COMPUTER SCIENCES DEPARTMENT
UNIVERSITY OF WISCONSIN
1210 West Dayton Street
Madison, Wisconsin 53706

ON THE COMPUTATION OF RIGOROUS BOUNDS FOR THE SOLUTIONS OF LINEAR INTEGRAL EQUATIONS WITH THE AID OF INTERVAL ARITHMETIC

by

C. W. Cryer

Computer Sciences Technical Report #70

August 1969

Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-31-124-ARO-D-462, and the Office of Naval Research under Contract No.: N 00014-67-A-0128-0004.

ABSTRACT

A method is given for approximately solving linear Fredholm integral equations of the second kind with non-negative kernels. The basis of the method is the construction of piecewise-polynomial degenerate kernels which bound the given kernel. The method is a generalization of a method suggested by Gerberich. When implemented on a computer, interval arithmetic is used so that rigorous bounds for the solution of the integral equations are obtained.

The method is applied to two problems: the equation considered by Gerberich; and the equation of Love which arises in connection with the problem of determining the capacity of a circular plate condenser.

CONTENTS

APPENDICES

§1 Introduction.

We consider the Fredholm integral equation of the second kind,

$$f(x) = s(x) + \int_a^b k(x,y)\, f(y)\, dy\ , \tag{1.1}$$

or, equivalently,

$$f = s + \mathfrak{k} f. \tag{1.2}$$

Here,

$$s(x) \in \mathcal{C}\,(a,b)\ , \tag{1.3}$$

$$k(x,y) \in \mathcal{C}\,([a,b] \times [a,b]), \tag{1.4}$$

and

$$s(x) \geq 0, \quad x \in [a,b]\ , \tag{1.5}$$

$$k(x,y) \geq 0,\ (x,y) \in [a,b] \times [a,b], \tag{1.6}$$

$$\| \mathfrak{k} \| = \int_a^b |k(x,y)|\,dy \leq \rho < 1,\ x \in [a,b]. \tag{1.7}$$

In 1956 Gerberich [4] showed how piecewise-constant functions $f^{(1)}(x)$ and $f^{(2)}(x)$ could be constructed such that

$$f^{(1)}(x) \geq f(x) \geq f^{(2)}(x),\ x \in [a,b], \tag{1.8}$$

where f(x) is the solution of (1.1). To obtain the functions $f^{(1)}$ and $f^{(2)}$ Gerberich approximated the kernel k(x,y) of (1.1) by degenerate piecewise-constant kernels $k^{(1)}$ and $k^{(2)}$. Because of round-off errors, it was of course impossible to compute the functions $f^{(1)}$ and $f^{(2)}$ exactly. Therefore, Gerberich implemented his method so as to obtain computable functions $\tilde{f}^{(1)}$ and $\tilde{f}^{(2)}$ such that

$$\tilde{f}^{(1)}(x) \geq f^{(1)}(x) \geq f(x) \geq f^{(2)}(x) \geq \tilde{f}^{(2)}(x),\ x \in [a,b]. \tag{1.9}$$

There are of course many methods for approximately solving equations of the type (1.1). The method of Gerberich is of interest for two reasons. Firstly, the method of Gerberich can be generalized to provide bounds for the solutions of certain nonlinear integral equations; this has been done by Rall [11] and Brown [1]. Secondly, Gerberich is one of the few workers to take round-off errors into account.

The major disadvantage of the method of Gerberich is its low accuracy, which is a consequence of the fact that the kernel $k(x,y)$ is approximated by piecewise-constant kernels $k^{(1)}$ and $k^{(2)}$. In the present work we generalize the method of Gerberich by approximating the kernel $k(x,y)$ by degenerate piecewise-polynomial kernels $k^{(1)}$ and $k^{(2)}$, thereby obtaining a more accurate method. The resulting approximating functions $f^{(1)}$ and $f^{(2)}$ satisfy (1.8) and are piecewise-polynomial.

Our approach also differs from that of Gerberich in two other respects. Firstly, we use direct methods, rather than iterative methods, to obtain $f^{(1)}$ and $f^{(2)}$. Secondly, we use interval arithmetic (Moore [9]) to compute the functions $\tilde{f}^{(1)}$ and $\tilde{f}^{(2)}$ of (1.9).

The basic theory is extremely simple and is described in sections 2 to 4. The remainder of the paper is devoted to a discussion of the application of the method to two specific problems.

The first problem is the problem treated by Gerberich, and the results illustrate the greater accuracy which results from using higher order approximations to the kernel $k(x,y)$.

The second problem is that of determining the capacity of a circular plate condenser, which, as was shown by Love (see Sneddon [13, p. 230]) is equivalent to solving a certain integral equation. We apply our method to obtain upper and lower bounds for the capacity of the condenser.

Acknowledgement

The author wishes to thank Judith Braun who did much of the programming.

§2 Basic Theory.

Let $\pi = \{z_1, \ldots, z_{n+1}\}$ be a partition of $[a,b]$, that is

$$\text{(i)} \quad a = z_1, \quad b = z_{n+1},$$

$$\text{(ii)} \quad z_i < z_{i+1}, \quad 1 \leq i \leq n.$$

(In the numerical applications we always set $z_i = a + (b-a)(i-1)/n$.)

We set

$$\left.\begin{aligned} I_i &= (z_i, z_{i+1}), \quad I_{ij} = I_i \times I_j, \\ I_x &= \bigcup_{i=1}^{n} I_i, \qquad I_{xy} = \bigcup_{i,j=1}^{n} I_{ij}. \end{aligned}\right\} \tag{2.1}$$

We denote by $\mathcal{C}_\pi(a,b)$ the space of functions which are defined and uniformly continuous on I_x. For example, if

$$\varphi(x) = \begin{cases} 0, & 0 < x < \frac{1}{2}, \\ x^2, & \frac{1}{2} < x < 1, \end{cases} \tag{2.2}$$

and $\pi = \{0, \frac{1}{2}, 1\}$, then $\varphi \in \mathcal{C}_\pi(0,1)$. It will be noted that if $\varphi \in \mathcal{C}_\pi(a,b)$ then φ is not defined at the points $z_i \in \pi$, but that, since φ is uniformly continuous, the limits $\varphi(z_i+0)$ $(1 \leq i \leq n)$ and $\varphi(z_i-0)$ $(2 \leq i \leq n+1)$ exist.

Similarly, we denote by $\mathcal{C}_{\pi}([a,b] \times [a,b])$ the space of functions which are uniformly continuous on I_{xy}.

We shall obtain bounds for the solution f(x) of (1.1) by approximating k(x,y) and s(x) by piecewise-polynomial functions $k^{(p)}(x,y)$ and $s^{(p)}(x)$, p = 1,2. More precisely, let m be a positive integer and let

$$\varphi(i,k;x) = \begin{cases} (x-z_i)^{k-1}, & z_i \leq x \leq z_{i+1}, \\ 0, \text{ otherwise}, & \end{cases} \tag{2.3}$$

$$\text{for } 1 \leq i \leq n,\ 1 \leq k \leq m.$$

If $k^{(p)} \in \mathcal{C}_{\pi}([a,b] \times [a,b])$ and $s^{(p)} \in \mathcal{C}_{\pi}(a,b)$ for p = 1,2, then we shall say that $k^{(p)}$ and $s^{(p)}$ are <u>piecewise-polynomial of degree m-1</u> if

$$k^{(p)}(x,y) = \sum_{i,j=1}^{n} \sum_{k,\ell=1}^{m} q^{(p)}_{ikj\ell}\ \varphi(i,k;x)\ \varphi(j,\ell;y), \tag{2.4}$$

and

$$s^{(p)}(x) = \sum_{i=1}^{n} \sum_{k=1}^{m} v^{(p)}_{ik}\ \varphi(i,k;x), \quad \text{for } p = 1,2, \tag{2.5}$$

where $q^{(p)}_{ikj\ell}$ and $v^{(p)}_{ik}$ are constants. For example, if φ is defined by (2.2) then φ is piecewise-polynomial of degree 2.

It is convenient to introduce matrix notation. If $q^{(p)}_{ikj\ell}$ and $v^{(p)}_{ik}$ are as in (2.4) and (2.5) we introduce the m × m matrices $\underline{q}^{(p)}_{ij} = (q^{(p)}_{ikj\ell})$, the m × 1 matrices $\underline{v}_i = (v_{ik})$, the mn × mn matrices $\underline{q}^{(p)} = (\underline{q}^{(p)}_{ij})$, and the mn × 1 matrices $\underline{v}^{(p)} = (\underline{v}_i)$.

It is necessary to have an explicit representation for the $mn \times mn$ Gram matrix $\underline{g}$ corresponding to the functions $\varphi(i,k;x)$ of (2.3). Let $\underline{g}_i = (g_{ik\ell})$ be $m \times m$ matrices, where

$$g_{ik\ell} = \int_a^b \varphi(i,k;x)\, \varphi(i,\ell;x)dx = (z_{i+1}-z_i)^{k+\ell-1}/(k+\ell-1). \tag{2.6}$$

Then $\underline{g}$ is the block-diagonal matrix,

$$\underline{g} = \text{diag.}\ (\underline{g}_1, \ldots, \underline{g}_n). \tag{2.7}$$

We can now state the basic theorem:

Theorem 2.1

Let $k, \mathfrak{k}$, f, and s, satisfy (1.1) through (1.7)

Let $s^{(p)}$ and $k^{(p)}$ be defined by (2.4) and (2.5) and be such that

$$s^{(1)}(x) \geq s(x) \geq s^{(2)}(x),\ x \in I_x, \tag{2.8}$$

$$k^{(1)}(x,y) \geq k(x,y) \geq k^{(2)}(x,y),\ (x,y) \in I_{xy}, \tag{2.9}$$

$$s^{(2)}(x) \geq 0,\ \ x \in I_x\ , \tag{2.10}$$

$$k^{(2)}(x,y) \geq 0,\ (x,y) \in I_{xy}, \tag{2.11}$$

$$\int_a^b k^{(1)}(x,y)dy \leq \rho_1 < 1,\ \ x \in I_x\ . \tag{2.12}$$

Then, there exist functions $f^{(p)}(x)$ such that

$$f^{(p)}(x) = s^{(p)}(x) + \int_a^b k^{(p)}(x,y) f^{(p)}(y)dy,\ x \in I_x\ , \tag{2.13}$$

or, equivalently,

$$f^{(p)} = s^{(p)} + \mathscr{k}^{(p)} f^{(p)}. \tag{2.14}$$

We have

$$f^{(1)}(x) \geq f(x) \geq f^{(2)}(x), \quad x \in I_x. \tag{2.15}$$

Finally,

$$f^{(p)}(x) = \sum_{i=i}^{n} \sum_{k=1}^{m} u_{ik}^{(p)} \varphi(i,k;x) , \tag{2.16}$$

where if $\underline{u}_i^{(p)}$ is the $m \times 1$ matrix $(u_{ik}^{(p)})$ and $\underline{u}^{(p)}$ is the $mn \times 1$ matrix $(\underline{u}_i^{(p)})$, then

$$\underline{u}^{(p)} = (\underline{h}^{(p)})^{-1} \underline{v}^{(p)}, \tag{2.17}$$

where,

$$\underline{h}^{(p)} = I - \underline{g}^{(p)} \underline{g}. \tag{2.18}$$

<u>Proof</u>: By virtue of(1.7) and (2.12), the Neumann series (Taylor [14, p. 164])

$$\left.\begin{aligned} (I - \mathscr{k})^{-1} &= I + \mathscr{k} + \mathscr{k}^2 + \dots , \\ (I - \mathscr{k}^{(p)})^{-1} &= I + \mathscr{k}^{(p)} + [\mathscr{k}^{(p)}]^2 + \dots , \end{aligned}\right\} \tag{2.19}$$

hold. The inequalities (2.15) follow easily from (2.8), (2.9), and (2.19). Equations (2.16), (2.17), and (2.18), are simply a restatement of wellknown results for integral equations with degenerate kernels (Mikhlin [8, p. 20])

The numerical method which we have used consists of a straight-forward implementation of Theorem 2.1. Two aspects of the implementation require further consideration:

1. It is necessary to develop methods for determining $\underline{v}^{(p)}$ and $\underline{g}^{(p)}$ so that (2.8) through (2.12) hold.

2. It is necessary to make allowance for the effects of round-off errors.

These questions are considered in the next two sections.

§3 Determination of $\underline{v}^{(p)}$ and $\underline{q}^{(p)}$.

The matrices $\underline{v}^{(p)}$ and $\underline{q}^{(p)}$ must be determined so that (noting (2.4) and (2.5)), equations (2.8) through (2.12) hold. We consider first the problem of satisfying (2.8) and (2.9), and then the problem of satisfying (2.10) through (2.12).

The following is a general algorithm for determining $\underline{q}^{(p)}$ and $\underline{v}^{(p)}$ so that (2.8) and (2.9) are satisfied:

Algorithm 3.1

Let $s \in \mathcal{C}^{(m-1)}(a,b)$ and $k \in \mathcal{C}^{(m-1)}([a,b] \times [a,b])$.

Hence, by Taylor's theorem,

$$s(x) = \sum_{k=1}^{m-1} \frac{(x-z_i)^{k-1}}{(k-1)!} \frac{\partial^{k-1} s(z_i)}{\partial x^{k-1}} + \frac{(x-z_i)^{m-1}}{(m-1)!} \frac{\partial^{m-1} s(\xi_i)}{\partial x^{m-1}},$$

for $x \in I_i$, and some $\xi_i \in I_i$,

and,

$$k(x,y) = \sum_{2 \le k+\ell \le m} \frac{(x-z_i)^{k-1}}{(k-1)!} \frac{(y-z_j)^{\ell-1}}{(\ell-1)!} \frac{\partial^{k+\ell-2} k(z_i, z_j)}{\partial x^{k-1} \partial y^{\ell-1}} +$$

$$+ \sum_{k+\ell=m+1} \frac{(x-z_i)^{k-1}}{(k-1)!} \frac{(y-z_j)^{\ell-1}}{(\ell-1)!} \frac{\partial^{m-1} k(\xi_i, \eta_j)}{\partial x^{k-1} \partial y^{\ell-1}},$$

for $(x,y) \in I_{ij}$ and some $(\xi_i, \eta_j) \in I_{ij}$.

Let

$$\left.\begin{aligned} v_{ik}^{(1)} &= v_{ik}^{(2)} = \frac{1}{(k-1)!} \frac{\partial^{k-1} s(z_i)}{\partial x^{k-1}}, \quad 1 < k < m-1, \\ v_{im}^{(1)} &= \frac{1}{(m-1)!} \sup_{x \in I_i} \frac{\partial^{m-1} s(x)}{\partial x^{m-1}}, \\ v_{im}^{(2)} &= \frac{1}{(m-1)!} \inf_{x \in I_i} \frac{\partial^{m-1} s(x)}{\partial x^{m-1}}, \end{aligned}\right\} \quad (3.1)$$

and

$$\left.\begin{aligned} q_{ikj\ell}^{(1)} &= q_{ikj\ell}^{(2)} = \frac{1}{(k-1)!\,(\ell-1)!} \frac{\partial^{k+\ell-2} k(z_i, z_j)}{\partial x^{k-1}\,\partial y^{\ell-1}}, \\ &\quad \text{for } 2 \le k+\ell \le m, \\ q_{ikj\ell}^{(1)} &= \frac{1}{(k-1)!\,(\ell-1)!} \sup_{(x,y) \in I_{ij}} \frac{\partial^{m-1} k(x,y)}{\partial x^{k-1}\,\partial y^{\ell-1}}, \\ q_{ikj\ell}^{(2)} &= \frac{1}{(k-1)!\,(\ell-1)!} \inf_{(x,y) \in I_{ij}} \frac{\partial^{m-1} k(x,y)}{\partial x^{k-1}\,\partial y^{\ell-1}}, \\ &\quad \text{for } k+\ell = m+1, \\ q_{ikj\ell}^{(1)} &= q_{ikj\ell}^{(2)} = 0, \quad m+1 < k+\ell \le 2m. \end{aligned}\right\} \quad (3.2)$$

Then $\underline{v}^{(p)}$ and $\underline{q}^{(p)}$ satisfy (2.8) and (2.9).

It should be noted that if $m = 1$ then (3.1) and (3.2) reduce to the approximations used by Gerberich.

If $k(x,y)$ and $s(x)$ are defined in terms of elementary functions, then bounds for the extrema occurring in (3.1) and (3.2) can be computed automatically with the aid of interval arithmetic (see section 4 for details). Unfortunately, the bounds obtained using interval arithmetic can be grossly pessimistic, and we have therefore preferred to compute the extrema in (3.1) and (3.2) analytically when possible.

There are two disadvantages, one major and one minor, in using Algorithm 3.1.

The major disadvantage is that if $k(x,y)$ is not smooth, and in particular if $k(x,y)$ is the Green's function for a boundary value problem, then m may only take on certain values. For example, in Problem 1 (see section 5) we must take $m = 1$ since $k(x,y)$ is not continuously differentiable.

The minor disadvantage is that the approximations (3.1) and (3.2) are far from being the best possible ones, as is obvious from the fact that $q_{ikj\ell}^{(p)} = 0$ for $m+1 < k+\ell \leq 2m$.

In view of the above-mentioned disadvantages we have not relied entirely upon Algorithm 3.1 and have used alternative methods for obtaining $\underline{q}^{(p)}$ and $\underline{v}^{(p)}$ when appropriate. It should, however, be emphasized that although we have not hesitated to use special methods for finding $\underline{q}^{(p)}$ and $\underline{v}^{(p)}$, these methods use only the most elementary techniques in calculus.

Having determined $\underline{q}^{(p)}$ and $\underline{v}^{(p)}$ so that (2.8) and (2.9) hold, it is easy to check directly whether (2.10), (2.11), and (2.12) hold. In view of (1.5), (1.6), and (1.7), the inequalities (2.10), (2.11), and (2.12), will, in general, hold for large n. We have used both analytical techniques and interval arithmetic to check (2.10), (2.11), and (2.12).

§4 Use of interval arithmetic.

The main purpose in using interval arithmetic is to make allowance for the fact that in computing $f^{(1)}$ and $f^{(2)}$ round-off errors occur. However, we have also used interval arithmetic to check conditions (2.10) through (2.12) and to obtain bounds for the extrema in (3.1) and (3.2).

We recall (Moore [9]) that a rounded interval number A is an ordered pair of machine numbers, $[\tilde{a}_1, \tilde{a}_2]$, where $\tilde{a}_1 \leq \tilde{a}_2$. If $A = [\tilde{a}_1, \tilde{a}_2]$ we set

$$\text{left}(A) = \tilde{a}_1, \quad \text{right}(A) = \tilde{a}_2. \tag{4.1}$$

The interval number $[\tilde{a}_1, \tilde{a}_2]$ may also be regarded as the set of real numbers x such that $\tilde{a}_1 \leq x \leq \tilde{a}_2$.

Arithmetic operations between interval numbers are defined in a natural way. For example, if A and B are rounded interval numbers we say that $C = A + B$ if C is a rounded interval number such that $xy \in C$ if $x \in A$ and $y \in B$.

Given a real-valued function $f(x)$, an interval - valued function $F(X)$ may be defined in a natural way: for any rounded interval X, $F(X)$ is a rounded interval such that if $x \in X$ then $f(x) \in F(X)$.

With the aid of the interval arithmetic package developed by Reiter [12] for the CDC 3600 it is possible to do rounded interval arithmetic within the framework of a FORTRAN program.

As a simple example, let $z = 1/3$ so that $z^2 = 1/9$. The program segment

```
ZSMALL = 1./3.
ZSMALL = ZSMALL * ZSMALL
```

will result in a machine number ZSMALL which is approximately equal to 1/9. On the other hand, suppose that ZINT, ONEINT, and THREEINT, have been declared to be interval numbers. Then the program segment

```
ONEINT = (1., 1.)
THREEINT = (3.,3.)
ZINT = ONEINT/THREEINT
ZINT = ZINT * ZINT
```

will result in a rounded interval number ZINT containing the real number 1/9.

After this brief introduction to interval arithmetic, we can proceed to describe how interval arithmetic is used in the present context.

In the first place, all the "real data", that is a, b, z_i, $s(x)$, $\varphi(i,k;x)$, $k(x,y)$, etc., is replaced by "rounded interval data" A, B, Z_i, $S(X)$, $\Phi(i,k;X)$, $K(X,Y)$, etc. Of course we require that the "interval data" should "include" the "real data" so that, for example $a \in A$, and $s(x) \in S(X)$ if $x \in X$.

Next, using (2.6), an interval-valued array $\underline{G}$ is computed such that $\underline{g} \in \underline{G}$. Similarly, if $\underline{v}^{(p)}$ and $\underline{q}^{(p)}$ are known analytically, interval-valued arrays $\underline{V}^{(p)}$, $\underline{Q}^{(p)}$, and $\underline{H}^{(p)}$ can be computed so that $\underline{v}^{(p)} \in \underline{V}^{(p)}$, $\underline{q}^{(p)} \in \underline{Q}^{(p)}$, and $\underline{h}^{(p)} \in \underline{H}^{(p)}$.

However, $\underline{V}^{(p)}$ and $\underline{Q}^{(p)}$ can also be computed by using interval arithmetic to bound the extrema occurring in (3.1) and (3.2). To illustrate the method, consider the problem of computing the extrema of

$$w(t) = t/(1+t^2)^2$$

for $t \in [-2,+2]$. Using interval arithmetic we find that

$$\begin{aligned} W([-2,+2]) &= [-2,+2]/(1+[0,4])^2 , \\ &= [-2,+2]/[1,25] , \\ &= [-2,+2] , \end{aligned}$$

so that

$$-2 \leq w(t) \leq 2, \quad t \in [-2, +2]. \tag{4.2}$$

Since it is easily shown that

$$-3\sqrt{3}/16 \leq w(t) \leq 3\sqrt{3}/16, \quad t \in [-2,+2],$$

the bounds in (4.2) are grossly pessimistic. We can improve upon (4.2) by sub-dividing the interval $[-2,+2]$. For example we find that

$$\begin{aligned} W([-2,-1]) &= [-.5, -.04], \\ W([-1,0]) &= [-1,0], \\ W([0,1]) &= [0,1], \\ W([1,2]) &= [.04,.5], \end{aligned}$$

so that

$$-1 \leq w(t) < 1, \quad t \in [-2,+2].$$

Before proceeding further it is necessary to check conditions (2.10) through (2.12). This can be done analytically, but is easily done using interval arithmetic. Indeed, noting (2.3), we see that (2.10) through (2.12) are satisfied if

$$\text{left}\left\{\sum_{k=1}^{m} V_{ik}^{(2)} \quad \Phi(i,k;[Z_i,Z_{i+1}])\right\} \geq 0, \text{ for } 1 \leq i \leq n, \tag{4.3}$$

$$\text{left}\left\{\sum_{k=1}^{m} \Phi(i,k;[Z_i,Z_{i+1}]) \sum_{\ell=1}^{m} Q_{ikj\ell}^{(2)} \quad \Phi(j,\ell:[Z_j,\ Z_{j+1}])\right\} \geq 0. \tag{4.4}$$

for $1 \leq i,\ j \leq n$,

and

$$\text{right}\left\{\sum_{j=1}^{n} \sum_{k=1}^{m} \quad \Phi\ (i,k;[Z_i,Z_{i+1}]) \sum_{\ell=1}^{m} Q_{ikj\ell}^{(1)} \quad (Z_{j+1} - Z_j)^{\ell}/\ell\right\} < 1, \tag{4.5}$$

for $1 \leq i \leq n$.

Next, we wish to find $\underline{U}^{(p)}$ such that

$$\underline{U}^{(p)} = (\underline{H}^{(p)})^{-1} \underline{V}^{(p)}. \tag{4.6}$$

To determine $\underline{U}^{(p)}$ we could use a standard elimination method, such as Gaussian elimination, programmed in interval arithmetic, but it seems better to use an a-posteori method based upon the following lemma:

Lemma 4.1

Let $\underline{H} = (H_{rs})$ and $\underline{C}$ be $N \times N$ interval-valued matrices, and $\underline{Y}$ and $\underline{V}$ be $N \times 1$ interval-valued matrices.

Let $\|\cdot\|$ denote the maximum row sum norm so that, for example,

$$\|\underline{H}\| = \max_{1 \leq r \leq N} \sum_{s=1}^{N} |H_{rs}|,$$

where $|H_{rs}| = \max\{|\text{left}\ (H_{rs})|,\ |\text{right}\ (H_{rs})|\}$.

Let $\underline{R} = I - \underline{H}\,\underline{C}$, $P = \underline{H}\,\underline{Y} - \underline{V}$.

If there are machine numbers $\widetilde{\|\underline{C}\|}$, $\widetilde{\|\underline{P}\|}$, $\widetilde{\|\underline{R}\|}$, and $\tilde{\rho}$,

such that

$$\widetilde{\|\underline{P}\|} \geq \|\underline{P}\|, \quad \widetilde{\|\underline{C}\|} \geq \|\underline{C}\|, \quad \|\underline{R}\| \leq \widetilde{\|\underline{R}\|} < 1,$$

and

$$\tilde{\rho} \geq \frac{\widetilde{\|\underline{P}\|}\ \widetilde{\|\underline{C}\|}}{1 - \widetilde{\|\underline{R}\|}},$$

then $\underline{H}^{-1}$ exists and

$$\underline{H}^{-1}\underline{V} = \underline{Y} + \underline{E}$$

where $\underline{E}$ is the $N \times 1$ interval matrix each of whose elements is the interval number $[-\tilde{\rho}, +\tilde{\rho}]$.

Proof:

Lemma 4.1 follows immediately from Theorem 8 in Isaacson and Keller [6, p. 48].

In determining $\underline{U}^{(p)}$ using Lemma 4.1, approximate inverses $\underline{C}^{(p)}$ and approximate solutions $\underline{Y}^{(p)}$ are needed. To obtain $\underline{C}^{(p)}$ and $\underline{Y}^{(p)}$ we choose machine-number matrices $\underline{\tilde{h}}^{(p)}$ and $\underline{\tilde{v}}^{(p)}$ such that $\underline{\tilde{h}}^{(p)} \in \underline{H}^{(p)}$ and $\underline{\tilde{v}}^{(p)} \in \underline{V}^{(p)}$. Then, using standard single-precision matrix subroutines we compute the approximate inverses $\underline{\tilde{c}}^{(p)} \approx (\underline{\tilde{h}}^{(p)})^{-1}$ and the approximate solutions $\underline{\tilde{y}}^{(p)} \approx (\underline{\tilde{h}}^{(p)})^{-1}\underline{\tilde{v}}^{(p)}$, and set $\underline{C}^{(p)} = \underline{\tilde{c}}^{(p)}$ and $\underline{Y}^{(p)} = \underline{\tilde{y}}^{(p)}$. Lemma 4.1 can now be used; of course the machine numbers $\widetilde{\|\underline{H}^{(p)}\|}$ etc. are computed using interval arithmetic.

It should be remarked that the above method of computing $\underline{U}^{(p)}$ is closely related to the method of Hansen (Moore [9, p. 32]).

The above method of computing $\underline{U}^{(p)} = (U_{ik}^{(p)})$ has been found to be very satisfactory. Typically, [right $(U_{ik}^{(p)})$ - left $(U_{ik}^{(p)})$] was less than 10^{-9}. It should be borne in mind that the computations were performed on the CDC-3600 which uses twelve-figure floating point arithmetic.

Having determined $\underline{U}^{(p)} = (U_{ik}^{(p)})$ we set

$$F^{(p)}(X) = \sum_{i=1}^{n} \sum_{k=1}^{m} U_{ik}^{(p)} \Phi(i,k;X), \; X \subset I_x. \tag{4.7}$$

By construction,

$$f^{(p)}(x) \in F^{(p)}(X), \text{ if } x \in X \subset I_x. \tag{4.8}$$

so that we may set

$$\left.\begin{aligned} \tilde{f}^{(1)}(x) &= \text{right } (F^{(1)}(X)), \\ \tilde{f}^{(2)}(x) &= \text{left } (F^{(2)}(X)), \text{ for } x \in X \subset I_x. \end{aligned}\right\} \tag{4.9}$$

Finally, we set

$$\left.\begin{aligned} F_{i+o}^{(p)} &= \sum_{k=1}^{m} U_{ik}^{(p)} \; \Phi(i,k;Z_i), \quad 1 \leq i \leq n, \\ F_{i-o}^{(p)} &= \sum_{k=1}^{m} U_{(i-1),k}^{(p)} \; \Phi(i-1,k;Z_i), \quad 2 \leq i \leq n+1. \end{aligned}\right\} \tag{4.10}$$

Remembering that $f \in \mathcal{C}(a,b)$ we have that

$$f(z_i) \leq \begin{cases} \text{right } (F_{1+o}^{(1)}), \text{if } i = 1, \\ \text{right } (F_{n+1-o}^{(1)}), \text{ if } i = n+1, \\ \min \{\text{right } (F_{i+o}^{(1)}, \text{ right } (F_{i-o}^{(1)}) \}, \text{ otherwise}, \end{cases} \tag{4.11}$$

$$f(z_i) \geq \begin{cases} \text{left}\,(F^{(1)}_{1+o}), & \text{if } i = 1, \\ \text{left}\,(F^{(1)}_{n+1-o}), & \text{if } i = n+1, \\ \min\{\text{left}\,(F^{(1)}_{i+o}),\ \text{left}\,(F^{(1)}_{i-o})\}, & \text{otherwise.} \end{cases} \tag{4.12}$$

§5 Problem 1.

As our first problem we consider the equation

$$f(x) = s(x) + \int_0^1 k(x,y)\, f(y)\, dy, \tag{5.1}$$

$$k(x,y) = \begin{cases} x(1-y), & 0 \leq x \leq y \leq 1, \\ y(1-x), & 0 \leq y \leq x \leq 1, \end{cases} \tag{5.2}$$

$$s(x) = x^2. \tag{5.3}$$

This equation was considered by Rall [10] and Gerberich [4].

It can be shown (see Appendix A) that the solution of (5.1) is given by,

$$f(x) = x^2 + x(1-x^3)/12 +$$

$$+ 2 \sum_{k=1}^{\infty} \left\{ \frac{\sin(k\pi x)(-1)^{k-1}}{(k\pi)^3\ (k^2\pi^2 - 1)} - \frac{4\sin([2k-1]\pi x)}{([2k-1]\pi)^5\ ([\{2k-1\}\pi]^2 - 1)} \right\}. \tag{5.4}$$

For $m = 1$ Algorithm 3.1 was used (the extrema in (3.1) and (3.2) being derived analytically). For $m = 2$ Algorithm 3.1 could not be used since $k(x,y)$ is not continuously differentiable. Therefore $\underline{g}^{(p)}$ and $\underline{v}^{(p)}$ were derived analytically by taking advantage of the fact that $s(x)$ is convex and that $k(x,y)$ is semi-linear (Cryer [3]). The analysis is given in Appendix C.

Numerical results are given in Table 5.1. It is of interest to note that, summing the first 100 terms of (5.4) using interval arithmetic and estimating the truncation error, it was found that $f(\frac{1}{2}) \in [2.907591, 2.907592]$.

n	Lower Bounds		Upper Bounds	
	m = 1	m = 2	m = 2	m = 1
2	.2499	.2784	.3076	.4584
4	.2591	.2876	.2954	.3549
8	.2710	.2899	.2921	.3188
16	.2797	.2905	.2912	.3039
32	.2849	.2907	.2909	.2972

Table 5.1.

Bounds for $f(\frac{1}{2})$ for Problem 1.

Finally, it should be remarked that the program was used to compute bounds for the case m = 1, n = 10, since results for this case had been given by Gerberich ([4, Figure 10]). Our bounds are consistently worse than those of Gerberich. For example, Gerberich finds that

$$f(x) \leq .2 \quad \text{for} \quad .3 < x < .4 \ ,$$

and

$$f(x) \leq .3 \quad \text{for} \quad .4 < x < .5 \ ,$$

while we find that

$$f(x) \leq .2136 \quad \text{for} \quad .3 < x < .4 \,,$$

and

$$f(x) \leq .3126 \quad \text{for} \quad .4 < x < .5 \,.$$

We find this puzzling since we believe that our bounds should always be slightly sharper than those of Gerberich because of our use of direct methods, as opposed to iterative methods, for computing $f^{(p)}(x)$. Of course, round-off errors could make our bounds worse but this is not the case in the present instance since the interval arithmetic computations show that

$$f^{(1)}(x) \in [.2135, .2136] \quad \text{for} \quad .3 < x < .4 \,,$$

and

$$f^{(1)}(x) \in [.3125, .3126] \quad \text{for} \quad .4 < x < .5 \,.$$

§6 Problem 2.

As our second problem we consider the equation

$$f(x) = 1 + \frac{\kappa}{\pi} \int_{-1}^{+1} \frac{f(y)}{(x-y)^2 + \kappa^2} \, dy. \tag{6.1}$$

Equation (6.1), which was derived by Love (see Sneddon [13, p. 230]) occurs in connection with the determination of the capacity of a circular plate condenser. Of particular interest is the quantity

$$\gamma(\kappa) = \frac{1}{2} \int_{-1}^{+1} f(x) \, dx, \tag{6.2}$$

which is proportional to the capacity of the condenser.

If $f^{(1)}$ and $f^{(2)}$ satisfy (1.8) then,

$$\int_a^b f^{(1)}(x)\,dx \geq \int_a^b f(x)\,dx \geq \int_a^b f^{(2)}(x)\,dx. \tag{6.3}$$

Hence, having computed $F^{(1)}$ and $F^{(2)}$ (see 4.7) we may compute the following bounds for $\gamma(\kappa)$:

$$\begin{aligned} &\text{right}\left\{ \sum_{i=1}^{n} \sum_{k=1}^{m} U_{ik}^{(1)} \; (Z_{i+1} - Z_i)^k / k \right\} > 2\,\gamma(\kappa) \\ &\geq \text{left}\left\{ \sum_{i=1}^{n} \sum_{k=1}^{m} U_{ik}^{(2)} \; (Z_{i+1} - Z_i)^k / k \right\}. \end{aligned} \tag{6.4}$$

Since $s(x) \equiv 1$ for equation (6.1),

$$v_{ik}^{(p)} = \begin{cases} 1, & \text{if } k = 1, \\ 0, & \text{otherwise.} \end{cases}$$

The arrays $\underline{q}^{(p)}$ were computed in three ways:

(a) Using Algorithm 3.1 with $m = 1$, the extrema in (3.1) and (3.2) being derived analytically. Details are given in Appendix D.

(b) Using a special method for $m = 2$ which took advantage of the fact that $k(x,y)$ is a function of $(x-y)$. The analysis is given in Appendix D.

(c) Using Algorithm 3.1 with $m = 1$, 2, and 3, the extrema in (3.1) and (3.2) being bounded automatically with the use of interval arithmetic.

Numerical results are presented in Tables 6.1 to 6.3. In these tables, an asterisk indicates that inequalities (4.3) to (4.5) were not satisfied.

n	Lower Bounds		Upper Bounds	
	m = 1	m = 2	m = 2	m = 1
1	1.145	1.145	2.752	2.752
2	1.286	*	*	2.752
4	1.453	1.690	1.887	2.308
8	1.605	1.786	1.838	2.064
16	1.706	1.812	1.825	1.942
32	1.762	1.818	1.822	1.881

Table 6.1.

Bounds for $\gamma(1)$ (using methods (a) and (b)).

n	Lower Bounds		Upper Bounds	
	m = 1	m = 2	m = 2	m = 1
1	*	*	*	*
2	*	*	*	*
4	*	*	*	*
8	*	2.587	3.714	*
16	2.136	2.909	3.271	5.679
32	2.515	3.056	3.140	4.031

Table 6.2.

Bounds for $\gamma(.4)$ (using methods (a) and (b)).

n	Lower Bounds			Upper Bounds		
	m = 1	m = 2	m = 3	m = 3	m = 2	m = 1
1	1.145	*	*	*	*	2.752
2	1.286	*	*	*	*	2.752
4	1.453	*	*	*	*	2.308
8	1.605	1.614	1.728	1.935	2.106	2.064
16	1.706	1.763	1.808	1.833	1.885	1.942
32	1.762	1.806	not run	not run	1.837	1.881

Table 6.3

Bounds for $\gamma(1)$ (using method (c)).

Comparing Tables 6.1 and 6.3 we see, as is to be expected, that it is better to use approximations $k^{(p)}$ specially tailored to the problem rather than approximations $k^{(p)}$ obtained automatically from (3.2). Nevertheless, reasonable bounds are obtained using the automated Algorithm 3.1 provided that sufficiently large values of n are used.

An approximate expression for $\gamma(\kappa)$ was derived by Maxwell in 1866, since when $\gamma(\kappa)$ has been estimated in many ways. Sneddon [13, p. 230] summarizes this work and quotes, among others, the following estimates:

$\gamma(1)$: 1.8208 (truncated series),

1.8138 (variational method).

$\gamma(.4)$: 3.1029 (truncated series),

3.0023 (variational method),

3.0846 (analytic method due to Maxwell),

3.1044 (analytic method due to Kirchhoff).

Remembering that the bounds in Tables 6.1 to 6.3 are rigorous, we see that for Love's equation the method described in the present paper is fully competitive with analytical techniques.

§7 Concluding remarks

The method which we have presented is substantially more accurate than the method of Gerberich, as the results of sections 5 and 6 show.

Interval arithmetic is too costly in terms of machine time for general computation. However, section 6 shows that when rigorous bounds are required, methods using interval arithmetic can compare favorably with classical analytical methods.

APPENDIX A

Analytic solution of Problem 1.

We consider the equation $f(x) = x^2 + \lambda \int_0^1 k(x,y)\, f(y)\, dy,$ (A. 1)

$$k(x,y) = \begin{cases} x(1-y), & 0 \leq x \leq y \leq 1, \\ y(1-x), & 0 \leq y \leq x \leq 1. \end{cases} \qquad \text{(A. 2)}$$

Now, it is easily shown that

$$\int_0^1 k(x,y) \sin(\pi j y)\, dy = \frac{1}{\pi^2 j^2} \sin(\pi j x),\ j = 1, 2, \ldots, \qquad \text{(A. 3)}$$

and (Byerly [2, p. 41]),

$$x^2 = 2 \left\{ \sum_{k=1}^{\infty} \frac{\sin(k\pi x)(-1)^{k-1}}{k\pi} - 4 \sum_{k=0}^{\infty} \frac{\sin[(2k+1)\pi x]}{[(2k+1)\pi]^3} \right\}, \qquad \text{(A. 4)}$$

for $0 \leq x \leq 1$.

We can solve (A. 1) by expressing f(x) as a Fourier sine series and using (A. 1), (A. 3), and (A. 4), to determine the Fourier coefficients of f(x). However, the resulting Fourier series for f(x) is slowly convergent and unsuitable for numerical computation. The reason for this difficulty is that the Fourier series (A. 4) is the Fourier series for the saw-tooth function s(x) shown in Figure A. 1,

$$s(x) = \begin{cases} x^2, & 0 \leq x \leq 1, \\ -x^2, & -1 < x < 0. \end{cases}$$

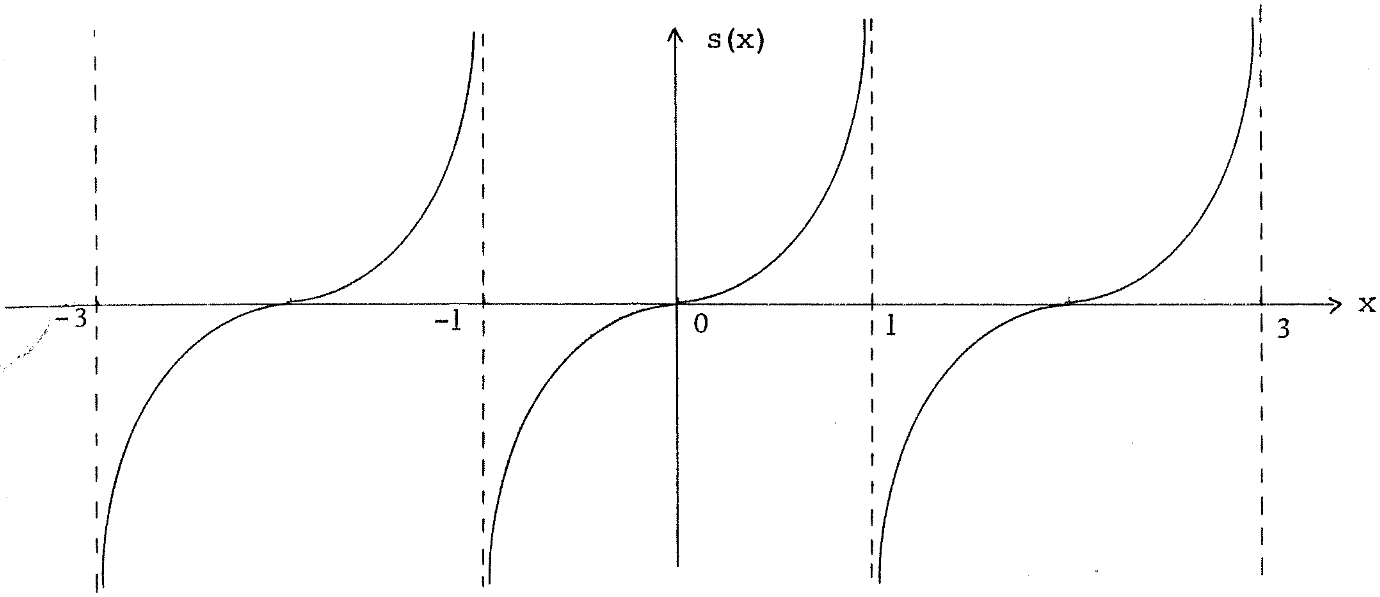

Figure A. 1.

The function s(x).

Since s(x) is discontinuous, the Fourier series (A. 4), and, consequently, the Fourier series for f(x), are slowly convergent.

To obtain a more rapidly convergent Fourier series for f(x) we proceed as follows. Setting

$$f(x) = x^2 + f_1(x) \,, \tag{A. 5}$$

and substituting in (A. 1) it is found that

$$f_1(x) = \lambda\, s_1(x) + \lambda \int_0^1 k(x,y)\, f_1(y)\, dy. \tag{A. 6}$$

where

$$\left. \begin{aligned} s_1(x) &= \int_0^1 k(x,y)\, y^2\, dy, \\ &= x(1-x^3)/\,12 \,, \quad 0 \le x \le 1 \,. \end{aligned} \right\} \tag{A. 7}$$

Setting

$$f_1(x) = \lambda s_1(x) + f_2(x) , \qquad (A. 8)$$

and substituting in (A. 6) it is found that

$$f_2(x) = \lambda^2 s_2(x) + \lambda \int_0^1 k(x,y) f_2(y) dy. \qquad (A. 9)$$

where

$$\left. \begin{aligned} s_2(x) &= \int_0^1 k(x,y) s_1(y) dy , \\ &= \frac{x(4-5x^2+x^5)}{360} , \quad 0 \le x \le 1 . \end{aligned} \right\} \qquad (A. 10)$$

From (A. 3), (A. 4), (A. 7), and (A. 10), it follows that

$$s_2(x) = 2 \left\{ \sum_{k=1}^{\infty} \frac{\sin(k\pi x)(-1)^{k-1}}{(k\pi)^5} - 4 \sum_{k=0}^{\infty} \frac{\sin[(2k+1)\pi x]}{[(2k+1)\pi]^7} \right\} . \qquad (A. 11)$$

We observe that the Fourier series (A.11) is far more rapidly convergent than the Fourier series (A. 4).

Let

$$\left. \begin{aligned} s_2(x) &= \sum_{k=1}^{\infty} c_k \sin(\pi k x), \\ f_2(x) &= \sum_{k=1}^{\infty} e_k \sin(\pi k x). \end{aligned} \right\} \qquad (A. 12)$$

Noting (A. 9) and (A. 3), it follows that

$$e_k = \lambda^2 \frac{\pi^2 k^2}{(\pi^2 k^2 - \lambda)} c_k . \tag{A. 13}$$

Combining (A. 5), (A. 8), (A. 11), (A. 12), and (A. 13), we obtain the desired result:

$$f(x) = x^2 + \lambda\, x(1-x^3)/12 +$$

$$+ 2\lambda^2 \left\{ \sum_{k=1}^{\infty} \frac{\sin(k\pi x)(-1)^{k-1}}{(k\pi)^3 (k^2\pi^2 - \lambda)} - 4 \sum_{k=1}^{\infty} \frac{\sin([2k-1]\pi x)}{[(2k-1)\pi]^5 \{[(2k-1)\pi]^2 - \lambda\}} \right\} . \tag{A. 14}$$

The series (A. 14) is quite suitable for numerical compution. For example, if $\lambda = 1$ and $f_{100}(x)$ denotes the value obtained by summing only the first 100 terms in (A. 14), then

$$|f(x) - f_{100}(x)|$$

$$\le 2 \sum_{k=101}^{\infty} \left[\frac{1}{(k\pi)^3 (k^2\pi^2 - 1)} + \frac{4}{[(2k-1)\pi]^5 [(2k-1)^2\pi^2 - 1]} \right] ,$$

$$\le 2 \sum_{k=101}^{\infty} \left[\frac{11/10}{(k\pi)^5} + \frac{4}{[(2k-2)\pi]^7} \right] ,$$

$$\le 2 \left[\frac{11}{10} \int_{t=100}^{\infty} \frac{dt}{(t\pi)^5} + \frac{1}{32} \int_{t=99}^{\infty} \frac{dt}{(t\pi)^7} \right] ,$$

$$= 2 \left[\frac{11/10}{6 . \pi^5 . (100)^6} + \frac{1}{32 . \pi^7 . 8 . (99)^8} \right] ,$$

$$< 10^{-8} .$$

APPENDIX B

Some remarks on the implementation of interval arithmetic.

The interval arithmetic computations in this report were performed using the interval arithmetic package INTERVAL developed for the CDC 3600 by Reiter [12]. In this appendix we make some general remarks arising from our experiences using INTERVAL. These remarks in no way constitute a criticism of INTERVAL which was developed as an experiment and which we have found to be very useful.

The most important features of INTERVAL are:

1. Programs using interval arithmetic are written in FORTRAN and compiled by the usual FORTRAN compiler.
2. INTERVAL uses the TYPE-other feature of the CDC 3600 FORTRAN compiler. The user designates interval-valued valued variables to be of TYPE INT5(2); the compiler then assigns two storage locations to each such variable.
3. The compiler arranges for arithmetic operations of the form A r B, where A and B are interval-valued variables, and r is an arithmetic operator, to be evaluated by the appropriate subroutine in INTERVAL.
4. Mixed-mode arithmetic is permitted: INTERVAL contains subroutines for converting real and integer variables into interval-valued variables.
5. INTERVAL has special subroutines, such as LOGINT, for evaluating elementary transcendental functions.

For example, the following FUNCTION subprogram computes $x^2 + 1/3$ in interval arithmetic:

```
    FUNCTION F(X)
    TYPE INT5(2) X, Z1, Z3, F, F1, F2, F3, F4
    Z1 = 1.
    Z3 = 3.
101 F1 = (1./3.) + X * X
102 F2 = Z1/Z3   + X * X
103 F3 = Z1/Z3   + X ** 2
104 F4 = Z1/Z3   + X ** 2.0
    F  = F4
    RETURN
    END
```

The above program illustrates certain idiosyncracies of INTERVAL:

(α) Since mixed-mode arithmetic statements are legal, statement 101 is correct. However, the quantity 1./3. is first evaluated in REAL arithmetic and then converted to an interval number.

(β) X * X and X ** 2 are both computed by direct multiplication so that F2 = F3. However, if $X = \{x:\tilde{x}_1 \leq x \leq \tilde{x}_2\}$ then $X * X = \{xy:\tilde{x}_1 \leq x, y \leq \tilde{x}_2\}$. For example, if X = [-1,+1] then X * X = X ** 2 = [-1,+1]. On the other hand, X ** 2.0 is evaluated by special subroutines in INTERVAL. In effect, $X ** 2.0 = \{x^2:\tilde{x}_1 \leq x \leq \tilde{x}_2\}$, so that for example, if X = [-1,+1] then X ** 2.0 = [0, 1].

Thus the program computes three different results, F1, F2 = F3, and F4, of which F4 is probably the desired result.

A further idiosyncracy of INTERVAL is:

(γ) Statements containing "relational expressions" involving interval-valued variables are accepted by the compiler, but are not correctly interpreted. For example, if $A = [\tilde{a}_1, \tilde{a}_2]$ and $B = [\tilde{b}_1, \tilde{b}_2]$, then, according to the definition of Moore [9, p. 7], $A \leq B$ iff $\tilde{a}_2 \leq \tilde{b}_1$. However, the compiler interprets the statement

IF (A. LE. B) GO TO 100

to mean

IF ($\tilde{a}_1$.LE.$\tilde{b}_1$) GO TO 100.

The purpose of using interval arithmetic is to ensure that the final results are rigorously correct. This requires that the following conditions be satisfied:

1. The interval arithmetic package should be correctly coded.
2. The computer should be working correctly and in accordance with the descriptions in the manufacturer's manuals.
3. The compiler should be error-free.
4. The program should be correctly coded.

Unfortunately it is impossible to guarantee even one of the above conditions. For

1. While developing the program used in the present report, two minor mistakes in the input/output subroutines of INTERVAL were found.

2. Of computers on which the author has worked, at least two have had errors in their circuits which remained undetected for over a year.

3. It is quite common for errors to be found in compilers which have been in use for several years. In addition, in order to optimize the object code, compilers sometimes introduce operations which are not implied by the FORTRAN code; a particular case of this was encountered while developing the program used in this report.

4. It is impossible to ensure that a program is fully debugged.

Incorrect proofs are of course not unknown in mathematics; for example, Lefschetz [7, p. 7] refers to the large number of incorrect proofs of the Jordan curve theorem. However, it is possible for mathematicians to arrive at a consensus of opinion on the correctness or otherwise of a mathematical proof. It does not seem that it will be possible to ensure similar certainty for results obtained using interval arithmetic.

Nevertheless, it seems to the author that the goal of certainty is worth striving for, even if it is not attainable. It is in accordance with this view that the following suggestions are made:

1. When possible, the techniques of "program-proving" should be used. (See Good and London [5]).

2. Since most mistakes in compilers arise because of attempts to optimize the resulting code, it is desirable for the user to have the option of using a simple compiler which has no optimization features.

3. Expressions such as X ** 2 which are not well-defined (see (β) and (γ)) should not be permitted.

4. Extensive debugging facilities should be provided. For example, the user's attention should be drawn to mixed-mode expressions (see (α)).

APPENDIX C

$\underline{v}^{(p)}$ and $\underline{q}^{(p)}$ for Problem 1.

m = 1

We use Algorithm 3.1.

Since s(x) is convex, $v_{i1}^{(1)} = s(z_{i+1})$, $v_{i1}^{(2)} = s(z_i)$.

The kernel k(x,y) has been considered previously (Cryer [3, Problem 6]); using these results, and noting that because of symmetry $q_{i1j1}^{(p)} = q_{j1i1}^{(p)}$, we find that

$$q_{i1j1}^{(1)} = k(z_i, z_{j+1}), \quad \text{if } j < i ,$$

$$q_{i1i1}^{(1)} = \begin{cases} k(\tfrac{1}{2}, \tfrac{1}{2}) = 1/4, & \text{if } \tfrac{1}{2} \in [z_i, z_{i+1}], \\ \max \{k(z_i, z_i), k(z_{i+1}, z_{i+1})\}, & \text{otherwise.} \end{cases}$$

$$q_{i1j1}^{(2)} = k(z_{i+1}, z_j) , \quad j \leq i.$$

m = 2

Since s(x) is convex, for $x \in [z_i, z_{i+1}]$ s(x) lies below the chord connecting z_i and z_{i+1} and above the tangent at z_i. Hence, we set

$$v_{i1}^{(1)} = s(z_i), \; v_{i2}^{(1)} = [s(z_{i+1}) - s(z_i)]/[z_{i+1} - z_i],$$

$$v_{i1}^{(2)} = s(z_i), \; v_{i2}^{(2)} = s'(z_i) .$$

To obtain $\underline{q}_{ij}^{(p)}$ we first note that, because of symmetry, we need only consider the case $j \leq i$.

If $j < i$ then we can choose $\underline{q}_{ij}^{(p)}$ so that $k^{(p)}(x,y) = k(x,y)$ for $(x,y) \in I_{ij}$. That is, we set

$$q^{(1)}_{i1j1} = q^{(2)}_{i1j1} = k(z_i, z_j) ,$$

$$q^{(1)}_{i1j2} = q^{(2)}_{i1j2} = k_y(z_i, z_j) ,$$

$$q^{(1)}_{i2j1} = q^{(2)}_{i2j1} = k_x(z_i, z_j) ,$$

$$q^{(1)}_{i2j2} = q^{(2)}_{i2j2} = k_{xy}(z_i, z_j).$$

Finally, to obtain $\underline{q}^{(p)}_{ii}$ we note that $k(x,y)$ is semi-concave (Cryer [3]). Hence, if $\underline{q}^{(1)}_{ii}$ is chosen so that $k^{(1)}(x,y)$ represents the tangent plane to $k(x,y)$ at the midpoint $(z_{i+\frac{1}{2}}, z_{i+\frac{1}{2}})$, $z_{i+\frac{1}{2}} = (z_i + z_{i+1})/2$, then $k^{(1)}(x,y) \geq k(x,y)$ for $(x,y) \in I_{ii}$. Also, if $\underline{q}^{(2)}_{ii}$ is chosen so that $k^{(2)}(x,y) = k(x,y)$ for $(x,y) =$ (z_i, z_i), (z_i, z_{i+1}), (z_{i+1}, z_i), and (z_{i+1}, z_{i+1}), then $k^{(2)}(x,y) \leq k(x,y)$ for $(x,y) \in I_{ii}$.

The detailed formulas for $\underline{q}^{(p)}_{ii}$ are:

$$q^{(1)}_{i1i1} = w(z_{i+\frac{1}{2}}) - (z_{i+\frac{1}{2}} - z_i)\, w'(z_{i+\frac{1}{2}}) ,$$

$$q^{(1)}_{i1i2} = q^{(1)}_{i2i1} = w'(z_{i+\frac{1}{2}})/2 ,$$

$$q^{(1)}_{i2i2} = 0,$$

where $z_{i+\frac{1}{2}} = (z_i + z_{i+1})/2$ and $w(t) = t - t^2$.

$$q^{(2)}_{i1i1} = k(z_i, z_i) ,$$

$$q^{(2)}_{i1i2} = q^{(2)}_{i2i1} = [k(z_{i+1}, z_i) - k(z_i, z_i)]/[z_{i+1} - z_i],$$

$$q^{(2)}_{i2i2} = \frac{k(z_{i+1}, z_{i+1}) + k(z_i, z_i) - k(z_i, z_{i+1}) - k(z_{i+1}, z_i)}{(z_{i+1} - z_i)^2} .$$

APPENDIX D

$g^{(p)}$ for Problem 2.

It is clearly only necessary to obtain approximations for the function $p(x-y)$, where

$$p(t) = \frac{1}{\kappa^2 + t^2} . \tag{D. 1}$$

The function $p(t)$ has the bell-shaped form shown in Figure D.1.

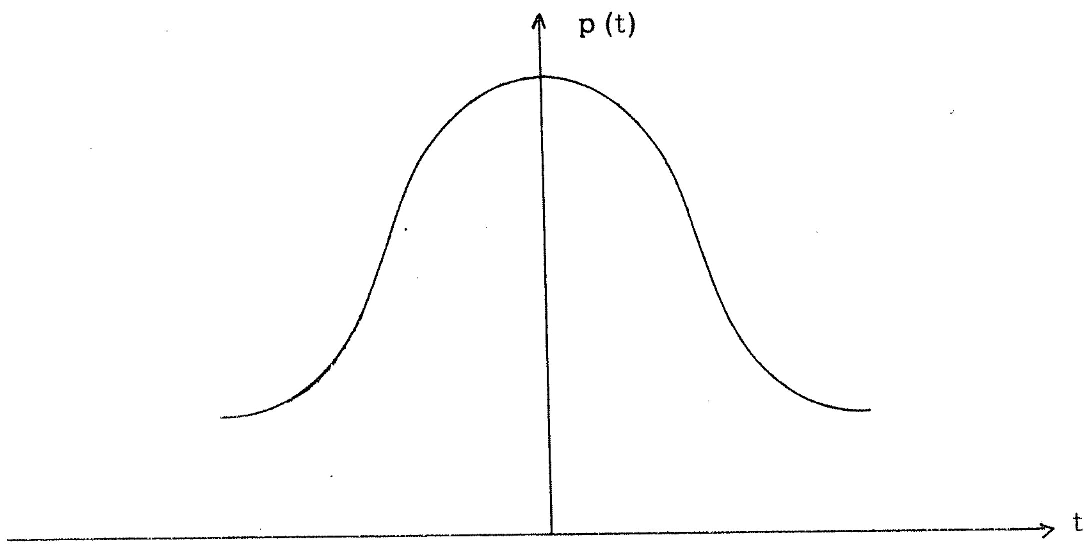

Figure D.1

The function p(t).

Because of symmetry we need only consider the case $(x,y) \in I_{ij}$, $j \leq i$, or

$$\left. \begin{aligned} & t_\ell \leq t \leq t_r , \\ & t_\ell = z_i - z_{j+1},\ t_r = z_{i+1} - z_j . \end{aligned} \right\} \tag{D. 2}$$

We introduce the points $P_\ell = (t_\ell, p(t_\ell))$, and $P_r = (t_r, p(t_r))$.

We shall obtain bounds in the form

$$\sum_{k=1}^{m} \alpha_{ijk}^{(1)} t^{k-1} \geq p(t) \geq \sum_{k=1}^{m} \alpha_{ijk}^{(2)} t^{k-1}, \quad t_\ell \leq t \leq t_r. \qquad \text{(D. 3)}$$

Since $t = x-y$, (D. 3) is easily brought into the form (2.4).

m = 1

From Figure D. 1, and (D. 2), we see that

$$\alpha_{ij1}^{(1)} = \begin{cases} p(0), & \text{if } i = j \\ p(t_\ell), & \text{otherwise.} \end{cases}$$

$$\alpha_{ij1}^{(2)} = p(t_r).$$

m = 2, i = j

Then $t_\ell = t_r$. Hence, we set

$$\alpha_{ii1}^{(1)} = p(0), \quad \alpha_{ii1}^{(2)} = p(t_r),$$

$$\alpha_{ii2}^{(1)} = \alpha_{ii2}^{(2)} = 0.$$

This approximation corresponds to bounding p(t) between the tangent at t = 0 and the chord between P_ℓ and P_r.

m = 2, j < i

This case is rather complicated.

First, we note that $t_r \geq t_\ell \geq 0$. Next, we note that for positive t p(t) has only one turning point, namely at $t = \kappa/\sqrt{3}$. Hence, if $\kappa/\sqrt{3}$ is not an interior point of $[t_\ell, t_r]$ then p(t) is either convex or concave on $[t_\ell, t_r]$ and is bounded by the chord joining P_ℓ and P_r,

$$u = p(t_\ell) + \frac{[p(t_r)-p(t_\ell)](t-t_\ell)}{[t_r-t_\ell]} , \qquad \text{(D. 4)}$$

and the tangent through the midpoint,

$$\left.\begin{aligned} u &= p(t_a) + p'(t_a)(t-t_a) , \\ t_a &= (t_\ell+t_r)/2 . \end{aligned}\right\} \qquad \text{(D. 5)}$$

If $\kappa/\sqrt{3} \leq t_\ell$, then $\alpha^{(1)}_{ij1}$ and $\alpha^{(1)}_{ij2}$ are determined by (D. 4) and $\alpha^{(2)}_{ij1}$ and $\alpha^{(2)}_{ij2}$ are determined by (D. 5); if $\kappa/\sqrt{3} \geq t_r$, the roles of (D. 4) and (D. 5) are reversed.

Finally, we must consider the case when $\kappa/\sqrt{3} \in (t_\ell, t_r)$. If this is so, then we can construct a line passing through P_ℓ which lies below p(t) and is tangent to p(t) at a point $t_{\ell c}$, $t_{\ell c} > t_\ell$ (see Figure D. 2). The equation for this line is easily found to be

$$\left.\begin{aligned} u &= p(t_\ell) + \frac{[p(t_{\ell c}) - p(t_\ell)]\,(t-t_\ell)}{[t_{\ell c} - t_\ell]} , \\ t_{\ell c} &= -t_\ell + \sqrt{t_\ell^2 + \kappa^2} . \end{aligned}\right\} \qquad \text{(D. 6)}$$

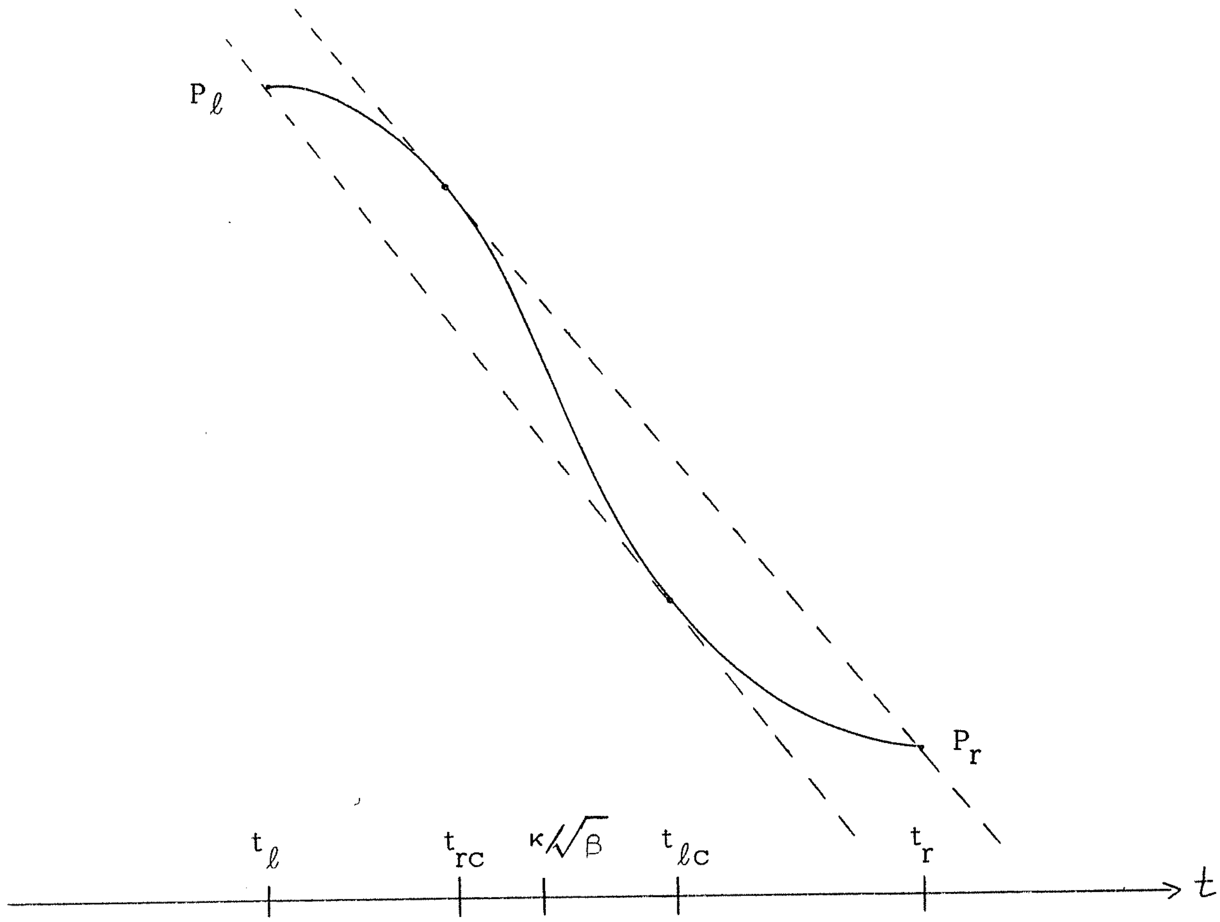

Figure D. 2

The case when $\kappa/\sqrt{3}$ is in (t_ℓ, t_r).

Similarly, there is a line passing through P_r which lies above $p(t)$ and is tangent to $p(t)$ at a point $t = t_{rc}$, $t_{rc} < t_r$ (see Figure D. 2). The equation for this line is easily found to be,

$$\left.\begin{aligned} u &= p(t_r) + \frac{[p(t_{rc}) - p(t_r)]\,(t-t_r)}{(t_{rc} - t_r)} \,, \\ t_{rc} &= -t_r + \sqrt{t_r^2 + \kappa^2} \quad . \end{aligned}\right\} \qquad \text{(D. 7)}$$

Then, $\alpha_{ij1}^{(1)}$ and $\alpha_{ij2}^{(1)}$ are determined by (D. 7), and $\alpha_{ij1}^{(2)}$ and $\alpha_{ij2}^{(2)}$ are determined by (D. 6).

References

1. Brown, R. W.: Upper and lower bounds for solutions of integral equations. In Error in Digital Computation, vol. 2, Rall, L. B. (editor). New York: Wiley, 1965.

2. Byerly, W. E.: Fourier's Series and Spherical, Cylindrical, and Ellipsoidal Harmonics. Boston: Ginn and Co., 1893.

3. Cryer, C. W.: On the calculation of the largest eigenvalue of an integral equation. Numer. Math., 10, p. 165, (1967).

4. Gerberich, C. L.: On the solution of a Fredholm integral equation, Report No. ORNL-2241, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1956.

5. Good, D. I. and R. L. London: Interval arithmetic for the Burroughs B5500: four ALGOL procedures and proofs of their correctness. Technical Report No. 26, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, 1968.

6. Isaacson, E. and H. B. Keller: Analysis of Numerical Methods. New York: Wiley, 1966.

7. Lefschetz, S.: Introduction to Topology. Princeton, N. J.: Princeton Univ. Press, 1949.

8. Mikhlin, S. G.: Integral Equations. New York: Macmillan, 1957.

9. Moore, R E.: Interval Analysis. Englewood Cliffs, N. J.: Prentice-Hall, 1966.

10. Rall, L. B.: Error bounds for iterative solutions of Fredholm integral equations. Pacific J. Math., 5, p. 977, (1955).

11. ___________: Numerical integration and the solution of integral equations by the use of Riemann sums. SIAM Review, 7, p. 55, (1965).

12. Reiter, A.: Interval arithmetic package (INTERVAL). Technical Summary Report No. 794, Math. Res. Center, U. S. Army, Univ. of Wisconsin, Madison, Wisconsin, 1968.

13. Sneddon, I. N.: Mixed Boundary Value Problems in Potential Theory. New York: Wiley, 1966.

14. Taylor, A. E.: Introduction to Functional Analysis. New York: Wiley, 1961.