

On Optimal Integration Methods for Volterra Integral Equations of the First Kind

By C. J. Gladwin

Abstract. Families of methods depending on free parameters are constructed for the solution of nonsingular Volterra integral equations of the first kind in [5]. These parameters are restricted to certain regions in order that a certain polynomial satisfies both a stability and a consistency condition. In this note an optimal choice of the free parameters is outlined in order that the l_2 -norm of the roots of the polynomial is minimized.

1. Introduction. Consider the linear Volterra integral equation of the first kind

$$(1.1) \quad \int_0^x K(x, t)y(t) dt = f(x),$$

where f, K are given functions on

$$S_1 = \{x | 0 \leq x \leq a\} \quad \text{and} \quad S_2 = \{(x, t) | 0 \leq t \leq x \leq a\}.$$

It will be assumed that sufficient regularity conditions are satisfied by f and K in order that a unique regular solution $y(x)$ exists on $[0, a]$. See also [12], for example.

On the mesh

$$I_N = \{x_n = nh | n = 0(1)N, Nh = a, h > 0\} \subset S_1$$

approximations y_n to the exact solution $y(x_n)$ are generated by the family of integration rules $Q(C, B, r)$ of the form:

$$(1.2) \quad h \sum_{i=0}^p c_i K(x_n, x_i)y_i + \sum_{j=p+1}^n h \sum_{i=0}^{p+1} b_i K(x_n, x_{j-i})y_{j-i} = f(x_n),$$

$$n = p(1)N.$$

The following brief remarks pertaining to (1.2) are noted here, see also [5].

- (1) p starting values y_0, y_1, \dots, y_{p-1} are needed before (1.2) can be implemented.
- (2) The vector $C \in R^{p+1}$ consists of the closed Newton-Cotes quadrature weights with step number p .
- (3) The vector $B \in R^{p+2}$ is (partially) determined by the linear system

$$(1.3) \quad \sum_{i=0}^{p+1} i^j b_i = \frac{1}{j+1}, \quad j = 0(1)r-1 \leq p+1; \quad 0^j = \begin{cases} 1, & j = 0, \\ 0, & j > 0. \end{cases}$$

Received October 6, 1981.

1980 *Mathematics Subject Classification*. Primary 65R05; Secondary 45L10.

©1982 American Mathematical Society
0025-5718/82/0000-0323/\$03.00

(4) The remaining $p + 2 - r$ parameters b_i ; $i = 0(1)p + 1 - r$, say, are constrained so that the polynomial

$$(1.4) \quad \rho(z) = \sum_{i=0}^{p+1} b_i z^{p+1-i}$$

is simple von Neumann. (The full rank, $r = p + 2$, case is only of use if $p = 0$. See also [7].)

We will be interested in the following classes of polynomials in this note.

Definition 1. Let $P(z)$ be a real polynomial of exact degree n . Then

(a) $P(z)$ is *simple von Neumann* if $P(z) = 0$ implies $|z| \leq 1$ with zeros on the unit circle being simple.

(b) $P(z)$ is *Schur* if $P(z) = 0$ implies $|z| < 1$.

(c) $P(z)$ is *Hurwitz* if $P(z) = 0$ implies $\operatorname{Re}(z) < 0$.

It is well known that $P(z)$ is Schur if and only if

$$Q(w) = (w - 1)^n P((w + 1)/(w - 1))$$

is Hurwitz; see [3] for an excellent survey of such results.

With these choices for C and B , and assuming the starting errors $e_n = y_n - y(x_n) = O(h^r)n = 0(1)p - 1$, the method (1.2) was shown to have an error $e_n = O(h^r)$, $n = 0(1)N$.

It should be mentioned here that other authors, [9], [10], [14], and [1], have considered cyclical multi-step methods recently with the latter even calling their methods “if not optimal, at least good practical methods”. The idea of their methods is that a number of *different* end formulae (the B of (1.2)) have to be used cyclically, necessitating an inordinate number of parameters. Many of these are used up in a “trivial” fashion to place as many roots as possible of their characteristic polynomial at the origin. However, they still maintain the necessity of solving a linear system as well as a *nonlinear* system to fulfill the accuracy and stability requirements, respectively.

2. Determination of the Free Parameters. The main result of this note is to display a system of $p + 2 - r$ linear equations to be solved for the free parameters, in addition to the r consistency conditions (1.3). We first give some motivation for this system.

The error e_n of the method (1.2) has an asymptotic expression, $n \rightarrow \infty$, $h \rightarrow 0$, $nh = x_n$,

$$(2.1) \quad e_n = h^r e(x_n) + h^r \sum_{k=0}^p z_k^n d_k(x_n) + O(h^{r+1}),$$

where

(a) $e(x_n)$ is the “magnified error” function,

(b) z_k are the zeros of $\rho(z)$ defined in (1.4),

(c) $d_k(x)$ are solutions of the initial value problems

$$K(x, x) d_k^{(1)}(x) - z_k K_t(x, x) d_k(x) = 0, \quad d_k(0) = d_k,$$

where K is given in (1.1), and

(d) the d_k satisfy the linear system

$$\sum_{k=0}^p z_k^n d_k = e_n - e(x_n), \quad n = 0(1)p.$$

See also [8], [6], and [11]. The latter only considered the case $r = p + 2$, $p = 0$ (i.e., the trapezoidal method applied to (1.1)). We also note that in the case that $\rho(z)$ has multiple roots, necessarily interior to the unit circle, only the system in (d) need be modified—i.e., m linearly independent column vectors must be added to the coefficient matrix. These take the form $(u^n, nu^n, \dots, n(n-1) \cdots (n-m+2)u^n)$ corresponding to a multiple root u of a fixed multiplicity, $m > 1$. We observe that if $\rho(z)$ is simple von Neumann, the phenomenon of marginal instability occurs if $|d_k(x)|$ are increasing functions, but can be eliminated by making $\rho(z)$ Schur.

Thus, it is natural to choose the free parameters of $\rho(z)$ so as to try to minimize some norm of the roots, in addition to the previously mentioned stability and consistency conditions. In particular, we choose the square of the l_2 -norm in view of the following observations:

(1) $\sum_{i=1}^n |z_i|^2 \geq \sum_{i=1}^n z_i^2$, where z_i are the zeros of a real polynomial of exact degree n . Also, equality only occurs if all the z_i are real.

(2) $\sum_{i=1}^n |z_i|^2 \geq n \prod_{i=1}^n |z_i|^{2/n}$ with equality only if $|z_i| = |z_1|$, say, for all i .

Thus we desire to try to make $\rho(z)$ have as many equal real roots as possible. Obviously we cannot make $\rho(z)$ have a single root of multiplicity $p + 1$, since we only have $p + 2 - r$ parameters at our disposal. Also, we choose the smallest p for a given r , compatible with stability, to ensure the use of the fewest free parameters as possible.

The following theorem is the main result of this note. It shows how to construct the system of linear equations for the free parameters.

THEOREM 1. (1) Assume $\rho(z)$ is as in (1.4) with its coefficients satisfying (1.3).

(2) Assume $\tau(w) = \sum_{i=0}^{p+1} \tau_i w^i$ is the transformed version of $\rho(z)$ using the Möbius map

$$w = \frac{z+1}{z-1}, \quad z = \frac{w+1}{w-1}.$$

Then

(1) τ_i : $i = 0(1)p + 1 - r$ depend linearly on the free parameters.

(2)
$$\frac{\partial \tau_i}{\partial b_j} = 0, \quad \begin{matrix} i = p + 2 - r(1)p + 1, \\ j = 0(1)p + 1 - r. \end{matrix}$$

Proof. The system (1.3) may be solved in terms of the free parameters as follows, [13],

$$(2.2) \quad b_{p+1-r+i} = \frac{(-1)^{r-i}}{r-1!} \binom{r-1}{i-1} \left\{ \int_0^1 s_i(x) dx - \sum_{k=0}^{p+1-r} s_i(k) b_k \right\}$$

with $s_i(x) = \prod_{j=1}^r [x - (p + 1 - r + j)]$, $i = 1(1)r$.

We shall denote the first term on the right of (2.2) by λ_{ri} for a given number of free parameters. A simple calculation shows

$$\frac{\partial b_{p+1-r+i}}{\partial b_{p+2-r-n}} = (-1)^i \binom{n+r-1}{n+i-1} \binom{n+i-2}{n-1} = v_{nri}, \quad \text{say,}$$

$$i = 1(1)r, n = 1(1)p + 2 - r \text{ (or } k = p + 1 - r(-1)0).$$

Now $\rho(z)$ may be written as

$$(2.3) \quad \rho(z) = \sum_{n=1}^{p+2-r} b_{p+2-r-n} \left(z^{n+r-1} + \sum_{i=1}^r v_{nri} z^{r-i} \right) + \sum_{i=1}^r \lambda_{ri} z^{r-i}.$$

Denote the first term of (2.3) by $q(z)$ which is a polynomial of degree $p+1$ whose coefficients are only (linearly) dependent on the free parameters. If we write $q(z)$ in powers of $z-1$, we assert that the coefficients of the terms $(z-1)^j$, $j = 0(1)r-1$ vanish. We return to this assertion later so as not to interrupt the flow of the proof.

We calculate $\tau(w) = (w-1)^{p+1} \rho((w+1)/(w-1))$. The powers $(z-1)^j$ become

$$(w-1)^{p+1} \left(\frac{w+1}{w-1} - 1 \right)^j = 2^j (w-1)^{p+1-j}, \quad j = r(1)p + 1,$$

so that only τ_i ; $i = 0(1)p + 1 - r$ depend linearly on the free parameters, while the remaining coefficients are independent of them.

Returning to our assertion, we replace z by $[1 + (z-1)]$ in $q(z)$, expand the binomials in powers of $z-1$, and interchange the order of summation in the second term of $q(z)$. This yields:

$$(2.4) \quad q(z) = \sum_{n=1}^{p+2-r} b_{p+2-r-n} \sum_{j=0}^{n-1} \binom{n+r-1}{j} (z-1)^{n+r-1-j} \\ + \sum_{n=1}^{p+2-r} b_{p+2-r-n} \sum_{j=0}^{r-1} \left\{ \binom{n+r-1}{j} + \sum_{i=1}^{r-j} v_{nri} \binom{r-i}{j} \right\} (z-1)^j.$$

We now show that the factor $\{\dots\}$ vanishes for all $j = 0(1)r-1$. It will suffice to consider only $n = p+2-r$. Insert a factor

$$\frac{i}{i} \frac{r-j!}{r-j!} \frac{p+1-j!}{p+1-j!}$$

inside the sum over i , cancel, and use

$$\frac{i}{p+1-r+i} = 1 - \frac{p+1-r}{p+1-r+i}.$$

The inner sum becomes

$$\sum_{i=0}^{r-j} (-1)^i \binom{r-j}{i} - (p+1-r) \sum_{i=0}^{r-j} \frac{(-1)^i}{p+1-r+i} \binom{r-j}{i}.$$

The first sum is 0 since $(1-1)^{r-j} = 0$. For the second sum we use the binomial identity

$$(2.5) \quad \sum_{i=0}^m \frac{(-1)^i}{x+i} \binom{m}{i} = \frac{m!}{x(x+1)\dots(x+m)}, \quad x > 0,$$

with $x = p+1-r$ and $m = r-j$. (2.5) may be found in [4], for example.

It should be noted that the use of (2.5) does not cover the case $r = p + 1$. However, returning to (2.3), it is easily seen that $q(z) = b_0(z - 1)^{p+1}$. This completes the proof.

Remarks. (1) The idempotent map $w = (z + 1)/(z - 1)$ transforms the unit circle of the z plane to the left half of the w plane in a 1-1 fashion. In particular, equal roots of $\rho(z)$ correspond to equal roots of $\tau(w)$. This suggests using the system of linear equations

$$(2.6) \quad \tau^{(i)}(w_0) = 0, \quad i = 0(1)p + 1 - r,$$

where w_0 is one of the at most $r - 1$ real roots of $\tau^{(p+2-r)}(w)$. We choose that w_0 which makes $\tau(w)$ a Hurwitz polynomial, if possible.

(2) Thus the $p + 2$ coefficients of $\rho(z)$ can be completely determined by (1.3) and (2.6), on a machine. Note that we must check that the coefficients of $\tau(w)$ yield a Hurwitz polynomial for each w_0 , but even this can be done mechanically when the coefficients are numerical. See some algorithms in [3], for example.

(3) Here we shall only illustrate the coefficient matrix of the system (2.6) in the case $r = p$, i.e.,

$$F = \begin{bmatrix} (p+1)2^p(w_0 - 1) + 2^{p+1} & 2^p(w_0 - 1) \\ (p+1)2^p & 2^p \end{bmatrix}, \quad \det(F) = 2^{2p+1}$$

to be solved for the vector $S = (b_0, b_1)^T$.

3. Examples and Conclusions. We first consider two methods, $Q(C, B, r)$, which are indeed optimal.

(1) $r = 2, p = 1$,

$$\tau(w) = (4b_0 - 2) + w + w^2, \quad \tau^{(1)}(w) = 1 + 2w.$$

Using $w_0 = -1/2$, we obtain $b_0 = 9/16$. In this case we actually get two equal real roots of $\rho(z)$; $z_{1,2} = -1/3$. $\tau(w)$ is Hurwitz for $b_0 > 1/2$.

(2) $r = 3, p = 2$,

$$\tau(w) = (8b_0 - 11/3) + (2/3)w + 2w^2 + w^3, \quad \tau^{(1)}(w) = (2/3) + 4w + 3w^2.$$

Only the root $w_0 = (-2 + 2^{1/2})/3$ makes $\tau(w)$ a Hurwitz polynomial. In this case we obtain $b_0 = (95 + 4(2)^{1/2})/216$. $\tau(w)$ is Hurwitz for b_0 such that $11/24 < b_0 < 15/24$; see also [5]. Again, in this case, we have all real roots with as many equal roots as possible.

Rather than use the irrational b_0 above, we shall use $b_0 = 7/15$ which agrees to 3 places of decimals. A slightly better approximation is $b_0 = 706/1515$, to 5 places of decimals.

We note that in [2], methods with $b_0 = 9/16$ $r = 2$ and $b_0 = 23/48$ $r = 3$ are considered, but no reason for these choices are given.

The following Volterra equation will be solved.

$$(3.1) \quad \int_0^x \exp[-(x-t)] y(t) dt = \sinh x, \quad 0 \leq x \leq 5,$$

which has the exact solution $y(x) = \exp(x)$. The phenomenon of marginal instability, recall (2.1), does not occur in this equation. Thus we shall solve it by our two optimal methods as well as by methods where $\rho(z)$ is simple Von Neumann in order to see how much improvement in the numerical solution we get between the "worst" and the "best" methods.

TABLE 1

Methods $Q(C, B, r)$				
r	p	b_0	zeroes of $\rho(z)$, z_j	$\ z_j\ _2^2$
2	1	1/2	0, -1	1.000
2	1	9/16	-1/3, -1/3	0.222
3	2	15/24	1/3, $(-1 \pm 2(6)^{1/2} i)/5$	2.111
3	2	7/15	0.234, -0.671 \pm 0.856i	0.969

The polynomials, $\rho(z)$, are:

$$p = 1: \rho(z) = b_0 z^2 + (3/2 - 2b_0)z + b_0 - 1/2,$$

$$p = 2: \rho(z) = b_0 z^3 + (23/12 - 3b_0)z^2 + (3b_0 - 16/12)z + 5/12 - b_0.$$

We also remark that for the actual optimal $b_0 = (95 + 4(2)^{1/2})/216$, in the case $p = 2$, the l_2 -norm is 0.961. Also, for the method corresponding to $r = 3$, $p = 2$, $b_0 = 11/24$, $\rho(z)$ has zeros $z_{1,2,3} = -1, (-1 \pm 2(3)^{1/2})/11$ with $\|z_i\|_2^2 = 1.215$, so that we took $b_0 = 15/24$ as the "worst" method.

(3.1) was solved with step sizes $h = 0.1$ and $h = 0.05$. Exact starting values were used when needed. All computations were performed on the Honeywell sigma 9 (in double precision, i.e., a 64 bit word) at the Communications Research Centre in Ottawa. A selection of numerical results is displayed in Table 2.

Remarks. (1) The "optimal" method of order 2 actually has larger errors than the "worst" method. However, the latter has a root of $\rho(z)$ at the origin, which means the summation in (2.1) has only one term. On the other hand, the errors of the "optimal" method of order 3 are less than those of the "worst" method.

(2) In any case, the importance of this method of construction is not to minimize the errors for a given order r , but rather to permit a mechanical, lexicographical search, for a given pair of positive integers (r, p) with $r \leq p + 2$, for polynomials $\rho(z)$ which are Schur.

(3) In [5] and [6], it was observed that for $r = p + 1$ (one free parameter) $\tau(w)$ has coefficients of opposite signs for $3 \leq p \leq 5$ and hence cannot be Hurwitz for any b_0 ; see [3], e.g.. Thus one might conjecture that $r = 3$ is the highest order one can obtain with one free parameter. However, using the crude and arbitrary choice, $b_i = 0$: $i = 0(1)p + 1 - r$, it can be shown that $\tau(w)$ is Hurwitz for $p = 19$, at least.

TABLE 2
Numerical Results

x_n	h	$y(x_n)$	$ e_n = y_n - y(x_n) $			
			$b_0 = 1/2$ $r = 2$	$b_0 = 9/16$ $r = 2$	$b_0 = 15/24$ $r = 3$	$b_0 = 7/15$ $r = 3$
1.0	0.10	2.7182818	1.2×10^{-3}	1.5×10^{-2}	1.7×10^{-3}	3.5×10^{-4}
	0.05		5.3×10^{-4}	3.9×10^{-3}	7.9×10^{-4}	2.3×10^{-4}
2.0	0.10	7.3890561	2.2×10^{-2}	4.1×10^{-2}	1.3×10^{-2}	4.7×10^{-3}
	0.05		5.5×10^{-3}	1.1×10^{-2}	1.8×10^{-3}	6.3×10^{-4}
3.0	0.10	20.0855369	6.6×10^{-2}	1.1×10^{-1}	3.4×10^{-2}	1.3×10^{-2}
	0.05		1.6×10^{-2}	2.9×10^{-2}	4.5×10^{-3}	1.7×10^{-3}
4.0	0.10	54.5981500	1.8×10^{-1}	3.0×10^{-1}	9.2×10^{-2}	3.5×10^{-2}
	0.05		4.5×10^{-2}	7.8×10^{-2}	1.3×10^{-2}	4.7×10^{-3}
5.0	0.10	148.4131591	4.9×10^{-1}	8.3×10^{-1}	2.5×10^{-1}	9.5×10^{-2}
	0.05		1.2×10^{-1}	2.1×10^{-1}	3.4×10^{-2}	1.3×10^{-2}

4. Acknowledgement. The author wishes to thank the Communications Research Centre, Department of Communications, Ottawa for the use of the necessary computing and secretarial facilities required in the completion of this paper.

c/o D. B. Muldrew
Communications Research Centre
Box 11490 Station 'H'
Ottawa, Ontario, Canada K2H 8S2

1. C. ANDRADE & S. MCKEE, "On optimal high accuracy linear multistep methods for first kind Volterra integral equations," *BIT*, v. 19, 1979, pp. 1-11.
2. H. BRUNNER, *The Approximate Solution of Integral Equations by Projection Methods Based on Collocation*, Mathematics and Computation No. 1/78, ISBN 82-7151-022-3, Dept. of Math., The University of Trondheim, Trondheim, Norway, 1978.
3. R. J. DUFFIN, "Algorithms for classical stability problems," *SIAM Rev.*, v. 11, 1969, pp. 196-213.
4. H. FREEMAN, *Finite Differences for Actuarial Students*, Cambridge Univ. Press, Cambridge, 1962, p. 113.
5. C. J. GLADWIN, "Quadrature rule methods for Volterra integral equations of the first kind," *Math. Comp.*, v. 33, 1979, pp. 705-716.

6. C. J. GLADWIN, *Numerical Solution of Volterra Integral Equations of the First Kind*, Ph.D. Thesis, Dalhousie Univ., 1975.
7. C. J. GLADWIN & R. JELTSCH, "Stability of quadrature rules for first kind Volterra integral equations," *BIT*, v. 14, 1974, pp. 144–151.
8. P. HENRICI, *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York, 1962.
9. P. A. W. HOLYHEAD, S. MCKEE & P. J. TAYLOR, "Multistep methods for solving linear Volterra integral equations of the first kind," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 698–711.
10. P. A. W. HOLYHEAD & S. MCKEE, "Stability and convergence of multistep methods for linear Volterra integral equations of the first kind," *SIAM J. Numer. Anal.*, v. 13, 1976, pp. 269–292.
11. M. KOBAYASI, "On numerical solutions of Volterra integral equations of the first kind by the trapezoidal rule," *Rep. Statist. Appl. Res. Un. Japan. Sci. Engrs.*, v. 14, 1967, pp. 1–14.
12. W. POGORZELSKI, *Integral Equations and their Applications*, Vol. 1, Pergamon Press, Oxford, 1966, p. 14.
13. A. RALSTON, *A First Course in Numerical Analysis*, McGraw-Hill, New York, 1965, p. 152.
14. P. J. TAYLOR, "The solution of Volterra integral equations of the first kind using inverted differentiation formulae," *BIT*, v. 16, 1976, pp. 416–425.
15. P. H. M. WOLKENFELT, *Linear Multistep Methods and the Construction of Quadrature Formulae for Volterra Integral and Integro-Differential Equations*, Report No. NW 76/79, Mathematisch Centrum, Amsterdam, 1979.