A FAST MULTILEVEL ALGORITHM FOR INTEGRAL EQUATIONS *

C. T. KELLEY†

Abstract. We show how the discretization of integral equations by composite Gauss rules can be related to approximations of integral operators that converge in the operator norm, rather than strongly. From this norm convergent formulation a two level approximate inverse can be constructed whose evaluation requires no fine mesh evaluations of the integral operator. The resulting multilevel algorithm, therefore, is roughly half as costly as the Atkinson-Brakhage iteration. The algorithm is applicable to both linear and nonlinear equations.

Key words. Integral equations, multilevel methods, Atkinson-Brakhage iteration, composite Gauss rule

AMS(MOS) subject classifications. 45G10, 65F10, 65J15,

1. Introduction.

The purpose of this paper is to show how the discretization of integral equations by composite Gauss rules can be related to approximations of integral operators that converge in the operator norm, rather than strongly. From this norm convergent formulation a two level approximate inverse can be constructed whose evaluation requires no fine mesh evaluations of the integral operator. The resulting multilevel algorithm, therefore, is roughly half as costly as the Atkinson-Brakhage iteration. The algorithm is applicable to both linear and nonlinear equations.

We begin by reviewing the Atkinson-Brakhage [2], [5] algorithm for integral equations. The important ideas are completely illustrated by consideration of the linear equation on $C[0,1]$,

$$u(x) = (Ku)(x) + g(x) = \int_0^1 k(x,y)u(y)\,dy + g(x). \tag{1.1}$$

In (1.1) k and g are given continuous functions and $u \in C[0,1]$ is to be found. We assume throughout this paper that the linear integral operator given in (1.1) is such that $I - K$ is a nonsingular map on $C[0,1]$. Continuity of k and the Fredholm alternative theorem imply that $I - K$ is also a nonsingular map on $L^p[0,1]$ for all $p \in [1,\infty]$. Throughout this paper we will let $\mathcal{L}(X)$ denote the space of bounded operators on a Banach space X endowed with the usual operator norm.

The Atkinson-Brakhage algorithm begins with a sequence of quadrature rules, indexed by m, with nodal points $\{x_j^m\}_{j=1}^{N_m}$ and weights $\{w_j^m\}_{j=1}^{N_m}$. If

$$\lim_{m\to\infty} \sum_{j=1}^{N_m} f(x_j^m)w_j^m = \int_0^1 f(x)\,dx \tag{1.2}$$

then the sequence of operators $\{K_m\} \subset \mathcal{L}(C[0,1])$ defined by

$$K_m(u)(x) = \sum_{j=1}^{N_m} k(x,x_j^m)u(x_j^m)w_j^m \tag{1.3}$$

* This document was printed on April 17, 1994.

† North Carolina State University, Center for Research in Scientific Computation and Department of Mathematics, Box 8205, Raleigh, N. C. 27695-8205, USA. This research was supported by National Science Foundation grant #DMS-9024622 and Air Force Office of Scientific Research grant #AFOSR-FQ8671-9101094. Computing was paritally supported by an allocation of time from the North Carolina Supercomputing Center.

is collectively compact [1] and converges strongly to the operator K.

Equations of the form

$$(I - K_m)u = f \tag{1.4}$$

for $f \in C[0,1]$ can be solved by first solving the finite dimensional system

$$u(x_i^m) - \sum_{j=1}^{N_m} k(x_i^m, x_j^m)u(x_j^m)w_j^m = f(x_i^m) \tag{1.5}$$

for the values of the solution at the nodal points and then recovering the values of u at an arbitrary $x \in [0,1]$ by Nyström interpolation

$$u(x) = f(x) + \sum_{j=1}^{N_m} k(x, x_j^m)u(x_j^m)w_j^m. \tag{1.6}$$

There are several important consequences of the collective compactness and strong convergence of K_m to K. First of all, there is l_0 such that if $m \geq l_0$ then then $I - K_m$ is nonsingular and

$$(I - K_m)^{-1} \xrightarrow{s} (I - K)^{-1}. \tag{1.7}$$

Equation (1.7) implies, of course, that the solutions of (1.4) converge uniformly to the solution of (1.1).

Another consequence of collective compactness and strong convergence is that that for any $\rho > 0$ there is l_0 such that if $L \geq l \geq l_0$ the operator

$$B_l^L = I + (I - K_l)^{-1}K_L \tag{1.8}$$

satisfies

$$\|I - B_l^L(I - K_L)\| \leq \rho. \tag{1.9}$$

(1.9) is the central estimate in the Atkinson-Brakhage two level method. This method was proposed in [2] and [5] as a method to solve

$$u - K_L u = g$$

for a given level L by using B_l^L as a preconditioner for a Richardson iteration. The two level iteration, described in terms of the transition from a current iterate u_c to a new iterate u_+, is

$$u_+ = u_c - B_l^L(u_c - K_L u_c - g). \tag{1.10}$$

L is usually called the fine mesh or fine level and l the coarse mesh or level.

One might think that $(I - K_l)^{-1}$ alone would make an effective approximate inverse. However the convergence in (1.7) is only strong and not in the operator norm. The effect of this is that convergence of the iteration with $(I - K_l)^{-1}$ as the approximate inverse is two-step q-linear [2] and therefore no more efficient than the iteration based on (1.8). The iteration based on (1.8) has a more predictible and regular behavior, being one step q-linearly convergent. The algorithm proposed here may be thought of as replacement of $(I - K_l)^{-1}$ by a norm convergent sequence that is no more expensive to evaluate than $(I - K_l)^{-1}$.

From (1.10) it is clear that two evaluations of K_L are required for a single Atkinson-Brakhage iteration. First $K_L u_c$ is computed to form the linear residual

$$r_c = g - (I - K_L)u_c$$

The second evaluation is for the computation of $B_l^L r_c$

$$B_l^L r_c = r_c + (I - K_l)^{-1} K_L r_c.$$

The approach suggested here eliminates this second evaluation of K_L.

Various extensions of the Atkinson-Brakhage iteration have been proposed recently. [18] advocates a nested iteration strategy in which the fine mesh L is refined as the iteration progresses. One may also use piecewise linear interpolation to move from the coarse to the fine mesh in the evaluation of B_l^L, rather than the Nyström interpolation that is implicit in (1.8). The Atkinson-Brakhage method can also be applied to linear and nonlinear compact fixed point problems in which the fixed point map is not given explicitly or with full accuracy [21], is not smooth [12], [20], [19], or (in a nonlinear setting) is singular at the solution [18]. In all of this previous work, strong convergence and collective compactness was central to proving that some form of the Atkinson-Brakhage iteration could be applied and in all cases the cost of a single iterate was two fine mesh evaluations of K_L. Second-kind multigrid methods [11] are known to have similar computational costs [3].

In a recent paper, [15], on application of the Atkinson-Brakhage iteration to the source iteration map for the linear transport equation, we used the fact that certain discretizations could be formulated either as collectively compact and strongly convergent sequences of maps [22] or as a norm convergent sequences [25]. In [15] we use show that the difference in the two viewpoints is, from the algorithmic point of view, equivalent to a change in the fine to coarse mesh intergrid transfer from a point evaluation to an average. This enables us to avoid the second evaluation of the fine mesh evaluation of the action of the linear source iteration map.

In this paper we show how the observation in [15] that this second evaluation of K_L can be avoided extends to integral equations if the coarse and fine mesh quadrature rules are composite Gauss rules in which the fine mesh intervals are nested within the coarse mesh intervals. There is no requirement that the coarse and fine mesh rules be based on the same order of Gaussian quadrature.

2. Description and Analysis of the Algorithm.

We assume the quadrature rule at level m is a composite Gauss rule with q_m subintervals $\{I_r^m\}_{r=1}^{q_m}$ with $p_m + 1 \geq 1$ points used in each interval. Hence $N_m = q_m(p_m + 1)$. Now consider the sequence of operators $\{\mathcal{K}_m\}$ defined by

$$\mathcal{K}_m u(x) = \int_0^1 k_m(x, y) u(y)\, dy \tag{2.1}$$

where k_m is the piecewise polynomial on $[0,1] \times [0,1]$ defined for $x \in \mathrm{Int}(I_r^m)$ and $y \in \mathrm{Int}(I_s^m)$ by

$$k_m(x, y) = \sum_{x_i^m \in I_r^m, x_j^m \in I_s^m} k(x_i^m, x_j^m) l_i^r(x) l_j^s(y). \tag{2.2}$$

In (2.2), l_i^r is the Lagrange interpolating polynomial on I_r^m of degree p_m. For $x \in I_r^m$

$$l_i^r(x) = \prod_{j \neq i, x_j^m \in I_r^m} \frac{x - x_j^m}{x_i^m - x_j^m}.$$

Here $\mathrm{Int}(S)$ denotes the interior of a set S. The kernels k_m are discontinuous and we define them at the endpoints of the intervals $\{I_r^m\}_{r=1}^{q_m}$ as the arithmetic means of their limiting values from the right and left. We use these kernels only as tools in the analysis and do not construct them in the implementation. $\mathcal{K}_m$ is a degenerate kernel operator. Such operators have been used as approximations for the purposes of solution and error estimation for many years [4], [10], where the kernels are typically constructed using orthonormal bases or operator product formulae for k. In this paper, the operator $\mathcal{K}_m$ has been constructed as a preconditioner, which is a quite different purpose.

We will require the following lemma, which is a direct consequence of elementary facts on polynomial interpolation [14].

LEMMA 2.1. *If k is continuous and the sequence of composite Gauss rules is such that* (1.2) *holds and there is P such that $p_m \leq P$ for all m, then $k_m \rightarrow k$ uniformly. In that case $\mathcal{K}_m \rightarrow \mathcal{K}$ in the operator norm.*

A consequence of Lemma 2.1 is that if $I - K$ is nonsingular, so is $I - \mathcal{K}_l$ for l sufficiently large. Moreover, $(I - \mathcal{K}_l)^{-1}$ is a good approximate inverse of $I - K$. We show that the action of $(I - \mathcal{K}_l)^{-1}$ on a function can be evaluated in a more efficient way than that the Atkinson-Brakhage preconditioner (1.8). To do this we seek solutions in $L^\infty[0,1]$, rather than $C[0,1]$. Continuous solutions may be recovered to full accuracy with a Nyström interpolation if needed.

2.1. Basic Algorithm. Let V^m be the subspace of L^∞ consisting of piecewise polynomials having degree $p_m + 1$ on each of the m subintervals I_r^m. We assume that the intervals are nested and the degrees are related in such a way that $V^m \subset V^n$ if $m \leq n$.

We can now propose the initial two level form of the iteration. We seek to find an approximate solution $u^L \in V^L$ of $(I - \mathcal{K}_L)u = f$ and employ the iteration

$$u_+^L = u_c^L - (I - \mathcal{K}_l)^{-1}(u_c^L - \mathcal{K}_L u_c^L - g). \tag{2.3}$$

This is simply Richardson iteration using $(I - \mathcal{K}_l)^{-1}$ as an approximate inverse for $I - \mathcal{K}_L$. Since

$$\mathcal{K}_L u = K_L u \text{ for all } u \in V^L \tag{2.4}$$

we may rewrite (2.3) as

$$u_+^L = u_c^L - (I - \mathcal{K}_l)^{-1}(u_c^L - K_L u_c^L - g), \tag{2.5}$$

which is identical to (1.10) except that B_l^L is replaced by $(I - \mathcal{K}_l)^{-1}$. This means that the additional computation of the action of K_L that was required in the computation of the action of B_l^L is not needed in the iteration given by (2.5).

The multilevel form of the algorithm refines the fine mesh as the iteration progresses. If the spaces V^m are nested, as we have assumed, then we may compute $u^{L+1} \in V^{L+1}$ from $u^L \in V^L \subset V^{L+1}$ by

$$u^{L+1} = u^L - (I - \mathcal{K}_l)^{-1}(u^L - K_{L+1} u^L - g). \tag{2.6}$$

Note that the computation of u^{L+1} requires only a single evaluation of the action of K_{L+1}, namely the computation of $K_{L+1}u^L$ in the right hand side of (2.6). The main convergence result in this paper is

THEOREM 2.2. *Assume that k is continuous, that $I-K$ is nonsingular, and that the sequence of composite Gauss rules is such that* (1.2) *holds, $p_m \leq P$ for some P, and that $V^m \subset V^{m+1}$. Then for l sufficiently large and $u^l \in V^l$ the iteration given by* (2.6) *converges uniformly to $u^* = (I-K)^{-1}g$.*

Proof. Lemma 2.1 implies that there are l and M such that if $L \geq l$ then $I-\mathcal{K}_L$ is nonsingular and

$$\|(I-\mathcal{K}_L)^{-1}\|_{\mathcal{L}(L^\infty)} \leq M.$$

We define

$$\nu_l^L = \|I-(I-\mathcal{K}_l)^{-1}(I-\mathcal{K}_L)\|_{\mathcal{L}(L^\infty)} \text{ and } \tau_L = \|(I-\mathcal{K}_L)u^* - g\|_\infty. \tag{2.7}$$

Our assumptions imply that

$$\lim_{l\to\infty, L\geq l} \nu_l^L = 0 \text{ and } \lim_{L\to\infty} \tau_L = 0. \tag{2.8}$$

Let $e^L = u^L - u^*$. Using (2.5) and the fact that $u^L \in V^L$ implies that $\mathcal{K}_L u^L = K_L u^L$, we obtain

$$e^{L+1} = e^L - (I-\mathcal{K}_l)^{-1}((I-\mathcal{K}_{L+1})e^L + (I-\mathcal{K}_L)u^* - g).$$

Hence

$$\|e^{L+1}\|_\infty \leq \nu_l^{L+1}\|e^L\|_\infty + M\tau_L. \tag{2.9}$$

This completes the proof if we require l to be large enough so that $\nu_l^L < 1/2$ for all $L \geq l$. □

Suppose that we seek to maintain accuracy to truncation error as the iteration progresses. By this we mean that

$$\|e^L\|_\infty \leq M\tau_L \tag{2.10}$$

for all L. While it is impossible to verify (2.10), one can test

$$\|u^L - \mathcal{K}_L u^L - g\|_\infty \leq M\tau_L, \tag{2.11}$$

provided an estimate can be found for τ_L. (2.11) implies (2.10) with a different choice of M. If there is μ such that

$$\lim_{L\to\infty} \tau_{L+1}/\tau_L \to \mu$$

then a way to maintain (2.11) is to approximate τ_{L+1} by $\mu\|u^L - \mathcal{K}_{L+1}u^L - g\|_\infty$. Hence, a useful strategy would be to set $N_{L+1} = N_L$ (*i. e.* do not refine the mesh) until the residual had been reduced by a factor of μ over the last time the mesh had been refined. In principle several iterations could be required for this criterion to hold. However, (2.9) implies that if l is sufficiently large only a single iteration will be needed to reduce the residual by a factor of μ.

2.2. Implementation. In this subsection we indicate how the iteration given by (2.6) can be implemented efficiently. We show how $\Pi_l r^L$ can be computed and give an example using a composite midpoint rule as the coarse mesh rule. We also discuss the computational costs of the iteration and compare these costs in detail to those of the Atkinson-Brakhage iteration.

Let P_m be the projection from $C[0,1]$ onto V^m defined by piecewise polynomial interpolation. Let Π_m be the orthogonal projection, relative to the inner product of $L^2[0,1]$, onto V^m. This projection is also a projection in $C[0,1]$ in the Banach space sense, but is not a norm one projection.

We begin the iteration by solving

$$u^l - \mathcal{K}_l u^l = g$$

either exactly or approximately. We defer our discussion of how equations of the form $(I - \mathcal{K}_l)w = f$ are solved until later in this subsection. Given u^L we must first compute

$$r^L = u^L - K_{L+1} u^L - g. \tag{2.12}$$

It is sufficient to compute r^L at the nodal points of the quadrature rules with indices l and $L+1$. This requires addition of vectors of length N_{L+1} and N_l and the evaluation of u^L and the action of K_{L+1} on $u^L \in V^L$ at the nodal points of the quadrature rules with indices l and $L+1$. As one can see from the definition of $\mathcal{K}_L$, the cost of these evaluations is $O(N_{L+1}N_L)$ floating point operations.

Next we solve

$$w - \mathcal{K}_l w = r^L. \tag{2.13}$$

Since

$$\mathcal{K}_m = \mathcal{K}_m \Pi_m = P_m \mathcal{K}_m \Pi_m$$

for all m. We may solve (2.13) by first computing $\Pi_l w$ by solving

$$\Pi_l w - \mathcal{K}_l \Pi_l w = \Pi_l r^L \tag{2.14}$$

and then recovering w by the Nyström interpolation

$$w = r^L + \mathcal{K}_l \Pi_l w. \tag{2.15}$$

Finally we compute $u^{L+1} = u^L - w$. The Nyström interpolation in (2.15) has a cost of $O(N_l N_{L+1})$ floating point operations which can be neglected when compared to the cost of the computation of r^L.

The remaining issue is the solution of (2.14). Note that the action of Pi_l on $r^L \in v^L$ requires integration of polynomials of degree $p_l p_L$ on each subinterval I_r^L. These integrals can be computed exactly using the quadrature rule at level L and hence need only values of r^L at the nodal points of the quadrature rule. For example, if $l = 0$ corresponds to a composite midpoint rule and $v \in V^L$ then

$$\begin{aligned} (\Pi_l v)(x) &= \sum_{i=1}^{q^l} \chi_i(x) \int_{I_i^l} v(x)\,dx \\ &= \sum_{i=1}^{q^l} \chi_i(x) \sum_{x_j^L \in I_i^L} v(x_j^L) w_j^L. \end{aligned} \tag{2.16}$$

In (2.16), χ_i^l denotes the characteristic function of the interval I_i^l. Note that the cost of the computation of the action of Π_l on r^L is $O(N_L)$ floating point operations, which can also be neglected when compared to the cost of the computation of r^L. This fine-to-coarse intergrid transfer by averaging is the key distinction between our approach and the classical one, where the fine-to-coarse intergrid transfer is by point evaluation.

The solution of (2.14) can be reduced to the solution of a finite dimensional system for the values of $\Pi_l w$ at the nodal points of the quadrature rule at level l.

$$u(x_i^l) - \sum_{j=1}^{N_l} k(x_i^l, x_j^l) u(x_j^l) w_j^l = (\Pi_l r^L)(x_i^l), \tag{2.17}$$

The action of Π_l on a function f can be computed by (2.16) if $f \in V^L$ or approximated by $\Pi_l P_L f$ if $f \notin V^L$. In our implementation we compute $\Pi_l r^L$ by applying (2.16) to compute $\Pi_l(u^L - K_{L+1}u^L)$ and approximating $\Pi_l g$ by $\Pi_L P_L g$. This approximation is accurate to within truncation error as the accuracy of $P_L g$ as an approximation of g partially determines the accuracy of $(I - K_L)^{-1} g$ as an approximation to the solution.

This is very similar to the system to the solved for the Atkinson-Brakhage iteration, the difference between them being only in the right hand sides. In fact, (1.5) with $m = l$ can be written as

$$u(x_i^l) - \sum_{j=1}^{N_l} k(x_i^l, x_j^l) u(x_j^l) w_j^l = (P_l f)(x_i^l) \tag{2.18}$$

since $(P_l f)(x_i^l) = f(x_i^l)$.

The direct solution of (2.17) would incur one time cost $O(N_l^3)$ floating point operations for a matrix factorization and a cost of $O(N_l^2)$ for each subsequent solve. For large problems, such as those considered in [15] and [19], this cost may not be negligible when compared to the $O(N_{L+1}N_L)$ cost for the evaluation of r^L. Moreover it may not be possible to store a matrix representation of $\mathcal{K}_l$. For these reasons we prefer an iterative method such as GMRES [26] which can give a sufficiently accurate solution at a cost of $O(N_l^2)$ floating point operations for each solve. This clearly can be neglected when compared to the cost of the computation of r^L.

The Atkinson-Brakhage approach requires the same computational effort that we require and an additional evaluation of the action of K_{L+1} on a vector. If we let C_m^n be the cost in floating point operations of the application of K_m (or $\mathcal{K}_m$) to a function in V^n, our approach is dominated by the C_{L+1}^L cost of the computation of r^L. The Atkinson-Brakhage iteration incurs the same cost and must also compute

$$K_{L+1} r^L$$

in order to compute the action of B_l^{L+1} on r^L. This adds a cost of C_{L+1}^{L+1} floating point operations. Hence our approach reduces the cost for each iterate by at least a factor of two. If the cost of an evaluation of the action of K_L on a function is $O(N_L^2)$, $N_m = 2N_{m-1}$ for all m, and the coarse mesh is fine enough so that only one iterate is needed at each mesh level, then the total cost of an iteration based on the algorithm proposed here is roughly $4/3$ that of a K_L evaluation at the finest mesh. This analysis is also valid for nonlinear problems. Contrast this with the estimates, also valid for nonlinear problems, of $8/3$ for the Atkinson-Brakhage approach and $7/3$ for the multigrid method of the second kind from [11].

Finally we remark that if we use $(I - K_l)^{-1}$ as an approximate inverse, a possibility that was investigated in [2], we would use pointwise evaluation (based on P_l and (2.18)) as the fine-to-coarse intergrid transfer. This leads to two step linear convergence for l sufficiently large and is no more efficient than the Atkinson-Brakhage algorithm as described in § 1. However, use of $(I - \mathcal{K}_l)$ as an approximate inverse will give one step linear convergence for sufficiently large l by Lemma 2.1. The finite dimensional equation to be solved is exactly the same as for $(I - K_l)^{-1}$ but the fine-to-coarse intergrid transfer is an averaging (based on Π_l and (2.17)).

2.3. Modifications. In this subsection we show how modifications that have been used successfully in the context of the Atkinson-Brakhage algorithm [12], [21], [19], [15], can also be used with the present algorithm.

We begin by considering (2.12). Often it is desirable to replace u^L in (2.12) by a more accurate interpolation, an interpolation that enforces continuity, or a less expensive interpolation into the finer mesh. Nyström interpolation as a coarse-to-fine intergrid transfer for solution information is an example of this. If we denote by I_L^{L+1} an prolongation operator for solution information and replace (2.12) by

$$\tilde{r}^L = I_L^{L+1} u^L - K_{L+1} I_L^{L+1} u^L - g \tag{2.19}$$

convergence will not be affected if there are C_I and $\sigma_L \to 0$ such that for all L and $u \in V^L$

$$\|I_L^{L+1} u - u^*\|_\infty \le C_I(\|u - u^*\| + \sigma_L) \tag{2.20}$$

For example, our assumptions imply that (2.20) is satisfied by Nyström interpolation,

$$I_L^{L+1} u = g + \mathcal{K}_L u,$$

with $\sigma_L = \tau_L$, where τ_L is given by (2.7).

In many cases, such as the very large problems considered in [15], there is no sufficient storage to solve even the coarse mesh problems by a direct method. In that event we replace the the computation of $\Pi_l w$ from (2.14) by any $\tilde{w} \in V^l$ that satisfies

$$\|\tilde{w} - \mathcal{K}_l \tilde{w} - \Pi_l \tilde{r}^L\|_{L^2} \le \rho_l^L \|\Pi_l \tilde{r}^L\|_{L^2}, \tag{2.21}$$

which is a typical termination criterion for Krylov methods such as GMRES [26]. If ρ_l^L remains bounded away from zero then the norm convergence of $\mathcal{K}_l$ implies that (2.21) will be satisfied after a number of GMRES iterations that is independent of L [24] and therefore the cost of the approximate solution of (2.14) is $O(N_l^2)$ floating point operations. The L^2 norms may be computed exactly in terms of the quadrature weights and nodes at level l. (2.21) implies that $\tilde{w}$ is an L^2-norm accurate approximation to $\Pi_l w$ and can be used in (2.15) to obtain an L^∞-norm accurate approximation to w.

As was pointed out in [18], the Nyström interpolation implicit in (2.15) may be replaced by any interpolation that converges strongly to the identity. This may be necessary in cases, such as those considered in [15] for which an explicit representation of the kernel of the integral operator is not available, or convenient for enforcement of continuity. In [15] we used a piecewise linear interpolation Q_m which we now define. For a given m we set $x_j = x_j^m$, $N = N_m$, and order the quadrature nodes as

$$x_1 < x_2 < \ldots < x_N.$$

For a given m, $x \in [0,1]$, and $j = 1, \ldots, N-1$,define

$$h_j = x_{j+1} - x_j \text{ and } l_j(x) = (x - x_j),$$

If $u \in C[0,1]$ we define $Q_m u$ to be the piecewise linear function given by

$$Q_m u(x) = \begin{cases} u(x_1), & 0 \le x \le x_1, \\ (-l_{j+1}(x)u(x_j) + l_j(x)u(x_{j+1}))/h_j^m, & x_j \le x \le x_{j+1}, \\ & j = 1, \ldots, N-1, \\ u(x_N), & x_N \le x \le 1. \end{cases} \tag{2.22}$$

The iteration that incorporates these changes is given by

$$\begin{aligned} &\tilde{r}^L = I_L^{L+1} u^L - K_{L+1} I_L^{L+1} u^L - g \\ &\text{Find } \tilde{w} \in V^L \text{ such that } \|\tilde{w} - \mathcal{K}_l \tilde{w} - \Pi_l \tilde{r}^L\|_{L^2} \le \rho_l^L \|\Pi_l \tilde{r}^L\|_{L^2} \\ &u^{L+1} = u^L - \tilde{r}^L + Q_l \mathcal{K}_l \tilde{w}. \end{aligned} \tag{2.23}$$

Q_l may be replaced by any map such that $Q_l \mathcal{K}_l \to K$.

We summarize these observations as a corollary of Theorem 2.2.

Corollary 2.3. *Assume that k is continuous, that $I - K$ is nonsingular, and that the sequence of composite Gauss rules is such that* (1.2) *holds, $p_m \le P$ for some P, and that $V^m \subset V^{m+1}$. Let Q_m be any map such that $Q_m \mathcal{K}_m \to K$ in $\mathcal{L}(L^\infty)$. Let I_L^{L+1} satisfy* (2.20). *Then for l sufficiently large and ρ sufficiently small, the iteration given by* (2.23) *converges uniformly to $u^* = (I - K)^{-1} g$.*

2.4. Nonlinear Problems. We can apply the ideas in this section to nonlinear integral equations of the form

$$u(x) = T(u)(x) = \Phi\big(x, \Psi(u)(x)\big), \tag{2.24}$$

where Ψ is a nonlinear integral operator of the form

$$\Psi(u)(x) = \int_0^1 \psi(x, y, u(y))\, dy.$$

We make the standard assumptions from nonlinear equations.

Assumption 2.1. *Here Φ is continuous and Lipschitz continuously differentiable in its second argument and ψ is continuous and Lipschitz continuously differentiable in its third argument. There is a solution $u^* \in C[0,1]$ to* (2.24) *and $I - T'(u^*)$ is nonsingular.*

Note that the differentiability assumptions on Φ and ψ imply that T is Fréchet differentiable in L^∞ and $C[0,1]$.

We approximate T at level m by

$$T_m(u)(x) = \Phi\big(x, \Psi_m(u)(x)\big)$$

where

$$\Psi_m(u)(x) = \sum_{j=1}^{N_m} \psi(x, x_j^m, u(x_j^m)) w_j^m.$$

The analog of the iteration (2.23), of which (2.6) is a special case, is an approximate Newton iteration. Here the role of the kernels k_l is played by partial derivatives of ψ. The Fréchet derivative of T_m is defined for $u, v \in V^m \cup C[0,1]$ by

$$T'_m(u)(v)(x) = \Phi_2\big(x, \Psi_m(u)(x)\big) \sum_{j=1}^{N_m} \psi_3(x, x_j^m, u(x_j^m)) v(x_j^m) w_j^m. \tag{2.25}$$

Here Φ_2 and ψ_3 denote differentiation with respect to the second argument in Φ and the third in ψ.

For a fixed $u \in C[0,1]$, $\{T_m(u)\}$ is a collectively compact strongly convergent sequence of operators and the classical Atkinson-Brakhage algorithm could be applied to the linear equations for Newton steps. We now define a norm convergent sequence $\{\mathcal{T}_m(u)\}$ for $u \in C[0,1] \cup V^l$. For $u, v \in C[0,1] \cup V^l$

$$\mathcal{T}_m(u)(v)(x) = \int_0^1 \psi_{3,m}(x, y; u) v(y)\, dy$$

where $\psi_{3,m}(x,y;u)$ is the piecewise polynomial on $[0,1]\times[0,1]$ defined for $x \in \mathrm{Int}(I_r^m)$ and $y \in \mathrm{Int}(I_s^m)$ by

$$\begin{aligned} &\psi_{3,m}(x,y;u) = \\ &\textstyle\sum_{x_i^m \in I_r^m, x_j^m \in I_s^m} \Phi_2\big(\big(x_i^m, \Psi_m(u)(x_i^m)\big) \psi_3(x_i^m, x_j^m, u(x_j^m)) l_i^r(x) l_j^s(y). \end{aligned} \tag{2.26}$$

The nonlinear analog of the iteration (2.23) is

$$\begin{aligned} &\tilde r^L = T_{L+1}(I_L^{L+1} u^L) \\ &\text{Find } \tilde w \in V^L \text{ such that } \|\tilde w - \mathcal{T}_l(u^l)\tilde w - \Pi_l r^L\|_{L^2} \le \rho_l^L \|\Pi_l r^L\|_{L^2} \\ &u^{L+1} = u^L - \tilde r^L + Q_l \mathcal{T}_l(u^l)\tilde w. \end{aligned} \tag{2.27}$$

Here we require that $Q_m \mathcal{T}_m(u) \to T'(u)$ in norm uniformly in u for u sufficiently near u^*. We require that the (possibly nonlinear) intergrid transfers I_L^{L+1} satisfy (2.20). Nonlinear Nyström interpolation will do this.

The convergence of the nonlinear algorithm is also a corollary of the proof of Theorem 2.2 and known results on Newton-like methods [18], [8].

Corollary 2.4. *Let Assumption 2.1 hold. Assume that the sequence of composite Gauss rules is such that* (1.2) *holds,* $p_m \le P$ *for some* P*, and that* $V^m \subset V^{m+1}$*. Let* Q_m *be any map such that* $Q_m \mathcal{T}_m(u) \to T'(u)$ *in norm uniformly in* u *for* u *sufficiently near* u^**. Let* I_L^{L+1} *satisfy* (2.20)*. Then for* l *sufficiently large,* u^l *sufficiently near* u^**, and* ρ *sufficiently small, the iteration given by* (2.27) *converges uniformly to* u^**.*

The final algorithmic modification that one can make in the nonlinear case is to replace exact computation of Φ_2 and ψ_3 by difference operators. The effects of such a change are have been described in [21] in the setting of compact fixed point problems and in [27] and [9] in the context of nonlinear equations. If the differencing is done carefully, there is no observable change in the convergence rates if the nonlinearities are not too severe. The papers cited above offer a complete analysis that carries over to the case here. In the nonlinear example in § 3 we computed the action of $\mathcal{T}_m(u)$ on a function with forward differences.

Finally we remark that no changes are needed in the analysis or algorithms to solve systems of integral equations, rather than single equations, over compact sets $\Omega \subset R^N$ rather than $[0,1]$.

3. Numerical Results. In the computations in this section the quadrature rules were composite midpoint rules on $[0,1]$ with $N_m = q_m = 2^m$,

$$x_i^m = (i - 1/2)/N_m, \text{ and } w_i^m = 1/N_m.$$

We report on two examples, one linear and the other nonlinear. All computations were performed on a SUN SPARCstation 1+ running SUN OS 4.1 in SUN FORTRAN version 1.3.1.

Coarse mesh linear equations were solved with GMRES using a modification of the Brown-Hindmarsh GMRES code [6] with changes made in the inner product to allow the approximate L^2 inner product to be used instead of the R^N inner product. We terminated the GMRES iteration when the relative residual was below 10^{-2}. Intergrid transfers of solution information (I_L^{L+1}) were done by Nyström interpolation in all cases. We terminated the iteration on the coarse mesh l when the residual was below 10^{-4}. For subsequent iterates we terminated when the residual had been reduced by a factor of $\mu = .25$, the expected reduction from the accuracy of the quadrature rule. In all cases a single iterate was required at levels $L > l$.

As the iteration progresses we tabulate, for $L \geq l$, the number of fine mesh points N_L, the iteration counter i, the number of GMRES iterations required for that iteration i_g, the norm of the residual R_i, and, for $i > 0$, the ratio of successive residuals at the current mesh level.

Our first computation is on the linear problem

$$u(x) - \frac{1}{\lambda}\int_0^1 \frac{u(y)\,dy}{1+(x-y)^2} = g(x)/\lambda. \tag{3.1}$$

g was set so that the solution was $u(x) = 1$. We report on two computations, one with $\lambda = 1$, an easy problem, and the other with $\lambda = .01$. The more difficult problem required a finer coarse mesh for convergence and the convergence from level to level took longer to settle down. However the performance of the approximate inverse was still good.

For a nonlinear example we consider the Chandrasekhar H-equation [7]

$$H(x) = \left(I - \frac{c}{2}\int_0^1 \frac{xH(y)}{x+y}\,dy\right)^{-1}. \tag{3.2}$$

We computed derivatives by forward differences.

The nonlinear equation has two solutions for $0 < c < 1$, only one of which has physical meaning. That solution will naturally be found by a preconditioned Richardson iteration [16] which preserves analyticity in the parameter c. There is a simple fold singularity [23], [17], [13], and therefore a singular Fréchet derivative at the solution when $c = 1$. The solution has a logarithmic singularity at $x = 0$ and hence only first order accuracy is observed. We report on two computations, one for $c = .5$ and the other at $c = .99$ near the singularity.

TABLE 3.1
Linear equation, $\lambda = 1$.

N	i	i_g	R_i	R_i/R_{i-1}
8	0		0.18D+00	
	1	2	0.45D-03	0.25D-02
	2	1	0.13D-05	0.29D-02
16	0		0.63D-03	
	1	2	0.64D-04	0.10D+00
32	0		0.15D-03	
	1	2	0.23D-04	0.15D+00
64	0		0.38D-04	
	1	2	0.68D-05	0.18D+00
128	0		0.95D-05	
	1	2	0.18D-05	0.19D+00
256	0		0.24D-05	
	1	2	0.47D-06	0.20D+00
512	0		0.60D-06	
	1	2	0.12D-06	0.20D+00
1024	0		0.15D-06	
	1	2	0.30D-07	0.20D+00

TABLE 3.2
Linear equation, $\lambda = .01$.

N	i	i_g	R_i	R_i/R_{i-1}
16	0		0.92D+02	
	1	1	0.23D-01	0.25D-03
	2	3	0.41D-04	0.18D-02
32	0		0.14D-01	
	1	3	0.21D-02	0.15D+00
64	0		0.25D-01	
	1	3	0.10D-02	0.41D-01
128	0		0.14D-01	
	1	3	0.49D-03	0.36D-01
256	0		0.71D-02	
	1	2	0.21D-03	0.29D-01
512	0		0.30D-02	
	1	3	0.16D-03	0.53D-01
1024	0		0.18D-02	
	1	2	0.32D-04	0.18D-01

TABLE 3.3
H-equation, $c = .5$.

N	i	i_g	R_i	R_i/R_{i-1}
8	0		0.10D+01	
	1	2	0.46D-01	0.46D-01
	2	2	0.42D-04	0.91D-03
16	0		0.15D-02	
	1	2	0.13D-04	0.85D-02
32	0		0.74D-03	
	1	2	0.61D-05	0.82D-02
64	0		0.36D-03	
	1	2	0.23D-05	0.62D-02
128	0		0.18D-03	
	1	2	0.75D-06	0.42D-02
256	0		0.90D-04	
	1	2	0.24D-06	0.26D-02
512	0		0.45D-04	
	1	2	0.72D-07	0.16D-02
1024	0		0.23D-04	
	1	2	0.21D-07	0.94D-03

TABLE 3.4
H-equation, $c = .99$.

N	i	i_g	R_i	R_i/R_{i-1}
8	0		0.10D+01	
	1	2	0.34D+00	0.34D+00
	2	2	0.83D-01	0.25D+00
	3	2	0.12D-01	0.14D+00
	4	2	0.44D-03	0.37D-01
	5	3	0.72D-06	0.16D-02
16	0		0.31D-02	
	1	3	0.76D-04	0.24D-01
32	0		0.15D-02	
	1	3	0.39D-04	0.26D-01
64	0		0.73D-03	
	1	3	0.15D-04	0.21D-01
128	0		0.36D-03	
	1	3	0.53D-05	0.15D-01
256	0		0.18D-03	
	1	3	0.17D-05	0.97D-02
512	0		0.90D-04	
	1	3	0.55D-06	0.62D-02
1024	0		0.45D-04	
	1	3	0.17D-06	0.38D-02

REFERENCES

[1] P. M. Anselone, *Collectively Compact Operator Approximation Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1971.

[2] K. E. Atkinson, *Iterative variants of the Nyström method for the numerical solution of integral equations*, Numer. Math., 22 (1973), pp. 17–31.

[3] ———, *A survey of numerical methods for solving nonlinear integral equations*, Journal of Integral Equations and Applications, 4 (1992), pp. 15–46.

[4] C. Baker, *The Numerical Treatment of Integral Equations*, Oxford University Press, Oxford, 1977.

[5] H. Brakhage, *Über die numerische Behandlung von Integralgleichungen nach der Quadraturformelmethode*, Numer. Math., 2 (1960), pp. 183–196.

[6] P. N. Brown and A. C. Hindmarsh, *Reduced storage matrix methods in stiff ODE systems*, Tech. Report UCRL-95088, Lawrence Livermore National Laboratory, 1987.

[7] S. Chandrasekhar, *Radiative Transfer*, Dover, New York, 1960.

[8] R. Dembo, S. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400–408.

[9] J. E. Dennis and H. F. Walker, *Inaccuracy in quasi-Newton methods: Local improvement theorems*, in Mathematical Programming Study 22: Mathematical programming at Oberwolfach II, North–Holland, Amsterdam, 1984, pp. 70–85.

[10] H. Bückner, *Numerical method for integral equations*, in Survey of Numerical Analysis, J. Todd, ed., New York, 1962, McGraw Hill, pp. 439–465.

[11] W. Hackbusch, *Multigrid methods of the second kind*, in Multigrid Methods for Integral and Differential Equations, Oxford University Press, Oxford, 1985.

[12] M. Heinkenschloss, C. T. Kelley, and H. T. Tran, *Fast algorithms for nonsmooth compact fixed point problems*, SIAM J. Numer. Anal., 29 (1992), pp. 1769–1792.

[13] H. B. Keller, *Lectures on Numerical Methods in Bifurcation Theory*, Tata Institute of Fundamental Research, Lectures on Mathematics and Physics, Springer-Verlag, New York, 1987.

[14] H. B. Keller and E. Isaacson, *Analysis of numerical methods*, Wiley, New York, 1966.

[15] C. T. Kelley, *Multilevel source iteration accelerators for the linear transport equation in slab geometry*. submitted for publication.

[16] ———, *Solution of H-equations by iteration*, SIAM J. Math. Anal., 10 (1979), pp. 844–849.

[17] ———, *Solution of the Chandrasekhar H-equation by Newton's method*, J. Math. Phys., 21 (1980), pp. 1625–1628.

[18] ———, *Operator prolongation methods for nonlinear equations*, in Computational Solution of Nonlinear Systems of Equations, E. L. Allgower and K. Georg, eds., vol. 26 of AMS Lectures in Applied Mathematics, American Mathematical Society, Providence, RI, 1990, pp. 359–388.

[19] C. T. Kelley and E. W. Sachs, *Multilevel algorithms for constrained compact fixed point problems*. SIAM J. on Sci. Comp., to appear.

[20] ———, *Multilevel algorithms for constrained optimal control problems*. proceedings of Copper Mountain Conference on Iterative Methods, Copper Mountain, Co., 1992.

[21] ———, *Fast algorithms for compact fixed point problems with inexact function evaluations*, SIAM J. on Sci. Stat. Comp., 12 (1991), pp. 725–742.

[22] E. W. Larsen and P. Nelson, *Finite difference approximations and superconvergence for the discrete ordinate equations in slab geometry*, SIAM J. Numer. Anal., 19 (1982), pp. 334–348.

[23] T. W. Mullikin, *Some probability distributions for neutron transport in a half space*, J. Appl. Prob., 5 (1968), pp. 357–374.

[24] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778–795.

[25] J. Pitkäranta and R. Scott, *Error estimates for the combined spatial and angular approximations of the transport equation in slab geometry*, SIAM J. Numer. Anal., 20 (1983), pp. 922–950.

[26] Y. Saad and M. Schultz, *GMRES a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 7 (1986), pp. 856–869.

[27] T. J. Ypma, *The effect of rounding errors on Newton-like methods*, IMA J. Numer. Anal., 3 (1983), pp. 109–118.