This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (https://dspace.lboro.ac.uk/) under the following Creative Commons Licence conditions.

![Creative Commons Licence](https://creativecommons.org/licenses/by-nc-nd/2.5/)

**Attribution-NonCommercial-NoDerivs 2.5**

You are free:
- to copy, distribute, display, and perform the work

Under the following conditions:

**By:** Attribution. You must attribute the work in the manner specified by the author or licensor.

**Noncommercial**. You may not use this work for commercial purposes.

**No Derivative Works**. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license).

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
<table>
<thead>
<tr>
<th>AUTHOR/FILING TITLE</th>
<th>FORBES, R C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESSION/COPY NO.</td>
<td>008026/01</td>
</tr>
<tr>
<td>VOL. NO.</td>
<td></td>
</tr>
<tr>
<td>CLASS MARK</td>
<td>T</td>
</tr>
</tbody>
</table>

- 1 JUL 1994
The Numerical Solution of Singular Non-Linear Integral Equations

by

Raymond C. Forbes

A Doctoral Thesis
Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of the Loughborough University of Technology.

August 1984.

Acknowledgements

The author wishes to extend his sincere thanks to Dr. J. Hyslop and Dr. G.A. Evans for their support and guidance throughout the duration of the work; also to Professor C. Storey, Director of Research, for the use of Departmental facilities.

It is a pleasure to acknowledge the support of the Science and Engineering Research Council who provided the studentship.
Abstract

This thesis investigates the direct application of quadrature methods to solve, iteratively, singular non-linear integral equations. This is divided into two main spheres of interest, firstly, the production of quadrature methods to evaluate integrals which contain singularities, and secondly, the production of generally convergent iterative techniques.

Chapter 2, is devoted to the production of quadrature methods, and involves two different philosophical approaches. The first, uses polynomial transformations to alleviate the singularities, and then employs a finite range high order quadrature method to evaluate the transformed integral. The second, uses variants of the TANH transformation to move the singularities to infinity and then uses a low order quadrature method to evaluate the integrals. Both, schemes have proved to be highly successful and compare favourably with earlier techniques.

Later, in chapter 4, iterative methods are obtained by defining various perturbations, each involving a scaling parameter. This parameter is then used to optimize the convergence of the method by using a variational measure of the error. The solutions of a number of linear equations are also presented, and it is clear that the most useful method is based on the residual function.

This residual perturbation is, in chapter 5, extended to solve non-linear equations, by obtaining a sequence of approximate linear equations. This is done by two different techniques: the first involves the use of Frechet derivative, and the second involves using finite differences between successive iterates. A number of
solutions of non-linear equations are also presented and are used to display the relative merits of each approach. Some of these examples are also singular, requiring the use of the quadrature methods discussed earlier.

Lastly, in chapter 6, the well-known variation-iteration technique for solving eigenvalue problems is extended using the residual perturbation method, and a considerable reduction in the required amount of computational work is achieved.
Contents

1. INTRODUCTION.
   1.1 Background.  2
   1.2 Introduction.  8

2. TRANSFORMATION QUADRATURE METHODS FOR THE EVALUATION
   OF SINGULAR INTEGRALS.
   2.1 Introduction.  12
   2.2 Polynomial Transformations.  15
   2.3 Further Tests and Discussion.  25
   2.4 TANH Transformations.  29
   2.5 Truncation Strategies and Error Estimates.  31
   2.6 Results for Singular Integrals Using
       Schwartz's Optimum h.  39
   2.7 Automatic Quadrature Procedure.  43
   2.8 Test Integrals.  48

3. THE SOLUTION OF SINGULAR INTEGRAL EQUATIONS.  51
   3.1 Introduction.  52
   3.2 Implementation of the Method.  55
   3.3 Numerical Examples.  57
   3.4 Conclusions.  63

4. VARIATION - ITERATION APPLIED TO FREDHOLM EQUATIONS.  65
   4.1 Introduction.  66
   4.2 Description of the Problem.  68
   4.3 The Residual Perturbation.  70
   4.4 Comparison with the Rayleigh-Ritz
       Variational.  73
   4.5 Neumann Series.  76
   4.6 Half-Iterate Extension.  78
   4.7 Other Perturbations.  82
   4.8 Numerical Examples.  89
   4.9 Conclusions.  105
CHAPTER 1

Introduction
1.1 BACKGROUND.

Integral equations fall into many different types. The most well-known are the 2nd kind linear Fredholm equations defined by:

\[ f(x) = g(x) + \lambda \int_{a}^{b} K(x,y) f(y) \, dy \quad a \leq x \leq b \quad (1.1.1) \]

and the 2nd kind linear Volterra equations, defined by:

\[ f(x) = g(x) + \lambda \int_{a}^{x} K(x,y) f(y) \, dy \quad a \leq x \leq b \quad (1.1.2) \]

Also, in the special case when \( g(x) = 0 \) the above equations reduce to the homogeneous forms. Such homogeneous problems only possess solutions for certain values of \( \lambda \) called eigenvalues, and are therefore, eigenvalue problems and are written as:

\[ f = \lambda K f \quad (1.1.3) \]

where \( K \) is an integral operator of either the Fredholm or Volterra type. Associated with the above equations are the 1st kind Fredholm equations, defined by:

\[ g(x) = \lambda \int_{a}^{b} K(x,y) f(y) \, dy \quad (1.1.4) \]

where \( g(x) \) is known and \( f(y) \) unknown, and the 1st kind Volterra equations, similarly defined by:

\[ \lambda \int_{a}^{x} K(x,y) f(y) \, dy = g(x) \quad (1.1.5) \]

Such equations are not considered in the present work, since in practice they are often solved by conversion into the 2nd kind equations, (1.1.1) and (1.1.2) above.
The above equations of the 2nd kind are themselves a special case of a more general non-linear equation called a Hammerstein equation:

\[ f(x) = g(x) + \int_{a(x)}^{b(x)} K(x,y) \psi(y, f(y)) \, dy \]  

(1.1.6)

which itself is a special case of the Urysohn equation defined by:

\[ f(x) = g(x) + \int_{a(x)}^{b(x)} U(x,y, f(y)) \, dy \]  

(1.1.7)

In the present work attention is given to all of the above types of 2nd kind problems, though the limits are taken to be of the constant Fredholm type. Volterra equations may be similarly solved by representing them as a Fredholm equation with a discontinuous kernel, such that

\[ K(x,y) = K(x,y) \text{ when } a \leq y \leq x \]

and \[ K(x,y) = 0 \text{ when } x < y \leq b. \]

Over the years a variety of numerical methods of solving integral equations have been developed. These methods fall into four general categories, which are:

1. The iterative methods.
2. The degenerate kernel methods.
3. The Nyström or approximate quadrature methods, and
4. The projection or basis set methods.

Recently, two authors, Atkinson 1976 and Baker 1977, reviewed the present state of the art of such methods and carried out critical comparisons between them. Generally their conclusions
were that the Nyström and projection methods prove to be superior when applied to the solution of linear Fredholm equations of the second kind, (1.1.1) above. However, even these methods suffer from quite severe difficulties when applied to non-linear equations of the more general Hammerstein or Urysohn type.

Considering each of these categories individually, the iterative methods, which include the well-known method of successive approximation, the direct pointwise application of the Newton-Kantrovich theorem of Noble 1964, and the matrix representation of the same theorem by Moore 1964, lend themselves very naturally to the solution of non-linear equations, since the Newtonian representation of the non-linear problem is implicit within the iterative schemes of the methods. These methods, also have the distinct advantage that any quadrature approximation to the integral operator is independent of the iterative scheme. This gives the user the freedom to use the most suitable quadrature method as and when required. However, these methods also suffer from quite severe disadvantages. For example the method of successive approximations only converges when the Neumann series converges, which may be difficult to determine. The pointwise Newton-Kantrovich theorem, may produce iterative trial functions which contain spurious essential singularities, resulting in divergence, and the iterative scheme due to Moore 1964, may result in the inversion of large nearly singular matrices. Such methods are therefore undesirable, since it may be difficult to determine when these hazards occur.

The degenerate kernel methods rely on the ability to approximate the kernel of an integral equation by a degenerate
kernel, which possess similar properties. Once this has been
done the degenerate kernel is then solved for its eigenvalues
and eigenfunctions and the resolvent kernel is formed. However,
since a small error in the approximation of the kernel may result
in a proportionately larger error in the resolvent kernel, this
method is not very easy to apply in general. Moreover, even
if a good approximation can be obtained to the resolvent kernel,
these methods must also be applied iteratively in the solution of
non-linear equations. These methods are not easily applied to
problems where the kernels contain weak singularities, discontinuities,
or convolutions, since such functions may not be approximated by
a degenerate kernel. Due to this lack of generality, no consideration
is given to these methods in the present work.

The Nyström or approximate quadrature methods, are based on
the idea of representing the unknown function by a grid of points,
arranged on the abscissae of a predetermined quadrature rule.
The problem is then expressed in terms of a matrix equation,
involving a combination of the weights of the quadrature rule and
the kernel calculated at the abscissae in both the x and y terms
independently. The major disadvantage with these methods is
their extreme dependence on one predetermined quadrature rule.
Thus, such methods do not lend themselves to the use of a variety
of quadrature rules, making it again very difficult to integrate
singularities, discontinuities, or convolutions in the kernel.
Again, this method must be applied iteratively to solve a sequence
of Newton approximate linear integral equations when applied to
the solution of the more general non-linear equations. For these
reasons, these methods cannot be used in the solutions of singular non-linear integral equations, as logarithmic displacement singularities cannot be handled, for example. Again due to this lack of generality these methods are not considered further.

Lastly, the projection methods, are based on the idea of choosing a basis set of functions, and then projecting the solutions onto this set by obtaining coefficients to the linear combination. Let the basis set be given by \{\phi_i\}, and let the solution be approximated by:

\[
f(x) = \sum_{i=0}^{N} a_i \phi_i(x)
\] (1.1.8)

The integral equation (1.1.1) may now be written as:

\[
\sum_{i=0}^{N} \left\{ \phi_i(x) - \lambda \int_{a}^{b} K(x,y) \phi_i(y) \, dy \right\} = g(x)
\] (1.1.9)

This equation is then solved for \(a_i\) using one of two philosophical approaches called Galerkin and collocation. In the first, Galerkin, the set \{\phi_i\} is orthogonal and inner products are taken on both sides of the above equation, (1.1.9), with each of the functions \{\phi_i\}. This reduces (1.1.9) to become a simple system of \(N+1\) linear equations which may be solved for \(a_i\). This method is optimal in that it always obtains the best solution in a least squares sense, but also suffers from the allied disadvantage of requiring \((N+1)^2\) double integrals involving the kernel and \((N+1)\) inner-products involving the driving function. This cost, though considerable, is acceptable when solving linear equations, but when the solution of non-linear equations is sought this cost becomes prohibitive, since the method must be applied iteratively.
to obtain a sequence of Newtonian approximations. It may also be noted that since the projection methods furnish the user with freedom of quadrature methods, they are highly appropriate to solve equations with singular and discontinuous kernels.

Secondly, the method of collocation solves (1.1.9) by using a judicial choice of grid points, \( x_i \), \( i = 0,1,2,\ldots,N \) at which the quadratures are evaluated, so as to produce a different set of \( N+1 \) equations, for \( \alpha_i \), from that obtained by Galerkin. This method may suffer from large errors when interpolation is carried out between the grid points. Also the errors incurred at these points are highly dependant on the choice of this grid. This scheme requires fewer quadratures and is therefore computationally much cheaper to implement. Hence, it is more appropriate to use this scheme to generate approximate solutions to non-linear equations, since iterative solution of Newtonian approximations are required. This iterative collocation scheme was originally suggested by Baker 1977, and used to solve a non-linear chemical reactor problem, by Evans, Hyslop and Morgan, 1982, who also made use of low order collocation in the early iterates since high accuracy in the early approximate solutions was not required. The number of grid points was progressively doubled to allow greater accuracy as the proximity of the actual solution was attained. This method also possesses the advantage of allowing variation of the quadrature technique used locally within the range. Hence the method is highly applicable to problems which exhibit singularities and/or discontinuities, either in the integrands or in the derivatives at the integrands, since natural
interpolation occurs for further quadrature function evaluations by direct use of the basis set. This motivates a preliminary study of quadrature methods, for the case when the integrands are known to contain singularities. In this study it was sought to minimize the number of function evaluations to obtain a given accuracy of integration.

Both of these projection methods, suffer from the major disadvantage of solving possibly large systems of linear equations. For this reason, the present work does not make much use of their direct application. Though an embellishment of the collocation method, due to Morgan, 1984, is considered and the results used for comparison.

1.2 INTRODUCTION.

In the present work, the method of successive approximations is extended to produce an iterative scheme, which forces convergence when the Neumann series diverges. This scheme is highly applicable and easily extended to handle non-linear equations.

However, as mentioned above, an efficient quadrature method was sought, to enable the rapid computation of integrals where the integrands contain singular functions. Therefore, extending the above method to make possible the direct solution of singular integral equations. The results of this quest are presented in chapter 2, where two differing philosophies are discussed. The first, uses polynomial transformations to alleviate the singularity, and then applies a powerful quadrature rule to integrate the resulting function. The second philosophy, was to consider a
generalization of the TANH transformation, to displace the singularities towards infinity, and then to use a low order quadrature rule to evaluate the resulting integrand.

In chapter 3, the applicability of these singular quadrature techniques, to the solution of singular non-linear integral equations, is illustrated. Making direct use of the iterative collocation techniques of Morgan, 1984.

Then, in chapter 4, the method of successive approximations is extended to produce a more generally convergent scheme. This scheme uses variational theory to optimize the rate of convergence of a perturbation, whose general form has previously been defined. This chapter, for the purpose of clarity and simplicity is confined to the solution of linear equations, and the solutions to a number of examples are presented and comparison is made with other methods.

Next, in chapter 5, the best of these perturbations, referred to as the residual perturbation is extended to the solution of non-linear equations. Two possible extensions are discussed, the first, using Fréchet derivatives to approximate the non-linearities in a Newtonian sense. The second, uses a Steffensen like approach to accelerate the convergence of the sequence of iterates. These methods are compared on a number of examples and their advantages and disadvantages noted.

Chapter 6, is used to display how this residual perturbation may be applied to improve the solution of eigenvalue problems. Critical comparison is made with the previously accepted method of variation-iteration using a number of examples, and significant improvement is obtained when reasonable trial functions are used.
Finally, chapter 7, is used to discuss the conclusions drawn from the present work, and a number of possible extensions are noted, which may form the basis of future work.
CHAPTER 2

Transformation Quadrature Methods for the Evaluation of Singular Integrals
2.1 INTRODUCTION.

The main theme of this chapter is the investigation of the use of transformations to produce quadrature rules, which will accurately evaluate integrals where the integrands contain weak singularities. This was motivated by the recent work carried out by Evans, Hyslop and Morgan 1983, who proposed a method for the evaluation of finite range singular integrals. Their method was based on the acceleration of a sequence of quadratures, carried out on a judicial choice of sub-intervals, which forced rapid convergence of the $\varepsilon$-algorithm of Wynn 1956. The quadratures on these sub-intervals were obtained by using the pseudo-Gaussian formulae of Patterson 1968. Therefore, combining the advantages of the high accuracy of the Patterson procedure, together with the power of the $\varepsilon$-algorithm, which produces rapid acceleration of the sequences of quadratures for many commonly occurring singularities. The authors made very favourable comparison with earlier work, which includes The Cautious Romberg Extrapolation technique of Cohen 1980, The Accelerated Quadrature Sequence techniques of Chisholm, Genz and Rowlands 1973, and The Special Gaussian Formulae of Harris and Evans 1977. Such comparisons lead to the accumulation of a large number of test integrals which are also used in the present work. However, although this acceleration method of Evans, Hyslop and Morgan 1983 proved to be generally more effective than the earlier procedures, there were still some cases where relatively high numbers of function evaluations were required. This was particularly true for the family of integrals $L_1 - L_6$ (see §2.8) which are singular at
both end points. In some cases it was necessary to utilize in excess of 100 function evaluations to obtain integrals accurate to 10 significant figures. Hence, although this represented a very reasonable achievement for the direct numerical evaluation of such integrals, the possibility of obtaining an alternative high accuracy procedure incorporating a reduction in function evaluations was highly desirable. The importance of such work is realized in the context of extending the work of Evans, Hyslop and Morgan 1981 and 1982 for the solution of integral equations by direct recursive numerical integration techniques, to include the case where the kernel is weakly singular. In such work it is essential to minimize the number of function evaluations required to produce a given accuracy, since a large number of integrations are likely to be required.

The method adopted here, is the well-known technique of variable transformation. A great deal of work has been carried out in the area and two particular variants have proved highly successful. Firstly, there are the types of transformation which make the transformed range finite, whilst alleviating the singularity, and then use a powerful open type quadrature rule like that of Patterson 1968. Such transformations are already widely used, but are usually tailored to the specific integrand in question. In the present context, it was found that simple polynomial transformations could be used which were sufficiently general to be applied in conjunction with Patterson's quadrature to many commonly occurring singularities. Somewhat surprisingly, the order of the polynomials could be increased sufficiently to
produce alleviation of even quite severe singularities without resulting in numerical instability. Care had to be exercised, for the higher order transformations, in evaluating the transformed integrands near the singularities, which was done by using simple expansion techniques. Also, more complex polynomial transformations could be obtained which could alleviate singularities at both ends of the range of integration simultaneously.

Secondly, there are the types of transformation which move the singular points of the integrand to infinity, thus converting a finite range integral, with an end point singularity, into an infinite range integral. A low order quadrature rule such as the trapezoidal rule then proves to be particularly effective in evaluating this infinite range integral. The justification is provided by the Euler-Maclaurin summation formula, since the transformed integral and all its derivatives fall off rapidly at infinity. This approach was also investigated by Schwartz 1969, and subsequently studied by a number of authors including Takahasi and Mori 1973, Stenger 1973, and Haber 1977. One of the transformations suggested by Schwartz has become known as the "TANH RULE" and has proved to be very effective over a wide class of integrals. The method has been developed by Squire 1976 and 1979 who demonstrated the generality and simplicity of the algorithm by presenting a very compact computer programme and testing it on a variety of integrals, of both analytic and singular nature. As pointed out by Evans, Hyslop and Morgan 1983, this implementation suffers from the serious defect of being non-progressive. Hence, if an attempt is made to use this as an automatic integrator by doubling up the
number of points and monitoring the convergence empirically, the function evaluations of the earlier cycles are lost. Therefore, approximately twice the number of function evaluations are computed than are required. In the present work a number of modifications are considered to the "TANH RULE", which preserve the simplicity of the low order quadrature rule used by Squire, and a different progressive implementation is discussed. The viability of this modified TANH RULE as an alternative to the earlier polynomial transformations is considered and critical comparisons made on a number of test integrals. The present chapter forms the basis of two papers published by Evans, Forbes and Hyslop 1983 and 1984.

2.2 POLYNOMIAL TRANSFORMATIONS.

Interest will be confined to singular integrals of the form:

$$S = \int_0^1 f(x) \, dx$$  \hspace{1cm} (2.2.1)

where \( f(x) \) and/or its derivatives are singular at one or both end-points, \( x = 0 \) and \( x = 1 \), of the normalized range. Since more general finite range singular integrals are easily reducible to this form. For example, integral equations frequently contain terms of the form:

$$\int_a^b \phi(t) K(x,t) \, dt, \quad a \leq x \leq b$$  \hspace{1cm} (2.2.2)

in which the kernel \( K(x,t) \) may be singular at \( t = x \) (an important class being the displacement kernels \( K(|x-t|) \)) and, in addition, singularities may exist at the end-points \( t = a \) and \( t = b \). Subdivision at \( t = x \) and subsequent linear transformation will produce
two integrals of the form (2.2.1).

In general the transformation

$$x = \psi(t)$$  \hspace{1cm} (2.2.3)

is utilized which, for convenience, is assumed to map the interval 
$(0, 1)$ on to itself. Various simple polynomial forms for $\psi(t)$ were 
tested for the transformed integral

$$S = \int_0^1 f(\psi(t)) \dot{\psi}(t) \, dt \hspace{1cm} (2.2.4)$$

It is the derivative $\dot{\psi}(t)$ which produces the alleviation of the 
singularity and it is of interest to investigate the effect of 
increasing the degree of the polynomial to achieve this end for 
various types of singularity. In the following sub-sections the 
results of extensive numerical tests are presented.

2.2.1 The Transformation $x = t^n$.

This simple change of variable is, of course, ideal for 
alleviating singularities of the form $F(x) = x^{-\alpha}$. Indeed, the 
choice $n = (1-\alpha)^{-1}$ will remove this type of singularity completely, 
but this approach is really only of value in analytic treatment of 
specific cases. In the present work, concentration is on using 
integer values of $n$ which are significantly large to alleviate 
many commonly occurring singularities, examples including 
$x^{-\alpha}(1-x)^{-\beta}$, $x^{-\alpha}(\ln x)^{\beta}$, $\ln \ln x^{-1}$ etc.

In cases where $f(x)$ is singular at $x = 1$ as well as at $x = 0$, 
the transformation used by Evans, Hyslop and Morgan 1983 is adopted. 
That is the singularity at $x = 1$ is mapped on to $x = 0$ and integral
(2.2.1) is hence expressed in the form:

\[ S = \frac{1}{2} \int_0^1 \left[ f(x/2) + f(1-x/2) \right] \, dx \quad (2.2.5) \]

where the integrand is now singular at \( x = 0 \) only and the transformation \( x = t^n \) may now be applied directly.

It will be noted that care must be exercised in computing the second term, \( f(1-x/2) \), in the integrand as \( x \to 0 \). This point has been commented on by many authors and the commonly adopted remedy is to express \( f(z) \) as \( z \to 1 \) in the form

\[ f(z) = f(1-\delta) \quad (2.2.6) \]

where the "residual" \( \delta \) as well as the variable \( z \) must be accurately available in the function routine. In the present context, \( \delta = t^n/2 \) and, even though \( t \) may be of order of \( 10^{-7} \) when implementing the highest order \( (N = 511) \) Patterson quadrature formula, underflow may be avoided, even for fairly large values of \( n \). Consequently, terms such as \((1-x)^{-\delta}\) may be evaluated without overflow occurring since cancellation is avoided. Note also that for certain "composite" singularities such as \( \ln \ln x^{-1} \), this device is not successful since \( \ln \ln x^{-1} \) cannot be evaluated without overflow under these extreme circumstances, unless arithmetic precision is increased prohibitively. It is necessary in these cases to use analytic expansions such as \( \ln (\delta + \delta^2/2 + \ldots) \) in the function routines.

In Table I results are presented for tests carried out on the various test integrals considered by Evans, Hyslop and Morgan 1983. These integrals arose from various sources and the
corresponding integrands are catalogued in §2.8 for easy reference. The number of quadrature points, $N$, required to obtain a relative error of $10^{-10}$ in each of the integrals is depicted for values of $n$ ranging from $n=1$ to $n=8$. Patterson quadrature is utilized with $N = 3, 7, 15, 31, \ldots, 511$ points, the $N$-point formula integrating exactly a polynomial of degree $(3N+1)/2$. (The corresponding Gaussian formula has polynomial precision $(2N-1)$).

For comparison purposes, the last column shows the number of points used in the acceleration procedure described by Evans, Hyslop and Morgan 1983.

The $n=1$ column corresponds to ignoring the singularity and it will be observed that even the highest order Patterson formula fails in many cases to produce convergence to the required accuracy. This is, indeed, a measure of the difficulty of the integrals considered. Nonetheless, the $x = t^n$ transformation rapidly smooths the integrands and convergence was attained in all cases (except $I_1$) with a maximum of $N = 31$ quadrature grid points. It was found to be necessary to use values of $n$ as large as $n=6$ to achieve convergence in general for the types of singularity tested here. Calculations were carried out for larger values of $n$ ($n=7$ and $n=8$ results are shown here) and it was found that numerical instability due to round-off only began to appear for values approaching $n=10$. There is slight evidence of this effect appearing at $n=8$ for $I_2$ and for $K_4$, so that $n=6$ may be a good practical limit.

It is, of course, possible to produce integrands which would defeat this transformation. For instance, the pathological
TABLE I

Number of Patterson quadrature points ($N$) required to attain relative error $E = 1.0(-10)$ for various integrals, using $x = t^n$ transformation.

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Evans et al. 1983</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>21*</td>
</tr>
<tr>
<td>$I_2$</td>
<td>63</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>$I_3$</td>
<td>-</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>111</td>
</tr>
<tr>
<td>$I_4$</td>
<td>-</td>
<td>15</td>
<td>255</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>127</td>
</tr>
<tr>
<td>$J_1$</td>
<td>127</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>57</td>
</tr>
<tr>
<td>$J_2$</td>
<td>31</td>
<td>3</td>
<td>15</td>
<td>7</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>29</td>
</tr>
<tr>
<td>$J_3$</td>
<td>511</td>
<td>63</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>59</td>
</tr>
<tr>
<td>$J_4$</td>
<td>127</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>51</td>
</tr>
<tr>
<td>$J_5$</td>
<td>-</td>
<td>3</td>
<td>255</td>
<td>3</td>
<td>31</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>45</td>
</tr>
<tr>
<td>$J_6$</td>
<td>-</td>
<td>511</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>133</td>
</tr>
<tr>
<td>$J_7$</td>
<td>-</td>
<td>127</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>89</td>
</tr>
<tr>
<td>$J_8$</td>
<td>-</td>
<td>511</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>75</td>
</tr>
<tr>
<td>$J_9$</td>
<td>-</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>79</td>
</tr>
<tr>
<td>$J_{10}$</td>
<td>-</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>79</td>
</tr>
<tr>
<td>$J_{11}$</td>
<td>-</td>
<td>255</td>
<td>63</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>141</td>
</tr>
<tr>
<td>$K_1$</td>
<td>255</td>
<td>3</td>
<td>15</td>
<td>3</td>
<td>15</td>
<td>7</td>
<td>15</td>
<td>7</td>
<td>45</td>
</tr>
<tr>
<td>$K_2$</td>
<td>-</td>
<td>511</td>
<td>3</td>
<td>31</td>
<td>31</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td>45</td>
</tr>
<tr>
<td>$K_3$</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>511</td>
<td>127</td>
<td>3</td>
<td>3</td>
<td>31</td>
<td>45</td>
</tr>
<tr>
<td>$K_4$</td>
<td>15</td>
<td>7</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>17</td>
</tr>
<tr>
<td>$K_5$</td>
<td>-</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>97</td>
</tr>
<tr>
<td>$K_6$</td>
<td>-</td>
<td>511</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>141</td>
</tr>
<tr>
<td>$K_7$</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>$L_1$</td>
<td>-</td>
<td>15</td>
<td>255</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>127</td>
</tr>
<tr>
<td>$L_2$</td>
<td>-</td>
<td>-</td>
<td>511</td>
<td>63</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>135</td>
</tr>
<tr>
<td>$L_3$</td>
<td>-</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>110</td>
</tr>
<tr>
<td>$L_4$</td>
<td>-</td>
<td>31</td>
<td>255</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>127</td>
</tr>
<tr>
<td>$L_5$</td>
<td>63</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>54</td>
</tr>
<tr>
<td>$L_6$</td>
<td>-</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>88</td>
</tr>
</tbody>
</table>

* 6 figures only.

For the doubly singular integrals $L_1 - L_6$ the number of function evaluations is $2N$, see (2.2.5).
integral $I_1$ would require a transformation of order $n = 10^6$ to remove the singularity. However, the efficiency and generality of the $n = 6$ transformation is remarkable.

2.2.2 The Transformation $x = (1-(1-t)^m)^n$.

When the simple transformation $x = t^n$ was applied to integrals which are singular at both end-points, the form shown in equation (2.2.5) was adopted. Hence, at each grid-point two function evaluations are required, namely $f(x/2)$ and $f(1-x/2)$. In an attempt to remedy this defect the power transformation was applied separately at each end of the interval. Thus in (2.2.1), to alleviate the singularity at $x = 0$ the transformation

\[ x = u^n \]  

(2.2.7)

is used and, if the integrand is still found to be singular at $x = u = 1$, then the further substitution

\[ 1-u = (1-t)^m \]  

(2.2.8)

is adopted. Hence, the function $\psi(t)$ of equation (2.2.4) was defined by:

\[ x = \psi(t) = (1-(1-t)^m)^n \]  

(2.2.9)

and the double singularity alleviating derivative $\dot{\psi}(t)$ by:

\[ \dot{\psi}(t) = nm(1-t)^{m-1}(1-(1-t)^m)^{n-1}, \quad (n,m \geq 1) \]  

(2.2.10)

In implementing these formulae care must be taken to avoid cancellation effects similar to those arising in equation (2.2.5). Thus the binomial expansion
\[ t - (1-t)^m = \binom{m}{1} t - \binom{m}{2} t^2 + \ldots + (-1)^{m+1} t^m \quad (2.2.11) \]

was employed to avoid cancellation as \( t \to 0 \).

The results obtained by applying this transformation to the doubly singular integrals \( L_1 - L_6 \) are shown in Table II. For illustration purposes the exponents \( n \) and \( m \) are taken to be equal. In practice, of course, it would be possible to adjust \( n \) and \( m \) separately to deal with singularities of differing severity at each end-point. However, such a procedure would be purely empirical and shows no generality.

Again, the transformation proves to be efficient in smoothing the doubly singular integrands and very economical results are obtained for order \( n = m = 4 \) and \( 6 \). There is once more some evidence of round-off fluctuations starting to appear for higher orders, due to the increasing amount of arithmetic involved in computing the transformed grid-points, and \( n = m = 6 \) is perhaps a reasonable practical limit for this reason. The results compare favourably with those presented in Table I especially when it is recalled that the number of function evaluations in Table I was really \( 2N \) because of the doubling effect in equation (2.2.5) when \( x = t^n \) was applied to doubly singular integrals.
2.2.3 The Beta Transformation.

The transformation \( x = (1-(1-t)^m)^n \) whilst proving more effective than \( x = t^n \) for dealing with doubly singular integrals suffers from the defect of being markedly skew for larger values of \( n \) and \( m \). This is immediately obvious from the results

\[
\dot{x}(t) \sim n^m t^{n-1} + \ldots \quad (t \rightarrow 0) \hspace{2cm} (2.2.12a)
\]

\[
\dot{x}(t) \sim n^m (1-t)^{m-1} + \ldots \quad (t \rightarrow 1) \hspace{2cm} (2.2.12b)
\]

Hence, for equal values of \( n \) and \( m \), the smoothing effect of the transformation is much greater at the end \( t = 1 \). In addition, the degree of the polynomial transformation is \( nm \) which may lead to round-off instabilities for large values of \( n \) and \( m \) as well as requiring increasing computer time.

### Table II

Number of Patterson points (N) required to attain \( E = 1.0 \times 10^{-10} \) using transformation (2.2.9) with \( n = m \).

<table>
<thead>
<tr>
<th>( n=m )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_1 )</td>
<td>-</td>
<td>15</td>
<td>255</td>
<td>15</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_2 )</td>
<td>-</td>
<td>-</td>
<td>511</td>
<td>63</td>
<td>127</td>
<td>63</td>
<td>127</td>
<td>31</td>
</tr>
<tr>
<td>( L_3 )</td>
<td>-</td>
<td>255</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>63</td>
<td>127</td>
<td>63</td>
</tr>
<tr>
<td>( L_4 )</td>
<td>-</td>
<td>31</td>
<td>255</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_5 )</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_6 )</td>
<td>-</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>
For this reason, the transformation based on

\[ x(t) = c \ t^{n-1} (1-t)^{m-1} \]  \hspace{1cm} (2.2.13)

is proposed, where \( c \) is a normalizing constant. It is clear that

\[ c = 1/B(n,m) \]  \hspace{1cm} (2.2.14)

where \( B(n,m) \) is the Beta function

\[ B(n,m) = \int_0^1 t^{n-1}(1-t)^{m-1} \, dt \]  \hspace{1cm} (2.2.15)

The transformation itself involves the incomplete Beta function quoted by Abramowitz and Stegun 1965, defined as

\[ \beta_t(n,m) = \int_0^t t^{n-1}(1-t)^{m-1} \, dt \]  \hspace{1cm} (2.2.16)

and is written in the form

\[ x(t) = I_t(n,m) = \beta_t(n,m)/\beta(n,m) \]  \hspace{1cm} (2.2.17)

which represents a distribution much used in statistical applications.

This transformation has the advantage of avoiding the skew effect exemplified by equation (2.2.12) and also having degree \((n+m-1)\) as opposed to \(nm\). Additionally, it also avoids the factor of 2 increase in function evaluations when \(x = t^n\) was applied to both end-points of a doubly singular integrand in the manner of equation (2.2.5).

For computational purposes, the binomial expansion

\[ x(t) = I_t(n,m) = t^n \sum_{i=0}^{m-1} (-1)^i (n+i-1)^{-1} \binom{m-1}{i} t^i / \beta(n,m) \]  \hspace{1cm} (2.2.18)
is arithmetically stable for small values of \( t \). To avoid the cancellation effects as \( t \to 1 \) it is suggested that the equation (2.2.18) is utilized for \( 0 \leq t \leq 0.5 \) and that for \( 0.5 \leq t \leq 1 \) the symmetry relation

\[
I_t(n,m) = 1 - I_\delta(n,m) \quad (\delta = 1-t)
\]  

(2.2.19)

should be employed (cf. equation (2.2.6)).

The advantages of the Beta function transformation are demonstrated in Table III. For comparison purposes \( n \) and \( m \) are again set to be equal.

<table>
<thead>
<tr>
<th>( n=m )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_1 )</td>
<td>-</td>
<td>15</td>
<td>255</td>
<td>15</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_2 )</td>
<td>-</td>
<td>-</td>
<td>511</td>
<td>127</td>
<td>63</td>
<td>63</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_3 )</td>
<td>-</td>
<td>127</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_4 )</td>
<td>-</td>
<td>31</td>
<td>255</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_5 )</td>
<td>63</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>( L_6 )</td>
<td>-</td>
<td>63</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>

The number of function evaluations required compares favourably with the requirements of the earlier tables, particularly when the \( 2N \) effect of Table I is recalled and also that the present transformation is of degree \( (2n-1) \) as opposed to \( n^2 \) for Table II, thus a significant reduction in computing time was obtained.
2.3 FURTHER TESTS AND DISCUSSION.

From the previous section it is apparent that the transformation \( x = t^n \) is particularly efficient for the alleviation of single singularity integrands provided values of \( n \) as large as 6 are used in conjunction with Patterson quadrature. The high order polynomial precision of the Patterson prescription means that the analytical part of the transformed integrand may still be represented accurately by a power series and hence accurate quadrature is possible for an economic number of grid points. For double end-point singularities, the Beta function transformation \( x = I_t(n,m) \) is similarly effective provided \( n \) and \( m \) are allowed to increase sufficiently.

As further evidence of these claims, some of the integrals arising from the singular integral equations studied by Delves et al. 1979 and 1981 are considered.

For example, the integral

\[
D_1 = \int_{-1}^{c} \exp(x)\left[(c-x-1)\ln(c-x) + (c-x-\frac{1}{2})(c-x)^{-\frac{1}{2}}\right]dx \quad (2.3.1)
\]

arises from a singular Volterra integral equation with solution \( \exp(x) \). The integrand is singular at the single end-point \( x = c \).

The linear transformation

\[
x + c - (1+c)x
\]

maps the range on to \((0,1)\) with the singularity at \( x = 0 \) and the transformation \( x = t^n \) is then used directly. The results obtained for various values of \( c \) are shown in Table IV.
Table IV

Number of Patterson quadrature points (N) required to attain \( E = 1.0 \times 10^{-10} \) for the integral \( D_n \) using \( x = t^n \). \((x + c - (1+c)x)\).

<table>
<thead>
<tr>
<th>c</th>
<th>n=2</th>
<th>n=4</th>
<th>n=6</th>
<th>n=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.8</td>
<td>127</td>
<td>31</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>-0.4</td>
<td>127</td>
<td>31</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>0.0</td>
<td>127</td>
<td>31</td>
<td>15</td>
<td>31</td>
</tr>
<tr>
<td>0.4</td>
<td>127</td>
<td>31</td>
<td>15</td>
<td>31</td>
</tr>
<tr>
<td>0.8</td>
<td>127</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>1.0</td>
<td>127</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>

The integral

\[
D_2 = \int_0^1 x \ln|x-c| \, dx, \quad 0 \leq c \leq 1 \tag{2.3.3}
\]

is now considered which arose from a Fredholm integral equation with a singular kernel \( \ln|x-c| \) and analytic solution \( x \). Here the range is sub-divided at \( x = c \) producing two integrals with a single end-point singularity at \( x = c \). This is mapped onto \( x = 0 \) of the normalized range for both integrals using linear transformations similar to (2.3.2) and \( x = t^n \) is then used on both integrals separately. The total number of points required for various values of \( c \) are shown in Table V.
Number of Patterson quadrature points \((N)\) required to attain \(E = 1.0(-10)\) for integral \(D_2\) using \(x = t^n\) after sub-division at \(x = c\).

<table>
<thead>
<tr>
<th>(c)</th>
<th>(n=2)</th>
<th>(n=4)</th>
<th>(n=6)</th>
<th>(n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>0.2</td>
<td>126</td>
<td>62</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>0.4</td>
<td>126</td>
<td>62</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>0.6</td>
<td>126</td>
<td>62</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>0.8</td>
<td>190</td>
<td>62</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>1.0</td>
<td>127</td>
<td>31</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Finally the integral

\[
D_3 = \int_{-1}^{1} |x-c|^{-\frac{1}{2}}(1-x^2)^{\frac{3}{4}} \, dx = 3\pi(2-c^2)/(4\sqrt{2}) \tag{2.3.4}
\]

with \(-1 \leq c \leq 1\) is treated. This originated in a Fredholm equation with singular kernel \(|x-c|^{-\frac{1}{2}}\) whose solution is \((1-x)^{3/4}\) which has singular derivatives at the end points \(x = \pm 1\). Subdivision at \(x = c\) produces two integrals with double end-point singularities. Linear transformations map the sub-ranges on to \((0,1)\) and the transformation \(x = \xi_t(n,m)\) is then invoked for both integrals. The total number of quadrature points required for various values of \(c\) are presented in Table VI. Only positive values of \(c\) need be considered since the results for the symmetrically placed negative values are identical. Once again the exponents \(n\) and \(m\) are chosen to be equal for convenience.
TABLE VI

Number of Patterson quadrature points required to attain $E = 1.0(-10)$ for integral $D_3$ using $x = I_t(n,m)$ after sub-division at $x = c$.

<table>
<thead>
<tr>
<th>c</th>
<th>n=2</th>
<th>n=4</th>
<th>n=6</th>
<th>n=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>62</td>
<td>30</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>0.2</td>
<td>62</td>
<td>30</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>0.4</td>
<td>62</td>
<td>46</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>0.6</td>
<td>62</td>
<td>30</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>0.8</td>
<td>62</td>
<td>46</td>
<td>46</td>
<td>62</td>
</tr>
<tr>
<td>1.0</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>31</td>
</tr>
</tbody>
</table>

In all cases considered, these polynomial transformations produce accurate values for the integrals with very appreciable saving in function evaluations compared with the earlier methods. The transformation $x = t^n$ is shown to be particularly suitable for integrands which are singular at a single point only, whilst doubly singular integrals are efficiently evaluated by the Beta function transformation $x = I_t(n,m)$. It has been demonstrated empirically that high order transformations ($n = m = 6$ in particular) are extremely effective in alleviating a wide class of commonly occurring singularities, when used in conjunction with the progressive high order polynomial precision of the Patterson quadrature prescription.
2.4 TANH TRANSFORMATIONS.

As before, attention is confined to singular integrals of the form:

\[ S = \int_{a}^{b} f(x) \, dx \]  \hspace{1cm} (2.4.1)

where \( f(x) \) may be singular at both end-points. Integrands possessing more singularities may be reduced to several integrals of this form by sub-division at the singularities. Generalizing the procedure of Squire 1976, the substitution

\[ x = \frac{1}{2} (b+a) + \frac{1}{2} (b-a) \tanh g(t) \]  \hspace{1cm} (2.4.2)

is proposed in which \( g(t) \to \pm \infty \) as \( t \to \pm \infty \). For example, the choice \( g(t) = t \) gives the classical tanh transformation employed by Squire. Consequently

\[ S = \frac{1}{2} (b-a) \int_{-\infty}^{\infty} f \left( \frac{1}{2} (b+a) + \frac{1}{2} (b-a) \tanh g(t) \right) \text{sech}^2 g(t) \, \dot{g}(t) \, dt \]  \hspace{1cm} (2.4.3)

in which \( \dot{g}(t) \) represents the derived function \( \frac{dg}{dt} \). With a suitable choice of \( g(t) \), the \( \text{sech}^2 g(t) \) factor falls off extremely rapidly as \( t \to \pm \infty \) and controls the behaviour of the transformed integrand even though the singular values \( f(a) \) and \( f(b) \) are being approached. For this reason, when the midpoint rule is applied by setting

\[ t = jh \quad \text{where} \quad j = 0, \pm 1, \pm 2, \ldots \]  \hspace{1cm} (2.4.4)

the summation may be truncated at \( j = \pm N \), say, yielding

\[ S \approx 2h(b-a) \sum_{j=-N}^{N} A_j \, f(x_j) \]  \hspace{1cm} (2.4.5)
where
\[ x_j = \frac{1}{2} (b+a) + \frac{1}{2} (b-a)(1-q_j)(1+q_j)^{-1} \]  \hspace{1cm} (2.4.6)

\[ A_j = q_j(1+q_j)^{-2} g(jh) \] \hspace{1cm} (2.4.7)

and \[ q_j = \exp[-2g(jh)] \] \hspace{1cm} (2.4.8)

As mentioned by Squire 1976, difficulty arises in evaluating \( f(x_j) \) for large values of \( j \). The reason for this difficulty is that cancellation effects are encountered in calculating \( x_j \), which give rise to large errors in computing the function values in close proximity to the singularities. This problem was also encountered, in section 2.2 where the function \( f(x) \) was programmed as a two parameter function \( f(x,\delta) \) in which the residual \( \delta \) represents the distance from the singularity. It is therefore recommended that \( x_j \) in equation (2.4.6) should be expressed as:

\[ x_j = a + (b-a) q_j(1+q_j) \quad j = 0,1,2,3,... \] \hspace{1cm} (2.4.9)
in the vicinity of \( x = a \) and as

\[ x_j = b - (b-a) q_j(1+q_j)^{-1} \quad j = 1,2,3,... \] \hspace{1cm} (2.4.10)

near, \( x = b \). It is also required for the function \( g(t) \) to map the infinite range onto itself bijectively. Therefore, the odd functions defined by:

\[ g(t) = t^v \quad , \quad v = 1,3,5 \] \hspace{1cm} (2.4.11)

and \[ g(t) = c \sinht \quad , \quad c > 0 \] \hspace{1cm} (2.4.12)

were used for the purpose of examining the behaviour of \( g(t) \).
In such circumstances, it is easily seen that (2.4.5) reduces to:

\[ S \approx 2h(b-a)\left\{ A_0 \left[ \frac{1}{2} (b+a) \right] + \sum_{j=1}^{N} A_j \left[ f(a+\delta_j) + f(b-\delta_j) \right] \right\} \tag{2.4.13} \]

in which

\[ \delta_j = (b-a)q_j \left( 1+q_j \right)^{-1} \tag{2.4.14} \]

Equation (2.4.13) is the form adopted for computational purposes, remembering that the functions are programmed as two-parameter procedures. As an example, consider the integral

\[ \int_{-1}^{+1} (x-2)^{-1}(1-x)^{-1/4}(1+x)^{-3/4} \, dx \tag{2.4.15} \]

which was one of the test cases treated by Takahasi and Mori 1973. The integrand \( f(x) \) is then programmed according to the algorithm

\[
\text{if } x < 0 \text{ then } f(x,\delta) = (x-2)^{-1}(1-x)^{-1/4}. \delta^{-3/4} \\
\text{else } f(x,\delta) = (x-2)^{-1} \delta^{-1/4}(1+x)^{-3/4} \tag{2.4.16}
\]

thus avoiding cancellation effects as \( x \to \pm 1 \).

\section*{2.5 TRUNCATION STRATEGIES AND ERROR ESTIMATES.}

In Squire's 1976 work on the tanh t transformation, the grid spacing, \( h \), is selected for a given \( N \) by means of

\[ h = \pi (2N)^{-\frac{1}{2}} \tag{2.5.1} \]

According to Stenger 1973, as \( N \to \infty \) the error of the quadrature rule then falls off to zero as

\[ O\left\{ \| f \| \exp \left\{ \left[ (N/2)^{4} \right] \right\} \right\} \tag{2.5.2} \]
demonstrating the efficiency of the method. However, the non-
progressive nature of the process is clearly indicated by
equation (2.5.1) since a completely new set of grid-points is
generated if \( N \) is doubled.

Some insight into the performance of the quadrature rule
for various choices of \( g(t) \) is provided by considering the
error estimate derived by Schwartz 1969 in his pioneer work on
the method. He demonstrated that the error involved in replacing
the infinite range integration of equation (2.4.3) by a mid-point
quadrature summation (2.4.5) at the points \( t = jh \) is given by:

\[
\varepsilon = O(\exp(-2\pi w/h)) \tag{2.5.3}
\]

where \( w \) represents the distance from the real axis of the nearest
singularity of the integrand in the complex plane. The truncation
error, \( \varepsilon_N \), in the \( N \)th term \( T_N \) of equation (2.4.13) is governed by
the behaviour of the exponential factor

\[
q_N = \exp(-2g(Nh)) \tag{2.5.4}
\]

As suggested by Takahasi and Mori 1973, the overall error is
minimized by equating the truncation and quadrature errors. The
equality of the controlling exponential factors in (2.5.3) and
(2.5.4) produces

\[
2g(Nh) \approx 2\pi w/h \tag{2.5.5}
\]

which would yield an optimum value of \( h \), say \( h_o \), for a given \( N \).
The truncation error must, of course, take into account, the
presence of the factor
appearing in $T_N$. Noting that $\delta_N \simeq (b-a) q_N$ it appears that the $q_N$ controlling factor should be modified to:

$$q_N \chi(q_N)$$  \hspace{1cm} (2.5.7)

where $\chi(\delta)$ characterizes the singular behaviour of the factor (2.5.7), typical examples being $\delta^{-\alpha}$ and $\ln \delta^{-1}$ as $\delta \to 0$. For example, the case $\delta^{-\alpha}$ clearly produces the control factor $q_N^{1-\alpha}$ and equation (2.5.5) is modified to:

$$2(1-\alpha) g(Nh) \simeq 2\pi w/h$$  \hspace{1cm} (2.5.8)

a result obtained by Takahasi and Mori 1973. Different types of singularities will require different modifications and it is for this reason that empirical estimation of the truncation point has been suggested (see section 2.7).

However, the crude, un-modified, estimate $h_0$ from equation (2.5.5) does provide a rough picture of the behaviour of the quadrature formula in practice and suggests in particular that the ratio $w/h$ provides a convenient performance index.

**Example 1.**

In the case

$$g(t) = t^v \hspace{1cm} (v = 1, 3, 5, \ldots ) \hspace{1cm} (2.4.11)$$

it may be shown that $\omega$, the distance of the nearest pole in the complex plane from the real axis may be expressed by:

$$\omega = \left(\frac{\pi}{2}\right)^{1/v} \sin \left(\frac{\pi}{2v}\right)$$  \hspace{1cm} (2.5.9)
and that the corresponding value of $h$ from solving (2.5.5) is given by:

$$ h_o = \left[ \pi \left( \frac{\pi}{2} \right)^{1/\nu} \sin \frac{\pi}{2\nu} N^{-\nu} \right]^{1/(\nu+1)} $$

(2.5.10)

while the appropriate error estimate is:

$$ \epsilon \propto \exp \left( -2\pi \omega / h_o \right) = \exp \left[ -\pi \left( 2N \sin \frac{\pi}{2\nu} \right)^{\nu/(\nu+1)} \right] $$

(2.5.11)

The case $\nu = 1$ corresponds to the classical tanh $t$ transformation and the well-known result,

$$ h_o = \pi / (2N)^{1/2} $$

(2.5.12)

as quoted in equation (2.5.1) is reproduced. The corresponding error estimate for this value is given by:

$$ \epsilon \propto \exp \left[ -\pi \left( 2N \right)^{1/2} \right] $$

(2.5.13)

This agrees with the formula quoted by Takahasi and Mori 1973, but is apparently at variance with the result of Stenger 1973 cited by Squire 1976. There appears to be some notational variation in the literature of Schwartz 1969, Takahasi and Mori 1973, Stenger 1973, Squire 1976 and 1979, and Haber 1977. The confusion seems to be centred around the value of $N$, which is occasionally used to denote the number of terms in the sum (2.4.13) and at other times used to denote the total number of points used, which in the present work, is given by $(2N+1)$ in all cases. It was therefore, considered desirable to test the validity of (2.5.13) for various values of $N$.

For analytic functions the factor involving $\|f\|$ may be omitted.
This was first considered by Schartz 1969 who showed that the asymptotic behaviour of the formula, was described by the exponential factor for large \( N \). Therefore, tests were carried out on the simple analytic function \( f(x) = 1 \). The integral

\[
\int_{-1}^{1} f(x) \, dx = 2
\]

was then evaluated using (2.4.13) for a given value of \( N \), with the prescribed \( h \) value supplied by equation (2.5.12). The actual errors obtained are compared with the estimate, \( \varepsilon \), from (2.5.13) in Table VII below.

<table>
<thead>
<tr>
<th>( N )</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>150</th>
<th>200</th>
<th>250</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta )</td>
<td>3.8(-3)</td>
<td>8.7(-5)</td>
<td>3.7(-7)</td>
<td>5.6(-12)</td>
<td>1.8(-17)</td>
<td>1.0(-21)</td>
<td>2.6(-25)</td>
<td>1.7(-28)</td>
<td>1.7(-31)</td>
</tr>
<tr>
<td>( \varepsilon )</td>
<td>4.8(-5)</td>
<td>7.9(-7)</td>
<td>2.3(-9)</td>
<td>2.3(-14)</td>
<td>5.1(-20)</td>
<td>2.3(-24)</td>
<td>5.2(-28)</td>
<td>3.1(-31)</td>
<td>3.8(-34)</td>
</tr>
</tbody>
</table>

\( \delta = |\text{Actual error}|, \quad \varepsilon = \text{error estimate from (2.5.13)} \)

It will be noted that the estimate \( \varepsilon \) follows the observed error \( \delta \) very closely, apart from a factor, which remains fairly constant over a large range of \( N \). (These calculations were carried out in double precision using the ICL 2900 computer at the University of Nottingham).
It would appear from equation (2.5.11) that the performance may be improved by increasing \( v \), from the classical value of \( v = 1 \). Indeed for the set of values \( v = 1, 3, 5, \ldots \) it follows that \( v = 3 \) should give the smallest error estimate for the values of \( N \) considered. This is due to the fact that as \( v \) is increased, the increasing value of the factor \( N^v/(1+v) \) is off-set by the decrease in the factor \( \sin \left( \frac{\pi}{2v} \right) \). As expected this fact is borne out by the results presented in Table VIII for the actual errors obtained, when the more general transformation \( x = \tanh t^v \) was applied.

**TABLE VIII**

Evaluation of \( \int_{-1}^{1} f(x) \, dx \) with \( f(x) = 1 \) using \( x = \tanh t^v \)

<table>
<thead>
<tr>
<th>( N )</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v=1 )</td>
<td>3.8(-3)</td>
<td>8.7(-5)</td>
<td>4.6(-6)</td>
<td>3.7(-7)</td>
<td>5.2(-9)</td>
<td>1.4(-10)</td>
<td>5.6(-12)</td>
</tr>
<tr>
<td>( v=3 )</td>
<td>3.9(-4)</td>
<td>1.2(-6)</td>
<td>9.3(-9)</td>
<td>3.9(-11)</td>
<td>1.3(-13)</td>
<td>8.3(-20)</td>
<td>1.0(-23)</td>
</tr>
<tr>
<td>( v=5 )</td>
<td>2.1(-3)</td>
<td>3.7(-5)</td>
<td>2.4(-7)</td>
<td>7.8(-10)</td>
<td>5.5(-14)</td>
<td>6.9(-18)</td>
<td>5.6(-22)</td>
</tr>
<tr>
<td>( v=7 )</td>
<td>4.9(-2)</td>
<td>7.8(-4)</td>
<td>4.8(-6)</td>
<td>9.0(-8)</td>
<td>1.9(-11)</td>
<td>4.5(-15)</td>
<td>7.9(-19)</td>
</tr>
</tbody>
</table>

The tabulated values are \( \delta = |\text{actual error}| \), where \( h \) is obtained by (2.5.10). The error estimates provided by equation (2.5.11) were found to follow the \( \delta \) values in a very similar way to the behaviour exhibited in Table I, therefore validating the procedure of Schwartz 1969. It would appear from the present consideration that the transformation

\[
x = \tanh t^3
\]

(2.5.14)
represents a considerable improvement over the classical tanh transformation, and is therefore worth serious consideration as a modification to the Squire procedure.

Example 2.

The alternative example for the function \( g(t) \) given by:

\[
g(t) = c \sinh t
\]  

(2.4.12)

is now considered. In this case the distance of the nearest singularity from the real axis after the application of the transformation is given by:

\[
\omega = \begin{cases} 
\frac{\pi}{2} & 0 < c \leq \frac{\pi}{2} \\
\sin^{-1} \frac{\pi}{2c} & c > \frac{\pi}{2}
\end{cases}
\]  

(2.5.15)

Clearly, a rapid decrease in the value of \( \omega \) ensues when \( c \) is increased beyond \( \pi/2 \). Therefore there is no advantage in pursuing the consideration of values of \( c \) outside the range \( 0 < c \leq \pi/2 \).

In fact the simple value \( c = 1 \), for which \( \omega = \pi/2 \), is found to be very useful and therefore suggested as a practical choice. It has been found in practice that the performance of this transformation is not particularly sensitive to variation in the value of \( c \), around unity. Thus the optimum value of \( h \) is now obtained by solving

\[
\exp(-2\pi\omega/h) = q_N \chi(q_N)
\]  

(2.5.16)

as before, where \( q_N \) is defined by (2.4.8).

Since this results in a non-linear equation, which is not easily solved as in Example 1. A crude estimate for \( h_0 \) may be
obtained by neglecting the singularity correction factor $\chi$ and
is therefore specified by the equation

$$2c \sinh N \, h_o = \frac{2\pi \omega}{h_o}.$$  \hfill (2.5.17)

An approximate solution to this equation for large $N$ is given by:

$$h_o \approx N^{-1} \ln(2\pi N \omega/c).$$  \hfill (2.5.18)

with a corresponding error estimate

$$\varepsilon \sim \exp\left[-2\pi N/\ln(2\pi N \omega/c)\right]$$  \hfill (2.5.19)

which indicates a very high level of performance by this transformation, in that the error is now of the form

$$\exp(-kN/\ln N)$$

where $k$ is a fixed constant.

A better approximation to the value $h_o$ is obtained if (2.5.17) is solved by the method of successive approximations, which yields

$$h_o \approx N^{-1} \left[L - \ln(L - \ln L)\right]$$  \hfill (2.5.20)

where $L = \ln(2\pi N \omega/c)$  \hfill (2.5.21)

with the corresponding error estimate given by

$$\varepsilon \sim \exp(-2\pi \omega/h_o) .$$  \hfill (2.5.22)

The results of which are presented in Table IX below.
2.6 RESULTS FOR SINGULAR INTEGRALS USING THE SCHWARTZ'S OPTIMUM h.

Tests were carried out on the doubly singular integrals used by Evans, Hyslop and Morgan 1983, together with the examples considered by Takahashi and Mori 1973, which are tabulated in
section 2.8. These doubly singular integrals together with their accurate results are catalogued below in Table X.

**TABLE X**

Test Integrals L₁ - L₈

<table>
<thead>
<tr>
<th>Integral</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₁</td>
<td>[\int_0^1 x^{-\frac{1}{2}} (1-x)^{-\frac{1}{2}} , dx = 3.141592653589793]</td>
</tr>
<tr>
<td>L₂</td>
<td>[\int_0^1 x^{-\frac{1}{2}} \ln n , \ln x^{-1} , dx = 0.2318630313168250]</td>
</tr>
<tr>
<td>L₃</td>
<td>[\int_0^1 (1+x)^{-\frac{1}{2}} \ln n , \ln x^{-1} , dx = -0.06281647980603901]</td>
</tr>
<tr>
<td>L₄</td>
<td>[\int_0^1 (1-ln x)^{-1} (-ln x)^{-\frac{1}{2}} , dx = 1.343293421646734]</td>
</tr>
<tr>
<td>L₅</td>
<td>[\int_0^1 \ln x \ln (1-x) , dx = 0.3550659331517735]</td>
</tr>
<tr>
<td>L₆</td>
<td>[\int_0^1 (1-x)^{-1} \ln x , dx = -1.644934066848227]</td>
</tr>
<tr>
<td>L₇</td>
<td>[\int_{-1}^1 (x-2)^{-1} (1+x)^{-3/4} , dx = 1.949054259166746]</td>
</tr>
<tr>
<td>L₈</td>
<td>[\int_{-1}^1 (1-x)^{-\frac{1}{2}} \cos \pi x , dx = -0.690494588746604]</td>
</tr>
</tbody>
</table>

For a specified value of N the Schwartz optimum h obtained from equation (2.5.17) is required to be corrected to include the term \(x\) in equation (2.5.16) to cater for the presence of end-point singularities. To avoid the problem of overflow in
calculating this value, it was found to be beneficial to express this equation in the form:

$$\frac{-2\pi\omega}{h} \approx -2g + \psi(g)$$  \hspace{1cm} (2.6.1)

where

$$\psi(g) = \ln \left[ \chi \left( \exp(-2g) \right) \right]$$  \hspace{1cm} (2.6.2)

and

$$g = g(Nh)$$  \hspace{1cm} (2.6.3)

Nh being the limit of the truncated integral.

For example, in the case where the singular behaviour of the integrand is characterized by the function $\chi(\delta) = \delta^{-\alpha}$, it appears that $\psi(g) = 2\alpha g$. Again, for the logarithmic singularity $\chi(\delta) = \ln \delta^{-1}$, $\psi(g)$ is given by $\ln(2g)$. As an alternative to solving (2.5.16) an approximate solution of the simpler equation (1) above was then obtained, using an initial approximation for the value $h_o$, which had previously been obtained from (2.5.20). It may be noted that it was not necessary to know the value of $h_o$ to any great degree of accuracy, since the decimal precision $P$ of the calculations is roughly described by:

$$10^{-P} = \exp\left(-\frac{\pi^2}{h_o}\right)$$  \hspace{1cm} (2.6.4)

from (2.5.22). Thus $P \approx 4.3/h_o$, so that a small change in $h_o$ will only produce a correspondingly small change in the precision attained.

The results of these calculations for values of $N = 5, 10, 15$ are presented in Table XI below.
TABLE XI

Absolute errors, $\varepsilon$, for integrals $L_1 - L_8$ using tanh (sinh $t$)

<table>
<thead>
<tr>
<th></th>
<th>$N = 5$</th>
<th>$N = 10$</th>
<th>$N = 15$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
</tr>
<tr>
<td>$L_1$</td>
<td>$4.5(-7)$</td>
<td>$1.7(-5)$</td>
<td>$1.2(-11)$</td>
</tr>
<tr>
<td>$L_2$</td>
<td>$3.9(-7)$</td>
<td>$4.3(-5)$</td>
<td>$1.5(-11)$</td>
</tr>
<tr>
<td>$L_3$</td>
<td>$6.3(-8)$</td>
<td>$1.4(-6)$</td>
<td>$4.3(-13)$</td>
</tr>
<tr>
<td>$L_4$</td>
<td>$4.5(-7)$</td>
<td>$4.5(-5)$</td>
<td>$1.2(-11)$</td>
</tr>
<tr>
<td>$L_5$</td>
<td>$1.6(-9)$</td>
<td>$4.9(-7)$</td>
<td>$2.0(-15)$</td>
</tr>
<tr>
<td>$L_6$</td>
<td>$6.3(-8)$</td>
<td>$8.1(-6)$</td>
<td>$4.3(-13)$</td>
</tr>
<tr>
<td>$L_7$</td>
<td>$3.4(-6)$</td>
<td>$1.5(-4)$</td>
<td>$2.7(-10)$</td>
</tr>
<tr>
<td>$L_8$</td>
<td>$4.5(-7)$</td>
<td>$2.8(-3)$</td>
<td>$1.2(-11)$</td>
</tr>
</tbody>
</table>

The number of function evaluations is $(2N+1)$.

The optimum $h, h_o,$ is obtained from equation (2.6.1).

The Schwartz error estimate is $\varepsilon \sim \exp(-\pi^2/h_o)$.

It is clearly seen from the above table that accurate results are obtained for all these integrals with comparatively few function evaluations, errors of order $10^{-10}$ or better being achieved with at most 31 quadrature points and in some cases as few as 21. These results compare extremely favourably with the existing methods of polynomial transformations, and also endorse the prediction made in the previous section using the Schwartz type error analysis. The error estimate, $\varepsilon$, above are optimistic in all cases, since the
factor involving $\|f\|$ has been omitted. Thus direct application of the theory due to Schwartz has little practical use in providing direct estimates of the error in a given quadrature. The theory is, however, very useful in predicting an optimum value of $h_o$ and hence the truncated range required for the transformed integral. Unfortunately even this prediction requires knowledge of the type of singularity which is present in the problem, which is a major drawback to an otherwise highly effective transformation. Therefore the following section will consider an automatic procedure, which seeks to extend the present work to cater for functions about which little may be known.

2.7 AUTOMATIC QUADRATURE PROCEDURE.

The optimum $h$ procedure while producing excellent results in particular cases, suffers from several defects when an attempt is made to utilize it in an automatic manner. As shown above the indeterminate nature of the error estimate $\varepsilon$ means that little confidence may be placed on a prediction of $N$ carried out before quadrature has been attempted. Also any attempt to increase $N$ after attempting some quadrature would result in the loss of the previously calculated function values, since a completely different $h$ would be obtained. Also the necessary information required to use the function $\chi$ in (2.6.1) may not, in general, be available, as would be the case, for example, in the solution of integral equations which possess singular solutions.

Hence the present section is concerned with the discussion of a simple strategy for automating this procedure. This is done
by selecting an initial value for $h$ and then continuing the summation (2.4.13) until the contribution from the $N^{th}$ term $T_N$ is negligible. If the specified truncation error is $\varepsilon_N$ then the summation is continued until

$$|T_N| < \varepsilon_N \sum_{j=0}^{N} T_j$$

(2.7.1)

with

$$T_j = A_j \left[ f(a+\delta_j) + f(b-\delta_j) \right] \quad j \neq 0$$

(2.7.2)

and

$$T_0 = A_0 \left( \frac{1}{2} (b-a) \right)$$

(2.7.3)

the quadrature limit is effectively

$$t = Nh = H.$$  

(2.7.4)

Then once $H$ has been reached the number of points is then doubled and $h$ halved systematically and the successive quadrature estimates are monitored empirically until convergence to a specified tolerance (taken here to also be $\varepsilon_N$) is achieved. At each stage the previously computed function values may be re-utilized, thus providing an economic progressive procedure.

In practice a convenient choice of $h$ is $h = 1$ resulting in the quadrature limit

$$t = N = H$$

(2.7.5)

so that the $h$ values used form the sequence

$$h = 1, \frac{1}{2}, \frac{1}{4}, \ldots$$

(2.7.6)

with corresponding $N$ values given by

$$N = H, 2H, 4H, \ldots$$

(2.7.7)
In the case of the transformation \( \tanh (c \sinh t) \), values of \( H \) of the order of 5.0 were found to be adequate to produce approximately 10 significant figures for all but the most violently singular integrals. In most cases convergence had been attained when \( h = 1/4 \) which resulted in very high accuracy for comparatively low numbers of function evaluations.

It should be emphasized that the ultimate \( h \) values used in this automatic procedure are not the optimum values in the Schwartz sense. Nonetheless, extremely rapid convergence has been observed in all cases tested and the suggested procedure has the advantage of empirical determination of the truncation limit, which takes the singularities into account. It also possesses the distinct advantage of being progressive and therefore economic in evaluation the function values.

Extensive tests have been carried out using the integrals \( L_1 - L_8 \) from section 2.6 and also using the examples from Section 2.2 quoted in Table I. These examples have been collected from the work of various authors and are tabulated in Section 2.8. Results are presented in Table XII below, depicting the total number of function evaluations \( (2N+1) \) required to produce a relative error of \( e_N = 10^{-10} \). For comparison purposes results from the polynomial transformations of Section 2.2 and the acceleration procedures of Evans, Hyslop and Morgan 1983 are also presented.
### TABLE XII

Number of function evaluations, $2N+1$, required to attain relative error $\varepsilon_N = 10^{-10}$

<table>
<thead>
<tr>
<th>Integral</th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I_3$</th>
<th>$I_4$</th>
<th>$J_1$</th>
<th>$J_2$</th>
<th>$J_3$</th>
<th>$J_4$</th>
<th>$J_5$</th>
<th>$J_6$</th>
<th>$J_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceleration Procedure</td>
<td>27</td>
<td>21</td>
<td>57</td>
<td>29</td>
<td>59</td>
<td>51</td>
<td>43</td>
<td>133</td>
<td>89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x = t^6$</td>
<td>-</td>
<td>15</td>
<td>31</td>
<td>15</td>
<td>31</td>
<td>15</td>
<td>15</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>tanh(sinh t)</td>
<td>-</td>
<td>25</td>
<td>31</td>
<td>29</td>
<td>33</td>
<td>15</td>
<td>21</td>
<td>19</td>
<td>21</td>
<td>21</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Integral</th>
<th>$J_8$</th>
<th>$J_9$</th>
<th>$J_{10}$</th>
<th>$J_{11}$</th>
<th>$K_1$</th>
<th>$K_2$</th>
<th>$K_3$</th>
<th>$K_4$</th>
<th>$K_5$</th>
<th>$K_6$</th>
<th>$K_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceleration Procedure</td>
<td>75</td>
<td>79</td>
<td>141</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>17</td>
<td>97</td>
<td>141</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>$x = t^6$</td>
<td>31</td>
<td>31</td>
<td>15</td>
<td>31</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>tanh(sinh t)</td>
<td>25</td>
<td>33</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>25</td>
<td>29</td>
<td>25</td>
<td>25</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Integral</th>
<th>$L_1$</th>
<th>$L_2$</th>
<th>$L_3$</th>
<th>$L_4$</th>
<th>$L_5$</th>
<th>$L_6$</th>
<th>$L_7$</th>
<th>$L_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceleration Procedure</td>
<td>127</td>
<td>135</td>
<td>110</td>
<td>127</td>
<td>54</td>
<td>88</td>
<td>150</td>
<td>105</td>
</tr>
<tr>
<td>Beta Function</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>63</td>
</tr>
<tr>
<td>tanh(sinh t)</td>
<td>21</td>
<td>25</td>
<td>29</td>
<td>25</td>
<td>17</td>
<td>17</td>
<td>33</td>
<td>41</td>
</tr>
</tbody>
</table>

It will easily be observed that the present method achieves the required accuracy in a very economic manner for all the functions tested, only around 30 function evaluations being necessary and somewhat fewer in a number of cases. The exceptions are the near singular integral $I_1$ and the last Takahasi example $L_8$.

In the case of $I_1$, the integrand is the pathological function...
\( x^{-0.999999} \) and all methods tested, with the exception of the acceleration procedure, failed here using 12 figure arithmetic. For \( L_8 \), a slightly greater number, 41, quadrature points are required, not because of the difficulty in dealing with the singular part \((1-x)^{-1}\) of the integrand, but because of the variation of the analytic factor \(\cos \pi x\) over the range.

Comparison with earlier work indicates a very considerable improvement over the acceleration techniques of Evans et al. 1983, particularly in the case of the more difficult doubly singular integrals \( L_1 \) - \( L_7 \). Indeed in most cases the method is comparable with the highly efficient polynomial transformations discussed in Sections 2.2 and 2.3. It may be noted that the spectacular success of the polynomial transformations in isolated instances (e.g. \( K_1 - K_3 \)) is explained by the simplicity of the integrands. For example, the function \( x^{-1/3} \) becomes trivial when the transformation \( x = t^6 \) is employed. The present method is in fact superior to the polynomial transformations in most cases where such simplification does not occur.

Hence it is concluded that the \( \tanh(\sinh t) \) transformation when implemented in the automatic procedure described above is a very efficient method for the evaluation of singular integrals, when economy and accuracy are essential. It is indeed a viable alternative to the Patterson Polynomial prescription described earlier in this chapter. One of the main advantages of this method is its simplicity and since no previously computed weights are required it may be implemented very simply on almost any computer.
However in practice there is very little difference between the two methods put forward in this chapter, since the Patterson polynomial prescription is more able to take account of a high order oscillation which is likely to be present when the integrand contains an unknown function. Such a function is present in the quadratures involved in solving integral equations. This oscillation is due to the interpolating polynomial used to approximate the unknown function and is of varying importance depending on the nature of singularities which it may contain.

### 2.8 Test Integrals.

In all cases, except L7 and L8, the integrals were carried out over the range \( x \in (0,1) \). The two integrals L7 and L8 being carried out over the range \( x \in (-1,1) \).

The test integrals used fall into four main categories, the first being a general set, labelled I₁ - I₄ which were used by Evans, Hyslop and Morgan 1983, and are represented by the following integrands.

- **I₁ - I₄: The Preliminary Integrands.**
  1. \( x^{-\alpha} (\alpha = 1 - 10^{-6}) \)
  2. \( x^{0.95} \exp(x) \)
  3. \( (1+x^2)^{-1} (\ln x)^2 \)
  4. \( x^{-\frac{1}{2}} (1+x)^{-1} \exp(-x) \)

It may be noted that 2. is analytic while 1., 3. and 4. contain singularities at one end of the range. The difficulty with 2. is that it contains a severe singularity in its derivative.
at the point $x = 0$.

The second set of test integrals were originally due to Chisholm et al. 1973, and are labelled $J_1 - J_{11}$. These were also used by Evans et al. 1983, and are represented by the following integrands.

$J_1 - J_{11}$ The Chisholm Examples.

1. $\ln(2\pi x) \sin(2\pi x)$
2. $x^{3/2}$
3. $x^{1/2} \ln x$
4. $x^{3/4} \cos x$
5. $x^{-1}$
6. $(x^{1/4} + x^{1/3})^{-1}$
7. $\ln(1-\cos x)$
8. $x^{-1/2} \ln x$
9. $u(1+u^2)^{-1}$
10. $(1+u)^{-1/2}$
11. $u^{7/2}$ ($u = -\ln x$)

It may also be shown that all of these examples possess singularities either in themselves or in their derivative at one end of the range $(0,1)$.

Thirdly, a set of examples employed by Harris and Evans 1977 were used, these examples are labelled $K_1 - K_7$ and were also used in the work of Evans et al. 1983.

$K_1 - K_7$ The Harris and Evans Examples.

1. $x^{1/4}$
2. $x^{-1/3}$
3. $x^{-2/3}$
4. $x^{7/2}$
5. $(\ln x)^2$
6. $(\ln x)^4$
7. $(1+x^2)^{-1}$
Again these examples possess singularities in their actual functions or in the derivatives near one end only of the range of integration. Lastly a set of examples were used which have either real or removable singularities at both ends of the range at integration, these were labelled $L_1 - L_6$, and also the two examples $L_7$ and $L_8$, due to Takahasi and Mori 1973 are included for comparison. The first Takahasi example, $L_7$, contains singularities at both ends of the range, where the second Takahasi example, $L_8$, only contains a singularity near $x = 1$, and therefore does not really lie in this class of examples.

$L_1 - L_8$ Doubly Singular Examples.

1. $x^{-\frac{1}{4}}(1-x)^{-\frac{3}{4}}$
2. $x^{-\frac{1}{4}} \ln \ln x^{-1}$
3. $(1+x)^{-2} \ln \ln x^{-1}$
4. $(1-\ln x)^{-1}(-\ln x)^{-\frac{1}{4}}$
5. $\ln x \ln(1-x)$
6. $(1-x)^{-1} \ln x$
7. $(x-2)^{-1} (1-x)^{-1/4} (1+x)^{-3/4}$, $x \in (-1,1)$
8. $(1-x)^{-\frac{1}{4}} \cos wx$, $x \in (-1,1)$. 
CHAPTER 3

The Solution of Singular Integral Equations
3.1 INTRODUCTION.

The purpose of this chapter is to illustrate the applicability of the transformation quadrature methods of the previous chapter, to solve integral equations where the kernels contain weak singularities. To demonstrate this applicability many methods of solving integral equations may have been used including the methods of successive approximations, extensions to the Nyström method, the well-known basis set projection methods of collocation and Galerkin and many more. However, each of these methods have their own drawbacks. The main drawback of using the method of successive approximation is the difficulty in guaranteeing convergence. The main problem with using the Nyström method is involved in linking the weights and abscissae of the solution matrix with the variable number of points which may be required in the singular quadratures, and the main difficulty with projection methods is the problem of choosing a basis set of functions which may easily be able to model solutions, where in certain portions of the range the solution is known to vary rapidly. A method was, therefore, sought which displayed more general applicability to solving equations with solutions which possess interpolatory difficulties. Also this method was required to be able to handle a variety of differing quadrature methods easily with a minimum of complications.

Recently, Morgan 1984, solved a non-linear Fredholm integral equation, by using an embellishment to the regular projection method of collocation. In his treatment the range was divided into a grid of points, partitioned into groups of eight. Within each of these groups the points were placed on the zeros of
Chebyshev polynomials, so that the method of Clenshaw-Curtis quadrature may be trivially implemented. Since each group of points was treated as being independent the groups were concentrated into regions of the solution which were known to be difficult to model. The problem was then reformulated into a residual function defined by:

$$ R(F; x) = g(x) + K(F, x) - F(x) $$

(3.1.1)

which provides an estimate of the error at a point $x$ with respect to the given trial function $F$. Hence a perturbation, $\delta F$, was sought such that $R(F + \delta F; x) = 0$ for all $x$. This would only be true when $F + \delta F$ was the exact solution to the Fredholm equation:

$$ f(x) = g(x) + K(f, x) . $$

(3.1.2)

To obtain, $\delta F$, the linearized Fredholm equation:

$$ \delta F(x) = R(F; x) + K'(F, x).\delta F $$

(3.1.3)

requires solution. Where, $K'(F, x).\delta F$ is the Frechet derivative of the non-linear integral operator $K(F, x)$ acting on the function $\delta F$. It is important to note that when $K$ is non-linear (3.1.3) will only provide a linearized approximation to the actual required perturbation, $\delta F$, so reducing the above idea into an iterative scheme. This idea was initially suggested by Baker 1977. Evans et al. 1982, also noted that since this was iterative, the solutions to (3.1.3) are only required to be approximate. Hence collocation was then applied to (3.1.3) using an eight point Chebyshev fit. Each group of points is then solved independently
from the other groups, continuously updating the points of the function and hence re-evaluating the residual function \((3.1.1)\) as consideration is given to each new group. It is important to realize that the equation \((3.1.3)\) will remain stable as convergence is achieved since \(\delta F\) will decrease in magnitude as \(R(F;x)\) decreases in magnitude. The only potential difficulty is that the quadratures required to evaluate \((3.1.1)\) for a given \(x\) will require increased accuracy as the value of \(R(F;x)\) decreases since cancellation will take place. Hence the equation \((3.1.3)\) may be solved approximately to provide an approximate sequence of functions \(\delta F_n\) such that the sum of \(\delta F_n\) tends towards the solution as \(n\) increases.

This method possesses the advantage of simple implementation of quadrature procedures since any procedure may be used to replace the method of Clenshaw-Curtis, and extra points may easily be evaluated by direct interpolation on the basis set, if points with differing positions from the grid points are required. Such interpolation is not found to be very costly since any interpolatory constants may be applied to a large number of function approximations and each interpolation is only carried out to low order. Closely related to this low order approximation is the second advantage of this method, which is the independence of each group of grid points. This enables regions of greater variations in the solution to be modelled accurately by concentration of the grid points. For these reasons this method was adopted to illustrate the direct applicability of the quadrature methods described in chapter 2, to solve singular non-linear integral equations.
3.2 IMPLEMENTATION OF THE METHOD.

The purpose of this section is to describe in more detail how the method of Morgan 1984, was implemented to solve singular problems. Firstly, the iterative sequence of trial functions is defined by the updating equation:

$$F_{n+1}(x) = F_n(x) + \delta F_n(x)$$  \hspace{1cm} (3.2.1)

where $\delta F_n$ is obtained by solving (3.1.3) over each group of points, by the collocation method. Thus, the function $\delta F(x)$ is approximated by the basis set given by:

$$\delta F_i(x) = \sum_{j=0}^{N} \alpha_{ij} y_j(x)$$  \hspace{1cm} (3.2.2)

where $\delta F_i(x)$ is $\delta F(x)$ inside the $i^{th}$ group of points where the range of $x$ has been partitioned into $M$ groups of $N$ points. The ends of these partitions have been connected such that $x_{1,0} = a$, $x_{i,N} = x_{i+1,0}$, and $x_{M,N} = b$, so that the interpolatory polynomials are connected across the range $(a,b)$.

The collocation method is now applied to this system to obtain $M$ sets of $(N+1)$ equations which are not independent, which may be written as:

$$\sum_{j=0}^{N} \alpha_{ij} \left\{ y_j(x_{i,k}) - \int_{x_{i,0}}^{x_{i,N}} \frac{\partial K}{\partial F_n} (x_{i,k}, t, F_n) y_j(t) \, dt \right\} = R(F_n; x_{i,k}) + \sum_{l=1 \atop l \neq i}^{M} \sum_{j=0}^{N} \alpha_{l,j} \int_{x_{l,0}}^{x_{l,N}} \frac{\partial K}{\partial F_n} (x_{l,k}, t, F_n) y_j(t) \, dt$$  \hspace{1cm} (3.2.3)

where the $i^{th}$ subscript denotes the partition presently being considered.
Since it is only required to solve this large set of equations approximately, the summation of integrals on the RHS of (3.2.3) is approximated, by neglecting the terms where $t > i$ and incorporating the terms where $t < i$ into the residual function after the $(i-1)^{th}$ partition has been considered. The above system then reduces to:

$$\sum_{j=0}^{N} \alpha_{ij} \left\{ y_j(x_i,k) - \int_{x_i,0}^{x_i,N} \frac{\partial K}{\partial F_n}(x_i,k,t,F_n) \, y_j(t) \, dt \right\} = R(\tilde{F}_{i,n},x_i,k)$$

(3.2.4)

where $i = 1, \ldots, M$ and $k = 0, \ldots, N$. The result is $M$ systems of $(N+1)$ independent sets of equations. The modified function $\tilde{F}_{i,n}$ which makes this independence possible is given by:

$$\tilde{F}_{i,n}(x) = \begin{cases} 
F_n(x) + 1/2 F_n(x) & \text{when } x \leq x_{i,0} \\
F_n(x) & \text{when } x > x_{i,0}
\end{cases}$$

(3.2.5)

Each of these sets at $(N+1)$ equations is now solved using the well-known method of Gauss elimination, and one full iteration is executed when $i$ has been allowed to range from one to $M$.

The process of partitioning the range has a distinct advantage when singularities are present, since they may be isolated and hence the generally more complex singular quadrature rules may only be used when necessary. However, it may be noted that the singular quadrature rules may be used when the functions do not contain singularities. Though in the present context such implementation is likely to be wasteful when the degree of interpolation required to obtain the function values is taken into
account. Moreover, it may be noted that since each of the above partitioned regions is made up of a small number of grid points, any interpolation will result in an oscillatory approximation. However, such approximations may be hard to integrate numerically when the points are weighted towards one end, as is the case in the singular quadrature rules. In practice, this is found to reduce their efficiency when applied to partitions where singularities are not present. Also, the previous authors, using well-behaved kernels, found that interpolation on the grid points for Clenshaw-Curtis quadrature was superfluous. The present consideration, using not so well behaved kernels, does require, either much larger values of N or, interpolation to enable the quadrature rule to go to higher order, even though the function and kernel may be analytic in the partition being considered.

Also to make maximum use of the fact that the accuracy is only required as the value of the residual function decreases, the present implementation starts the scheme with very few points doubling the value of M periodically to generate improved interpolation as the value of \( R(F_n;x) \) decreases. Also the accuracy of the integration is tuned up as \( R(F_n;x) \) decreases, so as not to waste computer time unnecessarily. In practice, the two philosophies of using transformations to approximate singular integrals (i.e. polynomial and TANH) are found to be almost identical. For this reason they are both used without comment in the following numerical examples.

3.3 NUMERICAL EXAMPLES.

3.3.1 Logarithmic Displacement Kernels: The Delves Examples.

In 1978 Delves solved the logarithmic displacement kernel
integral equation given by:

\[ f(x) = 1.5x + 0.25 - 0.5 \left\{ x^2 \ln x + (1-x^2) \ln(1-x) \right\} + \int_0^1 \ln |x-y| f(y) \, dy \]

(3.3.1)

to eleven figures of accuracy using a five point Chebyshev basis set. This was a dramatic improvement over the previous solution of Baker 1977, who took sixty-four points to guarantee four figures of accuracy. However, despite the apparent difficulties, the above equation is trivial since its solution is known to be \( f(x) = x \). It is, therefore, clear that the solution poses no interpolation difficulties when described by a grid of points since 2 points could interpolate all values of the solution exactly.

The previous authors encountered difficulties associated with integrating the singular kernel. This problem was particularly apparent in the work of Baker 1977 who required most of his points to approximate the singular function and then apply product integration numerically. In contrast, Delves 1978, using semi-analytic techniques was able to escape most of the difficulties, but still could only obtain errors of \( 10^{-2} \) using three points.

The present enhanced collocation method, was applied to the above equation with \( M = 1 \) and \( N = 2 \), it no longer suffers from any problems associated with the singular kernel since the methods of chapter 2 can produce results exact to any required accuracy using 3 point interpolation on \( f(x) \) and the actual singular function.

It may be noted that the initial trial function was taken to be the driving function and since \( M = 1 \) the method reduces to the standard collocation method. Also any errors imposed by
interpolating the fairly complex driving function will merely perturb the initial trial function and will not affect the methods' ability to obtain the solution. In practice, the method was able to obtain an accuracy of thirteen figures when the quadratures were evaluated to thirteen figures, using only 3 points (i.e. $M = 1$, $N = 2$). This example provides confidence that the singular quadrature techniques are highly applicable to such problems. Though in general, it is limited by its simplicity.

3.3.2 Logarithmic Displacement Kernel: The Cohen Example.

In 1980, the example given by:

$$f(x) = 1 + \int_0^1 \ln|x-y|f(y) \, dy$$

(3.3.2)

was solved in a paper on extrapolatory quadratures to evaluate singular integrals by Cohen 1980, using 32 equally spaced grid points. However, it is doubtful if the solution he presents at the value $f(0)$ is as accurate as he claims. Since in practice, it is found that the solution is symmetric and varies very rapidly towards the ends of the range, as $x$ approaches zero and one. Therefore, two variants of the present method were used, and the symmetry property utilized to economize on the amount of work required. Comparison may be made between the present method, and the work of Cohen to derive a better understanding of the problem. These variants are firstly the case when $M = 1$ and $N = 32$ and, secondly, to start the scheme with $M = 1$ and $N = 4$ and then doubling the number of grid points as the accuracy is obtained. To enable the points to be concentrated towards the ends of the range, the first case used grid points spaced on the zeros of
Chebyshev polynomials, and the second case made use of a high order polynomial transformation to concentrate the groups of five points in the region of \( x = 0 \). The first case, which was equivalent to a 65 point collocation method, approximates the value of \( f(0) \) to be 0.57407 which agrees with Cohen's result to three decimal places. The second case, is truly an iterative method, and rapidly achieves the same accuracy as the first case, using six iterations and 17 points. However, to obtain any greater accuracy becomes a very difficult problem. The following table depicts this quoting the iteration number accuracy and number of grid points for each figure obtained.

<table>
<thead>
<tr>
<th>Number of figures</th>
<th>( \max R(x_i) )</th>
<th>Number of grid points</th>
<th>Iteration number</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6.4537442(-5)</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>7.7929615(-6)</td>
<td>33</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>8.7090638(-8)</td>
<td>65</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>1.10966026(-8)</td>
<td>129</td>
<td>15</td>
</tr>
</tbody>
</table>

It is clear, since the number of grid points increases exponentially that the amount of work required to obtain each extra figure becomes prohibitively large. Also to have confidence in any result obtained, these results must be consistent upon further iterations and subsequent doubling of the grid size. Therefore, to check the eight figure results obtained in the fifteenth iteration,
it was required to double the grid to 257 points and iterate further. These 257 points were partitioned into groups of five, and transformed into the region of zero. For comparison purposes the value of the solution obtained at $x = 0$ was found to be 0.57407671, to eight significant figures. It is important to note that such calculations were only made possible by re-writing (3.3.2) as:

$$f(x) = 1 + \int_{0}^{1} \left\{2n|x-y| + 2n(1-x-y)\right\}f(y) \, dy \quad (3.3.3)$$

so making efficient use of the symmetry.

3.3.3 Logarithmic Displacement Kernel: Non-linear Examples.

To further illustrate the power of the quadrature techniques the general Hammerstein equation given by:

$$f(x) = 1 + \int_{0}^{1} \ln|x-y|\psi(f(y)) \, dy \quad (3.3.4)$$

is solved, where the function $\psi$ is given by the equations:

$$\psi = f^s, \quad s = 2, 3, 4, 5 \quad \text{and} \quad \psi = e^f \quad (3.3.5)$$

As before the solutions to these five examples prove to be very badly behaved towards the ends of the range (i.e. when $x = 0$ and $x = 1$). For this reason the above equation (3.3.4) is re-written to make economic use of the symmetry property, to become:

$$f(x) = 1 + \int_{0}^{1} \left\{2n|x-y| + 2n(1-x-y)\right\}\psi(f(y)) \, dy \quad (3.3.6)$$

These equations were then solved, using a scheme of grid points, starting with $M = 1$, $N = 4$, and doubling the value of $M$
periodically as greater accuracy is required. Also to take account of the difficulties in approximating the solution near $x = 0$ these groups of five points are transformed towards this region using a high order polynomial transformation. The results obtained are presented in Table II below, using a similar format to that of Table I above.

**TABLE II**

<table>
<thead>
<tr>
<th>$\psi = f^2$</th>
<th>Number of accurate figures</th>
<th>$\max R(x_i)$</th>
<th>$\text{Number of grid points}$</th>
<th>Iteration number</th>
<th>$f(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6.6747763(-5)</td>
<td>17</td>
<td>8</td>
<td>6.7643003(-1)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5.9792522(-6)</td>
<td>33</td>
<td>10</td>
<td>6.7642809(-1)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>6.6210020(-8)</td>
<td>65</td>
<td>13</td>
<td>6.7642758(-1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.3291028(-7)</td>
<td>129</td>
<td>14</td>
<td>6.7642759(-1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi = f^3$</th>
<th>Number of accurate figures</th>
<th>$\max R(x_i)$</th>
<th>$\text{Number of grid points}$</th>
<th>Iteration number</th>
<th>$f(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6.6521643(-5)</td>
<td>17</td>
<td>8</td>
<td>7.3173922(-1)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5.1185414(-6)</td>
<td>33</td>
<td>10</td>
<td>7.3173730(-1)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5.9699097(-8)</td>
<td>65</td>
<td>13</td>
<td>7.3173682(-1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.1298504(-7)</td>
<td>129</td>
<td>14</td>
<td>7.3173684(-1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi = f^4$</th>
<th>Number of accurate figures</th>
<th>$\max R(x_i)$</th>
<th>$\text{Number of grid points}$</th>
<th>Iteration number</th>
<th>$f(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6.5539381(-5)</td>
<td>17</td>
<td>9</td>
<td>7.6780772(-1)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.2696277(-6)</td>
<td>65</td>
<td>12</td>
<td>7.6780521(-1)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.3304402(-7)</td>
<td>65</td>
<td>13</td>
<td>7.6780539(-1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.0514951(-7)</td>
<td>129</td>
<td>15</td>
<td>7.6780536(-1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi = e^5$</th>
<th>Number of accurate figures</th>
<th>$\max R(x_i)$</th>
<th>$\text{Number of grid points}$</th>
<th>Iteration number</th>
<th>$f(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>8.7928740(-5)</td>
<td>33</td>
<td>10</td>
<td>7.9369531(-1)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4.1343874(-6)</td>
<td>33</td>
<td>11</td>
<td>7.9369531(-1)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.0406501(-7)</td>
<td>65</td>
<td>13</td>
<td>7.9369492(-1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>9.5607684(-8)</td>
<td>129</td>
<td>15</td>
<td>7.9369488(-1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi = f^6$</th>
<th>Number of accurate figures</th>
<th>$\max R(x_i)$</th>
<th>$\text{Number of grid points}$</th>
<th>Iteration number</th>
<th>$f(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.1711954(-4)</td>
<td>17</td>
<td>10</td>
<td>1.4168218(-1)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2.2259270(-6)</td>
<td>33</td>
<td>14</td>
<td>1.4167598(-1)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.1697615(-7)</td>
<td>65</td>
<td>18</td>
<td>1.4167609(-1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3.2083933(-8)</td>
<td>65</td>
<td>19</td>
<td>1.4167612(-1)</td>
<td></td>
</tr>
</tbody>
</table>
The first thing to be noted about these results is that though convergence has taken place, and the results are constant across the range \((0, \frac{1}{2})\) by iteration and doubling the number of points, they are not confirmed by any other method. This, however, is rectified later in chapter 5 where these same examples are solved to obtain identical solutions to those above. It is, therefore, clear that the transformation quadrature method of the previous chapter is very applicable to the solution of non-linear integral equations. It is also important to realize that the value \(\max R_{i}(x)\) is an extreme upper bound on the convergence as it relates to the residual function prior to the calculation of the \(n^{th}\) iteration. It is therefore only of academic interest, showing that the residual function does tend towards zero as the solution is reached. It is, however, also clear that this method does often require a large number of grid points coupled with a large number of iterations to obtain 8 figures of accuracy, indicating the expense of such a scheme.

3.4 CONCLUSIONS.

In conclusion it has been clearly shown by the numerical examples in this chapter that the transformation quadrature techniques of the previous chapter are highly applicable to solve singular integral equations. It may be noted that the examples considered all contain displacement type singularities in the kernel, which are generally extremely difficult to overcome. It is also seen that though the method of Morgan 1984 is particularly good at solving linear equations with solutions which are simple to
interpolate, it converges fairly slowly, when a large grid of points is required on highly non-linear problems. It is therefore proposed, in later chapters, to use the transformation methods of chapter 2 to explore the possibility of producing methods which may converge more rapidly.
CHAPTER 4

Variation-iteration applied to
Fredholm Equations
4.1 INTRODUCTION.

In this chapter the variational and iterative ideas are unified to produce a method for finding the solution of Fredholm equations. Previously, such ideas have been used to solve integral equation eigenvalue problems, in the now well-known variational-iteration method considered by Morse and Feshbach 1953. This method is known to be highly successful, and many applications have been considered. More notably the complicated equations involved in atomic self consistent field calculations, considered by Robinson et al. 1970, Hyslop and Rees 1976, Blakemore et al. 1977 and Tricomi 1957, have proved the validity of the method in the solution of physical problems. This motivated the present work to consider the extension of such ideas to the general solution of non-eigenvalue problems. The problem was to construct a functional, which may be used to produce an iterative sequence of trial functions, taking due account of the driving function which has not been present in the earlier work. The best known and most widely used method of iterative solution of Fredholm equations is the method of successive approximations, which suffers from the major disadvantage, that convergence exists, only when the Neumann series is convergent, and then it only exhibits linear convergence. Over the last two decades, a number of people have sought to improve this method. Firstly, Petryshyn 1963, considered a number of iterative schemes, which under certain restrictive conditions, resulted in convergence. However, such conditions often prove harder to obtain than solving the equations themselves. Next, Moore 1964, used matrix theory to obtain Newton-like perturbations.
However, this approach suffers from the disadvantage of requiring the inversion of comparatively large matrices. Later, Noble 1964, applied the Newton and secant methods directly to solve integral equations. Unfortunately, either method may break down, when division by an accidental zero occurs, at some point in the range. In Newton's Method, this occurs when the Fréchet derivative operator yields such a zero, and in the secant method when the second difference between the iterated functions is zero. In the next year, Rall 1965, used matrices of Riemann sums, acting iteratively in place of the integral operators, therefore obtaining a vector representation of the unknown function. Later, Anselone 1966 obtained a solution to an elasticity problem by using direct application of the Newton-Kantrovich theorem 1957. However, he avoided the problem of zeros in the Fréchet derivatives by changing the variable when such situations arose. More recently, Schelin 1973 used over-relaxation techniques to solve functional equations, which also suffered from the possibility of division by accidental zeros. Later, Corradi 1975, considered solutions to the Chandrasekhar equation by applying acceleration techniques to a pair of monotonically convergent sequences, which lay above and below the actual solution respectively. This method clearly has no general applicability as such sequences may not generally be easy to obtain. Most of the more generally applicable work comes from the original work of Kantrovich 1957 who generalised Newton's method to solve non-linear equations in functional space, of which integral equations are a special case.

However, few practical examples of the use of such techniques
are published, mainly due to the relatively low computational power available to such authors, and the fact that analytical examples prove to be very cumbersome. Much more recently, Morgan 1984 has considered using ε-algorithm acceleration to improve the convergence of the sequence of iterates produced by successive approximations in a point-wise fashion.

4.2 DESCRIPTION OF THE PROBLEM.

In the present chapter the difficulties associated with numerical quadrature are overcome by the application of the powerful quadrature methods of Clenshaw-Curtis and Patterson. In cases where such methods may not be used directly, due to the presence of weak singularities in the integrands or their derivatives, the transformation methods of chapter 2 are applied.

In this chapter consideration will be limited to the case of linear Fredholm equations given by:

\[ f(x) = g(x) + \int_{a}^{b} k(x,y) f(y) \, dy \]  

(4.2.1)

which is also written in operator notation as \( f = g + Kf \). The more general case of non-linear equations will be dealt with in a later chapter. The residual function, with respect to a given trial function \( f_n(x) \) is defined as:

\[ R_n(x) = g(x) + \int_{a}^{b} k(x,y) f_n(y) \, dy - f_n(x) \]  

(4.2.2)

so as to be identically zero in the case where the trial function \( f_n(x) \) is the exact solution of (4.2.1) above for all \( x \in (a,b) \). Hence the variational defined as:
\[ N[f_n] = (R_n, R_n) = (g + K f_n - f_n, g + K f_n - f_n) \] (4.2.3)

was used to obtain a measure of convergence for the sequence of trial functions \( f_n(x) \). The inner-products in (4.2.3) have the usual integral definition:

\[ (\phi, \psi) = \int_a^b \phi(x) \psi(x) \, dx \] (4.2.4)

making \( N[f_n] \) the square of the RMS norm of \( R_n \). The sequence of trial functions \( f_n(x) \), is defined by the recurrence relation:

\[ f_{n+1}(x) = f_n(x) + \delta_n(x) \] (4.2.5)

and possibilities are considered for the relationship between \( \delta_n(x) \) and \( R_n(x) \). It is interesting to note that the simplest relationship between \( \delta_n(x) \) and \( R_n(x) \) namely \( \delta_n(x) = R_n(x) \) for all \( x \in (a,b) \) reduces to the method of successive approximations.

It is required to express \( N[f_{n+1}] \) in terms of \( N[f_n] \), so as to determine convergence conditions, since the solution is obtained when \( N[f_n] \to 0 \) as \( n \) increases, and convergence occurs when \( N[f_{n+1}] < N[f_n] \) for \( n > n_0 \). Before this may be done, it is required to express \( R_{n+1}(x) \) in terms of \( R_n(x) \). From (4.2.2)

\[ R_{n+1}(x) = g(x) + \int_a^b k(x,y) \{ f_n(y) + \delta_n(y) \} \, dy - f_n(x) - \delta_n(x) \] (4.2.6)

which, in the case of linear equations may be written as:

\[ R_{n+1}(x) = R_n(x) + \int_a^b k(x,y) \delta_n(y) \, dy - \delta_n(y) \] (4.2.7)

or

\[ R_{n+1} = R_n + K \delta_n - \delta_n. \]
Hence, the variational $N[f_{n+1}]$ may be expressed as:

$$N[f_{n+1}] = N[f_n] - (R_n, \delta_n - K \delta_n) - (\delta_n - K \delta_n, R_n) + (\delta_n - K \delta_n, \delta_n - K \delta_n).$$

(4.2.8)

The problem reduces to minimizing $N[f_{n+1}]$ by obtaining an approximate solution to equation (4.2.7) above.

In the remaining part of the chapter consideration is given to the use of a number of perturbations to achieve this minimization.

The most important perturbation considered, is to express $\delta_n(x)$ as:

$$\delta_n(x) = \alpha_n \cdot R_n(x).$$

(4.2.9)

This was motivated by the possibility of accelerating the convergence of the method of successive approximations by scaling each term to minimize the error. Other perturbations are shown to be inferior to (4.2.9) since they produce more restrictive conditions under which the variational (4.2.3) converges towards zero. As a result of this work a constructive condition is also obtained which relates the convergence of the $n^{th}$ and $(n+1)^{th}$ iterations of the method of successive approximations.

### 4.3 THE RESIDUAL PERTURBATION

The perturbation $\delta_n(x) = \alpha_n \cdot R_n(x)$ equation (4.2.9) above is given detailed consideration in this section, and convergence conditions obtained with respect to the variational (4.2.3). Firstly, equation (4.2.7) becomes:

$$R_{n+1} = (1-\alpha_n) R_n + \alpha_n K R_n$$

(4.3.1)
under the above perturbation, so that the \((n+1)\)th iterate of

\[ N[f_n] \]

may be written as:

\[ N[f_{n+1}] = N[f_n] - 2\alpha_n (R_n, R_n - K R_n) + \alpha_n^2 (R_n - K R_n, R_n - K R_n) \quad (4.3.2) \]

from (4.2.8), which is a quadratic in \(\alpha_n\). Since the term

\((R_n - K R_n, R_n - K R_n)\)

is non-negative, the quadratic (4.3.2) has a minimum when this inner-product is non-zero. This occurs when

\( R_n = 0 \)

that is, when the solution is attained, or when \( R_n = K R_n \)

for all \( x \in (a,b) \), which is very unlikely and very apparent, as the value of the inner-product may be monitored as the method proceeds. It may be noted that when the inner-product of \( R_n - K R_n \) with itself becomes small away from the solution, that \( R_n \) is approaching the eigenfunction of the operator \( K \) associated with the eigenvalue \( \lambda = 1 \). Clearly the problem will only potentially exist in the class of operators which possess an eigenvalue of \( \lambda = 1 \), where \( \lambda \) represents the eigenvalue in the general equation \( f = \lambda K f \).

In the more general case when a minimum of equation (4.3.2) exists the optimal value of \( \alpha_n \) is given by

\[ \hat{\alpha}_n = \frac{(R_n, R_n - K R_n)}{(R_n - K R_n, R_n - K R_n)} \quad (4.3.3) \]

and since \( N[f_{n+1}] = N[f_n] \) only when \( \alpha_n = 0 \) this minimum is less than or equal to \( N[f_n] \). The value of \( \frac{N[f_{n+1}]}{N[f_n]} \) at \( \hat{\alpha}_n \) is easily shown to be given by:

\[ \left( \gamma - \beta^2 \right) (R_n, R_n) \left/ \left( R_n - K R_n, R_n - K R_n \right) \right. \]

(4.3.4)
where
\[ \gamma = \frac{(K_n, K_n)_{R, R_n}}{(R_n, R_n)} \]  (4.3.5)

and
\[ \beta = \frac{(R_n, K_n)_{R, R_n}}{(R_n, R_n)} \]  (4.3.6)

It may easily be shown by direct application of the Cauchy-Schwartz inequality to (4.3.4) that \( \gamma > \beta^2 \) implying \( N[f_{n+1}] > 0 \). Hence the ratio \( N[f_{n+1}]/N[f_n] \) may easily be calculated at each step in the iterative process and gives a measure of the rate of convergence at each iteration, since all the required inner-products are known and also required in the evaluation of \( \hat{a}_n \). This measure is also bounded by 0 and 1. Therefore, such a perturbation is always convergent, with an a-priori estimate of the convergence, except when the two conditions quoted above occur, when the convergence tends to become very slow. This perturbation has the distinct advantage of having a constructive measure of the rate of convergence at any iteration and since the inner-product may be monitored at each step the approach of difficulties may be easily seen and evasive action taken. In practice the calculation of the inner-products proves to be more computationally stable when the difference between the functions \( R_n \) and \( K_n \) is obtained before they are evaluated. Further, the problem of orthogonality between the functions \( R_n \) and \( R_n - K_n \), may be avoided by reverting to the method of successive approximations at the \( n^{th} \) iteration, which is simply achieved by setting \( \alpha_n = 1 \). However, the problem of obtaining a situation where \( R_n = K_n \) is much more difficult to avoid, as this actually means...
that the integral equation for the perturbation (4.2.7) has no solution in the region of $\delta_n$. In such a case no method based on using this equation would be expected to work, however, alternatives are discussed in the following section.

4.4 COMPARISON WITH THE RAYLEIGH-RITZ VARIATIONAL.

Baker 1977, quotes the use of the Rayleigh-Ritz variational defined by:

$$J[f_n] = (g, f_n) + (f_n, g) + (K f_n, f_n) - (f_n, f_n)$$

(4.4.1)

for the solution of the Fredholm integral equation (4.2.1), where the above variational converges to a maximal at $J[f] = (g, f)$ when $K$ is a negative definite self adjoint operator, as the sequence $f_n$ converges to $f$. It is also known that when the solution $f$ is perturbed by a small function $\delta$ then the variational (4.4.1) is perturbed by a second order displacement such that:

$$\delta J[f] = - ((1-K)\delta, \delta) .$$

(4.4.2)

It may also be noted from (4.4.2) that when $K$ is a positive definite self adjoint operator, possessing the property $(\delta K \delta) > (\delta \delta)$, then the variational (4.4.1) above is a minimum at the solution, where $J[f] = (g, f)$. This property is equivalent to the divergence of the Neumann series.

The condition for convergence is obtained by constraining the iterative sequence to have the form $f_{n+1}(x) = f_n(x) + \delta_n(x)$ with $\delta_n(x) = \alpha_n R_n(x)$, when applied to the variational (4.4.1) above.
This condition follows from $J[f_{n+1}] > J[f_n]$, and is given by:

$$\delta J[f_n] = 2a_n (R_n, R_n) + a_n^2 (K_n R_n - R_n, R_n) > 0. \quad (4.4.3)$$

This implies that the optimal value of $a_n$ which maximizes the above condition is given by:

$$a_n' = \frac{(R_n, R_n)}{(R_n - K_n R_n, R_n)} \quad (4.4.4)$$

and hence $(4.4.3)$ and $(4.4.4)$ yield:

$$\delta J[f_n] a_n' = \frac{(R_n, R_n)^2}{(R_n - K_n R_n, R_n)}. \quad (4.4.5)$$

It is important to note that this sequence will still suffer from problems when $R_n$ is orthogonal to the function $R_n - K R_n$ or when a solution of $R_n = K R_n$ is reached.

The variational discussed in section 4.2, which attains a minimum of the solution, may be written in terms of the difference between the $n$th and $(n+1)$th iterations, such that equation (4.3.2) may be written as a convergence condition defined by:

$$\delta N[f_n] = -2a_n (R_n - K_n R_n, R_n) + a_n^2 (R_n - K R_n, R_n - K R_n) < 0. \quad (4.4.6)$$

If the above optimal value for $a_n$ given by $(4.4.4)$ above is now substituted into this condition to obtain

$$\delta N[f_n] a_n' = -2(R_n, R_n) (R_n - K_n R_n, R_n - K R_n) + a_n^2 (R_n - K_n R_n, R_n - K R_n)^2. \quad (4.4.7)$$
Clearly, this expression may be either negative or positive. Therefore, the optimal value predicted by the Rayleigh-Ritz variational does not imply convergence when applied to the RMS residual variational. However, when the optimal value of $\alpha_n$ obtained in equation (4.3.3) is substituted in (4.4.3) above, the resulting condition given by:

\[
\delta J[f_n]_{\alpha_n} = \frac{2(R_n, R_n)(R_n - K R_n, R_n)}{(R_n - K R_n, R_n - K R_n)} - \frac{(R_n - K R_n, R_n)^3}{(R_n - K R_n, R_n - K R_n)^2} > 0
\]

(4.4.8)

is clearly true since the Cauchy-Schwartz inequality implies:

\[
\frac{(R_n, R_n)(R_n - K R_n, R_n)}{(R_n - K R_n, R_n - K R_n)} > \frac{(R_n - K R_n, R_n)^3}{(R_n - K R_n, R_n - K R_n)^2}
\]

(4.4.9)

when \( (R_n - K R_n, R_n) > 0 \).

Therefore the optimal value predicted by the RMS residual variational does imply convergence with respect to the Rayleigh-Ritz variational. In general the optimal predicted by the RMS residual variation is therefore superior, since it is convergent under both variations, where the optimal value of $\alpha_n$ in (4.4.4) above is only convergent under both variationals when the equation (4.4.7) is negative. It may be noted that this condition may be written as $1 - 2\beta + 2\beta^2 > \gamma$ where $\beta$ and $\gamma$ are defined from (4.3.5) and (4.3.6).

However, in practice, when a solution of the equation $R_n = K R_n$ is approached, causing the value of the inner-product of $R_n - K R_n$
with itself to become small, the measure of the rate of convergence from (4.3.4) becomes comparatively large. In this case it is occasionally found to be more useful to employ the value of $a_n$ predicted by $a_n = a'_n$, if the value of $(1-\beta)^2$ is greater than $\gamma - \beta^2$, since the convergence condition, given by equation (4.4.7) being negative, tends towards $\delta N[f_n] a_n = -2(R_n R_n)$, which is clearly always negative. Though this is found to be very useful when $(R_n, R_n)$ is of order $10^{-2}$ it cannot avoid the difficulty when $(R_n, R_n)$ is also small, since the residual equation has no solution very close to the actual solution of the integral equation.

In conclusion, the prescribed optimal value of $a_n$ according to the Rayleigh-Ritz variational is found to be inferior to the prescription considered in Section 4.3, except in the circumstance where, the convergence of this variational becomes unstable due to the approach of the solution to the homogeneous equation for the operator $K$. In this case the prescription $a'_n$ is found to be useful especially where the trial function is far enough away from the actual solution to make the situation easily recoverable. In practice, such problems are unlikely to occur near the actual solution as the perturbation will tend to become very small as the solution is approached. Where an approaching solution to the homogeneous problem occurs the solution to the residual problem becomes large, making the situations incompatible.

4.5 NEUMANN SERIES.

The purpose of this section is to consider the convergence of the Neumann series, in the light of the convergence condition
defined by the variational (4.3.2). The Neumann series, which forms the basis of the method of successive approximations, may be written in the non-standard form as:

\[ f(x) = g(x) + R_0(x) + K R_0(x) + K(K R_0(x)) + K(K(K R_0(x))) + \ldots. \]

Which implies that \( n = R_n(x) \) and \( n = 1 \) for all \( n \), when compared to the earlier perturbation. Direct application of this perturbation to the equation (4.2.7) implies that:

\[ R_{n+1} = K R_n, \]  

so the ratio for the convergence of the \( n^{th} \) iteration with respect to the \( (n+1)^{th} \) iteration may be written as:

\[ \frac{N[f_{n+1}]}{N[f_n]} = \frac{\langle K R_n, K R_n \rangle}{\langle R_n, R_n \rangle} = \gamma \text{ from (4.3.5)}. \]

Therefore, convergence exists with respect to the residual variational (4.3.4) for the method of successive approximations when \( \gamma < 1 \), and a measure of the rate of convergence is given by \( \sqrt{\gamma} \). A direct comparison exists between the method of successive approximations and the method of residual perturbation discussed in Section 4.3, for the rate of convergence of each iteration.

However, the above measure (4.3.5) is not a-priori like the measure in (4.3.4) in that the \( R_{n+1} \) term has to be calculated, to enable \( \gamma \) to be evaluated. The perturbation for the method of successive approximations can therefore be extended to become
\[ \delta_n^* = R_n + K R_n \] enabling both of the above functions to be used in the \( n \)th iteration, so that one perturbation is equivalent to two terms of the Neumann series (4.5.1) above. The measure \( \sqrt{\gamma} \) is no longer representative of the convergence rate between the \( n \)th and \((n+1)\)th iterations, but the fact that \( \gamma < 1 \) does still imply that the extended iteration, \( \delta_n^* \), is convergent. Moreover, if \( \gamma < 1 \) for the \( n \)th iteration the sequence \( \delta_{n+1}^* \) will still be convergent for all higher iterations, since repeated application of the operator \( K \) will imply that the sequence of values \( N[f_n] \) is uniformly convergent for \( n > N \). It may also be noted that the square of the ratio of convergence with respect to the perturbation \( \delta_n^* \) is given by:

\[
\frac{N[f_{n+1}^*]}{N[f_n]} = \frac{(K K R_n, K K R_n)}{(R_n, R_n)} = \frac{(R_{n+2}, R_{n+2})}{(R_n, R_n)} < \frac{(K R_n, K R_n)}{(R_n, R_n)} = \gamma < 1,
\]

(4.5.3)

since \[ \frac{(K K R_n, K K R_n)}{(R_n, R_n)} < \frac{(K R_n, K R_n)}{(R_n, R_n)} < 1. \]

The measure of convergence given by \( \gamma < 1 \) is less restrictive than the usual measure of \( \|K\|_2 < 1 \), as the definition for the norm of the operator \( \|K\|_2 \) states that \( \|K\|_2 \geq \gamma \), hence \( \gamma \) may be less than one and imply convergence even though \( \|K\|_2 \) is greater than one.

4.6 HALF-ITERATE EXTENSION.

So far, in this chapter, all the work has been centred around the perturbation of the general form \( \delta_n = \alpha_n R_n \). In the present
section consideration will be given to the possibility of extending the residual perturbation of section 4.3 in a comparable way to the extended perturbation $\delta^*$ of the previous section of the method of successive approximations. It may be noted that to evaluate the optimal scaling parameter $\alpha_n = \hat{\alpha}_n$ in section 4.3, two functions had to be obtained namely $R_n(x)$ and $K R_n(x)$. However, when the actual perturbation is applied only the residual function is used.

The equation for the perturbation given in (4.2.7) may be used as a predictor-corrector, using the right hand side of the equations which may be written as:

$$\delta^*_n = R_n + K \delta_n = R_n + \hat{\alpha}_n K R_n,$$  \hspace{1cm} (4.6.1)$$

for the ideal case when $R_{n+1} \sim 0$. It is now required to determine the conditions under which the above extension is convergent.

To obtain the value of the variational $N[f^*_{n+1}]$ where $f^*_{n+1} = f_n + \delta^*_n$, the extended perturbation is applied to the equation for the updated residual function to become:

$$R^*_{n+1} = (1-\hat{\alpha}_n) K R_n + \hat{\alpha}_n K K R_n,$$  \hspace{1cm} (4.6.2)$$

which may be rewritten in terms of (4.3.1) as:

$$R^*_{n+1} = K R_{n+1},$$  \hspace{1cm} (4.6.3)$$

Therefore, the condition for this extension to be convergent with respect to the variational (4.2.3) is given by:

$$\frac{N[f^*_{n+1}]}{N[F^*_n]} = \frac{(K R_{n+1} K R_{n+1})}{(R_{n+1} R_{n+1})} < 1.$$  \hspace{1cm} (4.6.4)$$
However, since this is the value of $\gamma$ with respect to the $(n+1)^{th}$ iterate, the exact value is unknown until the next iterate has been evaluated, as the functions $R_{n+1}$ and $K_{R_{n+1}}$ are unknown.

Despite this apparent difficulty, a sufficient condition may be implied when $\gamma < 1$ for the case of the $n^{th}$ iterate, then the terms $R_{n+1}$ and $K_{R_{n+1}}$ are convergent, since from the Neumann Series if the $n^{th}$ term is known to be convergent, then so are the higher order terms. A more formal proof is obtained by using the perturbation equations for $R_{n+1}$ and $K_{R_{n+1}}$ which may be written as:

$$R_{n+1} = (1-\alpha_n) R_n + \alpha_n K R_n,$$

from (4.3.1) above, and

$$K_{R_{n+1}} = (1-\alpha_n) K_{R_n} + \alpha_n K K_{R_n},$$

from (4.6.2) above, and formulating the ratio of inner-products (4.6.4) above.

Firstly, the ratio

$$\frac{(K_{R_{n+1}}, K_{R_{n+1}})}{(R_n, R_n)}$$

may be rewritten as:

$$\frac{(K_{R_{n+1}}, K_{R_{n+1}})}{(R_n, R_n)} = (1-\hat{\alpha}_n)^2 \frac{(K_{R_n}, K_{R_n})}{(R_n, R_n)} + 2(1-\hat{\alpha}_n) \hat{\alpha}_n \frac{(K_{R_n}, K K_{R_n})}{(R_n, R_n)} + \hat{\alpha}_n^2 \frac{(K K_{R_n}, K R_n)}{(R_n, R_n)}.$$

(4.6.5)

Substituting for $\alpha_n$ to obtain:

$$\frac{(K_{R_{n+1}}, K_{R_{n+1}})}{(R_n, R_n)} = \frac{(\gamma-\beta)^2}{(1-2\beta+\gamma)^2} + \frac{2(\gamma-\beta)(1-\beta)}{(1-2\beta+\gamma)^2} \frac{(K_{R_n}, K K_{R_n})}{(R_n, R_n)} + \frac{(1-\beta)^2 (K K_{R_n}, K R_n)}{(R_n, R_n)^2}.$$
where $\beta$ and $\gamma$ have their previously adopted meanings (4.3.5), (4.3.6).

Now assume that $\gamma$ is less than one, so that the Neumann series converges, then the recursive operation of $K$ on $R_n$ produces higher order terms in the Neumann series implying

$$
(K R_n, K R_n) < (R_n, K R_n)
$$

and

$$
(K K R_n, K K R_n) < (K R_n, K R_n).
$$

Therefore, equation (4.6.5) may be rewritten as the inequality:

$$
\frac{(K R_n, K R_n)}{(R_n, R_n)} < \frac{1}{(1-2\beta+\gamma)^2} \left\{ \gamma^3+\gamma-2\beta^3-2\beta^2-2\gamma \right\}, \quad (4.6.6)
$$

which since \( \frac{(R_{n+1}, R_n+1)}{(R_n, R_n)} = \frac{\gamma-\beta^2}{(1-2\beta+\gamma)} < 1 \), becomes:

$$
\frac{(K R_{n+1}, K R_{n+1})}{(R_{n+1}, R_{n+1})} < \frac{(\gamma^3+\gamma-2(\beta^3+\beta^2+\beta\gamma))}{(1-2\beta+\gamma)(\gamma-\beta^2)},
$$

which must be less than one for convergence (4.6.4). It is therefore required to prove that

$$
\frac{(\gamma^3+\gamma-2(\beta^3+\beta^2+\beta\gamma))}{(1-2\beta+\gamma)(\gamma-\beta^2)} < 1, \quad (4.6.7)
$$

which may easily be reduced to the initial assumption that $\gamma < 1$.

In conclusion, the extended perturbation is known to be an improvement to the standard perturbation when $\gamma$ is less than one. However, it may still improve the perturbation when $\gamma$ is greater than one, since the Neumann series may still converge, though
such convergence cannot be guaranteed. Therefore, when $\gamma$ is less than one, it is clearly beneficial to use the extended perturbation in preference to the standard residual perturbation. However, when $\gamma$ is greater than one, it is safer to employ the standard residual perturbation.

4.7 OTHER PERTURBATIONS.

The purpose of this section is to compare the residual perturbation with a number of other possible choices of forms for the perturbation.

Firstly, another possible form is to vary the trial function, as opposed to varying the residual function. This would produce a perturbation of the form:

$$\delta_n = \alpha_n f_n$$

which, by the application of similar analysis to that of section 4.3, the variational $N[f_{n+1}]$ may be minimized with respect to $N[f_n]$ where $f_{n+1} = f_n + \alpha_n f_n$, to obtain an optimal value of $\alpha_n = \hat{\alpha}_n$ upon which bounds for the convergence may be obtained. Under this perturbation the residual equation becomes:

$$R_{n+1} = \alpha_n (R_n - g) + R_n = \alpha_n (Kf_n - f_n) + R_n,$$

and by taking inner-products and optimizing $\alpha_n$ as before, the value $\hat{\alpha}_n$ is now found to be given by:

$$\hat{\alpha}_n = \frac{(f_n - K f_n, R_n)}{(f_n - K f_n, f_n - K f_n)}$$

(4.7.3)
where the ratio of inner-product is given by:

$$\frac{N[f_{n+1}]}{N[f_n]} = \frac{(g, g)(R_n, R_n) - (R_n, g)^2}{(g-R_n, g-R_n)(R_n, R_n)}$$

where this ratio must lie between 0 and 1 for convergence. It may easily be shown that (4.7.4) is always greater or equal to zero by direct application of the Cauchy-Schwartz inequality, it therefore remains to prove that the value of equation (4.7.4) is less than one. This is easily shown to be true when 

$$(R_n, g) \neq (R_n, R_n),$$

which only occurs when $f_n$ is such that the homogeneous problem is solved, i.e. $f_n = Kf_n$. The main disadvantage of this as a method of solving linear equations is that it may only be applied once since any recursive calculation can show no improvement, since once the optimal scaling of $f_n$ is attained no further improvement may ensue. The advantage of this perturbation is that it may be used to enable the residual perturbation to progress when problems are encountered since no extra functions are required, and also that the amount of work required to evaluate each iteration is approximately half that in the residual, though as a general method it will not converge for more than one iteration. As before this perturbation may be extended to become:

$$\delta_n^* = R_n + \hat{\alpha}_n Kf_n$$

which may be shown to be usefully convergent when $\gamma < 1$, and also possesses the advantage of enabling the method to be used as a general iterative procedure.
In conclusion this perturbation is inferior to the residual perturbation, since it may only be used as an iterative scheme on problems whose Neumann series converges. However, it may be applied once to aid the method of residual perturbations since it is always convergent when the residual perturbation method runs into difficulties.

Secondly, a perturbation based on the equation, which has the form:

\[ \delta_n = R_n + \alpha_n f_n, \]  

(4.7.6)

is now considered. This was motivated by the idea that it forms a compromise between the method of successive approximations and the function perturbation discussed above, and similar analysis was applied to that of section 4.3, obtaining the residual update equation as:

\[ R_{n+1}(x) + \frac{K R_n}{n} + \alpha_n (K - I) f_n = K R_n + \alpha_n (R - g). \]  

(4.7.7)

The ratio of inner-products now prescribes an optimal value of \( \alpha_n \) to be:

\[ \alpha_n = \hat{\alpha}_n = -\frac{(K R_n, R_n - g)}{(R_n - g, R_n - g)}, \]  

(4.7.8)

which at this point attains the value of

\[ \frac{N[f_{n+1}]}{N[f_n]} = -\frac{(K R_n, R_n - g)^2}{(g - R_n, g - R_n)(R_n, R_n)}. \]  

(4.7.9)

In the usual way for convergence to exist, equation (4.7.9)
must be bounded between 0 and 1. Direct application of the Cauchy-Schwartz inequality implies that this ratio is greater than or equal to zero, but the ratio given by (4.7.9) above is only convergent when

$$\gamma < 1 + \frac{\left(K R_n^2 - R_{n-1}^2\right)^2}{(g-R_n^2)(g-R_{n-1})(R_n^2)}$$

which is not generally true. This is an extreme disadvantage since both of the perturbations considered previously have been generally convergent except under a few relatively unlikely circumstances. This perturbation also has the disadvantage that it requires two functions to be obtained per iteration by integration across the range of the problem when applied to linear equations, and requires three functions to be obtained when applied to non-linear equations, therefore requiring a reasonable degree of extra work in each iteration over and above the earlier perturbations. It is, therefore, not considered worthwhile to pursue the present investigation.

Thirdly, perturbations of the general form:

$$\delta_n = a_n R_n^m$$  \hspace{1cm} (4.7.10)

are considered, where $m$ is chosen to take the values $m = 3$ and $m = 5$. It may be noted that even values of $m$ are not considered as this would constrain the value of $R_n^m$ to be totally positive across the range of the equation, and would therefore not allow $\delta_n$ to model the case when $R_n$ takes both negative and positive values across the range. Also for reasons of desiring the Neumann
series to converge in terms of $\delta_n$, negative values of $m$ are
excluded. Clearly as $m$ becomes large the value of $a_n$ will also
be required to become large, since $R_n(x)$ tends to zero for all $x$
as the iteration progresses. Therefore, large values of $m$ cease
to be suitable in modelling the perturbation $\delta_n$. Being therefore
left with comparatively small positive odd values of $m$ as the
only useful consideration, from which the above two were chosen.
The updated residual function from (4.2.3) may now be written as:

$$R_{n+1} = R_n - a_n R_n^m + a_n K (R_n^m) .$$  \hspace{1cm} (4.7.11)

This formula predicts an optimal value of $a_n$ by similar analysis
to that carried out above, yields:

$$a_n = \frac{(R_n^m - K R_n^m)}{(R_n^m - K R_n^m, R_n^m - K R_n^m)} ,$$ \hspace{1cm} (4.7.12)

where the ratio between successive iterates of the residual
variational at the value $a_n$ is given by:

$$\frac{N[f_{n+1}]}{N[f_n]} = \frac{(R_n^m - K R_n^m, R_n^m - K R_n^m)(R_n^m, R_n^m) - (R_n^m, R_n^m - K R_n^m)^2}{(R_n^m - K R_n^m, R_n^m - K R_n^m)(R_n^m, R_n^m)} ,$$ \hspace{1cm} (4.7.13)

which is required to be bounded by 0 and 1 for convergence to
exist. The Cauchy-Schwartz inequality clearly implies that this
is bounded below by zero. It, therefore, remains to obtain the
conditions under which this equation is bounded above by one.
It may easily be shown that this reduces to the condition that
$0 < (R_n, K R_n^m - R_m)^2$, which holds for all $R$ except where $R_n$ is.
orthogonal to $K R_n - R_n$, or where, $R_n^m = K R_n^m$ for all $x$ in the range of the problem. Therefore this perturbation may be considered as being directly comparable with the residual perturbation, which is not surprising since the above perturbation reduces to the residual perturbation when $m = 1$.

Also, in general, the amount of work required to calculate each iteration is comparable, since two functions are required, namely $R_n$ and $K R_n^m$. The function $R_n^m$ is easily computed with relatively little extra work, by simply raising the residual function to the power of $m$. The main drawback of this as a technique is the accurate evaluation of the value of $\hat{\alpha}_n$ (4.7.12) above, since this term is of order $R_n^{2m}$ in the denominator and only of order $R_n^{m+1}$ in the numerator. Therefore, as $R_n$ tends to zero for all $x$ the value of $\hat{\alpha}_n$ will become extremely large since $\hat{\alpha}_n \sim R_n^{1-m}$.

Though this presents no problem if the values of the functions $R_n$ and $K R_n^m$ are accurately known, it must be remembered that any errors in the residual $R_n$ are likely to be dramatically propagated in the function $K R_n^m$. Such propagation being partly due to the raising of $R_n$ to a power $m$ and partly due to the application of the integral operator $K$. In practice, this will have an increasingly disastrous effect on the value $\hat{\alpha}_n$ as the residual becomes small, since to add to the problem the difference $R_n^m - K R_n^m$ will suffer from significant truncation errors. It may, therefore, be concluded that as the value of $m$ is increased the perturbation will become increasingly unstable as the residual function $R_n$ tends to zero across the range of $x$. So that the best value of $m$ is the one which causes the least possibility of instability
which is clearly given by \( m = 1 \), since in this case \( a_n \) is of order unity.

The perturbation (4.7.10) may also be extended as considered in Section 4.6 for the residual perturbation. The extended form is given by:

\[
\delta_n^* = R_n + \hat{\alpha}_n K R^n m.
\] (4.7.14)

It may also be shown by similar argument to be convergent when the ratio \( \gamma < 1 \), where \( \gamma \) is taken from (4.3.5). However, the value \( \gamma \) has not been calculated in the present consideration for values of \( m = 3 \) and 5. Indeed to calculate the value, the function \( K R_n \) would need evaluating, which would increase the amount of work required by a factor of about 50% per iteration. It is clearly impractical to evaluate this value since it would cost so much simply to obtain a bound on the convergence of the extended perturbation. The value

\[
\frac{(K R^n m, K R^n m)}{(R_n, R_n)}
\]

may not be used as an approximate measure of \( \gamma \) since it bears no relation to the Neumann series. This may, therefore, be considered as a secondary disadvantage of perturbation of the form (4.7.10), since the special case when \( m = 1 \) also produces a confidence limit on the applicability of using the extended form (4.6.1) for no extra cost.

In conclusion to this section, it is clear that an infinite number of possible perturbations exist. However, the few which
have been considered bear the closest resemblance to the well-known method of successive approximations. Moreover, the present investigation would appear to demonstrate that the residual perturbation is the best, since it possesses more advantages than the others considered, with the possible use of the function perturbation, when difficulties are encountered.

4.8 NUMERICAL EXAMPLES.

The later sections of this chapter are concerned with demonstrating this theory on a number of practical examples. Initially, a simple analytic example is considered, and comparison made between the various perturbations. Secondly, a similar comparison is made using a more complex displacement kernel, and finally a number of singular integral equations are considered, to demonstrate the residual perturbation method.

4.8.1 Analytical Example.

For the purpose of this demonstration the following example with a simple degenerate kernel is considered:

\[ f(x) = x + 4 \int_0^1 xy f(y) \, dy. \]  

The actual solution is:

\[ f(x) = -3x, \]

but when the method of successive approximations is applied to the above equation, with \( f(x) = x \), no solution is obtained as the Neumann series diverges. This is clearly seen from the 1st three
terms which are:

\[ f_1(x) = \frac{7}{3}x \quad , \quad f_2(x) = \frac{37}{9}x \quad , \quad \text{and} \quad f_3(x) = \frac{175}{27}x \ . \]

Firstly, applying the residual perturbation method of section 4.3 with \( f_0(x) = x \); the optimal value of \( \hat{a}_0 = -3 \) and the corresponding rate of convergence is zero, implying the exact solution is attained immediately. Therefore, the perturbation with respect to \( f_0(x) \) is given by:

\[ \delta_0 = \hat{a}_0 R_0(x) = -4x \]

where

\[ R_0(x) = \frac{4}{3}x \ . \]

As predicted the value of \( f_1(x) = -3x = f(x) \), which is the exact solution. It may also be noted that since \( \gamma = \frac{16}{9} \) convergence for the extended perturbation cannot be guaranteed. However, since

\[ \delta_0^*(x) = R_0(x) + \hat{a}_n K R_0(x) = -4x \ , \]

the extended perturbation also converges to the exact answer, but this could not be predicted beforehand.

Secondly, the function perturbation was applied to (4.8.1) above. The optimal value of \( \hat{a}_0 = -4 \), therefore implying that \( \delta_0 = -4x \) and \( f_1 = -3x \) the exact solution. This is not too surprising since the solution is proportional to the original trial function and the measure of convergence given by (4.7.4) is also zero. Hence this method is ideal for this particular problem. It may be noted in its favour that only one function was required, namely \( R_0(x) \) to obtain the solution by this method, where two were required by the previous method.
Thirdly, the perturbation given in (4.7.6) was applied to the problem. Again the trial function \( f_0(x) = x \) and now \( \hat{a}_0 = -\frac{16}{3} \). It will be remembered that this method was not proved to be generally convergent, but was only shown to be convergent for particular values of the inner-products. However, for this perturbation \( \delta_0 \) also equals \(-4x\), which results in the first iterate being \( f_1(x) = -3x \), which is again the exact solution.

This example, would appear to be too trivial for any distinction to be made between the methods.

Finally, the non-linear perturbations, described in (4.7.10) were applied to the equation (4.8.1) above, with the initial trial function \( f_0(x) = x \), as was previously considered. In both forms of this perturbation used, when \( m = 3 \) and \( m = 5 \) the initial value of the parameter \( \hat{a}_n \) was found to be zero, and the corresponding measure of convergence was found to be one implying no convergence. As no change occurs in the iterated function \( f_0(x) \), since \( \hat{a}_0 \) is zero, the iterative scheme cannot proceed any further. However, though such a failure is unfavourable, this example is regarded as too simplistic to provide sufficient grounds to reject this perturbation in general, and a more complex example appears below. It may also be noted that these perturbations suffer from an extreme disadvantage in comparison to the previous perturbations as they cannot model the exact solution directly as the earlier cases could.

### 4.8.2 Displacement Kernel Example

As an example of a more complex problem, Love's equation, considered by Baker 1977, was used. The actual form of this
The equation used in this example is given by:

\[ f(x) = 2 - \frac{1}{\pi} \int_{-1}^{1} \frac{1}{(1 + (x-y)^2)^2} f(y) \, dy \]  

(4.8.3)

The Neumann Series for this example is known to be convergent. Therefore, the method of successive approximations, starting from the trial function \( f_0(x) = 2 \), converges to the correct answer. In 24 iterations this answer has been achieved to eight figures across the range from \( x = -1 \) to \( x = 1 \). Table I below shows how the variational used in section 4.2 converges towards zero under the method of successive approximations.

**TABLE I**

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>( N[f_n] )</th>
<th>( (f_n, f_n) )</th>
<th>( \gamma_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.62667</td>
<td>8.0</td>
<td>0.20571</td>
</tr>
<tr>
<td>2</td>
<td>6.8842(-2)</td>
<td>4.56019</td>
<td>0.20573</td>
</tr>
<tr>
<td>6</td>
<td>1.23319(-4)</td>
<td>3.85887</td>
<td>0.20573</td>
</tr>
<tr>
<td>10</td>
<td>2.209111(-7)</td>
<td>3.83051</td>
<td>0.20573</td>
</tr>
<tr>
<td>16</td>
<td>1.673382(-11)</td>
<td>3.829267</td>
<td>0.20573</td>
</tr>
<tr>
<td>24</td>
<td>3.33204(-16)</td>
<td>3.829256</td>
<td>0.12600</td>
</tr>
</tbody>
</table>

The change in the value of \( \gamma_n \) when \( n = 24 \) is probably due to the fact that cancellation had made the residual inaccurate, and has nothing to do with the convergence of the Neumann series. This gives sufficient justification for the extended perturbations.
to be used in all the later comparisons using this example as $\gamma < 1$ for all $n$. In this case the integration was performed using Clenshaw–Curtis quadrature on 33 grid points.

Firstly, the residual perturbation is used to solve this equation. This method achieved convergence to 8 figures in only 3 iterations. However, each iteration involves slightly more than twice the amount of work required to perform each iterate with the method of successive approximations in Table I. Since $N(f_3)$ is a measure of the error before the third iteration is performed, the residual perturbation method converges to the same accuracy after about a quarter of the work has been carried out. The rows of Table II below may, therefore, be directly compared to the first 3 rows of Table I above.

**TABLE II**

<table>
<thead>
<tr>
<th>Iteration No. $(n)$</th>
<th>$N[f_1]$</th>
<th>$(f_n^2 f_n')$</th>
<th>$(\gamma_n - \beta_n^2) / (1 - 2\beta_n + \gamma_n)$</th>
<th>$\alpha_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.42667</td>
<td>8.0</td>
<td>6.43217(-6)</td>
<td>0.687971</td>
</tr>
<tr>
<td>1</td>
<td>1.499(-8)</td>
<td>3.829213</td>
<td>9.16444(-4)</td>
<td>0.96052</td>
</tr>
<tr>
<td>2</td>
<td>2.799(-12)</td>
<td>3.829260</td>
<td>1.313285(-5)</td>
<td>0.68798</td>
</tr>
<tr>
<td>3</td>
<td>1.6638(-16)</td>
<td>3.829256</td>
<td>5.47098(-3)</td>
<td>0.97832</td>
</tr>
</tbody>
</table>

It may be noted that since $\frac{(\gamma_n - \beta_n^2)}{(1 - 2\beta_n - \gamma_n)}$ is small and $\alpha_n$ is of order unity, then $R_n$ is a reasonable measure of the error at each
iteration. The variational $N[f_n]$ also provides a good measure of the square of the absolute error before the $n^{th}$ iteration is performed. It is important to note that this value may be approximated before the $n^{th}$ residual function has been calculated by using the relationship:

$$N[f_n] \approx N[f_{n-1}] \cdot \frac{(\gamma_{n-1} - \beta_{n-1}^2)}{(1 - 2\beta_{n-1}^2 + \gamma_{n-1})}.$$  \hspace{1cm} (4.8.4)

Though, this approximation reduces to an upper bound when $\gamma_{n-1} < 1$ and the extended perturbation is used. Hence, by updating the subscripts equation (4.8.4) may be used to approximate the error after the $n^{th}$ iteration.

Secondly, the function perturbation from (4.7.1) is applied to equation (4.8.3) above. However, since the original form of (4.7.1) may only be applied once before convergence ceases, the extended form given by (4.7.5) is used. This perturbation has the advantage that each iteration only requires slightly more work than the method of successive approximations. Moreover, results were obtained to an accuracy of 8 figures, using only 5 iterations making the method better than the residual-perturbation method when applied to this example. These results are presented in Table III below, along with the results obtained when the perturbation defined by (4.7.6) was applied to the same example.

These results are comparable to those obtained by the method of residual perturbations, when the amount of work carried out in each iteration is considered.
\[ \delta_n = \alpha_n f_n \quad , \quad \delta^*_n = R_n + \alpha_n K f_n \]

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>N[( f_n )]</th>
<th>(( f_n, f_n ))</th>
<th>Rate of convergence from (4.7.4)</th>
<th>( \alpha_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.62667</td>
<td>8.0</td>
<td>4.7384(-3)</td>
<td>-0.1310380</td>
</tr>
<tr>
<td>1</td>
<td>1.98019(-5)</td>
<td>3.82288</td>
<td>6.2860(-1)</td>
<td>9.5972(-4)</td>
</tr>
<tr>
<td>3</td>
<td>8.39379(-11)</td>
<td>3.829239</td>
<td>4.8183(-1)</td>
<td>2.33169(-6)</td>
</tr>
<tr>
<td>5</td>
<td>5.93580(-16)</td>
<td>3.829256</td>
<td>7.4359(-1)</td>
<td>4.17937(-9)</td>
</tr>
</tbody>
</table>

\[ \delta_n = R_n + \alpha_n f_n \quad , \quad \delta^*_n = R_n + K R_n + \alpha_n K f_n \]

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>N[( f_n )]</th>
<th>(( f_n, f_n ))</th>
<th>Rate of convergence from (4.7.9)</th>
<th>( \alpha_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.62667</td>
<td>8.0</td>
<td>1.20937(-3)</td>
<td>1.46929(-1)</td>
</tr>
<tr>
<td>1</td>
<td>5.097193(-6)</td>
<td>3.829270</td>
<td>3.266198(-3)</td>
<td>2.45819(-4)</td>
</tr>
<tr>
<td>2</td>
<td>7.000482(-11)</td>
<td>3.829256</td>
<td>3.780540(-3)</td>
<td>1.05689(-6)</td>
</tr>
<tr>
<td>3</td>
<td>9.460399(-16)</td>
<td>3.829256</td>
<td>3.439135(-3)</td>
<td>3.637466(-9)</td>
</tr>
</tbody>
</table>

Considering the functional perturbation in the earlier part of the table, it is clearly demonstrated that the first iteration gives rapid acceleration towards the solution. The value of \((f_1, f_1)\) is compatible with the value \((f_8, f_8) = 3.83533\) since both are in error by approximately \(6 \times 10^{-3}\), therefore the function perturbation has achieved in one iteration, a compatible result to that achieved in 8 iterations by the method of successive
approximations, giving an improvement of almost a factor of eight. The later iterations of this perturbation are only made possible since $\gamma < 1$, which implies that convergence of the extension $\delta_n^*$ is guaranteed. However, this extended perturbation achieves in the next four iterations what the method of successive approximations took roughly a further sixteen iterations to achieve. This method is, therefore, a dramatic improvement over the method of successive approximations, but unfortunately has little generality when $\gamma > 1$.

In the later part of the table similar conclusions may be drawn from the mixed perturbation. However, the power of attaining a good first approximation in the zeroth iteration, is extremely dramatic as the measure $(f_1, f_1)$ is accurate to $1.4 \times 10^{-5}$, this is comparable with the sixteenth iteration of successive approximations in Table I, which also predicts an improvement of about a factor of eight, since this perturbation takes slightly more than twice the amount of work to evaluate each iteration when compared to the work of Table I.

In conclusion, the two perturbations discussed in Table III are very powerful for solving situations where the trial function is a reasonable distance from the actual solution, and have been shown to be superior to the residual perturbation method in the zeroth iteration. However, they suffer from the disadvantage of not being generally convergent, and are, therefore, less useful than may be first imagined.

Finally, for this example, the higher order perturbations from (4.7.10) are considered, and comparable results are presented in Table IV below.
\[ \delta_n = \alpha_n R_n^3, \quad \delta_n^* = R_n + \alpha_n K R_n^3 \]

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>(N[f_n])</th>
<th>((f_n, f_n))</th>
<th>Rate of convergence from (4.7.13)</th>
<th>(\alpha_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.62667</td>
<td>8.0</td>
<td>1.57285254(-2)</td>
<td>8.023959(-1)</td>
</tr>
<tr>
<td>1</td>
<td>1.07283(-4)</td>
<td>3.803473</td>
<td>4.8696676(-2)</td>
<td>8.068811(+3)</td>
</tr>
<tr>
<td>2</td>
<td>2.39244(-7)</td>
<td>3.830542</td>
<td>2.5607945(-2)</td>
<td>5.163856(+6)</td>
</tr>
<tr>
<td>3</td>
<td>6.305311(-11)</td>
<td>3.8292346</td>
<td>3.3706269(-2)</td>
<td>1.8113049(+10)</td>
</tr>
</tbody>
</table>

\[ \delta_n = \alpha_n R_n^5, \quad \delta_n^* = R_n + \alpha_n K R_n^5 \]

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>(N[f_n])</th>
<th>((f_n, f_n))</th>
<th>Rate of convergence from (4.7.13)</th>
<th>(\alpha_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.62667</td>
<td>8.0</td>
<td>4.943024(-2)</td>
<td>8.914382(-1)</td>
</tr>
<tr>
<td>1</td>
<td>8.6458531(-4)</td>
<td>3.75075</td>
<td>1.12397717(-1)</td>
<td>2.0699167(+6)</td>
</tr>
<tr>
<td>2</td>
<td>7.647489(-6)</td>
<td>3.83644</td>
<td>9.8525972(-2)</td>
<td>3.0430256(+10)</td>
</tr>
</tbody>
</table>

In comparison, with Tables II and III the perturbations considered here exhibit few beneficial qualities. Clearly they do converge towards the correct answer, but with no great rapidity.

It may also be noted that the reason for the above table to appear incomplete is that, when implemented on an eight digit computer, cancellation was encountered when evaluating the inner-products which renders any further calculation as impractical. Such instability is clearly an extreme disadvantage. It is also valuable to note that
after a comparable amount of work the residual perturbation method has achieved an accuracy of about six decimal places, where the cubic perturbation has only achieved about four decimal places, and the quintic perturbation about 3 decimal places. Such comparison demonstrates the inferiority of such higher order perturbations.

In conclusion to this example, it is clear that the higher order perturbation exhibit disadvantages due to instabilities, which makes them unfavourable. However, there appears to be little to choose between the other 3 perturbations, which all show similar and dramatic improvements over the earlier method of successive approximations.

4.8.3 Singular Displacement Kernel Examples.

Next, a number of more difficult problems which exhibit displacement singularities in the kernel are considered. These examples demonstrate the use of the singular quadrature rules described in chapter 2.

The first such example is described by the equation:

\[
f(x) = x - \frac{1}{2} \left( x^2 \ln(x) + (1-x^2) \ln(1-x) - (x + \frac{1}{2}) \right) + \int_0^1 \ln(|x-y|) f(y) \, dy ,
\]

which was also considered by Delves et al. 1978, and Baker 1977. It may easily be shown that the exact solution of this equation is given by \( f(x) = x \). Therefore, despite the difficulty imposed by the singular kernel, it is clearly known how accurate the
integrations are and the solution is apparent when obtained, since interpolating the unknown function poses few problems. The results obtained by applying the method at residual perturbations are presented in Table V below.

### TABLE V

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>$N[f_n]$</th>
<th>$(f_n f_n')$</th>
<th>Rate of convergence $N[f_{n+1}]/N[f_n]$ (4.3.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.7697713</td>
<td>1.8278033</td>
<td>3.6228859(-3)</td>
</tr>
<tr>
<td>1</td>
<td>1.3657372(-2)</td>
<td>3.7040267(-1)</td>
<td>5.0157402(-3)</td>
</tr>
<tr>
<td>2</td>
<td>9.8206824(-5)</td>
<td>3.29130656(-1)</td>
<td>7.8974025(-3)</td>
</tr>
<tr>
<td>4</td>
<td>6.5980314(-9)</td>
<td>3.33350372(-1)</td>
<td>2.2450656(-2)</td>
</tr>
<tr>
<td>7</td>
<td>4.5236178(-14)</td>
<td>3.33333340(-1)</td>
<td>8.9478188(-3)</td>
</tr>
<tr>
<td>9</td>
<td>1.8579139(-17)</td>
<td>3.33333333(-1)</td>
<td>5.5700183(-2)</td>
</tr>
</tbody>
</table>

which shows that 9 iterations are required to achieve an accuracy of 8 figures. The same example was also solved using the function perturbation from (4.7.1) and the comparable results to Table V are given below in Table VI. It must be noted that since each iteration requires approximately half the work required for the residual perturbation that the 4th iterate in Table V is comparable to the 8th iterate below,
which clearly shows that the method is inferior to that of residual perturbation as expected. Since a comparable amount of work only achieved 2 figure accuracy to that which previously achieved 8 figure accuracy. It may, therefore, be concluded that except in a few examples where a simple application of the functional perturbation achieves very good approximation to the solution, that the residual perturbation method is superior and more generally applicable. It may also be noted that such an example allows little comparison to be made between the work of Delves and Baker and the present method, since the solution is obtained almost trivially. This is due to simplicity of interpolating the solution, which most basis set methods should be able to do to reasonably high accuracy in very few points.

Secondly, in this section, an apparently very similar example given by:

\[ f(x) = 1 + \int_{0}^{1} \ln(|x-y|)f(y) \, dy \]  

(4.8.6)
is considered. This example was also considered by Cohen 1980 who predicted that the value of the solution at zero was \( f(0) = 0.574205 \) by using an equispaced grid of 33 points. However, since the direct application of the collocation method with 65 points spaced on the Chebyshev zeros predicts \( f(0) = 0.574075 \), the earlier result was regarded as subject to errors, and a more accurate result sought. This quest encountered difficulties since the unknown functions was found to vary rapidly as \( x \) approached zero and one. Therefore, for the purpose of economizing on the number of grid points the above equation was rewritten to make use of the symmetry exhibited in (4.8.6) to become:

\[
f(x) = 1 + \int_0^\frac{1}{4} \left\{ \ln(|x-y|) + \ln(1-x-y) \right\} f(y) \, dy \tag{4.8.7}
\]

and solved over the range \((0, \frac{1}{4})\) instead. The method of residual perturbation, and function perturbation were then applied to (4.8.7) and the results are presented in Table VII below.
TABLE VII

Residual Perturbation.

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>$N[f_n]$</th>
<th>$(f_n, f'_n)$</th>
<th>Rate of convergence from (4.3.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.142510988</td>
<td>0.5</td>
<td>3.608525088(-4)</td>
</tr>
<tr>
<td>1</td>
<td>7.2957233457(-5)</td>
<td>8.140378810(-2)</td>
<td>4.913749558(-3)</td>
</tr>
<tr>
<td>2</td>
<td>7.607881203(-7)</td>
<td>8.33613139(-2)</td>
<td>7.425440819(-4)</td>
</tr>
<tr>
<td>3</td>
<td>5.309219574(-11)</td>
<td>8.314035460(-2)</td>
<td>1.944692406(-2)</td>
</tr>
<tr>
<td>4</td>
<td>1.678826184(-12)</td>
<td>8.314125087(-2)</td>
<td>2.813887513(-3)</td>
</tr>
<tr>
<td>5</td>
<td>3.375851988(-16)</td>
<td>8.314096903(-2)</td>
<td>6.4315591484(-2)</td>
</tr>
<tr>
<td>6</td>
<td>3.488727379(-17)</td>
<td>8.314096907(-2)</td>
<td>2.89960036(-3)</td>
</tr>
</tbody>
</table>

Function Perturbation.

<table>
<thead>
<tr>
<th>Iteration No. (n)</th>
<th>$N[f_n]$</th>
<th>$(f_n, f'_n)$</th>
<th>Rate of convergence from (4.7.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.142510988</td>
<td>0.5</td>
<td>2.438635909(-3)</td>
</tr>
<tr>
<td>2</td>
<td>1.138442009(-4)</td>
<td>8.357241968(-2)</td>
<td>7.914933390(-1)</td>
</tr>
<tr>
<td>4</td>
<td>4.391201917(-6)</td>
<td>8.318114321(-2)</td>
<td>8.647888739(-1)</td>
</tr>
<tr>
<td>6</td>
<td>1.846936902(-7)</td>
<td>8.314752374(-2)</td>
<td>8.788894411(-1)</td>
</tr>
<tr>
<td>8</td>
<td>7.877430122(-9)</td>
<td>8.314248602(-2)</td>
<td>8.819181236(-1)</td>
</tr>
<tr>
<td>10</td>
<td>3.369336057(-10)</td>
<td>8.314152760(-2)</td>
<td>8.825972506(-1)</td>
</tr>
<tr>
<td>12</td>
<td>1.442039837(-11)</td>
<td>8.314133330(-2)</td>
<td>8.827530473(-1)</td>
</tr>
</tbody>
</table>
Since the 12th iteration of function perturbation is comparable to the 6th iterate of the residual perturbation, it is again clearly seen that the residual perturbation method is superior. It is important to note that such perturbation methods are independent of the grid of points used to approximate the unknown function. Therefore, the results were checked by increasing the number of points and carrying out more iterations.

The results in Table VII were checked in such a way, and proved to be consistent when the number of interpolation points were doubled. The value of the function at \( x = 0 \), was found to be given by \( f(0) = 0.57407671_2 \) as opposed to the values proposed by Cohen of 0.57420, the reason for this discrepancy is due to the rapid variation of the function near \( x = 0 \) and \( x = 1 \), and the resulting interpolation difficulties.

Finally, the example described by:

\[
f(x) = (1-x^2)^{3/4} - \frac{\pi \sqrt{2}}{4} (2-x^2) + \frac{2}{3} \int_{-1}^{1} |x-y|^{-\frac{1}{2}} f(y) \, dy \quad (4.8.8)
\]

was solved using the method of residual-perturbation. This example was also considered by Delves et al. 1978 and Baker 1977. In a similar way to the example (4.8.6) above, equation (4.8.8) was rewritten to make use of the symmetry property to become:

\[
f(x) = (1-x^2)^{3/4} - \frac{\pi \sqrt{2}}{4} (2-x^2) + \frac{2}{3} \int_{-1}^{0} \left\{ |x-y|^{-\frac{1}{2}} + (-x-y)^{\frac{1}{2}} \right\} f(y) \, dy
\]

(4.8.9)

The solution of this equation is known to be \( f(x) = (1-x^2)^{3/4} \), which presents major difficulties to any attempt to interpolate
this as an unknown function, since singularities exist in the
derivatives of \( f(x) \) at \( x = -1 \) and \( x = 1 \). It may also be noted
that whilst solving this example, both of the inherent difficulties
associated with the method occurred. Namely, \( R_n \approx K R_n \) for all
\( x \in (-1, 1) \), and \( R_n - K R_n \) became nearly orthogonal to \( R_n \).
Such problems had the effect of impeding the rate of convergence.
Such difficulties were overcome by the use of either, the method
of successive approximations or, the use of the Rayleigh-Ritz
perturbation where \( \alpha_n \) was described by \( \alpha_n = (1-\beta)^{-1} \) from section
4.4, which, though these alternative procedures were not always
convergent, they resulted in comparatively rapid overall convergence.
It was found in practice that the constraint for \( (1-\beta)^2 > \gamma-\beta^2 \)
slowed the overall convergence down unnecessarily, though without
it each iteration was no longer convergent with respect to the
previous iterate. The results of applying the method of residual
perturbation to equation (4.8.8) above are presented in Table VIII
below.

<table>
<thead>
<tr>
<th>Iteration No. ( n )</th>
<th>( N[f_n] )</th>
<th>( (f_n, f_n) )</th>
<th>Rate of convergence from (4.3.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8.21431220</td>
<td>1.2866181</td>
<td>1.767673833(-4)</td>
</tr>
<tr>
<td>2</td>
<td>1.13849309(-4)</td>
<td>5.7031762(-1)</td>
<td>9.9988207(-1)</td>
</tr>
<tr>
<td>4</td>
<td>3.5702590(-5)</td>
<td>5.7591064(-1)</td>
<td>5.1494204(-1)</td>
</tr>
<tr>
<td>8</td>
<td>2.4180642(-8)</td>
<td>5.8882869(-1)</td>
<td>4.0290067(-1)</td>
</tr>
<tr>
<td>12</td>
<td>3.2195186(-10)</td>
<td>5.8899853(-1)</td>
<td>6.8973735(-1)</td>
</tr>
<tr>
<td>16</td>
<td>4.0133637(-13)</td>
<td>5.8904760(-1)</td>
<td>8.2754720(-2)</td>
</tr>
<tr>
<td>20</td>
<td>1.3205717(-15)</td>
<td>5.8904856(-1)</td>
<td>3.1489512(-1)</td>
</tr>
<tr>
<td>exact</td>
<td>( 3/16 \pi = 5.8904862(-1) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Therefore in 20 iterations 7 figures of accuracy have been obtained. It is important to note that this was only made possible by using a low order rational interpolation in the region of the singular derivatives. It may be clearly seen by comparing Tables VII and VIII that the convergence of the sequence is much slower when applied to equation (4.8.9) than equation (4.8.7). One may expect that a better way of overcoming the problems in the present example, would be to form a hybrid of the residual perturbation and the function perturbation, which were presented independently in Table VII. So that the function perturbation could be utilized to provide convergence when the residual perturbation is unstable. However, in practice this idea, though possible, is not practical since in such situations the function perturbation is found to be incredibly slow in its convergence, making it more worthwhile to use the above prescription.

4.9 CONCLUSIONS.

In conclusion to the chapter, it is clearly seen that the possibility of using variation-iteration techniques to solve linear Fredholm integral equations is very real. The method of residual perturbation has also been shown to be the best of the possibilities, and has been demonstrated to be very effective in the solution of a number of examples. There are, however, occasions when this method encounters difficulties, which are apparent to the user by monitoring the magnitude of the inner-products. In such cases, alternative forms of the scaling parameter $\alpha_n$ exist which may be adopted to avoid the difficulties. The alternative
perturbation based on varying the trial function may also be used to overcome difficulties though its convergence may well be prohibitively slow.
CHAPTER 5

Extension to non-linear Equations
5.1 INTRODUCTION.

The main consideration of this chapter is the extension of the perturbation method of chapter 4, to obtain the solution of the more general class of non-linear integral equations.

Such problems were considered by many of the authors of iterative methods referred to in the Introduction of the previous chapter. It is clearly seen from chapter 3, where the method of Morgan 1984 is extended to solve problems involving singular kernels, that such iterative methods lend themselves very naturally to the solution of non-linear equations. This results from the fact that such iterative schemes require the non-linear problem to be reduced to a sequence of linear equations. These linear equations then, only require approximate solution to the accuracy required by the present term in the sequence, since further accuracy will be obtained by the higher iterates. This sequence of approximate linear equations is derived from the application of the Newton-Kantrovich theorem derived by Kantrovich in 1957. This theorem is an extension to Newton's method for the solution of non-linear functional equations, of which integral equations are one possibility. Altman 1957-1961 considered expressing this sequence as inner products, which bears some similarity to the section on the Rayleigh-Ritz variational in the previous chapter. Such ideas were followed by Moore 1963 who used matrices to represent the perturbations between iterates of the sequence. More recently Morgan 1984 considered representing the sequence as residual linear integral equations and then approximately solved such equations by using low order piecewise collocation. In the present discussion
two possible extensions are considered. Firstly, to apply the perturbation method of chapter 4 directly to approximate the Newton-Kantrovich sequence, and secondly, to consider the extension from the angle of using low order accelerations to produce a different sequence, which has some similarity to Steffensen's method, for accelerating a sequence to its limit.

Kantrovich 1957, proposed that if Newton's method was applied to obtain the zeros of the general functional equation given by:

\[ P(f) = 0, \]  

(5.1.1)

then the resulting sequence of iterates for the solution \( f \) is given by \( f_1, f_2, f_3, \ldots, f_n \) and may be obtained by:

\[ f_{n+1} = f_n - \left[ P'(f_n) \right]^{-1} P(f_n) \quad n = 0, 1, 2, \ldots \]  

(5.1.2)

where \( P'(f_n) \) is the Fréchet derivative of (5.1.1) at \( f_n \). However, this formulation possesses many difficulties when an attempt is made to invert this Fréchet derivative. The main problem being that the derivative is a functional operator which according to (5.1.2) requires inverting before it operates on the function \( P(f_n) \). This inversion is equivalent to solving a linear functional equation at each step in the iteration. However, it may be noted that only approximate solutions to these linear functional equations are sought, since any errors may be accounted for in the inversion required for higher iterates of the sequence (5.1.2) above.
5.2 APPLICATION OF THE RESIDUAL PERTURBATION METHOD.

In the present discussion, interest is confined to the class of non-linear equations represented by the Hammerstein equation given by:

\[ f(x) = g(x) + \int_a^b K(x,y,f) \, dy \quad (5.2.1) \]

where the Fredholm equation of chapter 4 is a special case of (5.2.1) when:

\[ K(x,y,f) = K(x,y) \cdot f(y) \quad (5.2.2) \]

The residual function is now defined as:

\[ R_n(x) = g(x) + \int_a^b K(x,y,f_n) \, dy - f_n(x) \quad (5.2.3) \]

having clear similarities to the residual function in chapter 4, (4.2.2). By similar analysis to chapter 4, \( R_{n+1} \) may be written as:

\[ R_{n+1} = R_n(x) - \delta_n(x) + \int_a^b \left\{ K(x,y,(f_n+\delta_n)) - K(x,y,f_n) \right\} dy \quad (5.2.4) \]

However, the integral in (5.2.4) may no longer be simplified as in chapter 4, but may be written as:

\[ R_{n+1} \equiv R_n - \delta_n + K'_n(\delta_n) + \frac{1}{2} K''_n(\delta_n^2) + \frac{1}{6} K'''_n(\delta_n^3) + \ldots \quad (5.2.5) \]

where the operator \( K^m_n(\gamma) \) is the \( m \)th Frechet derivative of \( K \) with respect to \( f_n \) acting on the function \( \gamma \), defined by:

\[ K^m_n(\gamma) = \int_a^b \frac{\partial^m K}{\partial f_n^m} (x,y,f_n) \, \gamma(y) \, dy \quad (5.2.6) \]
It is clear from (5.2.5) that the \((n+1)\)th iterate is largely dependent on the terms involving the high order Fréchet derivatives of the kernel of the non-linear equation. For this reason it is now more difficult to determine criteria for the convergence of certain perturbations. Therefore, in this chapter, consideration is limited to the case of the residual perturbation. The convergence criteria discussed in chapter 4, Section 2 for \(N[f_{n+1}] < N[f_{n}]\) is now applied to the residual equation given by (5.2.5) above in a similar way. The inner-product \((R_{n+1}, R_{n+1})\) may now be considered as:

\[
(R_{n+1}, R_{n+1}) = (R_n, R_n) + 2\alpha_n (R_n, K' R_n - R_n) + \alpha_n^2 \left\{ (K' R_n - R_n, K' R_n - R_n) + (R_n, K'' R_n^2) \right\} + \left\{ \text{terms of order } (\alpha_n R_n)^3 \right\}.
\]

Since \(\alpha_n R_n\) decreases rapidly when convergence takes place the terms of order \((\alpha_n R_n)^3\) may be neglected, to leave a quadratic which may easily be solved for its minimum with respect to \(\alpha_n\). This results in a method which is very good at finding the solutions to equations when a good initial approximation is provided, or when the first few terms in the Neumann series converge rapidly to perhaps, one figure of accuracy. Cases have been experienced where \((\alpha_n R_n)\) does not become small rapidly. However, such equations are very hard to solve by any method, as the Newton-Kantrovich sequence converges slowly. It may also be noted that as the perturbation \((\alpha_n R_n)\) becomes small the residual function \(R_n\) will also become small, making it superfluous to carry out the extra
work required to evaluate the inner-product \((R_n \cdot K_n R_n^2)\), as this is of order \(R_n^3\) compared with \(R_n^2\) for the other terms in (5.2.7). The resulting error in the value of \(\alpha_n\) will, in general, be far out weighed by the large saving in not computing \(K_n R_n^2\). Again, similar to the work of chapter 4, on the Rayleigh-Ritz variational when the term \((K_n R_n - R_n, K_n R_n - R_n)\) becomes small, it is proposed to use the simplified form for \(\alpha_n\). It is, therefore, proposed, in general to set:

\[
\alpha_n = \alpha_n' = \frac{(R_n, K_n R_n - R_n)}{(K_n, K_n R_n - R_n)}
\]

(5.2.8)

to predict the optimum for \(\alpha_n\), except when \((K_n R_n - R_n, K_n R_n - R_n)\) is small. Then, when the Rayleigh-Ritz variational converges, it is proposed to predict \(\alpha_n\) from

\[
\alpha_n = \alpha_n' = \frac{(R_n, R_n)}{(R_n, R_n - K_n R_n)}
\]

(5.2.9)

In the case when neither of these conditions is true, it is found to be useful to set \(\alpha_n = 1\) and allow successive approximations to vary the trial function. Even though the sequence of iterates may diverge, it is found to be useful in changing the initial trial function to a position where more rapid convergence will ensue. Such a process clearly has its limitations and it is not advised to apply it more than once.

The possibility of extending the perturbation considered in chapter 4, section 6, may also be applied to the case of non-linear equations. Though, the criteria for the convergence of this extension
is now greatly more complex, and no simple constructive limit
may be found as terms involving \( R_n^2 \) may not, in general, be
neglected. However, when the residual function \( R_n \) is small, the
method of successive approximations converges, and the rate of
convergence given by

\[
\frac{N[f_{n+1}]}{N[f_n]} = \frac{(R_{n+1}, R_{n+1})}{(R_n, R_n)},
\]

from (5.2.7) is small, say, less than 0.2. Then since the terms of \( \delta_n^2 \) and above
in (5.2.5) may be neglected, it may be shown that the extended
perturbation given by \( \delta_n = R_n + \alpha_n K'R_n \) will converge better
than \( \delta_n = \alpha_n R_n \) the original residual perturbation. In practice
the convergence of the method of successive approximations may be
empirically determined from the condition:

\[
\gamma = \frac{(K' R_n, K' R_n)}{(R_n, R_n)} < 1
\]

which is comparable with the \( \gamma \) defined in (4.3.5), and the residual
function being small may be taken as \( N[f_n] = (R_n, R_n) < 10^{-1} \), so
the high order Fréchet derivatives tend rapidly to zero for all \( x \).
In such proximity to the solution the proof of convergence is the
same as that in section 4.6 except the kernel operators become
Fréchet derivatives, since (5.2.5) simplifies to become:

\[
R_{n+1} = R_n - \delta_n + K'(\delta_n)
\]

Therefore, when

\[
\delta_n = \alpha_n R_n, \quad R_{n+1} = (1-\alpha_n) R_n + \alpha_n K'(\delta_n),
\]

and when

\[
\delta_n^* = R_n + \alpha_n K' R_n
\]
then
\[ \hat{R}_{n+1} = (1-a_n) K'_n R_n + a_n K'_n K'_n R_n, \]  \hspace{1cm} (5.2.13)
so that
\[ \hat{R}_{n+1} = K'_n R_{n+1} \]  \hspace{1cm} (5.2.14)
which is comparable to (4.6.3), and the proof follows by similar analysis.

5.3 SOLUTION OF NON-LINEAR EXAMPLES BY RESIDUAL PERTURBATION.

5.3.1 Analytical Example.

Initially, the simple case of an analytical example will be considered. For this reason the equation:
\[ f(x) = x + x \int_0^1 y f^2(y) \, dy \]  \hspace{1cm} (5.3.1)
is solved by direct application of the above discussion. The Fréchet derivative of the above integral operator is given by:
\[ K'_n \delta_n = 2x \int_0^1 y f_n(y) \delta_n(y) \, dy. \]  \hspace{1cm} (5.3.2)
Applying the method of residual perturbation to the trial function
\[ f_0(x) = x \] results in
\[ R_o(x) = \frac{1}{4} x, \quad a = 2 \quad \text{and} \quad \frac{N[f_{n+1}]}{N[f_o]} = 0, \]
implying that the linear residual equation has been solved exactly.
Therefore, the 2nd term in the sequence for \( f_n \) is given by
\[ f_1(x) = \frac{3}{2} x. \]
A second application gives rise to a similar convergence criteria at zero and
\[ f_2(x) = \frac{7}{4} x. \] Since the actual solution to the above equation is
\[ f(x) = 2x \] it is seen that the error is halved at each
iteration, making it possible to predict that in 26 iterations
the error will be in the 9th figure.

However, since $\gamma$ from (5.2.10) is always less than one and
the higher order Fréchet derivatives are well behaved, in fact
$k^m_n$ are all zero for $m \geq 3$ over all values of $x$. This example
is a comparatively simple one. It has the advantage of producing
an exact solution to the linear residual equation at each iteration
and therefore displays very well the quadratic convergence of the
Newton-Kantrovich sequence. Since the linear equation is solved
effectively at each iteration the extended perturbation carries no
advantage and will produce the exact results to those presented
above. The above example illustrates the methods applicability
to solve what are in iterative terms very difficult problems.
For instance the method of successive approximations would produce
the convergent sequence of functions given by:

$$f_0(x) = x, \quad f_1(x) = 1.25x, \quad f_2(x) = 1.39x, \quad f_3(x) = 1.48x \ldots$$

(5.3.3)

which clearly converges to $f(x) = 2x$ extremely slowly. Therefore,
the above method may be expected to have advantages on more
difficult problems.

5.3.2 Displacement Kernel Examples.

Its direct comparison with the earlier chapter, the generalized
example given by:

$$f(x) = 2 - \frac{1}{\pi} \int_{-1}^{1} \frac{1}{(1+(x-y)^2)^{1/2}} (f(y))^m dy$$

(5.3.4)
is now discussed and solutions obtained, for various values of $m$, by direct application of the method of residual perturbation with $f_o(x) = 2$. The parameter $m$ being allowed to take integer values between 2 and 4 inclusive, use was made of the extended perturbation when $\gamma$ is less than 1. The results are presented in Table I below, it may be noted that comparison is not made with the method of successive approximations, for the present examples, since it is found to be divergent.

### Table I

<table>
<thead>
<tr>
<th>No. of iterations $(n)$</th>
<th>$(f_n, f')_n$</th>
<th>$N[f_n]$</th>
<th>Rate of convergence $N[f_{n+1}]/N[f_n]$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$m = 2$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8.0</td>
<td>6.50667715</td>
<td>2.74702902(-5)</td>
</tr>
<tr>
<td>2</td>
<td>3.27313709</td>
<td>4.86527606(-5)</td>
<td>1.08159562(-1)</td>
</tr>
<tr>
<td>4</td>
<td>3.26321828</td>
<td>8.92529639(-9)</td>
<td>1.45336653(-2)</td>
</tr>
<tr>
<td>6</td>
<td>3.26327100</td>
<td>1.18611346(-14)</td>
<td>6.17318736(-4)</td>
</tr>
<tr>
<td>8</td>
<td>3.26327088</td>
<td>3.98966481(-16)</td>
<td>6.91732578(-2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(machine accuracy)</td>
</tr>
<tr>
<td><strong>$m = 3$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8.0</td>
<td>2.60267086(+1)</td>
<td>4.71466142(-5)</td>
</tr>
<tr>
<td>2</td>
<td>3.07181028</td>
<td>1.28566832(-2)</td>
<td>4.90244536(-3)</td>
</tr>
<tr>
<td>4</td>
<td>2.94498026</td>
<td>2.12277175(-5)</td>
<td>9.96008108(-4)</td>
</tr>
<tr>
<td>6</td>
<td>2.93971714</td>
<td>1.93856084(-9)</td>
<td>7.9033052(-5)</td>
</tr>
<tr>
<td>8</td>
<td>2.93976760</td>
<td>1.45966016(-14)</td>
<td>9.02286230(-3)</td>
</tr>
<tr>
<td>10</td>
<td>2.93976745</td>
<td>2.85972842(-16)</td>
<td>1.38692616(-1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(machine accuracy)</td>
</tr>
<tr>
<td>No. of iterations (n)</td>
<td>$\left(f_n, f_n\right)$</td>
<td>$N[f_n]$</td>
<td>Rate of convergence $N[f_{n+1}]/N[f_n]$</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m = 4$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8.0</td>
<td>1.04106834(+2)</td>
<td>5.78205481(-5)</td>
</tr>
<tr>
<td>2</td>
<td>3.22006625</td>
<td>3.49961001(-1)</td>
<td>7.56455622(-4)</td>
</tr>
<tr>
<td>4</td>
<td>2.73824877</td>
<td>1.17868977(-4)</td>
<td>5.33057321(-2)</td>
</tr>
<tr>
<td>6</td>
<td>2.73790053</td>
<td>4.82328943(-9)</td>
<td>7.97964935(-3)</td>
</tr>
<tr>
<td>8</td>
<td>2.73790023</td>
<td>1.22619396(-15)</td>
<td>6.74761564(-3)</td>
</tr>
<tr>
<td>10</td>
<td>2.73790020</td>
<td>2.28849783(-16)</td>
<td>2.84746055(-1)</td>
</tr>
</tbody>
</table>

These calculations were carried out using a Honeywell Multics computer which has 8 figures of accuracy. The results are therefore, open to question in the 8th figure due to rounding on the computer. Despite this, it is clearly demonstrated that the method of residual perturbation may be used to solve non-linear equations accurately in comparatively little work, as in all 3 cases between 7 and 8 figures of accuracy are achieved for between 8 and 10 iterations, since the above calculations were carried out using a grid of 33 points placed on the zero's of Chebyshev polynomials. Therefore, each iteration requires only 66 quadratures using the 33 point Clenshaw-Curtis formula, to obtain the functions $R_n(x)$ and $K_n R_n(x)$, plus an extra four 33 point Clenshaw-Curtis quadratures to obtain the inner product. It may be noted that no further
function evaluations are required when the inner-products are computed since all the points used have previously been calculated in $R_n(x)$ and $K'R_n(x)$. Unfortunately the exact solutions to the above equations are not known analytically, so the above results may only be compared to other approximate solutions obtained by other numerical procedures. The method of iterative residual collocation due to Evans et al. 1982, has also been used to solve the above problems yielding identical solutions. However, it is difficult to compare their method with the present work, since the work involved in each iterate may not be compared directly. Also comparison of the errors is made difficult since the earlier method does not lend itself to the calculation of the inner-product $(R_n, R_n)$, which is used to measure the error in the present work. However, since the present method does not require the solution of matrices, as the earlier method does, it appears, by comparing the computation time that the present method requires marginally less work when applied to the above example, though this difference is not great enough to deduce any real conclusions.

A further example of a displacement kernel integral equation is:

$$f(x) = -\frac{2}{\pi} \cos x + \frac{\lambda}{\pi} \int_0^{\pi/2} \cos(x-f) \psi(f(t)) \, dt , \quad (5.3.5)$$

which unlike the previous example does not have a solution which exhibits symmetry about the mid-point at the range. The parameter $\lambda$ was allowed to take the values $\lambda = 1, 2, 4, 10$ and the function $\psi(f)$ was varied so that $\psi(f) = f^m$ where $m = 1, 2, 4, 6$. Hence some cases are considered where the Neumann series converges,
and others where it does not. For this reason general comparison cannot be made with the method of successive approximations. It may also be noted that the function \( \cos(x-t) \) is continuous when \( x = t \), so that no difficulties exist in the integration. Therefore, the highly powerful method of Clenshaw-Curtis may be used to perform the required quadratures. Also in the case \( m = 1 \), the analytical solution to (5.3.5) is known to be

\[
\psi(x) = \frac{8}{4\lambda^2 - (4-\lambda)^2 \pi^2} \left\{ 2\lambda \sin x + (4-\lambda)\pi \cos x \right\}
\]

which allows the solutions to be checked directly and hence the number of grid points required to obtain an accuracy of 8 figures could be predetermined. The results obtained when this example was solved using residual perturbations are presented in Table II below, for the number of iterations required to reach 8 figures of accuracy, which, for the purpose of the table is defined as \( N[\psi_n] \leq 10^{-16} \) and \( (\psi_n, \psi_n) \) being consistent to 8 figures.

**TABLE II**

<table>
<thead>
<tr>
<th>( m )</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>13</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
It is clear from the table that the method of residual perturbation is very powerful when applied to such problems. The worst cases being in the linear case when $\lambda = 2$ or 4. This was due to the presence of characteristic values, which had the effect of slowing down the rate of convergence. In the case where $\lambda = 10$ the Neumann series is known to be strongly divergent and the method of successive approximations yields no result for the cases $m = 1, 2$ and 4. However, when $m = 6$ the Neumann series is only slowly divergent yielding a value of $\gamma_0 = 1.02$, which though the method of successive approximations still yields no result, goes some way towards explaining the rapidity of convergence. It may be noted that for this example the value of $\gamma_1$ is 0.41, indicating that the method of successive approximations would converge if a better initial approximation other than the driving function had been used. Though this is of little practical use since the residual perturbation converges many times faster.

The method of piecewise residual collocation of Morgan 1984, was applied to the examples quoted above in Table II. However, the results were not very good, except in the linear case where due to the simplicity of the solutions the results were attained immediately using 5 point interpolation, yielding between 3 and 5 figures of accuracy. Indeed, all the examples presented in Table II, could only be solved to a maximum of 6 or 7 figures of accuracy, since as more points were added the collocation method proved unstable for approximating the solution to the residual equation which resulted in no convergence. In practice many of the collocation matrices became singular, giving no reliance on the perturbations obtained.
Solutions have also been obtained, to (5.3.5), for $\lambda = 1$, when the function $\psi(f)$ is given by $f^3$, $f^5$ and $e^f$. However, these were not included in Table II above since, there is reason to believe that solutions may not exist when $\lambda$ takes the values 2, 4 or 10 and the table would therefore, appear incomplete. In these cases the results were found to be comparably good to those in Table II and are presented below in Table III.

\begin{center}
\textbf{TABLE III}
\end{center}

\begin{center}
\begin{tabular}{|c|c|c|c|}
\hline
\$A = n$ & $N[f_n]$ & $(f_n, f_n)$ \\
\hline
$\psi(f) = f^3$ & 4 & $3.6840364(-19)$ & $4.1494260(-1)$ \\
$\psi(f) = f^5$ & 3 & $7.9958007(-21)$ & $3.4228991(-1)$ \\
$\psi(f) = e^f$ & 6 & $6.4825819(-20)$ & $7.0669884(-2)$ \\
\hline
\end{tabular}
\end{center}

Also for $\lambda = 2$

\begin{center}
\begin{tabular}{|c|c|c|}
\hline
where $\psi(f) = f^5$ & 4 & $8.6210922(-20)$ & $3.8214139(-1)$ \\
\hline
\end{tabular}
\end{center}

where $n$ is the number of the iterate where both $N[f_n] < 10^{-16}$ and $(f_n, f_n)$ is consistent to 8 figures at relative accuracy.

In conclusion to this example, it is clear that this method of residual perturbation is very powerful when applied to nonlinear examples of the form (5.3.5), and also that it performs much better than the method of piecewise residual collocation, which has extreme difficulties when applied to the above example.
5.3.3 Singular Displacement Kernel Example.

To conclude this section on examples of non-linear equations solved by the method of residual perturbations, an example involving singularities in the kernel is now investigated. The example is given by the equation:

\[ f(x) = 1 + \int_{0}^{1} \ln|x-y| \, \psi(f(y)) \, dy \] (5.3.7)

where the function \( \psi \) is allowed to take the forms:

\[ \psi(f) = f^m, \quad m = 2,3,4,5 \] (5.3.8)

and

\[ \psi(f) = e^f. \] (5.3.9)

To enable the singular integrals to be evaluated use was made of the quadrature transformations discussed in chapter 2. It may be noted that this equation bears similarity to the example considered by Cohen 1980 and discussed in chapter 4 as (4.8.5).

The above examples also exhibit, similar difficulties with interpolating the unknown functions, as the functions are known to vary rapidly towards the end of the range. To alleviate this problem, the above equation, (5.3.7) was rewritten:

\[ f(x) = 1 + \int_{0}^{1} \left\{ \ln|x-y| + \ln(1-x-y) \right\} \, \psi(f(y)) \, dy , \] (5.3.10)

making use of the symmetry property, so that interest may be confined to handle the difficulties near \( x = 0 \). Then transformations were applied to the grid points to concentrate points towards the lower end of the range. As an extra precautionary measure a rational interpolation was applied to the points at the end of the
range to account for any possible singularity in the derivatives of \( f(x) \).

The results obtained when the method of residual perturbation was applied to these examples are presented below in Table IV, where as previously, convergence is defined to 8 figures when \( N[f_{\tilde{n}}] < 10^{-16} \) and the inner-product of \((f_{\tilde{n}}, f_{\tilde{n}})\) is consistent to 8 figures and \( \tilde{n} \) is the number of iterations taken to achieve this accuracy.

<table>
<thead>
<tr>
<th>( \psi(f) )</th>
<th>( \tilde{n} )</th>
<th>( N[f_{\tilde{n}}] )</th>
<th>( (f_{\tilde{n}}, f_{\tilde{n}}) )</th>
<th>( f_{\tilde{n}}(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f^2 )</td>
<td>9</td>
<td>1.2204282((-18))</td>
<td>1.5253830((-1))</td>
<td>6.7642759((-1))</td>
</tr>
<tr>
<td>( f^3 )</td>
<td>8</td>
<td>1.3141644((-17))</td>
<td>1.9884390((-1))</td>
<td>7.3173684((-1))</td>
</tr>
<tr>
<td>( f^4 )</td>
<td>10</td>
<td>7.9755059((-19))</td>
<td>2.3225581((-1))</td>
<td>7.6780536((-1))</td>
</tr>
<tr>
<td>( f^5 )</td>
<td>11</td>
<td>8.0112569((-18))</td>
<td>2.5776009((-1))</td>
<td>7.9369488((-1))</td>
</tr>
<tr>
<td>( e^f )</td>
<td>7</td>
<td>1.1943449((-17))</td>
<td>2.8292606((-2))</td>
<td>1.4167612((-1))</td>
</tr>
</tbody>
</table>

It is clearly seen that such examples which involve many potential difficulties, are solved with comparative ease by the present method. It may also be noted, that the value of \( f(x) \) at \( x = 0 \), presented above in the right hand column, is in agreement to the results obtained using the method of piecewise residual collocation, discussed in chapter 3 when applied to the same examples.

In conclusion the method of residual perturbation is clearly
seen to be very powerful, when applied to non-linear equations, and possesses many advantages, due to the progressive monitoring of the convergence which may be used to increase the number of points in the grid, to confirm the accuracy of the solution.

The main limitations of this method are involved in the difficulties associated with interpolation over this grid. However, the method gives the user freedom to use different methods of interpolation over different portions of the grid to accommodate the unknown function. The instabilities discussed in chapter 4 may still be encountered when applied to non-linear equations. However, in the more general case of non-linear equations other forms of the perturbation exist which may be used to overcome such difficulties. These other forms are derived from the discussion on extensions to the use of low order accelerators which follows in the next section.

5.4 COMPARISONS WITH THE USE OF LOW ORDER ACCELERATORS TO SOLVE NON-LINEAR EQUATIONS.

Originally Steffensen considered the successive application of low order accelerators to obtain the solutions of non-linear equations in two dimensional Euclidian space which resulted in his well-known method. This method makes use of the powerful Aitken $\delta^2$ accelerator, and is known to converge rapidly to the correct zeros of a one dimensional function even though the direct iterative process diverges. Recently, Morgan 1984, considered the direct pointwise application of low order accelerators to accelerate iterative sequences of functions, in the solution of
integral equations. However, such direct application suffers from instabilities when the second differences between three successive iterates in the sequence of functions, possesses zeros in the range of the solution. This may easily be seen when it is realized that, when the Aitken \( \epsilon^2 \) accelerator is applied in such a manner to a sequence of functions the resulting iterative scheme is defined by:

\[
\begin{align*}
\hat{f}_n^*(x_m) &= f_n(x_m) - \frac{(f_{n+1}(x_m) - f_n(x_m))^2}{f_{n+2}(x_m) - 2f_{n+1}(x_m) + f_n(x_m)}
\end{align*}
\]

(5.4.1)

where \( x_m \) \( m = 0, \ldots, M \) are the grid points over which the solution is being considered, such that, \( a \leq x_0 \leq x_1 \leq \ldots \leq x_m \leq b \).

Clearly, if a zero exists in the denominator of the perturbation term, between \( a \) and \( b \), on the right hand side of (5.4.1) above, any resulting iterative scheme based on the "improved" set of function values \( f^*_n(x_m) \) will suffer from instabilities. This instability may exist either in reality or simply in the numerical evaluation of such a perturbation, depending on the possible existence of a comparable zero in the numerator. However, in either case this presents a major drawback to the use of (5.4.1) as a general method of solving integral equations. Though, for the cases where the first difference is, continuous over the range \((a,b)\), and is monotonically increasing or decreasing such a consideration forms the basis of a very good method.

The above equation (5.4.1) may be rewritten in terms of the previously defined residual function, (5.2.3), as:
where the term \( R_n^*(x) \) is the residual function for the iterated function \( f_n(x) + R_n(x) \) given by:

\[
R_n^*(x) = g(x) + \int_a^b K(x, y, f_n(y) + R_n(y)) \, dy - f_n(x) - R_n(x).
\]

This equation may now be seen to bear some similarity with the Newton-Kantrovich iteration (5.1.2), since the term:

\[
\frac{R_n(x)}{R_n^*(x) - R_n(x)}
\]

is an approximation to the inverse Fréchet derivative previously represented by \([R_n'(x)]^{-1}\). Previously in chapter 4, and also in the earlier part of this chapter, §5.2 this inverse operator has been approximated by a constant. Hence, by similar application of inner-products, the potential instability, commented on above, may be incorporated into a constant, producing a different means of obtaining an extension of residual perturbation to non-linear equations. For the sake of definition the present method will be referred to as, Steffensen perturbation, in the remainder of this chapter.

The question remains as to how to apply inner-products to the above expression (5.4.4), in such a way as to obtain an optimal perturbation for the convergence of the variational (4.2.3) towards zero. For this reason, it is valuable to understand the relationship between the functions, \( K_n R_n(x) \) and \( R_n^*(x) \), and how they relate to
the simpler function $\mathcal{K}_n R_n(x)$ of chapter 4. It has already been seen that for linear equations the functions $\mathcal{K}_n R_n$ and $\mathcal{K}'_n R_n$ are identical, in section 5.2. Hence from the definition of $R^*_n$ (5.4.3) above, it may also be shown that:

$$R^*_n(x) = \int_a^b \mathcal{K}(x, y, f_n(y) + R_n(y)) - \mathcal{K}(x, y, f_n(y)) \, dy,$$  \hspace{1cm} (5.4.5)

which clearly illustrates that $R^*_n(x)$ is also identical to $\mathcal{K}_n R_n$ for linear equations. In the case where the operator $\mathcal{K}$ is non-linear the function $R^*_n$ is now related to the previous work by the Neumann series, given by:

$$R^*_n(x) = \mathcal{K}'_n R_n(x) + \frac{1}{2} \mathcal{K}''_n R^2_n(x) + \frac{1}{6} \mathcal{K}'''_n R^3_n(x) + \ldots,$$  \hspace{1cm} (5.4.6)

which clearly tends towards $\mathcal{K}'_n R_n$ as the residual function $R_n$ becomes small, if the individual terms are non-singular over the range $(a, b)$. It may be noted that this approximation is not necessarily valid when $R_n(x) > 1$ for some $x \in (a, b)$ as the above sequence (5.4.6) will bear no relation to the earlier work. However, in the proximity of the solution, the above sequence forms a valid alternative to calculating the Fréchet derivative directly, and in the case when instabilities occur $R^*_n$ supplies a valid model of the higher Fréchet derivatives.

Clearly, an infinite number of different possibilities exist to reduce the function (5.4.4) to a constant, using inner-products. The remaining task is, therefore, to obtain such forms of this reduction, as to provide analogy with, and direct relation to, the earlier discussion on residual perturbation. Firstly,
inner-products were taken, both in the numerator and the
denominator, of (5.4.4) with the function $R_n^*(x) - R_n(x)$, as this
reduces to the residual perturbation when $K$ is linear. Under
this scheme the resulting iterative sequence becomes:

$$f_{n+1}(x) = f_n(x) + \frac{(R_n' R_n - R_n^*)}{(R_n^- R_n^- R_n^-)} R_n(x), \quad (5.4.7)$$

which is clearly directly related to using $\tilde{a}_n$ from (5.2.8),
replacing $\sum_n R_n$ by $R_n^*$. Also, since the previous iterative
sequence (5.4.2) may be written as:

$$f_n^*(x) = f_n(x) + R_n^*(x) + \frac{R_n(x) R_n^*(x)}{R_n(x) - R_n^*(x)}. \quad (5.4.8)$$

The above discussion may also give rise to the iterative scheme
defined by:

$$f_{n+1}(x) = f_n(x) + R_n(x) + \frac{(R_n' R_n^- R_n^*)}{(R_n^- R_n^- R_n^-)} R_n^*(x), \quad (5.4.9)$$

which also bears direct analogy to the extended perturbation of
(5.2.13). Therefore, the present discussion gives rise to the
two possible formulae (5.4.7) and (5.4.9) which, when applied to
linear equations, reduces to the standard residual perturbation
and the extended residual perturbation previously seen in
chapter 4.

Secondly, inner-products may also be taken in both the
numerator and denominator of (5.4.4) with the residual function
$R_n(x)$, which results in the perturbations of the form:
\[ f_{n+1}(x) = f_n(x) + \frac{(R_n^* - R_n)}{(R_n - R_n^*)} R_n(x) \quad (5.4.10) \]

and

\[ f_{n+1}(x) = f_n(x) + R_n(x) + \frac{(R_n^* - R_n)}{(R_n^* - R_n^*)} R_n^*(x) \quad (5.4.11) \]

by application with (5.4.2) and (5.4.8) respectively. It may also be noted that such perturbations are analogous to those obtained in chapter 4, when the Rayleigh-Ritz variational was used to define convergence, section 4.4.

In the earlier part of this chapter, section 5.2, conditions were established for the convergence of the Kantrovich type of residual perturbations for non-linear equations. It, therefore, remains to carry out a similar analysis on the Steffensen type of perturbations defined as (5.4.7), (5.4.9), (5.4.10) and (5.4.11) above. However, since it is clear that (5.4.9), (5.4.10) and (5.4.11) bear similar relations to the standard perturbation (5.4.7), as the analogous perturbations of chapter 4 and section 5.2, compare with the standard residual perturbation. It, therefore, remains to give detailed consideration to the perturbation defined (5.4.7) above. The \((n+1)\)th residual function is now related to the \(n\)th residual function by the equation:

\[ R_{n+1} = (1 - \alpha_n) R_n + \alpha_n (R^*_n - R_n) + \frac{1}{2} (\alpha^2_n - \alpha_n) K'' R_n^2 + \frac{1}{6} (\alpha^3_n - \alpha_n) K''' R_n^3 \quad (5.4.12) \]

where \(\alpha_n\) is given by:

\[ \alpha_n = \frac{(R_n^* - R_n)}{(R_n^* - R_n^*)} \quad (5.4.13) \]
As before, convergence is defined by \( N[f_{n+1}] < N[f_n] \), which requires the evaluation of the ratio:

\[
\frac{(R_{n+1}, R_{n+1})}{(R_n, R_n)},
\]

which from (5.4.12), may be written as:

\[
\frac{N[f_{n+1}]}{N[f_n]} = (1-\alpha_n)^2 + 2\alpha_n (1-\alpha_n) \frac{(R_n, R_n^*) - (R_n, R_n)}{R_n, R_n} + \alpha_n^2 \frac{(R_n, R_n^*, R_n^*, R_n)}{R_n, R_n,
\]

\[
+ (1-\alpha_n) (\alpha_n^2 - \alpha_n) \frac{(R_n, K_n^*, R_n^2)}{R_n, R_n} + \alpha_n (\alpha_n^2 - \alpha_n) \frac{(R_n^*, R_n, R_n^2)}{R_n, R_n,
\]

\[
+ \left\{ \text{terms of order } R_n^2 \text{ and above} \right\}. \quad (5.4.14)
\]

Since the sequence (5.4.6) is only valid in the proximity of the solution, it is reasonable to neglect the higher order terms involving inner-products containing \( R_n^4 \) and greater. Also, under such an assumption the inner-product \((R_n^*, R_n, K_n^*, R_n^2)\) becomes \((K_n^*, R_n, K_n^*, R_n^2)\). Resulting in a cubic equation which may not be easily solved since the inner-products \((R_n, K_n^*, R_n^2)\) and \((K_n^*, R_n, K_n^*, R_n^2)\) are not known. Hence similar to the results of section 5.2, it is no longer possible to obtain the exact rate of convergence without carrying out a great deal of work in each iteration.

However, considering the case when we are in close proximity to the solution, the terms involving inner-products of \( R_n^3 \) may also be neglected since it is only required to find an approximate solution to the residual equation at each iteration. When this is true the optimum value of \( \alpha_n \) in equation (5.4.14) is clearly
given by (5.4.13) above, therefore, substantiating the above
choice of inner-products to obtain the Steffensen perturbations.
The extension of (5.4.7) given by (5.4.9) may be shown by similar
consideration to that of section 4.6 and section 5.2 to be a
definite improvement over (5.4.7) when the inequality:
\[
\gamma = \frac{(R^*_n - R_n, R^*_n - R_n)}{(R^*_n, R_n)} < 1
\]
(5.4.15)
holds. However, when \( \gamma > 1 \), (5.4.9) may still be an improvement over
(5.4.7) though such cases may only be practically ascertained. In general
it is not suggested to calculate the term \( K''_n R^2_n \), as to do this
requires that the value of the second Fréchet derivative be known,
so resulting in an effective increase in work by approximately
50\% per iteration. The computational stability of this term is
also open to question, as errors in the residual function may be
propagated, by integrating the square of the function. The other
perturbations defined by (5.4.10) and (5.4.11) may be shown to be
optimally convergent when the Rayleigh-Ritz variational is used
to define convergence, similar to section 4.4, though in general
such convergence may be shown to be inferior. The question now
arises as to the possibility of these Steffensen perturbations
having any distinct advantages over the previously considered
residual perturbations. One advantage is clearly apparent:
the earlier perturbations have all required the user to provide
the value of the Fréchet derivative as part of the description
of the integral equation, to be present. Steffensen perturbations,
do not require this Fréchet derivative, and are, therefore, simpler
to use. It may also be noted that such an advantage allows the direct solution of equations, where, for various reasons, the Frechet derivatives may not be easily determined. Such an example is the Chandrasekhar equation, defined by:

\[ f(x) = 1 + \frac{1}{2} \int_0^1 \frac{x}{x+y} f(x) f(y) \, dy \quad \text{(5.4.16)} \]

Hence, the class of non-linear equations which possess a Urysohn kernel defined by:

\[ f(x) = g(x) + \int_a^b f(x) U(x,y,f(y)) \, dy \quad \text{(5.4.17)} \]

may be solved, though the Frechet derivatives may be difficult to obtain.

Another possible advantage of the Steffensen perturbations over the residual perturbations occurs when the parameter \( a_n \) is greater than a half, and especially so when it is also less than one. Since under these circumstances, the 2nd Frechet derivative term from (5.4.12) is clearly less than the comparable term from (5.2.5) in modulus, for all \( x \) in the range of the problem.

However, the Steffensen perturbations also suffer from a disadvantage over the previous perturbations. This is the failure of the scheme to necessarily converge when the trial function is not in the close proximity to the solution, i.e. when \( (R_o, R_o) \geq 10 \).

It may also be noted that though the trial function is required to be in the proximity of the solution, there is no requirement that it should model the solution. The disadvantage is, therefore, not as severe as it may first appear.
It may also be noted that by taking due regard to the equations for the \((n+1)^{\text{th}}\) variational, given by (5.2.7) for the residual perturbation, and obtained from (5.4.14) for the Steffensen perturbations, the Steffensen perturbation has more rapid convergence when:

\[
|a_n^2 (K_n R_n, K_n R_n^2)| > |(a_n^2 - a_n)(R_n, K_n R_n^2)| \tag{5.4.18}
\]

However, this is of little value, since, in general, the user will not know the value of the function \(K'' R_n\). Therefore, these perturbations will not be combined into a general scheme, except to note that, when the inner-products in one scheme are unstable the other may be used as a stable alternative, as long as the form of the operator \(K\) allows such interchangeability.

5.5 EXAMPLES OF THE STEFFENSEN PERTURBATION.

5.5.1 Analytical Example.

For the purpose of comparison the examples, which were used to demonstrate the earlier perturbation in section 5.3, are now applied to the Steffensen perturbations, starting with the initial analytical example given by:

\[
f(x) = x + x \int_0^1 y f^2(y) \, dy . \tag{5.3.1}
\]

The Steffensen perturbations of the last section were found to yield,

\[
R_o(x) = \frac{1}{4} x, \quad a_o = \frac{16}{7} \quad \text{and} \quad \frac{N[f_{n+1}]}{N[f_n]} = 0 ,
\]
implying that the linear approximation to the problem has been
solved exactly in one iteration. Successive application of
this perturbation gives rise to the sequence defined by:

\[ f_0(x) = x , \quad f_1(x) = \frac{11}{7} x , \quad f_2(x) = \frac{667}{371} x , \quad (5.5.1) \]

which may be shown to be convergent towards the exact result
\( f(x) = 2x \). Also by comparing the sequence (5.5.1) with
section 5.3.1 it may easily be seen that the present method
converges more rapidly, though as the solution is approached the
rate of convergence will become the same as that of the earlier
work. This more rapid convergence is not surprising when it is
remembered that the variational for the \((n+1)\)th iterate from
equation (5.4.14) is smaller than the variational given by (5.2.7)
when \( a_n > 1 \). In conclusion the applicability of the Steffensen
perturbation is clearly seen from this example, though detailed
comparison with the earlier work is of little value on such a
simple problem.

5.5.2 Displacement Kernel Examples.

In direct comparison with the examples solved in section 5.3.2,
the same equation is now solved using Steffensen perturbation.
This equation is given by:

\[ f(x) = 2 - \frac{1}{\pi} \int_{-1}^{1} \frac{1}{1+(x-y)^2} (f(y))^m \, dy , \quad (5.3.4) \]

where the parameter \( m \) was again allowed to vary over the values
\( m = 2, 3 \) and 4. The initial trial function was set to be \( f_0(x) = 2, \)
which was found to result in very good convergence in the first two cases, though divergence was experienced for \( m = 4 \). This was due to the trial function \( f_0(x) = 2 \) not being in the close proximity of the solution, and may have been easily predicted by the value \( N[f_0] = (R_o R_o) = 52.053416 \). In this case the trial function was set to be \( f_0(x) = 1 \), which though it bears no relation to the solution, is in the close proximity of the solution since \( N[f_0] = (R_o R_o) = 0.30596813 \).

Again, use was made of the extended perturbation (5.4.9) when the value of \( \gamma < 1 \), where \( \gamma \) is now defined from (5.4.15) and the results obtained are presented in Table V below.

| No. of iterations (n) | \((f_n, f'_n)\) | \(N[f_n]\) | Rate of convergence 
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( n+1/N[f_n])</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( m = 2 ) \  ( f_0(x) = 2 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 8.0</td>
<td>6.5066770</td>
<td>7.9412949(-7)</td>
<td></td>
</tr>
<tr>
<td>2 3.2251516</td>
<td>5.2668716(-4)</td>
<td>5.0394212(-7)</td>
<td></td>
</tr>
<tr>
<td>4 3.2631980</td>
<td>3.0565159(-15)</td>
<td>8.5419174(-2)</td>
<td></td>
</tr>
<tr>
<td>6 3.263270</td>
<td>3.7314274(-21)</td>
<td>1.057275(-1)</td>
<td></td>
</tr>
<tr>
<td>( m = 3 ) \  ( f_0(x) = 2 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 8.0</td>
<td>2.6026780</td>
<td>1.023678529(-4)</td>
<td></td>
</tr>
<tr>
<td>2 2.04606740</td>
<td>5.8147778(-1)</td>
<td>6.27959941(-5)</td>
<td></td>
</tr>
<tr>
<td>4 2.93281516</td>
<td>3.664632(-5)</td>
<td>1.4401300(-5)</td>
<td></td>
</tr>
<tr>
<td>6 2.93976534</td>
<td>3.1764289(-12)</td>
<td>1.6073343(-2)</td>
<td></td>
</tr>
<tr>
<td>8 2.93976746</td>
<td>4.521975(-16)</td>
<td>7.9972963(-4)</td>
<td></td>
</tr>
<tr>
<td>9 2.93976745</td>
<td>2.894781(-20)</td>
<td>1.3278225(-1)</td>
<td></td>
</tr>
</tbody>
</table>
To obtain the greater accuracy displayed above the results were computed on an ICL 2900 computer using 16 figure arithmetic. By comparing the above Table with Table I earlier in the chapter, it may be easily seen that the present results are better since the value of $N[f_n]$ is smaller for all $n$. It is important to note that since the trial function $f_0(x) = 1$ instead of $f_0(x) = 2$ was used for $m = 4$, the last section of the above table cannot be compared with the same example in Table I. Therefore, for direct comparison in the case when $m = 4$ the method of residual perturbation was also applied to the example using the trial function $f_0(x) = 1$. The results of which are presented below in Table Va.

<table>
<thead>
<tr>
<th>No. of iterations $(n)$</th>
<th>$(f_n^*, f_n)$</th>
<th>$N[f_n]$</th>
<th>Rate of convergence $N[f_{n+1}]/N[f_n]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 4$, $f_0(x) = 1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2.0</td>
<td>6.1193625(-1)</td>
<td>1.80581963(-2)</td>
</tr>
<tr>
<td>2</td>
<td>2.70160520</td>
<td>1.7632871(-3)</td>
<td>1.74363680(-3)</td>
</tr>
<tr>
<td>4</td>
<td>2.73752783</td>
<td>1.8665207(-7)</td>
<td>4.2633214(-3)</td>
</tr>
<tr>
<td>6</td>
<td>2.73790109</td>
<td>2.8389712(-12)</td>
<td>1.7979975(-1)</td>
</tr>
<tr>
<td>8</td>
<td>2.73790018</td>
<td>1.8269024(-15)</td>
<td>2.1723284(-1)</td>
</tr>
<tr>
<td>9</td>
<td>2.73790021</td>
<td>7.6467030(-17)</td>
<td>1.2379684(-1)</td>
</tr>
</tbody>
</table>
It is clear, by comparing the last section of Table V with Table Va above, that the method of residual perturbation is superior for finding a close approximation to the solution. However, the Steffensen perturbations are better for obtaining accurate results, once the sequence of iterates is in very close proximity to the solution.

In conclusion, the differences exhibited between Tables I and V imply that the Steffensen perturbations are superior with a saving of nearly 3 iterations for the case \( m = 2 \). However, it may be noted that in this case the parameter \( a_n \) lies between 0.4 and 0.6 and the Fréchet derivatives of order greater than three do not exist. This accounts for the superior convergence of the present method close to the solution, since the convergence ratio given by (5.4.14) is less than that of (5.2.7).
To provide further comparisons between these perturbations the example:

\[ f(x) = -\frac{2}{\pi} \cos x + \frac{\lambda}{\pi} \int_0^{\pi/2} \cos(x-t) \psi(t) \, dt \quad (5.3.5) \]

is now solved using the Steffensen perturbations, for the same values of \( \lambda \) and \( \psi \) which were considered earlier.

\[ \left\{ \text{i.e. } \lambda = 1,2,4,10 \text{ and } \psi(f) = f^m, \quad m = 1,2,4,6 \right\} \]

It may also be noted that this example is very useful to demonstrate the difficulties encountered when applying the Steffensen accelerators directly in the original pointwise sense, from (5.4.2). To illustrate this, let, \( \lambda = 2 \) and \( \psi(f) = f \). Then if the initial trial function is \( f_0(x) = -\frac{2}{\pi} \cos x \), it may be shown that the residual functions \( R_0(x) \) and \( R_0^*(x) \) are given by the equations:

\[ R_0(x) = -\frac{1}{\pi} \cos x - \frac{2}{\pi^2} \sin x \quad (5.5.2) \]

\[ R_0^*(x) = -\frac{1}{2\pi} \cos x - \frac{2}{\pi^2} \sin x - \frac{2}{\pi^3} \cos x \quad (5.5.3) \]

which when applied to the original iterative scheme of (5.4.2) results in the function:

\[ f_1(x) = -\frac{2}{\pi} \cos x + \frac{2\pi}{(4-\pi^2)} \cos x + \frac{8}{(4-\pi^2)} \sin x + \frac{8}{\pi(4-\pi^2)} \sin x \tan x. \quad (5.5.4) \]

It is now clear that the above function \( f_1(x) \) is singular at \( x = \pi/2 \), which causes the next residual function \( R_1(x) \) to be
unobtainable, since the integral of $\tan x$ between zero and $\pi/2$
does not exist. It is therefore seen that the original pointwise
Steffensen scheme fails to solve a comparatively simple example,
whose solution is known to be:

$$f(x) = \frac{4}{4-\pi^2} \left\{ 2 \sin x - \pi \cos x \right\}.$$  \hspace{1cm} (5.5.5)

The results obtained when the Steffensen perturbations were
applied to the above examples, are presented below in Table VI,
which displays the number of iterations required for both the
variational $N[f_n]$ to be less than $10^{-16}$ and the inner-product
$(f_n, f_n)$ to be consistent to eight figures of accuracy. The
numerical integrations were carried out using the powerful Clenshaw-
Curtis quadrature method, so the table may be directly compared to
Table II earlier in the chapter.

| TABLE VI |
| $\psi(f) = \ell^m$, \hspace{1cm} $m = 1, 2, 4, 6$ |

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>13</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
It may be clearly seen by comparing the above table with the work on residual perturbation in Table II that there is little difference between the methods in the number of iterations taken to achieve 8 figures of accuracy. However, as may be expected from the theoretical discussion of the previous section, Section 5.4, the Steffensen perturbations have a marginal advantage when the parameters $a_n$ are greater than a half, as is the case when $\lambda = 1, 2$. In these cases, the method of successive approximations also converges, which explains the extremely rapid convergence of the above perturbations. Also, it may be seen that when the values delivered by $a_n$ are not all positive, the method of Steffensen perturbations are at a disadvantage, which is the case when $m = 2$ and $\lambda = 10$. It may be noted that this case has the most divergent Neumann series of the non-linear cases considered above, and therefore, would be expected to be the hardest to solve. As expected the linear cases yield identical results to those of Table II since both perturbations simplify to the case of chapter 4.

Solutions have also been obtained to the equation (5.3.5) for the cases given by $\psi(f) = f^3, f^5, e^f$, when $\lambda = 1$, and for $\psi = f^5$ when $\lambda = 2$, though no solutions have been obtained for the other values of $\lambda$ under such functions for $\psi$. The results of these tests are presented below in Table VII for comparison with the results of Table III in section 5.3.
TABLE VII

\[ \lambda = 1 \]

\[
\begin{array}{|c|c|c|c|}
\hline
\psi(f) = f^3 & n & N[f_n] & (f_n, f_n) \\
\hline
\psi(f) = f^5 & 4 & 3.5554950(-18) & 4.1494260(-1) \\
\psi(f) = e^f & 3 & 7.6691542(-22) & 3.4228991(-1) \\
\psi(f) = e^f & 5 & 1.3978542(-18) & 7.0669884(-2) \\
\hline
\end{array}
\]

\[ \lambda = 2 \]

\[
\begin{array}{|c|c|c|c|}
\hline
\psi(f) = f^5 & 4 & 2.3229914(-19) & 3.8214139(-1) \\
\hline
\end{array}
\]

where \( n \) is the number of iterations required for both \( N[f_n] < 10^{-16} \) and \( (f_n, f_n) \) to be consistent to 8 figures of relative accuracy.

Comparing this table with Table III of section 5.3, it is hard to draw any overall conclusions about the methods. Clearly there are cases when the Steffensen perturbations converge more rapidly than the residual perturbations, as in the case when \( \psi = e^f \). However, there are also cases when the residual perturbation converges more rapidly, as is the case when \( \lambda = 2 \) and \( \psi = f^5 \), where \( N[f_n] \) is about half that obtained by Steffensen perturbation.

In conclusion, it may be noted that the method of Steffensen perturbation is very comparable to the earlier method of residual perturbation, and is, in general, an improvement when the parameter \( \alpha_n \) is consistently greater than a half. However, it is clear from the present example that any differences are small when applied to examples where the solutions are easily modelled.
5.5.3 Singular Displacement Kernel Example.

To establish that both the residual and Steffensen perturbations are equally valid, where the solution is known to contain interpolatory difficulties, the example involving a singular displacement kernel used in section 5.3.3, is now solved using the Steffensen perturbations. This example is given by the equation:

\[ f(x) = 1 + \int_0^1 \ln|x-y| \psi(f(y)) \, dy \quad (5.3.7) \]

where the function \( \psi \) is allowed to vary over the possibilities given by:

\[ \psi(\epsilon) = \epsilon^m \quad m = 2, 3, 4, 5 \quad (5.3.8) \]

and

\[ \psi(\epsilon) = e^\epsilon \quad (5.3.9) \]

As before, the quadrature transformation methods of chapter 2 were used to evaluate the integrals which contained singularities. Also, in a similar way to before, the above equation was rewritten as:

\[ f(x) = 1 + \int_0^1 \left\{ \ln|x-y| + \ln(1-x-y) \right\} \psi(f(y)) \, dy \quad (5.3.10) \]

to allow more detailed consideration to be given to the rapid variation in the function where \( x \) is zero, and making use of the symmetry of the problem. Again, grid points were transformed into the region around zero and rational interpolation was used to model the last few grid points near zero.

The results obtained by the present method are presented in Table VIII below where \( n \) is the number of iterations required to
obtain 8 figures of relative accuracy, using the same definition for accuracy as that previously considered in this chapter.

<table>
<thead>
<tr>
<th>( \psi(f) )</th>
<th>( n )</th>
<th>( \mathbf{N}[f_n] )</th>
<th>( (f_n, f_n) )</th>
<th>( f_n(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f^2 )</td>
<td>9</td>
<td>8.5549186(-19)</td>
<td>1.5253830(-1)</td>
<td>6.7642759(-1)</td>
</tr>
<tr>
<td>( f^3 )</td>
<td>8</td>
<td>5.9546496(-17)</td>
<td>1.9884390(-1)</td>
<td>7.3173648(-1)</td>
</tr>
<tr>
<td>( f^4 )</td>
<td>8</td>
<td>1.2964843(-17)</td>
<td>2.3225581(-1)</td>
<td>7.6780536(-1)</td>
</tr>
<tr>
<td>( f^5 )</td>
<td>10</td>
<td>7.6884339(-19)</td>
<td>2.5776009(-1)</td>
<td>7.9369488(-1)</td>
</tr>
<tr>
<td>( e^f )</td>
<td>8</td>
<td>5.3185778(-19)</td>
<td>2.8292606(-2)</td>
<td>1.4167612(-1)</td>
</tr>
</tbody>
</table>

It may be clearly seen, by comparing Tables IV and VIII that the Steffensen perturbations converge more rapidly, in the cases when \( \psi \) is given by \( f^2, f^4 \) and \( f^5 \) though the earlier residual perturbations converge more rapidly for the cases when \( \psi \) is given by \( f^3 \) and \( e^f \). These results bear close similarity with those already obtained in the previous examples in this section. Where the Steffensen perturbations proved to be superior when the trial function was in close proximity to the solution and the parameter \( \alpha_n \) was in the range (1, 1), and when the parameter \( \alpha_n \) lay between zero and a half, little difference exists between the perturbations. This accounts for the differences between the tables when, it is noted that in all the cases values of \( \alpha_n \) were encountered ranging from 0.3 to 0.6, notably more often 0.6 when \( \psi = f^4 \) and \( f^5 \). However, the main differences are due to the distances between
the trial function \( f_0(x) = 1 \) and the solutions, which in direct relation to the table were greater when \( \psi \) was given by \( t^3 \) or \( e^f \), when the residual perturbation converged more rapidly.

In conclusion, there is little difference between such perturbation techniques when applied to examples with singular kernels. Since both perturbations prove to be equally valid when applied to the above set of examples. Indeed the Steffensen approach, proves to be highly successful in cases where the original trial function is in close proximity to the actual solution.

5.5.4 Urysohn Kernel Example.

As previously mentioned in section 5.4, the Steffensen perturbations may also be used to solve integral equations which possess kernels of the Urysohn type, where the kernel is both a function of \( f(x) \) and \( f(y) \). Such problems may otherwise prove hard to solve since the Fréchet derivatives, are less easily obtained. As an example of such a problem, the Chandrasekhar equation given by:

\[
f(x) = 1 + \frac{1}{2} w f(x) \int_0^1 \frac{x}{x+y} f(y) \, dy
\]

(5.4.16)
is now solved using the Steffensen perturbations, when \( w = 0.95 \) and \( w = 1.0 \). The solutions to this problem were originally presented in a paper on atomic scattering theory by Stibbs and Weir 1959, whose results are reproduced by the present method. The number of iterations required to obtain these results are presented in Table IX below, where the initial trial function was taken to be

\[
f_0(x) = 1 + 2x
\]

(5.5.6)
### Table IX

<table>
<thead>
<tr>
<th>No. of iterations (n)</th>
<th>N[f_n]</th>
<th>(f_n,f_n)</th>
<th>f_n(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>w = 0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>5.1526135(-3)</td>
<td>4.3333333</td>
<td>2.325644</td>
</tr>
<tr>
<td>2</td>
<td>1.9802998(-4)</td>
<td>2.9568848</td>
<td>2.082254</td>
</tr>
<tr>
<td>4</td>
<td>1.5886480(-7)</td>
<td>2.7645336</td>
<td>2.077127</td>
</tr>
<tr>
<td>6</td>
<td>3.0540707(-11)</td>
<td>2.7588264</td>
<td>2.077113</td>
</tr>
<tr>
<td>8</td>
<td>6.0423504(-14)</td>
<td>2.758902</td>
<td>2.077123</td>
</tr>
<tr>
<td>w = 1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>5.8381921(-4)</td>
<td>4.3333333</td>
<td>2.926536</td>
</tr>
<tr>
<td>1</td>
<td>7.0397653(-6)</td>
<td>4.2906835</td>
<td>2.910291</td>
</tr>
<tr>
<td>3</td>
<td>1.1092318(-9)</td>
<td>4.2874362</td>
<td>2.907875</td>
</tr>
<tr>
<td>5</td>
<td>4.6478549(-13)</td>
<td>4.2874166</td>
<td>2.907811</td>
</tr>
<tr>
<td>7</td>
<td>7.2587968(-16)</td>
<td>4.2874157</td>
<td>2.907809</td>
</tr>
</tbody>
</table>

For the purpose of comparison, the following values (Table IXb) were calculated at the same nodes as those quoted by Stibbs and Weir. It may be noted that they are identical to the 7 significant figures, originally quoted.
It may be noted that the present method may be applied to the above example using any of the values for $w$, quoted by Stibbs and Weir. The two cases above were chosen as they represent two of the most difficult cases and therefore, illustrate the present method at its best. When considering the above table, it is important to note that the solution of the above example when $w = 1$ poses a very difficult problem which has normally been solved previously by rewriting equation (5.4.16) as:

$$\theta(x) = 1 + \frac{1}{2} w \int_0^1 \frac{x}{x+y} \cdot \frac{1}{\theta(y)} \, dy,$$

(5.5.7)

where $f(x) = 1/\theta(x)$. This equation becomes seriously unstable when $w = 1$, as the sequence of Fréchet derivatives diverges when $x = 1$. It is therefore, extremely advantageous to be able to solve equation (5.4.16) directly without the need of Fréchet derivatives.
5.6 CONCLUSIONS.

In the present chapter the possibilities of extending the residual perturbation method to handle the case of non-linear equations has been clearly seen. It has also been shown that such extension may make use of the Newton-Kantrovich theorem for the iterative solution of non-linear functional equations. Though in practice the method of residual perturbation requires very few embellishments, when applied to the more general problem. The only major change being the need to deliver the 1st Fréchet derivative of the non-linear kernel operator, in addition to the usual description of the problem. The method has also been shown to be highly useful in the solution of non-linear equations with a wide variety of examples being considered.

It has also been shown that extension of the earlier work may also be pursued from the viewpoint of accelerating a sequence of trial functions to its limit. This generalized acceleration technique bears direct relation to the work of chapter 4, but is distinctly different to that of the Newton-Kantrovich extension. It has the advantage of not requiring any previous knowledge of the Fréchet derivatives of the kernel operator, but also suffers from the disadvantage of requiring closer proximity to the solution before convergence ensues. It has been shown to be comparable in its ability to solve problems as that of the earlier work, though care is sometimes required to adjust the initial trial function so that convergence is obtained. However, probably the greatest potential advantage of this method lies in its ability to solve problems where the Fréchet derivatives may be hard to find.
In conclusion the evidence illustrates the possibilities of using variational perturbation techniques to solve non-linear equations. Also since more than one distinct extension of these techniques exist, the various methods may be used to complement each other when possible instabilities in the convergence of the perturbation occurs, since such instabilities are inherent in the method not the problems or overall technique.
CHAPTER 6

Application of the Residual Perturbation Method to the Eigenvalue Problem
6.1 INTRODUCTION.

This chapter is mainly concerned with the application of the method of residual perturbations discussed in chapter 4, to the solution of eigenvalue problems. Equations of this type have previously been solved by the method of variation-iteration which was mentioned in chapter 4 and motivated much of the present work. This method was discussed by Morse and Feshbach 1953, and other authors have applied the basic technique to solve a large number of problems. These authors include: Robinson et al. 1976, Tricomi 1957, Blakemore et al. 1977 and Hyslop 1973. In 1973 Hyslop claimed that this method had many advantages over other methods for solving eigenvalue problems, especially when the problem involved integral operators, and demonstrated this by applying the method to solve a number of problems involving Green's functions. Later Blakemore et al. 1977 applied this technique to the integral equation eigenvalue problems which result from the Hartree-Fock equations for the self consistent field atomic model. His results are compared in the present chapter, with those obtained by using the residual perturbation method.

In 1948, Kantrovich proposed an extension to the variation-iteration technique by applying the basic theory of perturbations used in chapter 4, to solve the eigenvalue problem. The basis of the idea was to use the usual variational ideas to obtain the first approximation to the eigenvalue. Then using the first approximation as the basis of a residual function, he attempted to improve the convergence towards the eigenfunction by introducing a perturbation variable as a factor so scaling the residual function.
Since the approximate eigenfunction was now closer to the actual eigenfunction, the second approximation for the eigenvalue should also be an improvement over the previous method. However, Kantrovich did not consider how the scaling parameter would influence the residual function obtained by the next iteration, therefore making his perturbation non-optimal in the sense considered in chapter 4. But, he was content to prove that his method was an improvement when the scaling parameter was bounded between the largest and smallest eigenvalues. However, such values are not known before the method is applied. In the present chapter, it is proposed to extend the variation-iteration scheme in a similar way to that of Kantrovich. However, in a similar way to that of chapter 4, the scaling parameter is optimized in such a way as to minimize the residual function of the next iteration. Use is also made of the possible half iterate improvement proposed by Hyslop in 1970. The process of approximating the eigenvalue by variational techniques and then applying residual-perturbation to the approximate equation to estimate the eigenfunction is then repeated iteratively, and the examples of Blakemore 1977 are reconsidered and the subsequent improvement noted.

6.2 DESCRIPTION OF THE METHOD.

The variation-iteration method, when applied to integral equation eigenvalue problems, given by:

\[ f(x) = \lambda \int_a^b K(x,y) f(y) \, dy, \quad \text{or} \quad f = \lambda K f \quad (6.2.1) \]
yields an initial approximation to the eigenvalue \( \lambda \), given by
\[
\lambda^{(0)} = \frac{(f_0, f_0)}{(f_0, K f_0)}
\] (6.2.2)

where the inner-products have their usual meaning, and then uses direct iteration to obtain the next approximation to the eigenfunction, use being made of the equation:
\[
f_1(x) = \lambda^{(0)} \int_a^b K(x, y) f_0(y) \, dy.
\] (6.2.3)

The process is then repeated iteratively. In 1970 Hyslop proved that a better half iterate improvement may be obtained to \( \lambda^{(n)} \) which is given by:
\[
\lambda^{(n+\frac{1}{2})} = \frac{(f_n, K f_n)}{(K f_n, f_n)}
\] (6.2.4)

and is easily found at the cost of one quadrature on the square of \( K f_n \). Use is made of this facility as it will naturally accelerate the convergence of any method.

In the present discussion, it is proposed to improve the iteration (6.2.3) above by using the method of residual perturbation considered in chapter 4. To implement this extension the residual function for a eigenvalue problem is defined by:
\[
R_n(x) = \lambda^{(n)} \int_a^b K(x, y) f_n(y) \, dy - f_n(x)
\] (6.2.5)

and similar to the earlier chapter the perturbation takes the form:
\[ \delta_n = \hat{a}_n \mathbf{R}_n, \text{ where } \hat{a}_n \text{ is described by:} \]

\[ \hat{a}_n = \frac{(R_n R_n - \lambda(n) K R_n)}{(R_n - \lambda(n) K R_n, R_n - \lambda(n) K R_n)} \quad (6.2.6) \]

which bears close similarity with the optimal form proposed in chapter 4. Clearly the case \( \hat{a}_n = 1 \) is equivalent to the earlier technique mentioned above. Also since the variation-iteration technique is known to converge, it may be shown that the value of

\[ \frac{(\lambda(n) K R_n, \lambda(n) K R_n)}{(R_n, R_n)} < 1, \]

which implies that the extended method from chapter 4, section 5 may be used to improve the rate of convergence still further. However, it may be seen that each of the present residual-perturbation iterations requires approximately twice the amount of work involved in one iteration of the earlier technique. This results from the work involved in calculating the function \( \lambda(n) K R_n \) across the range of \( x \) values involved in the problem. However, this function is very similar to the \( (n+1) \)th iterate of the eigenfunction approximation in the variation-iteration method which is given by \( \lambda(n+1) K R_n \). Therefore, two iterations of the variation-iteration method are calculated in every iteration of residual-perturbation.
6.3 NUMERICAL EXAMPLES.

6.3.1 Kantrovich; Green's Function Example.

Kantrovich 1948 solved the eigenvalue integral equation problem defined by:

\[ f(x) = \lambda \int_0^1 K(x,y) f(y) \, dy \tag{6.3.1} \]

where

\[
K(x,y) = \begin{cases} 
  x(1-y) & 0 \leq x \leq y \leq 1 \\
  y(1-x) & 0 \leq y \leq x \leq 1 
\end{cases} \tag{6.3.2}
\]

The solution to this problem is well-known to be given by:

\[
\lambda_k = k^2 \pi^2 \quad \text{and} \quad f^{(k)}(x) = \sqrt{2} \sin k \pi x , \quad k = 1,2,3,\ldots \tag{6.3.3}
\]

Table I shows comparable results for the application of the methods of variation-iteration and residual perturbation. It may be noted that the results for variation-iteration correspond to those of half the number of iterations for the present method.

It is clear from Table I that the perturbation method has taken three iterations to obtain an eigenfunction correct to 8 significant figures across the range of \( x \), since \( N[f_3] \approx 10^{-16} \) and \( (f_3,f_3) \) is exact to 8 figures, where the variation-iteration method took 8 iterations to obtain an eigenfunction correct to 8 significant figures. But since each iteration of the residual perturbation method is roughly twice as expensive as one for the variation-iteration method, it may be concluded that the present method displays a slight advantage when applied to this example. However,
<table>
<thead>
<tr>
<th>No. of iterations ( (n) )</th>
<th>( (f_n, f'_n) ) ( \rightarrow N[f_n] )</th>
<th>( \lambda^{n+1} )</th>
<th>No. of iterations ( (n) )</th>
<th>( (f_n, f'_n) ) ( \rightarrow N[f_n] )</th>
<th>( \lambda^{n+1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00000000</td>
<td>1.66666667</td>
<td>10.00000000</td>
<td>0</td>
<td>1.00000000</td>
</tr>
<tr>
<td>1</td>
<td>0.85559503</td>
<td>3.2407614(-6)</td>
<td>9.8696090</td>
<td>2</td>
<td>0.85427690</td>
</tr>
<tr>
<td>2</td>
<td>0.999997251</td>
<td>2.5527179(-13)</td>
<td>9.8696044</td>
<td>4</td>
<td>0.99998965</td>
</tr>
<tr>
<td>3</td>
<td>1.00000000</td>
<td>2.45698675(-16)</td>
<td>9.8696044</td>
<td>6</td>
<td>1.00000007</td>
</tr>
<tr>
<td></td>
<td>(machine accuracy)</td>
<td></td>
<td></td>
<td>8</td>
<td>1.00000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>1.00000000</td>
<td>0.0( \cdot 10^{-16} )</td>
<td>9.8696044</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the eigenvalues were obtained correct to 8 figures after only
2 iterations of the residual perturbation method and required a
comparable amount of work to the earlier method.

Therefore, it may be concluded that this example, is of
little interest in comparing the methods since the standard direct
iteration scheme converges so rapidly. Kantrovich using his
perturbation method presented an analytical result for \( \lambda^{(1)} = 9.876 \),
which is directly comparable with the amount of work expended above
to obtain the values \( \lambda^{(1)} = 9.86961 \) and \( \lambda^{(2)} = 9.86962 \) for each of
the methods respectively, displaying a significant improvement over
his work. The above results for variation-iteration using the half
iterate improvement were originally obtained by Hyslop 1973.

6.3.2 Further Green's Function Example.

As a further case of an integral eigenvalue Green's function
problem the equation given when the kernel in (6.3.1) is given by:

\[
K(x, y) = \begin{cases} 
  x & 0 \leq x \leq y \leq 1 \\
  y & 0 \leq y \leq x \leq 1 
\end{cases}
\]  

(6.3.4)
is now solved, using two trial functions on both the method of
residual perturbations and the method of variation-iteration.

These initial trial functions were given by:

(a) \( f_0(x) = 1 \) and (b) \( f_0(x) = (1+x)^{-2} \)

and the results are presented in Table II.
<table>
<thead>
<tr>
<th>No. of iterations (n)</th>
<th>$(f_n, f'_{n})$</th>
<th>Residual-perturbation $N[f_n]$</th>
<th>$\lambda^{n+1}$</th>
<th>No. of iterations (n)</th>
<th>$(f_n, f'_{n})$</th>
<th>Variation-iteration $N[f_n]$</th>
<th>$\lambda^{n+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$f_0(x) = 1$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00000000</td>
<td>1.66666667</td>
<td>2.50000000</td>
<td>0</td>
<td>1.00000000</td>
<td>1.66666667</td>
<td>2.50000000</td>
</tr>
<tr>
<td>1</td>
<td>0.85559504</td>
<td>3.2407477(-6)</td>
<td>2.4674025</td>
<td>2</td>
<td>0.85427690</td>
<td>1.1514533(-5)</td>
<td>2.4674052</td>
</tr>
<tr>
<td>2</td>
<td>0.99999719</td>
<td>2.5456809(-13)</td>
<td>2.4574011</td>
<td>4</td>
<td>0.99998966</td>
<td>2.0395274(-9)</td>
<td>2.4674011</td>
</tr>
<tr>
<td>3</td>
<td>1.00000000</td>
<td>1.4592627(-16)</td>
<td>2.4674011</td>
<td>6</td>
<td>0.99999993</td>
<td>3.1436544(-13)</td>
<td>2.4674011</td>
</tr>
<tr>
<td></td>
<td>(machine accuracy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>$f_0(x) = (1+x)^{-2}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.29166667</td>
<td>1.4237578(-1)</td>
<td>2.6259188</td>
<td>0</td>
<td>0.29166667</td>
<td>1.4237578(-1)</td>
<td>2.6259188</td>
</tr>
<tr>
<td>1</td>
<td>0.57967622</td>
<td>1.2419327(-5)</td>
<td>2.4674076</td>
<td>2</td>
<td>0.57575838</td>
<td>3.7383789(-5)</td>
<td>2.4674210</td>
</tr>
<tr>
<td>2</td>
<td>0.99998424</td>
<td>6.8712803(-13)</td>
<td>2.4674011</td>
<td>4</td>
<td>0.99995021</td>
<td>9.8240148(-9)</td>
<td>2.4674011</td>
</tr>
<tr>
<td>3</td>
<td>1.00000000</td>
<td>7.7064475(-17)</td>
<td>2.4674011</td>
<td>6</td>
<td>1.00000000</td>
<td>1.4816148(-12)</td>
<td>2.4674011</td>
</tr>
<tr>
<td>Exact</td>
<td>1.00000000</td>
<td>$0.0\cdot 10^{-16}$</td>
<td>$2.4674011 \approx \pi^2/4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II**
The results in Table II illustrate how the residual perturbation method converges more rapidly when the initial trial function differs greatly from the actual solution which is:

\[ f(x) = \sqrt{2} \sin \frac{\pi}{2} x \quad \text{(6.3.5)} \]

Since, in the latter part of the table when \( f_0(x) = (1+x)^{-2} \), the eigenvalue in the 2nd iteration of residual perturbation (when \( n = 1 \)) is clearly more accurate than the comparable result for variation-iteration. Also in the third and fourth iterations (when \( n = 2 \) and 3) the eigenfunctions are clearly more accurate in the present method. However, the above Green’s function also converges very rapidly since it only takes eight iterations to obtain eight figures of accuracy, in both of the cases presented above. Therefore, it may be seen that the method of residual perturbation displays a slight advantage over the earlier method as it converges to an accuracy of 8 figures before the fourth iteration is performed, requiring only about 75% of the computational work of the earlier method.

It is also important to note that the inner-product \( (f_n, f_n) \) may equal 1.0 before every point on \( f_n(x) \) tends to \( f(x) \). So the inner-product \( (f_n, f_n) \) is not an exact measure of convergence. Therefore, both \( \mathbf{N}[f_n] = (R_n, R_n) \) and \( (f_n, f_n) \) must be used to provide a measure of convergence, as is born out by the sixth iteration of the variation-iteration scheme in the latter part of Table II, where \( \mathbf{N}[f_n] \sim 10^{-12} \) indicating that the magnitude of the residual function \( R_6 \) is of order \( 10^{-6} \) and not of \( 10^{-8} \) as may be assumed from \( (f_6, f_6) \). In conclusion, the application of the residual perturbation method does appear to improve the convergence of the variation-iteration.
scheme, though in the present examples the number of iterations involved is too small to come to an adequate appreciation of this improvement. Therefore, the following sections apply the same techniques to more difficult eigenvalue problems which arise in the theory of atomic models.

6.3.3 The Hydrogen Atomic Model.

The model of the electrostatic potential around a hydrogen nucleus is given by the following integral equation:

\[ f(r) = \frac{2\lambda}{r} \int_{0}^{\infty} k(r,r') f(r') \, dr' \quad , \tag{6.3.6} \]

where the kernel \( k(r,r') \) is given by:

\[ k(r,r') = \begin{cases} -r' \sinh(r') \, V(r') \, e^{-r} & 0 \leq r' \leq r \leq \infty \\ -r' \sinh(r) \, V(r) \, e^{-r'} & 0 \leq r \leq r' \leq \infty \end{cases} \tag{6.3.7} \]

The inner-products are now considered as integrals over all three dimensional spherical space using the extended form given by:

\[ (\phi,\psi) = 4\pi \int_{0}^{\infty} \phi(r) \, V(r) \, \psi(r) \, r^2 \, dr \tag{6.3.8} \]

where \( V(r) \) is the radial potential function given by:

\[ V(r) = -\frac{1}{r} \tag{6.3.9} \]

In practical terms this extended form involves a weight function in every inner-product, which in this example may be written:

\[ w(r) = -4\pi r. \tag{6.3.10} \]

The equation was solved starting from two initial
trial functions, which represented the actual solution extremely poorly. These functions are given by:

\[
f_0(x) = \begin{cases} 
1 - \frac{1}{3} x & x \leq 3 \\
0 & x > 3 
\end{cases} 
\tag{6.3.11}
\]

and

\[
f_0(x) = \begin{cases} 
[x] & x \leq 16 \\
0 & x > 16 
\end{cases} 
\tag{6.3.12}
\]

where \([x]\) represents the staircase function of the largest integer less than \(x\). In both cases the range of integration was truncated, to facilitate practical numerical computations. Since the solution is known to be of order \(10^{-12}\) when \(x = 25\), the inclusion of values for \(x > 25\) is insignificant. Therefore, the upper limit of the range was set to be \(x = 25\) and the solution computed for values ranging from zero to twenty-five. The results of these calculations are presented in Table III comparing the residual perturbation and variation-iteration methods.

It is clear from Table III that the residual perturbation scheme, yields an eight figure result with much less work than the variation-iteration method, in both of the cases considered.

In the first case using \(f_0(x)\) from (6.3.11) both functions may be observed to converge towards the solution very rapidly, though the present method is seen to be more effective at obtaining the eigenfunction. However, the trial function though very different in shape from the eigenfunction, predicts a very good first approximation to the eigenvalue which is immediately accurate to two figures.
<table>
<thead>
<tr>
<th>No. of iterations (n)</th>
<th>((f_n^*, f_n^-))</th>
<th>Residual-perturbation</th>
<th>(\lambda^{n+1})</th>
<th>No. of iterations (n)</th>
<th>((f_n^*, f_n^-))</th>
<th>Variation-perturbation</th>
<th>(\lambda^{n+1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9.4221737</td>
<td>0.33845089</td>
<td>1.01648667</td>
<td>0</td>
<td>9.4221737</td>
<td>0.33845089</td>
<td>1.01648667</td>
</tr>
<tr>
<td>2</td>
<td>1.0001247</td>
<td>3.8693369(-7)</td>
<td>1.0000003</td>
<td>4</td>
<td>0.9997734</td>
<td>1.4859135(-5)</td>
<td>1.0000131</td>
</tr>
<tr>
<td>4</td>
<td>1.0000000</td>
<td>6.1741937(-12)</td>
<td>1.0000000</td>
<td>8</td>
<td>1.0000002</td>
<td>4.1185964(-8)</td>
<td>1.0000000</td>
</tr>
<tr>
<td>6</td>
<td>1.0000000</td>
<td>4.5681747(-16)</td>
<td>1.0000000</td>
<td>12</td>
<td>1.0000000</td>
<td>1.7426895(-10)</td>
<td>1.0000000</td>
</tr>
<tr>
<td>(machine accuracy)</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>1.0000000</td>
<td>6.7756086(-13)</td>
<td>1.0000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>1.0000000</td>
<td>2.7510303(-15)</td>
<td>1.0000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td>1.0000000</td>
<td>2.8009326(-16)</td>
<td>1.0000000</td>
</tr>
<tr>
<td>(machine accuracy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III (contd.)
<table>
<thead>
<tr>
<th>No. of iterations ( (n) )</th>
<th>Residual-perturbation ( (f_n, f_n) )</th>
<th>( N[f_n] )</th>
<th>( \lambda^{n+\frac{1}{2}} )</th>
<th>No. of iterations ( (n) )</th>
<th>Variation-iteration ( (f_n, f_n) )</th>
<th>( N[f_n] )</th>
<th>( \lambda^{n+\frac{1}{2}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial Function (6.3.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.9136625(+5)</td>
<td>3.2432159(+4)</td>
<td>5.5790452</td>
<td>0</td>
<td>1.9136625(+5)</td>
<td>3.2422159(+4)</td>
<td>5.5790452</td>
</tr>
<tr>
<td>2</td>
<td>0.68645027</td>
<td>1.4207232(-1)</td>
<td>3.0288896</td>
<td>4</td>
<td>1.0966054</td>
<td>1.5891126(-2)</td>
<td>1.0127788</td>
</tr>
<tr>
<td>4</td>
<td>0.83241885</td>
<td>1.3409468(-1)</td>
<td>1.2883375</td>
<td>8</td>
<td>1.0001637</td>
<td>4.3597913(-5)</td>
<td>1.0000434</td>
</tr>
<tr>
<td>6</td>
<td>1.0167584</td>
<td>3.6349551(-6)</td>
<td>1.0000020</td>
<td>12</td>
<td>1.0000007</td>
<td>1.6858967(-4)</td>
<td>1.0000002</td>
</tr>
<tr>
<td>8</td>
<td>0.99999982</td>
<td>4.82365298(-12)</td>
<td>1.0000000</td>
<td>16</td>
<td>1.0000000</td>
<td>6.5830504(-10)</td>
<td>1.0000000</td>
</tr>
<tr>
<td>10</td>
<td>1.0000000</td>
<td>3.2511006(-16)</td>
<td>1.0000000</td>
<td>20</td>
<td>1.0000000</td>
<td>2.5801117(-12)</td>
<td>1.0000000</td>
</tr>
<tr>
<td></td>
<td>(machine accuracy)</td>
<td></td>
<td></td>
<td>24</td>
<td>1.0000000</td>
<td>9.860920(-15)</td>
<td>1.0000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26</td>
<td>1.0000000</td>
<td>7.7663513(-16)</td>
<td>1.0000000</td>
</tr>
<tr>
<td>Exact</td>
<td>1.000000000</td>
<td>0.0 ( \sim 10^{-16} )</td>
<td>1.00000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE III (Contd.)**
In the second case when \( f_0(x) \) is defined by (6.3.12) the variation-iteration method is clearly seen to converge very much more rapidly even as far as the twelfth iteration. This difference is due to the fact that the first approximation to the eigenvalue is very poor, mainly due to the magnitude of the trial function. In such cases the variation-iteration method is known to do very well as it is an optimal variation for predicting the eigenvalue. However, the residual perturbation technique assumes the eigenvalue to be reasonably accurate and then optimizes the variation to obtain the eigenfunction. Therefore producing very poor convergence when the initial eigenvalue is not a good approximation, since it tries to optimize the solution to what is basically the wrong equation.

However, once the residual perturbation method obtains a reasonable approximation to the eigenvalue, during the fourth iteration, the convergence improves dramatically, and from here on it takes five more iterations to obtain an eight figure result. In comparison the variation-iteration method reduces to direct iteration, once the eigenvalue is reasonably obtained, and therefore, is at a distinct disadvantage. This results in a larger amount of work to obtain the eigenfunctions correct to eight figures. Clearly this example demonstrates some of the advantages and disadvantages of both methods. Though it may be noted that both of the trial functions used above were totally unreasonable approximations to an exponentially decaying curve. In general, therefore, where reasonable trial functions are used the problem of obtaining very poor initial estimates to the eigenvalue is very unlikely. It is very clear, in the above example, that large errors exist initially, and the trial function is very
unsuitable for the problem since the magnitude of $N[f_o]$ is of order $10^5$. The poor convergence is also predicted since the measure of the rate of convergence is approximately one.

### 6.3.4 The Screened Coulomb Field Atomic Model; Zero Energy

In this example the model for the zero energy of a particle in a screened Coulomb field is solved. This model is given by the equation:

\[
f(x) = \lambda \int_{0}^{\infty} k(x, y) e^{-y} f(y) \, dy \tag{6.3.13}
\]

where

\[
k(x, y) = \begin{cases} 
  x & 0 \leq y \leq x \leq \infty \\
  y & 0 \leq x \leq y \leq \infty 
\end{cases} \tag{6.3.14}
\]

This classical eigenvalue problem was solved using an initial trial function $f_o(x) = e^{-x}$, with the range of numerical quadrature truncated to $x = 80$. This value resulted from the smallest exponential the computer used could calculate without experiencing underflow. The inner-products are now defined using the weight function

\[
w(x) = 4\pi x e^{-x} \tag{6.3.15}
\]

a term derived from translating the three dimensional atomic integrals into one dimension and applying the potential function for a screened Coulomb field defined by:

\[
V(x) = -\frac{1}{x} e^{-Bx} \tag{6.3.16}
\]

where the screening parameter $B$ is set as one in the present case.
Comparison is made between the residual perturbation method and the variation-iteration method in Table IV.

Clearly from Table IV the method of residual perturbation converges to machine accuracy in fewer iterations and requiring less work than the method of variation-iteration, when applied to this example. Firstly, it may be noted that the initial trial function is a good approximation to the actual eigenfunction since the value of \( N[f] \sim 10^{-2} \) indicating that the residual function is of order \( 10^{-1} \) across the range of the problem. However, when the actual solution is compared with \( f(x) \) the trial function is found to decay much more rapidly for large values of \( x \). Also the initial eigenvalue \( \lambda^{1/2} \) is a reasonably good approximation providing about two figures of accuracy before any iterations have been carried out. Therefore, the residual perturbation method is optimizing the solution to a fairly accurate approximate equation and would be expected to converge very rapidly. Also as a result of this rapid convergence, the residual perturbation method provides better approximations to the eigenvalue, than the earlier method, where the linear convergence clearly restricts the power of the variational. To conclude this example, it is clear that when a reasonable trial function is used the residual perturbation method is superior in obtaining both the eigenvalue and eigenfunction.

6.3.5 The Screened Coulomb Field Atomic Model; Ground State.

As a further example of comparison between the two methods the eigenvalue integral equation for the ground state of an atomic particle in a screened Coulomb field is now solved. This equation
<table>
<thead>
<tr>
<th>No. of iterations (n)</th>
<th>((f_n, f_n'))</th>
<th>Residual-perturbation (N[f_n])</th>
<th>(\lambda^{n+1})</th>
<th>No. of iterations (n)</th>
<th>((f_n, f_n'))</th>
<th>Variation-iteration (N[f_n])</th>
<th>(\lambda^{n+1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.3962684</td>
<td>6.0067446(-2)</td>
<td>1.7012975</td>
<td>0</td>
<td>1.3962634</td>
<td>6.0067446(-2)</td>
<td>1.7012975</td>
</tr>
<tr>
<td>1</td>
<td>0.98393538</td>
<td>1.35339181(-5)</td>
<td>1.6798151</td>
<td>2</td>
<td>0.97782432</td>
<td>1.4564074(-4)</td>
<td>1.6798952</td>
</tr>
<tr>
<td>2</td>
<td>0.99999626</td>
<td>1.5598518(-9)</td>
<td>1.6798078</td>
<td>4</td>
<td>0.99993809</td>
<td>6.7781960(-7)</td>
<td>1.6798082</td>
</tr>
<tr>
<td>3</td>
<td>1.00000000</td>
<td>4.8135130(-13)</td>
<td>1.6798078</td>
<td>6</td>
<td>0.99999971</td>
<td>3.1228836(-9)</td>
<td>1.6798078</td>
</tr>
<tr>
<td>4</td>
<td>1.00000000</td>
<td>3.0160545(-16)</td>
<td>1.6798078</td>
<td>8</td>
<td>0.99999997</td>
<td>1.4371437(-11)</td>
<td>1.6798078</td>
</tr>
<tr>
<td></td>
<td>(machine accuracy)</td>
<td></td>
<td></td>
<td></td>
<td>(machine accuracy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>1.00000000</td>
<td>0.0 (\sim 10^{-16})</td>
<td>1.6798078</td>
<td></td>
<td>1.00000000</td>
<td>6.4741152(-14)</td>
<td>1.6798078</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.4066414(-16)</td>
<td>1.6798078</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(machine accuracy)</td>
<td></td>
</tr>
</tbody>
</table>
may be written as:

\[ f(x) = \frac{2\lambda}{x} \int_0^\infty k(x, y) e^{-\alpha y} f(y) \, dy \]  

(6.3.17)

where the kernel is given by the equation:

\[ k(x, y) = \begin{cases} 
\sinh(x) e^{-y} & 0 \leq x \leq y \leq \infty \\
\sinh(y) e^{-x} & 0 \leq y \leq x \leq \infty 
\end{cases} \]  

(6.3.18)

The parameter \( \alpha \) in the above equation (6.3.17) is set equal to \( (0.313)^{-1} \) to make this work compatible with that of previous authors. In practice it is found that the solution \( f(x) \) decays so rapidly for large values of \( x \) that the upper limit of the range of integration may be taken at \( x = 20 \), without loss of accuracy to 8 significant figures. Since the change in the screening parameter in (6.3.16) will change the value of the potential function \( V(x) \).

The weight function used in evaluating the inner-products becomes:

\[ w(x) = 4\pi x e^{\frac{x}{0.313}} \]  

(6.3.19)

The results are presented in Table V in a similar way to before for the purpose of comparison.

The exact results quoted in Tables IV and V are due to Blakemore et al. 1977 who used the variation-iteration procedure to calculate them to six figures of accuracy. For the purpose of the above tables, his method has been implemented to obtain the results accurate to eight figures, though the last two figures are not substantiated by any other author. In Table V, it is clearly seen that the residual perturbation method, again, attains 8 figures.
<table>
<thead>
<tr>
<th>No. of iterations ((n))</th>
<th>((f_{n+1}^nf_n))</th>
<th>Residual-perturbation (N[f_n])</th>
<th>(\lambda_{n+1}^n)</th>
<th>No. of iterations ((n))</th>
<th>((f_{n+1}^nf_n))</th>
<th>Variation-perturbation (N[f_n])</th>
<th>(\lambda_{n+1}^n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.6564769</td>
<td>3.4644011(-2)</td>
<td>3.8827900</td>
<td>0</td>
<td>4.6564769</td>
<td>3.4644011(-2)</td>
<td>3.8827900</td>
</tr>
<tr>
<td>1</td>
<td>0.9772658</td>
<td>6.5276399(-5)</td>
<td>3.7850141</td>
<td>2</td>
<td>0.96528652</td>
<td>4.0556771(-4)</td>
<td>3.7855821</td>
</tr>
<tr>
<td>2</td>
<td>0.99999779</td>
<td>3.2753069(-9)</td>
<td>3.7849111</td>
<td>4</td>
<td>0.99986573</td>
<td>3.3001904(-6)</td>
<td>3.7849165</td>
</tr>
<tr>
<td>3</td>
<td>0.99999997</td>
<td>3.9197988(-12)</td>
<td>3.7849111</td>
<td>6</td>
<td>0.99999899</td>
<td>2.6363224(-8)</td>
<td>3.7849112</td>
</tr>
<tr>
<td>4</td>
<td>1.00000000</td>
<td>5.8066748(-16)</td>
<td>3.7849111</td>
<td>8</td>
<td>0.99999997</td>
<td>2.10714137(-10)</td>
<td>3.7849111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(machine accuracy)</td>
<td></td>
<td>10</td>
<td>1.00000000</td>
<td>1.6809663(-12)</td>
<td>3.7849111</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>1.00000000</td>
<td>1.3212952(-14)</td>
<td>3.7849111</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>1.00000000</td>
<td>3.3214361(-16)</td>
<td>3.7849111</td>
</tr>
<tr>
<td><strong>Exact</strong></td>
<td>1.00000000</td>
<td>0.0 (\approx 10^{-16})</td>
<td>3.7849111</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of accuracy in both the eigenvalue and eigenfunction having carried out less work, than is required to achieve the same result using the variation-iteration method. The eigenvalue is seen to converge extremely rapidly by both methods, as in the third row of the table the methods have obtained eigenvalues exact to 8 figures and 5 figures respectively, which, since the original trial function was reasonably close to the solution, born out by the one figure accuracy achieved by \( \lambda^{\frac{1}{4}} \), is as would be expected. Therefore, since the method of residual perturbations contains accurate eigenvalues, as is also expected, the convergence towards the eigenfunction is quadratic in nature, compared to the linear convergence of the earlier method. This is consistent with the original hypothesis of each method, which states that the residual perturbation method optimizes the convergence of the eigenfunction, where variation-iteration optimizes the convergence of the eigenvalue.

6.4 CONCLUSIONS.

In conclusion, it may be noted that in each of the examples shown the residual perturbation method displays some improvement over the method of variation-iteration. Though in some of the earlier examples the improvement does not appear to be very significant. Since in the first three examples, using Green's functions, the improvement is only a matter of two iterations of the variation-iteration method. It must be said in defence of the method of residual perturbation that such problems are known to converge very quickly using variation-iteration since the Neumann series converges to eight figures in only a few terms. Hence the residual
perturbation method can do little to accelerate the convergence as the optimal values of the scaling parameter \( \alpha_n \) are very close to one, so making the residual perturbation iterates very similar to those of variation-iteration. However, in relative terms it may be seen that the residual perturbation method reduces the amount of work involved in calculating the eigenfunction by about 25%.

When the methods were applied to the more difficult atomic problems the differences became more noticeable, especially when consideration is limited to the three examples where the initial approximation to the eigenvalues was very good, (\( \lambda^1 \) having about 1 figure accuracy), where it was seen that the residual perturbation method, required about 33% less work than the method of variation-iteration. As expected the values predicted by \( \alpha_n \) were found to be very different from one demonstrating that the residual perturbation method is now very different from the direct iteration of the variation-iteration method.

Considering the example where the initial approximation to the eigenvalue is very poor. It is clearly seen, in the early iterates, that the variation-iteration method is at a clear advantage. This results from its theoretical background, which enables it to have optimal convergence towards the eigenvalue from very poor approximations. However, once the eigenvalue has been roughly approximated, the method is at a disadvantage over the present technique, since it is hindered by the slower convergence of the eigenfunction. It may be noted that since the residual perturbation technique assumes the eigenvalue possesses a degree of accuracy, it is at an extreme disadvantage in the early iterations where this assumption is invalid. However,
despite this disadvantage, the residual perturbation technique converges in fewer iterations than the earlier method, which clearly demonstrates its power of optimizing the eigenfunction. Though clearly the possibility of this problem being encountered is not desirable when a general technique is sought for, a possible solution to such problems is to monitor the values of $\lambda^n$ and $\lambda^{n+1}$ at each iterate and only once they agree to one figure, to implement the more powerful residual perturbation method. Therefore, making use of the variation-iteration technique to provide good initial approximations, enhancing its power when the initial trial functions may be particularly poor. Such a hybrid is clearly very useful since it uses both techniques at their best. Though in general where reasonable trial functions are used such considerations are unnecessary, and the residual perturbation method is an improvement over the earlier method of variation-iteration.
CHAPTER 7

Conclusions
The purpose of this chapter is to draw together the conclusions of the previous chapters and note possible directions in which further investigations may be made.

Initially in chapter 2, the use of transformations to provide efficient methods of evaluating integrals, where the integrands contain weak singularities, was investigated. This work was divided into two major divisions, namely polynomial transformations, and TANH transformations, and extensive tests were carried out to establish their validity and to carry out comparisons with earlier work. Both techniques were found to be highly effective and ultimately achieved results which compare together very closely and also display considerable improvement over the previous work of Evans, Hyslop and Morgan 1982. In practice the polynomial transformation methods proved to be more complex to implement since the Patterson quadrature required the use of a large number of weights and abscissae, though it proved to be slightly more adept at evaluating functions where the integrands were oscillatory. In contrast the TANH transformations methods are extremely simple to implement, and often take slightly fewer function evaluations to obtain the same accuracy, especially when the integrands are free from any oscillations. It was, therefore, concluded that these methods are a significant step forward in the efficient integration of singular functions. Though clearly further work may be carried out to broaden the present understanding of transformation methods in general. Further advancement will probably be most easily made by reconsideration of the infinite range transformations, since qualitative error estimates are fairly easily obtained for a particular choice of transformation.
The basic conclusions of chapter 3 are fairly clear, since the chapter is included by way of practical illustration of the quadrature methods developed in chapter 2. It is clearly seen in this chapter that these methods are very useful in the extension of the previous work of Morgan 1984 to solve equations possessing singular displacement kernels. The example considered by Cohen 1980, and the non-linear variations thereof, are seen to provide an excellent test of the methods since near singular derivatives are found to exist in the solutions of the equations at the end-points of the ranges of integration. However, it is also seen that the embellished low order piecewise collocation procedure of Morgan converges very slowly, illustrating the need to obtain new procedures which will be more easily applied to such problems.

Chapter 4, forms the basis of just such a procedure, the basic approach is to use variational theory to obtain an iterative scheme which has optimal forced convergence of any iteration, with respect to a chosen form for the perturbation. In general a large number of different forms of this perturbation may be applied to the scheme, though in the present chapter the best form was based on the direct scaling of the residual function, and trivially reduces to the method of successive approximations when this scaling is omitted. This residual perturbation was found to be generally convergent, when applied to linear equations, except when one of two particular cases were approached. These cases are, when the function $R_n - K R_n \equiv 0$, which indicates that no solution exists, as the equation reduces to the eigenvalue problem, and when the two functions $R_n$ and $R_n - K R_n$ are orthogonal. In which case, the method was reduced to that of
successive approximations for one iteration to perturb the iterative sequence and so avoid the apparent difficulty. This method was shown to be very successful on a number of integral equations, and displayed considerable advantages over the direct application of successive approximations. It was also shown to be highly successful at solving the singular displacement kernel problem due to Cohen 1980, which was also used in chapter 3, though in the present context as few as six iterations were found to be necessary to obtain eight figure convergence. However, it was also seen in the Abel singular displacement kernel example (4.8.8) that situations also exist where even these methods take a large number of iterations to converge, though this may be understood when it is realised that the kernel in this example is extremely close to an eigenvalue making the solution particularly hard to find.

Next, in chapter 5, two possibilities are explored to extend this method of residual perturbation to handle non-linear equations. The first uses Fréchet derivatives to obtain a Newtonian approximation to the non-linear equations, and may be considered as an approximation to the Newton-Kantrovich theory. This is the natural extension to the work of chapter 4, and is, therefore, found to possess the same properties as the earlier method. However, it also has the additional requirement for the Newton-Kantrovich sequence of Fréchet derivatives to also converge for all values of $x$ in the range of the problem. The disadvantage with this as a method is that it requires the user to analytically obtain and deliver the 1st Fréchet derivative. This may be difficult in the case of equations involving Urysohn kernels, and in any case requires a basic knowledge of functional analysis.
The second possible extension, is based on the well-known Steffensen acceleration procedure, and therefore, uses finite differences to provide approximations to the Fréchet derivatives. This frees the user from any previous analysis, and makes the solution of equations involving Urysohn kernels simpler. However, this method, referred to as Steffensen perturbation, also suffers from similar possible difficulties in convergence to the previous method, along with the added difficulty that convergence only occurs when the trial function lies in the proximity of the solution. When this proximity condition occurs, this Steffensen perturbation method was also shown to be equally effective at obtaining the solution to a wide variety of integral equations, including the extremely difficult Chandrasekhar equation. It may be noted that a possible improvement to these perturbation methods is to consider more involved forms of the scaling parameter, which was an optimized constant in the present work. More rapid convergence may be achieved when the perturbation has the form:

$$\delta_n = (\alpha_n + \beta_n x) R_n(x)$$

instead of the previous form, $$\delta_n = \alpha_n R_n(x)$$, where both $$\alpha_n$$ and $$\beta_n$$ are optimized with respect to each iteration. Clearly to do this a small matrix of inner products will require solution, and it would be interesting to see how the extra work may influence the convergence. In general, perturbations of the form:

$$\delta_n = \left( \sum_{i=1}^{M} \alpha_{i,n} \phi_i(x) \right) R_n(x)$$

(7.2)
may be considered though the cost may rapidly become prohibitive.

Finally, in chapter 6, the residual perturbations of chapter 4, are used to improve the iterative sequence involved in the well-known variation-iteration scheme for solving eigenvalue problems. In practice, when the initial approximation to the eigenvalue is in reasonable proximity to the exact value, the method is seen to accelerate the convergence of both eigenvalue and eigenfunction very dramatically, reducing the amount of work required by a factor of two. However, when the initial eigenvalue is not in the proximity of the exact value, the improvement is reduced significantly. It is, therefore, suggested as an improvement to the present work, to produce a hybrid method, which may use the optimality of the variation to provide a rough approximation to the eigenvalue (about one figure) in the early iterations and then to change over to the variation-residual perturbation method of chapter 6 to provide rapid convergence to the exact eigenvalue and eigenfunction.

It is, therefore, seen that the present thesis, presents significant advances in the fields of: quadrature involving singular functions, iterative solution of linear and non-linear equations, and improving the variation-iteration technique to solve eigenvalue problems. This work together provides tools and ideas which has made possible the solution of singular non-linear integral equations.
Bibliography

ABRAMOWITZ M. and STEGUN I.A.

Handbook of Mathematical Functions.
(Dover, New York 1965).

ALTMAN M.

On the Approximate Solution of Non-linear Functional Equations.

On the Generalization of Newton's Method.

A Note on the Approximate Solution of Non-linear Operator Equations in a Hilbert Space.

On the Approximate Solution of Non-linear Functional Equations in Banach Spaces.

An Iterative Method of Solving Functional Equations.

ANSELONE P.M. and MOORE R.H.


ATKINSON K.E.

(SIAM, 1976).
BAKER C.T.H.

The Numerical Treatment of Integral Equations.

BLAKEMORE M., EVANS G.A. and HYSLOP J.

A Numerical Integration Scheme for Iterative Calculations on Atomic Systems.

CHISHOLM J.S.R., GENZ A. and ROWLANDS G.E.

Accelerated Convergence of Sequences of Quadrature Approximations.

COHEN A.M.

Cautious Romberg Extrapolation.

CORRADI C.

On the Solution of the Ambartsumian-Chandrasekhar Equation by Monotone Iteration Processes.

DELVES L.M. and ABD-ELAL L.F.

Experience the Integral Equations Package.
University of Liverpool Research Report, CSS/78/8/1, (May 1978).

DELVES L.M., ABD-ELAL L.F. and HENDRY J.A.


A Set of Modules for the Solution of Integral Equations.
EVANS G.A., FORBES R.C. and HYSLOP J.

Polynomial Transformations for Singular Integrals.
International Journal of Computer Mathematics,
(1983), 14, 157-140.

The TANH Transformations for Singular Integrals.
International Journal of Computer Mathematics,
(1984), (awaiting publication).

EVANS G.A., HYSLOP J. and MORGAN A.P.G.

Iterative Solution of Volterra Integral Equations
using Clenshaw-Curtis Quadrature.

The Iterative Solution of a Non-linear Fredholm
Equation Arising from a Chemical Reactor Problem.
International Journal of Computer Mathematics,
(1982), 11, 73-83.

An Extrapolation Procedure for the Evaluation of
Singular Integrals.
International Journal of Computer Mathematics,
(1983), 12, 251-265.

HABER S.

The TANH Rule for Numerical Integration.

HARRIS C.G. and EVANS W.A.B.

Extension of Numerical Quadrature Formulae to Cater
for End-Point Singular Behaviour Over Finite Intervals.
International Journal of Computer Mathematics,
(1977), 6, 219-227.

HYSLOP J.

The Use of Discontinuous Trial Functions in
Variational Calculations.
Loughborough University of Technology: Mathematics
HYSLOP J. and REES D.

An Iterative Scheme for Electronic Systems Using One-Electron Green's Function.

KANTROVICH L.V.

Functional Analysis and Applied Mathematics.

On Some Further Applications of the Newton Approximation Method.
Leningrad University, (1957),

MOORE R.H.

Newton's Method and Variations.
Non-linear Integral Equations (ed. Anselone, P.M.),
(The University of Wisconsin Press, 1964).

MORSE P.M. and FESHBACK H.

Methods of Theoretical Physics.

MORGAN A.P.G.


NOBLE B.

The Numerical Solution of Non-linear Integral Equations.
Non-linear Integral Equations (ed. Anselone, P.M.),
(The University of Wisconsin Press 1964).

PATTERSON T.N.L.

The Optimal Addition of Points to Quadrature Formulae.
PETRYSYN W.V.
On a General Iterative Method for the Approximate Solution of Linear Operator Equations.

RALL L.B.
Numerical Integration and the Solution of Integral Equations by the Use of Riemann Sums.
SIAM Review, (1965), 7, 55-64.

ROBINSON P.D., WARNOCK T.T. and ANDERSON N.

ROBINSON P.D. and EPSTEIN S.T.

SCHELIN C.W.
Monotone Convergence of the SOR-Newton Iterative Technique.

SCHWARTZ C.
Numerical Integration of Analytic Functions.

SQUIRE W.
An Efficient Iterative Method for the Numerical Evaluation of Integrals over a Semi-Infinite Range.

A Quadrature Method for Finite Intervals.

A Comment on Quadrature in the Presence of End-Point Singularities.
STENGER F.


STIBBS D.W.N. and WEIR R.E.


Takahasi H. and Mori M.


TRICOMI F.G.


Wynn P.
