we see at once the way in which a small "a" (i.e. a large critical "u") yields a good $\tilde{g}^*(y)$. Here then is the crux of the resolution problem. It is all bound up in the constant "a". The size of "a" will then depend upon just how well we can damp out the "noise" in the evaluation of F(u). That it can be made to be as near zero as we wish is a consequence of the uniqueness of the solution.

8. The Fredholm Integral Equation of the Second Kind:

1. The Existence Theorem:

The equation to be considered is:

$$w(x) = g(x) + \lambda \int_a^b K(x,t)\, w(t)\, dt.$$

The functions w(x), f(x), and K(x,t) are assumed to be real functions of the real variables x,y. As in the preceeding section, if K(x,t) is not symmetrical it may be made to yield a symmetric kernel. The above equation may be written in symbolic form:

$$w(x) = g(x) + \lambda T w(x)$$

or

$$g(x) = (E - \lambda T)\, w(x) \qquad (*)$$

where E is the identity operator and T is the integral operator. Obviously the equation is solved if $(E - \lambda T)^{-1}$ can be obtained. In general, it will be shown that such an inverse exists, subject to certain conditions, as an infinite series in powers of T(i.e. as sums of iterated integrals).

Left multiplication of both sides of (*) by $(E - \lambda T)'$ (the transpose) yeilds

$$u(x) = (E - \lambda L)\, g(x) \qquad (\dagger)$$

where L is a symmetric integral operator and hence has only real eigenvalues. The equation to be considered then is either (*) or (†). For these equations the following existence theorem is fundamental:

If the known functions g(x) and K(x,t) are integrable squared over (a,b), then is w(x) integrable squared over (a,b). Furthermore if λ is such that the homogeneous equation

$$\phi(x) = \lambda T \phi(x)$$

has only the solution $\phi(x) = 0$, then there exists a unique solution w(x). If, however, the homogeneous equation has m independant solutions then there exists a w(x) if and only if g(x) is orthogonal to each of the m solutions of

$$\psi(x) = \lambda T' \psi(x)$$

The solution w(x) is not unique for there can be adjoined to it any linear combination of the m solutions of the homogeneous equation.

2. The Simple Iteration Procedure

We form, from:

$$u(x) = f(x) + \lambda \int_a^b L(x,t)\, u(t)\, dt$$

the iteration procedure:

$$u^{(n)}(x) = f(x) + \lambda T u^{(n-1)}(x).$$

then

$$u_r^{(n)}(x) = f(x) + \lambda T f(x) + \lambda^2 T^2 f(x) + \ldots$$
$$\ldots + \lambda^{n-1} T^{n-1} f(x) + R_n$$

where

$$R_n = \lambda^n T^n u_0(x).$$

let $|u_0(x)| \leq U$, then

$$\|R_n\| \leq |\lambda^n| |U| \|T^n\|$$

a necessary and sufficient condition for convergence is that $\lim_{n\to\infty} \|R_n\| = 0$ so that:

$$\lim_{n\to\infty} |\lambda^n| \|T^n\| \to 0 ;$$

a sufficient condition is, of course,

$$|\lambda| \|T\| < 1$$

or:

$$|\lambda| \|L\| < 1$$

We define the norm of T by the following equations:

$$\|L\| = \left(\int_a^b \int_a^b |L(x,y)|^2 dx dy \right)^{1/2}$$

and

$$\int_a^b L(x,y) u(y) dy = T u(y)$$

We define:

$$\frac{\|T u(x)\|}{\|u(x)\|} \leq \|T\|$$ i.e. the least number for which the inequality is true. It follows that

$$\|Tu\|^2 \leq \|L\|^2 \|u(x)\|^2$$

from the Schwartz inequality. Hence:

$$\|T\| \leq \|L\|$$

Furthermore

$$\|T^K u(x)\| \leq \|T\| \|T^{K-1} u(x)\|.$$

Hence, if

$$S_n = 1 + \sum_{n=1}^{\infty} T^n$$

we have

$$\|S_n\| \leq 1 + \|T\| + \|T^2\| +$$

The solution so obtained is unique. A major problem, then, is how to make use of an iterative process when, say, $\|\lambda L\| \geq 1$.

Let $\|\lambda L(x,y)\|$ be bounded. Then an algebraic procedure can be adjoined to the iteration procedure to provide a solution when $\|\lambda L(x,y)\| \geq 1$. It is always possible to find a set of functions $A_i(x)$, $B_i(x)$ such that:

$$\| L(x,y) - \sum_{i=1}^{N} A_i(x) B_i(y) \| < 1$$

or

$$L(x,y) = \sum_{i=1}^{N} A_i(x) B_i(y) + M(x,y)$$

where N is chosen so that $\| M(x,y) \| < 1$. Hence $M(x,y)$ has an inverse, J, consisting of convergent infinite series of integral operators. Then:

$$u(x) = J f(x) + \sum_{i=1}^{N} (JA_i(x)) \int_a^b B_i(y)\, u(y)\, dy,$$

or, using an obvious notation:

$$(B_j(x), u(x)) = (B_j(x), J f(x)) + \sum_{i=1}^{N} (B_j(x), JA_i(x)) (B_i(y), u(y)); \qquad j = 1, 2, \cdots, N$$

$$(B_j(x), u(x)) = \int_a^b B_j(x)\, u(x)\, dx$$

The solution this set of linear equations for $(B_j(x), u(x))$ when substituted in the equation for u(x) yields the desired solution u(x). This method, in essence, can be regarded as a double iteration scheme, if the algebraic equations are so solved.

9. The Method of Steepest Descent

The equation

$$f(x) = g(x) + \lambda \int_a^b K(x,y) f(y)\, dy$$

in operator form is

$$(E - \lambda T) f(x) = g(x)$$

and T will be assumed to be a positive operator. Then, it follows that the inequalities

$$0 < m \int_a^b f(x)^2 dx \leq \int_a^b f(x) dx \int_a^b K(x,y) f(y) dy \leq M \int_a^b f(x)^2 dx$$

are satisfied, the first from the positive nature of the operator, the second from the bounded nature of the operator and Schwartz's inequality.

The functional w(f) is defined by

$$W(f) = 1/2 \int_a^b \{ f(x)^2 - \lambda f(x) \int_a^b K(x,y) f(y) dy \} dx + \int_a^b f(x) g(x) dx$$

Squaring the expression

$$\int_a^b \{ f^2(x) - \lambda f(x) \int_a^b K(x,y) f(y) dy \} dx + \int_a^b f(x) g(x) dx,$$

yields the inequality

$$2 \left(\int_a^b \{ f^2(x) - \lambda f(x) \int_a^b K(x,y) f(y) dy \} dx \right) W(f) + \left(\int_a^b f(x) g(x) dx \right)^2 \geq 0$$

from which follows

$$2m \left(\int_a^b f^2(x) dx \right) W(f) + \int_a^b f^2(x) dx \int_a^b g^2(x) dx \geq 0.$$

Therefore w(f) has a finite lower bound, independant of f(x), given by

$$W(f) \geq - \frac{\int_a^b g^2(x) dx}{m}$$

If $f^{(v)}(x)$ is an approximation, then an improvement, $f^{(v+1)}(x)$ is obtained by considering

$$W(f^{(v+1)}) = W(f^{(v)}(x) + \alpha \pi^{(v)}(x))$$

where α, a real number, and $r^{(v)}(x)$ are to be determined. Then

$$W(f^{(v+1)}) = W(f^{(v)}) + \alpha \tfrac{1}{2} \int_a^b \{ \pi^{(v)}(x) f^{(v)}(x) - \lambda \pi^{(v)}(x) \int_a^b K(x,y) f^{(v)}(y) dy \} dx$$

$$+ \alpha \int_a^b g(x) \pi^{(v)}(x) dx + \alpha^2 \int_a^b \{ \pi^{(v)}(x)^2 - \lambda \pi^{(v)}(x) \int_a^b K(x,y) \pi^{(v)}(y) dy \} dx.$$

or, in shorter notation:

$$W(f^{(\nu+1)}) = W(f^{(\nu)}) + \alpha \textcircled{H}_1 + \alpha^2 \textcircled{H}_2 .$$

Completing the square in α:

$$W(f^{(\nu+1)}) - W(f^{(\nu)}) = \textcircled{H}_2 \left(\alpha + \tfrac{1}{2} \frac{\textcircled{H}_1}{\textcircled{H}_2} \right)^2 - \tfrac{1}{4} \left(\frac{\textcircled{H}_1}{\textcircled{H}_2} \right)^2 .$$

w(f) decreases most rapidly when

$$\alpha = -\tfrac{1}{2} \frac{\textcircled{H}_1}{\textcircled{H}_2}$$

as before, the function $r^{(\nu)}(x)$ is taken as the residue, i.e.

$$\pi^{(\nu)}(x) = f^{(\nu)}(x) - \lambda \int_a^b K(x,y) f^{(\nu)}(y) dy - g(x).$$

The next approximation is

$$f^{(\nu+1)}(x) = f^{(\nu)}(x) + \alpha \pi^{(\nu)}(x)$$

and this is the iteration algorithm. The sequence, $\{ w(f^{(\nu)}(x)) \}$, being monotonic decreasing, has a finite lower bound and so converges to a unique limit, from which it follows that

$$\lim_{\nu \to \infty} [W(f^{(\nu+1)}) - W(f^{(\nu)})] = 0$$

which by a previous inequality implies

$$\lim_{\nu \to \infty} \int_a^b \pi^{(\nu)}(x)^2 dx = 0$$

from which it follows that $\{ f^{(\nu)}(x) \}$ converges to the solution f(x), almost everywhere.

10. The Method of Back-Substitution

The equation to be considered is:

$$f(x) = g(x) + \lambda \int_a^b K(x,y) f(y) dy$$

The operator in question may be made "definite by reduction to:

$$f(y) = \left[g(y) + \lambda \int_a^b K(x,y) g(x) dx \right] + \lambda^2 \int_a^b K(x,y) dx \int_a^b K(x,y) f(y) dy.$$

or

$$f(y) = u(y) + \lambda^2 \int_a^b L(x,y) f(x) dx.$$

The iteration scheme is:

$$f^{(0)}(y) = -\alpha \, g(y)$$

and

$$f^{(\nu+1)}(y) = f^{(\nu)}(y) + \alpha \, r^{(\nu)}(y)$$

where

$$r^{(\nu)}(y) = f^{(\nu)}(y) - \lambda \int_a^b L(x,y) f^{(\nu)}(x) dx - g(y)$$

and α is to be chosen so that $\lim_{\nu \to \infty} | r^{(\nu)}(y) | = 0$

Then:

$$r^{(\nu+1)}(y) = f^{(\nu+1)}(y) - \lambda \int_a^b L(x,y) f^{(\nu+1)}(x) dx - g(y)$$

$$= f^{(\nu)}(y) + \alpha \, r^{(\nu)}(y) - \lambda \int_a^b L(x,y) \left[f^{(\nu)}(x) + \alpha \, r^{(\nu)}(x) \right] dx - g(y).$$

and

$$r^{(\nu+1)}(y) = r^{(\nu)}(y) + \alpha \, r^{(\nu)}(y) - \lambda \alpha \int_a^b L(x,y) \, r^{(\nu)}(x) dx$$

or, in operator notation:

$$r^{(\nu+1)}(y) = (E + \alpha(E - \lambda T)) \, r^{(\nu)}(y).$$

Since the operators are bounded and commute this becomes:

$$r^{(\nu+1)}(y) = \left[E + \alpha(E - \lambda T) \right]^{\nu} r^{(0)}(y)$$

and α must now be chosen so that $\lim_{\nu \to \infty} r^{(\nu+1)}(y) = 0$. Since the functions considered are $L^2(a,b)$ and T is a bounded "positive" operator, it will be assumed that

$$0 < m \int_a^b f^2(x) dx \leq \int_a^b f(x) dx \int_a^b L(x,y) f(y) dy \leq M \int_a^b f^2(x) dx$$

for any $f(x)$ in $L^2(a,b)$. From the above it follows that

$$\| \eta^{(\nu+1)}(y) \| \leq \| E + \alpha(\epsilon - \lambda T) \|^{\nu} \| \eta^{(1)}(y) \|$$

Hence $\lim_{\nu \to \infty} \| r^{(\nu+1)}(y) \| = 0$

if and only if

$$\| E + \alpha(\epsilon - \lambda T) \| < 1$$

or

$$\int_a^b f^2(x)dx + \alpha \int_a^b f^2(x)dx - \alpha\lambda \int_a^b f(x)dx \int_a^b L(x,y)f(y)dy < \int_a^b f^2(x)dx$$

Then,

$$| 1 + \alpha - \alpha\lambda m | < 1$$

yields a relation for α. At this point it might be mentioned that the convergence criterion is the same as for Newton's method, with $\alpha = -\gamma$ Hence, subject to this change in sign, the two methods converge or diverge together. This method also fails, then, for $\lambda = \lambda_i$, a characteristic value of the homogeneous equation. Newton's method will, in general, however, converge more rapidly.

11. The Solution by Newton's Method

The equation

$$f(x) = g(x) + \lambda \int_a^b K(x,t)\, f(t)\, dt$$

can be represented in operator formas

$$(E - \lambda T) f(x) = g(x)$$

defined on the interval (a,b), where E is the identity operator and T is a positive integral operator. An operator $(E - \lambda T)^{-1}$ exists and in the case of $\| \lambda K(x,t) \| < 1$ is representable as an infinite series in iterated integrals operating on g(x). Newton's method may be employed to obtain $(E - \lambda T)^{-1}$, in those cases where $\| \lambda K(x,t) \| \geq 1$.

By analogy to the Newton algorithm for obtaining reciprocals of numbers, the equation

$$p^{(\nu+1)} = p^{(\nu)} [2E - (E - \lambda T) p^{(\nu)}]$$

is formed where the operator $p^{(\nu)}$ is such that

$$\lim_{\nu \to \infty} p^{(\nu)} = (E - \lambda T)^{-1}$$

Considerations of commutativity may be ignored if $p^{(0)} = \gamma$, a scalar. Then $p^{(0)}$, and hence, by induction, $p^{(\nu)}$ $(\nu = 1,2,\ldots)$ commute with the integral operator λT. This is quite obviously a sufficient, and even convenient, but not necessary constraint on $p^{(\nu)}$ $(\nu = 0, 1,2,\ldots)$. By applying the recurrence relation it is seen that

$$E - (E - \lambda T) p^{(\nu)} = (E - (E - \lambda T) p^{(0)})^{2^\nu}$$

Hence

$$\lim_{\nu \to \infty} p^{(\nu)} = (E - \lambda T)^{-1}$$

if and only if

$$\lim_{\nu \to \infty} \| E - (E - \lambda T) p^{(0)} \|^{2^\nu} = 0$$

A sufficient condition that this be so is that

$$\| E - (E - \lambda T) p^{(0)} \| < 1$$

By a proper choice of $p^{(o)} = \gamma$ this can always be arranged. Hence Newton's method wil always yield a solution by a purely iterative scheme (providing of course the operator T is bounded, which will be assumed at all times). Obviously if $\lambda = \lambda_i$ is a characteristic value of the equation, the inequality above is not satisfied. Likewise such a λ violates a necessary condition, i.e. that all characteristic values of the operator $E - (E - \lambda T) p^{(o)}$ lie outside the unit circle in the complex plane for with a $\lambda = \lambda_i$ the operator has a characteristic value on the unit circle.

12. Solution by Approximate Quadrature

The Fredholm equation

$$f(x) = g(x) + \lambda \int_a^b K(x,y)\, f(y)\, dy$$

may be regarded (and was so treated originally by Fredholm) as the limiting case of n simultaneous linear algebraic equations

$$f(x_i) = g(x_i) + (\lambda/n) \sum_{j=1}^{n} K(x_i, y_j) f(y_j) \qquad i = 1, 2, \cdots, n.$$

as $n \to \infty$. If λ is not a characteristic value of the homogeneous equation, the above, for n large enough, may be assumed to yield a solution $f(x_i)$ $i = 1,2,\ldots,n$, which closely approximates $f(x)$ at least at n points. Since a solution can only be obtained for a finite number of points, in any case, when numerical methods are employed the above appears to be an attractive way to dispose of the Fredholm equation by reducing the transcendental to an algebraic problem.

The above equation represents a solution at "n" distinct points equidistantly spaced. However, it involves the most primitive quadrature method. A linear quadrature scheme which best approximates the integral is desired. One method is to approximate the functions involved by interpolation polynomials, and prevail upon an exact integration to give the algebraic system. Lagrangian or orthogonal polynomials are most commonly employed. Another method involves the use of non equidistantly spaced points and the concept of Gaussian mechanical quadrature. Firstly the interval (a,b) is mapped into (-1,1). The integral is approximated

$$\int_{-1}^{1} K(x,y)\, f(y)\, dy = \sum_{j=1}^{n} \alpha_j^{(n)} K(x, y_j^{(n)})\, f(y_j^{(n)}).$$

with a given fixed n. The $\alpha_j^{(n)}$ are weights which vary with j and each n. The points $y_j^{(n)}$, $j = 1,2,\ldots n$ are the n real, distinct, zeroes of the nth Legendre polynomial and lie in (-1,1). Tables of the $y_j^{(n)}$ are available for $n \leq 10$ and, in time, will be no doubt available for larger n. Such a quadrature yields an exact evaluation of the integral if the integrand is a polynomial of degree $2n - 1$ by the use of only n points.

The method may be generalized to treat any interval providing the points and weights are chosen to depend upon that polynomial orthogonal over (a,b) with respect to a unit weight function. The more common orthogonal functions may be used providing the proper weight functions are introduced. For example, on (-1,1) the Tchebichef polynomials may be used if the integral is written as:

$$\int_{-1}^{1} \left[\frac{K(x,y)\, f(y)}{\sqrt{1-y^2}} \sqrt{1-y^2} \right] dy = \sum_{i=1}^{n} \beta_i^{(n)} K(x, y_i^{(n)})\, f(y_i^{(n)}) \sqrt{1 - y_i^{(n)2}}$$

Here the $\beta_i^{(n)}$ are identical for all i and a given n, though the $y_i^{(n)}$ are not equidistantly spaced being the n zeroes of the Tchebichef polynomial $T_n(y)$. Likewise use may be made of the Laguerre polynomials over $(0,\infty)$ with weight function $e^{-x/2}$ and Hermite polynomials over $(-\infty, \infty)$ with weight $e^{-x^2/2}$.

Bibliography

T. Aberoni, "Antennae, An Introduction to Their Theory", Oxford, 260 pp. (1946).

H. Bateman, "On definite functions," Mess. Math., 37, 91-95 (1908).

H. Bateman, "A formula for the solving function of a certain integral equation of the second kind", Mess. Math., 37, 179-186 (1908).

H. Bateman, "Notes on Integral Equations", Mess. Math., 38, 8-13, 70-75, (1909)

H. Bateman, "The Theory of Integral Equations," Proc. Lond. Math. Soc. 4, 90-115 (1907).

H. Bateman, "On the Inversion of a Definite Integral," Proc. Lond. Math. Soc. 4, 461-498, (1907).

P. D. Crout, "An Application of Polynomial Approximation to the Solution of Integral Equations Arising in Physical Problems", Jour. Math. Phys. 19, 34-92 (1940).

J. B. Diaz and Alexander Weinstein, "Schwarz Inequality and Methods of Rayleigh-Ritz and Trefftz", Journ. Math. and Phys., 26, 133-136 (1947).

E. Picard, "Sur un Theoreme general relatif aux equations integrales de fremiere espece et sw quelques problemes de Physique mathematique," Rend. Circ. Math. Palermo, 29, 79-97 (1910).

E. Hellinger and O. Toeplitz, "Integral gleichungen und Gleichungen mit unendlichirelen Unbekannten", Encyl. Math. Wiss. 2 (3), 1335-1597 (1927).

Maurice Janet, "Equations integrales et applications a certains problemes de la physique mathematieque", Mem. Sci. Math., 101-102, 150 pp. (1941).

W. V. Lovitt, "Linear Integral Equations" McGyrais-Hill, (1924).

E. J. Mystrom, "Uber die Praktische Aublosung von Integral gleichungen mit Arrivendungen auf Randwertaufgaben", Acta. Math. 54, 185-204 (1930).

G. Prasad, "On the numerical solution of integral equations," Proc. Edin. Math. Soc. 42, 46-59 (1924).

A. Reiz, "On the Numerical Solution of Certain types of Integral Equations" Arkiv. for Math. 29A, 21 (1943).

Erhard Schmidt, "Zur Theorie der linearen und nicht linearen integral gleichungen," Math. Ann. 63, 64, (1907).

M. J. Soula, "L'Equation Integrale de Premiere Espece a Limites et Les Fonctions Permutables a Limites Fixes," Mem. Sci. Math., 80, 63 pp. (1936).

G. Temple, "The General Theory of Relaxation Methods Applied to Linear Systems" Proc. Roy Soc. Lond. 169A, 476-500 (1939).

E. C. Tichmarsh, "Theory of Fourier Integrals", Oxford (1937).

E. T. Whittaker, "On the Numerical Solution of Integral Equations," Proc. Roy. Soc. Lond. 94A, 367-383 (1918).

Signed: Alan J Perlis
Alan J. Perlis

Approved: Philip Franklin
Philip Franklin

AJP:bno