A HISTORICAL OVERVIEW OF ITERATIVE METHODS

by

David M. Young

August 1988

CNA-225

An invited talk based on this paper will be presented at the Workshop on Practical Iterative Methods for Large Scale Computation to be held on October 23-25 at the Minnesota Supercomputer Institute, The University of Minnesota, Minneapolis, Minnesota.
A Historical Overview of Iterative Methods

David M. Young
Center for Numerical Analysis,
The University of Texas at Austin

Abstract

The object of this paper is to present a historical overview of the development of iterative methods for the solution of large sparse systems of linear equations. The emphasis is on methods which are applicable to linear systems arising in the numerical solution of partial differential equations. Aspects to be covered include: early work, including the methods of L. F. Richardson and of Liebmann as well as relaxation methods used by Southwell and others; the SOR method and extensions such as block SOR methods, and p-cyclic matrices; Chebyshev polynomial methods; alternating direction implicit methods; the SSOR method; approximate matrix factorization methods including the strongly implicit method (SIP) and the incomplete Cholesky method (ICC); fast direct methods; conjugate gradient methods; adaptive methods for the automatic determination of iteration parameters; multigrid methods; methods for nonsymmetric systems; and iterative software. Future developments will be discussed with emphasis on the use of vector and parallel processors.

†The research was supported in part by the National Science Foundation under Grant DCR-8518722 and by the Department of Energy under Grant DE-FG05-87ER25048 with The University of Texas at Austin.
1. Introduction

The purpose of this paper is to present a historical overview of iterative methods for solving systems of linear algebraic equations of the form

\[(1.1) \quad Au = b\]

where \(A\) is a given nonsingular square matrix and \(b\) is a given column vector. We are primarily interested in the case where the coefficient matrix \(A\) is very large and very sparse. Such systems frequently arise from the numerical solution of partial differential equations by finite difference methods or by finite element methods.

For the benefit of those readers who are not familiar with iterative methods a brief review of some basic iterative methods such as the Jacobi, Gauss-Seidel, SOR methods, etc., is given in Section 2. Then a brief chronological review of the development of iterative methods is given in Section 3. Following this, individual topics will be discussed in more detail in the later sections. These topics include: the SOR method and related methods; Chebyshev methods; conjugate gradient methods; alternating direction implicit methods; the SSOR method; algorithms and software; and nonsymmetric systems. A final section will include a brief guide to the literature.

Since the paper is intended to correspond roughly to a one hour talk, it is clear that the discussion cannot be in any sense complete and authoritative. The topics discussed and the amount of emphasis devoted to each
will be based on my own perspective and biases and will be somewhat ar-
bbitrary. Forgiveness is requested in advance from anyone whose work has
been misrepresented or slighted.

Any comments and suggestions would be welcomed. Perhaps before too
long a very comprehensive history of iterative methods will be written with
a depth of coverage similar to that of the excellent review prepared recently
2. Some Basic Iterative Methods

In this section we give a brief review of iterative methods. For further details see, e.g., Young [1971], Chapter 3. To illustrate the methods we consider the following model problem: find the solution of the Poisson equation

\[(2.1) \quad u_{xx} + u_{yy} = -1\]

in the unit square \(0 \leq x \leq 1, \ 0 \leq y \leq 1\) subject to the condition \(u = 0\) on the boundary. The region is covered by a grid with grid size \(h = M^{-1}\) for some integer \(M\). The 5-point difference equation

\[(2.2) \quad h^{-2}[u(x + h, y) + u(x - h, y) + u(x, y + h) + u(x, y - h) - 4u(x, y)] = -1\]

is used for each interior grid point. For the case \(h = \frac{1}{3}\) and with the grid points shown in Figure 2.1, we obtain, after multiplying by \(-h^2\), the linear system

\[(2.3) \quad \begin{cases} 
4u_1 - u_2 - u_3 = \frac{1}{9} \\
- u_1 + 4u_2 - u_4 = \frac{1}{9} \\
- u_1 + 4u_3 - u_4 = \frac{1}{9} \\
- u_2 - u_3 + 4u_4 = \frac{1}{9}
\end{cases}\]

This system can be written in the matrix form

\[(2.4) \quad \begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\]
or, equivalently, in the form

\[(2.5) \quad Au = b\]

To define several standard iterative methods, we rewrite the system \[(2.4)\) in the form

\[(2.6) \quad \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} \frac{1}{36} \\ \frac{1}{36} \\ \frac{1}{36} \\ \frac{1}{36} \end{bmatrix}\]

or, in matrix form

\[(2.7) \quad u = Bu + c\]

The Jacobi method is defined by

\begin{align*}
(2.8) \quad 
\begin{cases}
  u_1^{(n+1)} = \frac{1}{4}u_2^{(n)} + \frac{1}{4}u_3^{(n)} + \frac{1}{36} \\
  u_2^{(n+1)} = \frac{1}{4}u_1^{(n)} + \frac{1}{4}u_4^{(n)} + \frac{1}{36} \\
  u_3^{(n+1)} = \frac{1}{4}u_1^{(n)} + \frac{1}{4}u_4^{(n)} + \frac{1}{36} \\
  u_4^{(n+1)} = \frac{1}{4}u_2^{(n)} + \frac{1}{4}u_3^{(n)} + \frac{1}{36}
\end{cases}
\end{align*}
The Gauss-Seidel method is the same as the Jacobi method except that one uses new values as soon as available. Thus we have

\[
\begin{align*}
    u_1^{(n+1)} &= \frac{1}{4} u_2^{(n)} + \frac{1}{4} u_3^{(n)} + \frac{1}{36} \\
    u_2^{(n+1)} &= \frac{1}{4} u_1^{(n+1)} + \frac{1}{4} u_4^{(n)} + \frac{1}{36} \\
    u_3^{(n+1)} &= \frac{1}{4} u_1^{(n+1)} + \frac{1}{4} u_4^{(n)} + \frac{1}{36} \\
    u_4^{(n+1)} &= \frac{1}{4} u_2^{(n+1)} + \frac{1}{4} u_3^{(n+1)} + \frac{1}{36}
\end{align*}
\]

The successive overrelaxation method (SOR) is defined by

\[
\begin{align*}
    u_1^{(n+1)} &= \omega \left\{ \frac{1}{4} u_2^{(n)} + \frac{1}{4} u_3^{(n)} + \frac{1}{36} \right\} + (1 - \omega) u_1^{(n)} \\
    u_2^{(n+1)} &= \omega \left\{ \frac{1}{4} u_1^{(n+1)} + \frac{1}{4} u_4^{(n)} + \frac{1}{36} \right\} + (1 - \omega) u_2^{(n)} \\
    u_3^{(n+1)} &= \omega \left\{ \frac{1}{4} u_1^{(n+1)} + \frac{1}{4} u_4^{(n)} + \frac{1}{36} \right\} + (1 - \omega) u_3^{(n)} \\
    u_4^{(n+1)} &= \omega \left\{ \frac{1}{4} u_2^{(n+1)} + \frac{1}{4} u_3^{(n+1)} + \frac{1}{36} \right\} + (1 - \omega) u_4^{(n)}
\end{align*}
\]

Here \( \omega \) is a parameter, called the relaxation factor. If \( \omega = 1 \) we get the Gauss-Seidel method.

The symmetric SOR method (SSOR method) is defined as follows. Given \( u^{(n)} \), compute \( u^{(n + \frac{1}{2})} \) using the SOR method and then compute \( u_4^{(n+1)} \), \( u_3^{(n+1)} \), \( u_2^{(n+1)} \) and \( u_1^{(n+1)} \) based on the \( \{ u_i^{(n+\frac{1}{2})} \} \), taking the equations in reverse order (a backward sweep).

We now consider the more general class of basic (one-step) iterative methods of the form

\[
u^{(n+1)} = G u^{(n)} + k
\]

Such a method can be defined by a splitting of the matrix \( A \) of the form

\[
A = Q - (Q - A)
\]
where $Q$ is an easily inverted matrix. The basic iterative method is defined by

\begin{equation}
Q u^{(n+1)} = (Q - A)u^{(n)} + b
\end{equation}

which is equivalent to (2.12) with

\begin{align}
G &= I - Q^{-1}A \\
k &= Q^{-1}b
\end{align}

Let us represent $A$ in the form

\begin{equation}
A = D - C_L - C_U
\end{equation}

where $D$ is a diagonal matrix and $C_L$ and $C_U$ are strictly lower and strictly upper triangular matrices, respectively. The standard methods defined above can also be defined by the following matrix splittings:

- **Jacobi** $Q = D$
- **Gauss-Seidel** $Q = D - C_L$
- **SOR** $Q = \frac{1}{\omega}D - C_L$
- **SSOR** $Q = \frac{\omega}{2 - \omega}(\frac{1}{\omega}D - C_L)D^{-1}(\frac{1}{\omega}D - C_U)$

We note that for the basic iterative method (2.12) corresponding to the splitting (2.13) the linear system

\begin{equation}
(I - G)u = k
\end{equation}

or, equivalently, the system

\begin{equation}
Q^{-1}Au = Q^{-1}b
\end{equation}
is called the preconditioned system.

We mention here one iterative method defined by a splitting, namely, the incomplete Cholesky method. Here \( Q = LU \) where \( L \) and \( U \) are lower and upper triangular matrices with the same sparsely as \( A \) such that \((A - LU)_{i,j} = 0 \) if \( a_{i,j} \neq 0 \). Thus for the matrix \( A \) of (2.4), we have

\[
(2.18) \quad A \approx \begin{bmatrix}
1 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 \\
-\frac{1}{2} & 0 & 1 & 0 \\
0 & -\frac{4}{15} & -\frac{4}{16} & 1
\end{bmatrix}
\begin{bmatrix}
4 & -1 & -1 & 0 \\
0 & \frac{15}{4} & 0 & -1 \\
0 & 0 & \frac{15}{4} & -1 \\
0 & 0 & 0 & -\frac{52}{15}
\end{bmatrix}
\]

Note that

\[
(2.19) \quad R = A - LU = 
\begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{4} & 0 \\
0 & \frac{1}{4} & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

All of the elements of \( R \) vanish except possibly for \((i,j)\) such that \( a_{i,j} = 0 \).

The rate of convergence of a basic iterative method depends in the spectral radius \( S(G) \) of the iteration matrix \( G \). Roughly speaking, the error is reduced each iteration by a factor of \( S(G) \). To reduce the error by a factor of \( 10^{-6} \) we must use \( n^* \) iterations where

\[
(2.20) \quad (S(G))^{n^*} \leq 10^{-6}
\]

or

\[
(2.21) \quad n^* \geq \frac{\log 10^6}{-\log S(G)}
\]

The quantity \( -\log S(G) \) is referred to as the rate of convergence.
3. A Brief Chronological Review

In this section we will give a brief chronological review of some of the work which has been done on iterative methods. We also give a table listing certain “milestones” in the field. Some of the topics covered will be treated in more detail in later sections of this paper.

In view of the intensive work on iterative methods which has taken place over the past 40 years, it is perhaps hard to understand why there was relatively little progress made prior to the late 1940s. This can probably be explained, at least in part, by the fact that high speed computers did not begin to appear for general use until the late 1940s and the early 1950s.

A review of the work done on iterative methods prior to 1949 was given by Geiringer [1949]. The following methods were discussed: the Jacobi method, (Jacobi [1845]), the Gauss-Seidel method (Gauss [1823] and Seidel [1874]), Richardson's method (Richardson [1910]), the method of steepest descent and relaxation methods, (Southwell [1946]). The Jacobi and Gauss-Seidel methods were referred to by Geiringer as the “method of simultaneous displacements” and the “method of successive displacements,” respectively. Richardson's method is defined by

\[ u^{(n+1)} = u^{(n)} + \gamma_{n+1}(b - Au^{(n)}) \]

where \( \gamma_1, \gamma_2, \ldots \) are parameters which are frequently referred to as “extrapolation factors.” They are used in order to improve the convergence of the method.
Liebmann [1918] applied the Gauss-Seidel method to solve 5-point difference equations based on Laplace's equation. Hence the Gauss-Seidel method is sometimes referred to as the "Liebmann method."

Shortley, Weller and Fried [1940] and Shortley and Weller [1938] gave an analysis of the convergence properties of the Liebmann method in terms of the eigenvalues, eigenvectors, and principal vectors of the associated iterative matrix. This was done both for point methods and for block methods.

Another important early paper concerning iterative methods is the paper of Temple [1938-39]. In this paper it was shown that if the coefficient matrix $A$ of (1.1) is SPD then the problem of solving (1.1) is equivalent to the problem of minimizing the quadratic form

$$F(u) = \frac{1}{2}(u, Au) - (b, u)$$

(3.2)

Among the several iterative methods based on this idea is the method of steepest descent. Another is the conjugate gradient method discussed in Section 6.

The main tools for obtaining numerical solutions to actual problems were the relaxation methods which were used by Southwell [1946] and his associates. Relaxation methods were applied to the solution of linear systems arising from the solution of partial differential equations by finite difference methods. Given an initial approximation $u^{(0)}$ to the true solution one would record the approximate value of the solution at each grid
point together with the component of the residual vector, $b - A u^{(0)}$, at the point. By inspection of the residuals, one would introduce displacements at one or more grid points and appropriately modify the residuals. The process was continued until all of the residuals became small. The solution at each grid point was the original value corrected by the sum of all of the increments which had been introduced at that point. With this process the increments were introduced at the various grid points in an irregular manner. The Gauss-Seidel method can be regarded as a relaxation process wherein the increments are introduced in a systematic manner.

It should be noted that use of the relaxation methods suggested many ideas, such as overrelaxation, or overcorrection, and block displacement, which have proved useful for later systematic methods.

Beginning in the late 1940s there was intensive activity in the field of iterative methods. This included the development and analysis of the SOR method as well as a number of modifications of the method and extensions of the theory. There was a parallel development of Chebyshev methods starting with a paper of Flanders and Shortley [1950] on the application of Chebyshev polynomials of the matrix eigenvalue problem. This led to the development of Chebyshev acceleration procedures, or to "semi-iterative methods" they were then called. Another important development was the conjugate gradient method developed by Hestenes and Stiefel [1952]. In 1955 the SSOR method was developed by Sheldon [1955]. At about the same
time alternating direction implicit methods (ADI methods) were developed by Peaceman and Rachford [1955] and by Douglas [1955]. Further discussions of the SOR method, Chebyshev methods, the conjugate gradient method, the SSOR method, and ADI methods are given in Sections 4 – 8.

A new class of iterative methods, known as “strongly implicit methods” were introduced by Stone [1968] and others in the mid 1960s. These methods led to approximate factorization of the coefficient matrix for a linear system arising from the solution of a partial difference equation. They were a forerunner of incomplete matrix factorization methods, developed by Meijerink and Van der Vorst [1977] and others, which are applicable to more general matrices.

The basic idea of strongly implicit methods is to represent a five-point difference operator involving the five points \((x, y), (x \pm h, y)\) and \((x, y \pm h)\) as the approximate product of two three-point operators — one involving the points \((x - h, y), (x, y - h)\) and \((x, y)\) and the other involving the points \((x + h, y), (x, y + h)\) and \((x, y)\). The former leads to a lower triangular matrix and the latter leads to an upper triangular matrix. The product of the three-point operators can be made to agree with the five-point operator at the five points, though this is not always done. Various parameters are often introduced to speed the convergence.

Another important development in the mid 1960s was the introduction of fast direct methods by Hockney [1965] and others. Subsequent work
was done by Buzbee, Golub and Nielson [1970], Buneman [1969], Schwart-
strauber [1974], Sweet [1974] and many others. Proskurowski and Widland
[1976] used capacitance matrix techniques based on fast direct methods for
rectangular regions, to obtain solutions for more general regions.

The mid and late 1960s marked the beginning of the resurrection of the
conjugate gradient method which had been seldom used since its auspicious
introduction in the early 1950s. Papers by Daniel [1965, 1967] and by Reid
[1971, 1972] sparked the resurgence. A number of papers appeared later,
including papers of Axelsson [1974] and Concus, Golub and O'Leary [1976].

Multigrid methods were introduced by Brandt [1977]. Basically, with
multigrid methods one tries to speed up the convergence of an iterative
method on a given grid by carrying out certain auxiliary calculations in-
volving iterations on a series of coarser grids. Multigrid methods have been
the focus of considerable research activity ever since Brandt's paper ap-
peared. An excellent recent reference is the book edited by McCormick
[1987]. For an introduction to the subject, the tutorial by Briggs [1987] is
also recommended.

Beginning with the mid 1970s, there were increased efforts to develop
iterative algorithms, as opposed to iterative methods, which could be in-
corporated into portable software. Software packages, such as the ITPACK
2C software package were developed which included automatic procedures
for determining any necessary iteration parameters and also included proce-
dures for using accurate and realistic tests for stopping the iterative process. Iterative algorithms and software are discussed further in Section 9.

Since the mid 1970s there has been intensive activity on methods for solving nonsymmetric systems. This activity was sparked by papers by Vinsome [1977], Concus and Golub [1976] and by Widlund [1978]. Various procedures for solving nonsymmetric systems are described in Section 10.

With the advent of vector computers in the early 1980s and, more recently, of parallel computers, much of the current research activity on iterative methods has been focused on the use of advanced computer architectures. The paper by Ortega and Voigt [1985] and the recent book by Ortega [1988] are recommended for anyone wishing to start work in this area.
Some Milestones in the Development of Iterative Methods

The following table lists some of the "milestones" in the development of iterative methods. While the choice of the milestones is somewhat arbitrary, it is hoped that the list will give the reader some picture of how the subject has developed.

Table 3.1

<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1823</td>
<td>Gauss</td>
<td>Gauss-Seidel method</td>
</tr>
<tr>
<td>1845</td>
<td>Jacobi</td>
<td>Jacobi method</td>
</tr>
<tr>
<td>1874</td>
<td>Seidel</td>
<td>Gauss-Seidel method</td>
</tr>
<tr>
<td>1910</td>
<td>Richardson</td>
<td>Richardson's method</td>
</tr>
<tr>
<td>1918</td>
<td>Liebmann</td>
<td>Gauss-Seidel method for Laplace's equation</td>
</tr>
<tr>
<td>1938-39</td>
<td>Temple</td>
<td>Method of Steepest Descent</td>
</tr>
<tr>
<td>1939</td>
<td>Shortley and Weller</td>
<td>Analysis of the Liebman method</td>
</tr>
<tr>
<td>1945</td>
<td>Southwell</td>
<td>Book on relaxation methods</td>
</tr>
<tr>
<td>1950</td>
<td>Young, Frankel</td>
<td>Analysis of the SOR method</td>
</tr>
<tr>
<td>1950</td>
<td>Flanders, Shortley</td>
<td>Chebyshev methods for the matrix eigenvalue problem</td>
</tr>
<tr>
<td>1952-54</td>
<td>Shortley, Lanczos, Young</td>
<td>Use of Chebyshev methods for linear systems</td>
</tr>
<tr>
<td>1932</td>
<td>Hestenes, Stiefel</td>
<td>Conjugate gradient method</td>
</tr>
<tr>
<td>1954</td>
<td>Arms, Gates, Zondek</td>
<td>Block SOR</td>
</tr>
<tr>
<td>1955</td>
<td>Sheldon</td>
<td>SSOR method</td>
</tr>
<tr>
<td>1957</td>
<td>Varga</td>
<td>Semi-iterative methods</td>
</tr>
<tr>
<td>1958</td>
<td>Kahan</td>
<td>Extension of the SOR theory</td>
</tr>
<tr>
<td>1959</td>
<td>Blair et al [1959]</td>
<td>Chebyshev acceleration</td>
</tr>
<tr>
<td>1961</td>
<td>Golub, Varga</td>
<td>Chebyshev methods and relation to second-degree and SOR methods</td>
</tr>
<tr>
<td>1965</td>
<td>Hockney</td>
<td>Fast direct methods</td>
</tr>
<tr>
<td>1968</td>
<td>Stone; Dupont, Kendall, R</td>
<td>Strongly implicit methods</td>
</tr>
<tr>
<td></td>
<td>achford</td>
<td></td>
</tr>
</tbody>
</table>
1965-72 Daniel, Reid
1974-75 Axelsson, Concus, Golub, O'Leary
1977 Brandt
1977 Meijerink, Manteuffel, Van der Vorst
1978-78 Vinsome, Concus, Golub, Widlund
1977 Manteuffel
1976-85 Rice et al
1976-82 Kincaid et al [1982]
1980 Young, Jea [1980]
1981 Hageman and Young [1981]
1982 Elman [1982]
1985 Ortega and Voigt [1985]
1988 Ortega

Resurgence of the conjugate gradient method
Preconditioned conjugate gradient methods
Multigrid methods
Incomplete Cholesky and Shifted Cholesky method
Nonsymmetric systems
Nonsymmetric Chebyshev
ELLPACK
ITPACK 2C
Generalized conjugate gradient methods
Adaptive procedure for choosing iterative parameters and for terminating iterative processes
Generalized conjugate gradient methods
Survey on work on vector and parallel computers for partial differential equations
Book on use of vector and parallel computers
4. The SOR Method and Related Methods

Young [1950, 1954] presented an analysis of the SOR method for the case where the coefficient matrix of the system (1.1) is consistently ordered. Such a matrix can be obtained from a matrix with Property A by a suitable permutation of the rows and corresponding columns. For a matrix with Property A, one can permute the rows and corresponding columns to obtain a "red-black" matrix, which has the form

\[
A = \begin{bmatrix}
D_R & H \\
K & D_B
\end{bmatrix}
\]

(4.1)

where \( D_R \) and \( D_B \) are square diagonal matrices. It should be noted that any matrix derived from the 5-point difference equation representation of an elliptic partial differential equation has Property A.

If \( A \) is consistently ordered then there is a relation between the eigenvalues of the iteration matrix \( L_\omega \) for the SOR method and the eigenvalues of the iteration matrix \( B \) for the Jacobi method. Moreover, if \( A \) is SPD then the eigenvalues of \( B \) are real and less than one and the optimum value of omega is given by

\[
\omega_b = \frac{2}{1 + \sqrt{1 - S(B)^2}}
\]

(4.2)

Moreover, the spectral radius \( S(L_\omega_b) \) of \( L_\omega_b \) is given by

\[
S(L_\omega_b) = \omega_b - 1
\]

(4.3)

It can be shown that for the model problem, the number of iterations required for convergence using the SOR method is \( O(h^{-1}) \) whereas the
number of iterations required using the Gauss-Seidel method or the Jacobi method is $O(h^{-2})$. Thus the SOR method is faster by an order-of-magnitude.

Frankel [1950] carried out an independent analysis of the SOR method for linear systems derived from the solution of Laplace's equation in the rectangle. He referred to the SOR method as the "extrapolated Liebmann method." Frankel also developed a stationary second-degree method which he called the "second-order Richardson method" and whose rate of convergence for the model problem is of the same order as that of the SOR method. This method will be discussed further in Section 5.

A number of modifications and extensions have been made to the SOR theory. Arms, Gates and Zondek [1956] developed block SOR methods. For block methods the unknowns are grouped into blocks and all values within a block are modified simultaneously using direct methods. Usually the matrices for the block are easily invertible. In the case of line SOR methods the matrices are tri-diagonal. The SOR theory applies for the block SOR method and faster convergence is obtained than for the point SOR method. Cuthill and Varga [1959] showed how the extra work per iteration for line SOR can be essentially eliminated.

Kahan [1958] extended the SOR theory to the case where $A$ is an $L$-matrix (i.e., $a_{ii} > 0$ for all $i$ and $a_{ij} \leq 0$ for $i > j$). He showed that if $\omega_b$ is determined by (4.2) then $\omega_b$, though not optimal, is "good" in the sense
that

\begin{equation}
(4.4) \quad \omega_b - 1 \leq S(L_{\omega_b}) \leq \sqrt{\omega_b - 1}
\end{equation}

Other extensions of the SOR method include \( p \)-cyclic matrices, see Varga [1959] and generalized consistently ordered matrices, see Broyden [1964, 1968], Verner, and Bernal [1968] and Varga [1959].

Another extension of the SOR method is the modified SOR method (MSOR method) for a linear system with red-black matrix. The MSOR method involves the use of relaxation factors \( \omega_1, \omega_1', \omega_2, \omega_2', \ldots \) where for each \( i \), \( \omega_i \) is used for the "red" components, corresponding to \( D_R \), and \( \omega_i' \) is used for the "black" points. It can be shown, see e.g., Young [1971], Chapter 10, that using suitable values of \( \omega_i \) and \( \omega_i' \), the norm of the error is considerably less than for the case where \( \omega_1 = \omega_1' = \omega_2 = \omega_2' = \cdots = \omega_n \).

The MSOR method is equivalent to the "cyclic Chebyshev semi-iterative method" (CCSI method) considered by Golub and Varga [1961]. The CCSI method will be discussed further in Section 5.
5. Chebyshev Methods

The SOR method can be regarded, in some sense, as speeding up the convergence of the Jacobi method. Another, more general, way of speeding up the convergence of the Jacobi method, and many other methods as well, is to use Chebyshev polynomial methods (Chebyshev acceleration).

Possibly the first use of Chebyshev polynomials for problems in numerical linear algebra was the work of Flanders and Shortley [1950] wherein Chebyshev polynomials were used with the power method to find the largest eigenvalue of an SPD matrix. To apply the method it was necessary to choose an interval $I = [\alpha, \beta]$ such that all eigenvalues of the matrix except the desired largest eigenvalue lie in $I$.

It did not take very long for several numerical analysts to realize that the technique could also be used to speed up the convergence of Richardson’s method for solving linear systems; see papers of Lanczos [1952], Shortley [1953] and Young [1954]. Thus from (3.1) if one applies Richardson’s method with extrapolation factors $\gamma_1, \gamma_2, \ldots, \gamma_n$ one obtains

$$u^{(n)} - \bar{u} = (I - \gamma_1 A) \cdots (I - \gamma_n A)(u^{(0)} - \bar{u})$$

$$= P_n(A)(u^{(0)} - \bar{u}) \quad (5.1)$$

where $\bar{u}$ is the true solution of (1.1). One chooses $\gamma_1, \gamma_2, \ldots, \gamma_n$ to minimize the maximum of $|P_n(x)|$ over the interval $[m(A), M(A)]$ where $m(A)$ and $M(A)$ are the smallest and largest eigenvalues of $A$, respectively. It can be
shown using a theorem of Markoff [1892] that the $P_n(x)$ should be chosen as normalized Chebyshev polynomials. The $\{\gamma_i\}$ are simply the reciprocals of the zeros of these polynomials. It can be shown that the use of this procedure speeds up the convergence of Richardson's method by an order of magnitude.

We remark that the above procedure can be applied to variable extrapolation of a basic iterative method. We assume that the basic iterative method has eigenvalues which are real and less than one. The variable extrapolation procedure is defined by

$$u^{(n+1)} = u^{(n)} + \gamma_{n+1}(Gu^{(n)} + k - u^{(n)})$$

(5.2)$$= \gamma_{n+1}(Gu^{(n)} + k) + (1 - \gamma_{n+1})u^{(n)}$$

Unfortunately the above procedure suffers from two drawbacks. First as $n$ changes all of the roots of the normalized Chebyshev polynomials change. Thus, the optimum values of the $\{\gamma_i\}$ corresponding to $n + 1$ are all different from those corresponding to $n$. A second drawback is stability. Unless the ordering of the $\{\gamma_i\}$ is chosen carefully, roundoff may cause the process to "blow up," if $n$ is large. A remedy for the problem was proposed by Anderson and Golub [1972].

A simple way to overcome the above problems is to make use of the three-term relation between the Chebyshev polynomials; see Blair et al [1959], and Golub and Varga [1961]. Thus since $T_n(x) = \cos(n \cos^{-1} x)$ is
the Chebyshev polynomial of degree \( n \), we have

\[
T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)
\]

Using this one can derive the following expression for \( u^{(n+1)} \)

\[
u^{(n+1)} = \rho_{n+1}\left[\gamma(Gu^{(n)} + k) + (1 - \gamma)u^{(n)}\right] + (1 - \rho_{n+1})u^{(n-1)}
\]

where

\[
\gamma = \frac{2}{2 - M(G) - m(G)}
\]

\[
\rho_{n+1} = \begin{cases} 
1, & n = 0 \\
(1 - \frac{\sigma^2}{4}), & n = 1 \\
(1 - \frac{\sigma^2}{4}\rho_n), & n \geq 2
\end{cases}
\]

Here \( m(G) \) and \( M(G) \) are the smallest and largest eigenvalues of \( G \), respectively. The quantity \( \sigma \) is given by

\[
\sigma = \frac{M(G) - m(G)}{2 - M(G) - m(G)}
\]

A serious problem with Chebyshev acceleration is the need to choose estimates for \( m(G) \) and \( M(G) \). The behavior of the acceleration procedure is often very sensitive to the estimates used, especially to the estimate for \( M(G) \). We will discuss this further in Section 9.

We remark that Varga [1957, 1962] used the term "semi-iterative method" when referring to the above procedure.

If one applies Chebyshev acceleration to the Jacobi method for a linear system with a red-black matrix, it turns out that the computation can be
simplified. Essentially, either the red part or the black part of each iteration is not needed and can be discarded. If one does this, the resulting procedure, which is referred to as the “cyclic Chebyshev semi-iterative method” (CCSI method) by Golub and Varga [1961] is equivalent to a special case of the MSOR method described in Section 4.

It can be shown, see Golub [1959], Golub and Varga [1961] and Young [1972] that one can obtain almost as much convergence acceleration as with Chebyshev acceleration by the use of a stationary second degree method as defined by

\[
\begin{align*}
    u^{(n)} &= \gamma(Gu^{(0)} + k) + (1 - \gamma)u^{(0)} \\
    u^{(n+1)} &= \rho[\gamma Gu^{(n)} + k] + (1 - \gamma)u^{(n)}] + (1 - \rho)u^{(n-1)}
\end{align*}
\]

As a matter of fact, the “second order Richardson method” considered by Frankel is a special case of (5.7) where the basic iterative method is Richardson’s method.
6. The Conjugate Gradient Method

In the early 1950s a paper by Hestenes and Stiefel [1952] appeared which presented a new method for solving linear systems. The method is based on the use of the quadratic form

\[ F(u) = \frac{1}{2}(u, Au) - (b, u) \]  

(6.1)

It was assumed that the coefficient matrix \( A \) of the system (1.1) is SPD. It is easy to show that the problem of solving (1.1) is equivalent to that of minimizing \( F(u) \).

One method based on the use of the above minimizing principle is the method of steepest descent. The residual vector \( r^{(n)} = b - Au^{(n)} \), which is the negative of the gradient of \( F(u^{(n)}) \), defines the direction of steepest descent. One proceeds in the direction of steepest descent by letting \( u^{(n+1)} = u^{(n)} + \lambda_n r^{(n)} \) and choosing \( \lambda_n \) so as to minimize \( F(u^{(n)} + \lambda_n r^{(n)}) \).

Unfortunately, the method of steepest descent converges extremely slowly if the condition number of \( A \) is large.

The conjugate gradient method is similar to the method of steepest descent in that, given \( u^{(n)} \), one chooses a direction vector, \( p^{(n)} \), and lets \( u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)} \). However, the direction vector \( p^{(n)} \) is not \( r^{(n)} \), as with the method of steepest descent, but rather \( p^{(n)} = r^{(n)} + \alpha_n p^{(n-1)} \) where \( \alpha_n \) is a constant so chosen that \( p^{(n)} \) is conjugate to \( p^{(n-1)} \), i.e., so that \( (p^{(n)}, Ap^{(n-1)}) = 0 \). Once \( \alpha_n \), and hence \( p^{(n)} \), have been determined, one chooses \( \lambda_n \) to minimize \( F(u^{(n)} + \lambda_n p^{(n)}) \).
The method thus defined has a number of remarkable properties. One property is that, in the absence of roundoff, \( u^{(t)} = \bar{u} \) for some \( t \leq N \), where \( N \) is the order of the system. Thus the method converges, in theory, in at most \( N \) steps. For this reason, the conjugate gradient method was considered by some to be a direct method, perhaps competitive with Gaussian elimination. It turned out, however, that in many cases, because of roundoff, the method did not converge in \( N \) steps. Also, it did not seem to be generally recognized that the method could be considered as an iterative method and that, as such, it had excellent convergence properties. In any case, for a long time the method was not widely used and little was heard about it for many years. Indeed, at a conference on sparse matrices held in 1971, see Rose and Willoughby [1971] there was scarcely any mention of the conjugate gradient method.

Beginning in the mid 1960s, there began to be a resurgence of interest in the conjugate gradient method with the appearance of papers by Daniel [1965, 1967], J.K. Reid [1971, 1972], Axelsson [1974], and Concus et al [1976].

There are several aspects of the conjugate gradient method, which, though included in the papers of Hestenes and Stiefel [1952] and of Hestenes [1956], do not seem to have been fully appreciated. First of all, the conjugate gradient method is not just a single method but rather it can be regarded as a whole family of methods. Thus, consider any basic iterat-
tive method of the form (2.12) which is symmetrizable in the sense that 
$I - G$ is similar to an SPD matrix. (If $A$ is SPD then many of the stan-
dard methods such as Richardson's method, the Jacobi method, the SSOR 
method and the ICC method (but not the SOR method) are symmetriz-
able.) If the conjugate gradient method is applied to the preconditioned 
system corresponding to the basic iteration then the convergence of the 
basic iterative method is greatly accelerated. We refer to the resulting 
procedure as a preconditioned conjugate gradient method or as conjugate 
gradient acceleration of the basic iterative method. If the basic iterative 
method is Richardson's method (3.1) with $\gamma_1 = \gamma_2 = \cdots = 1$, then the asso-
ciated preconditioned conjugate gradient method reduces to the (ordinary) 
conjugate gradient method.

We remark that the term "preconditioning" appears to have been in-
troduced by Evans [1967]; see Axelsson [1974].

Conjugate gradient acceleration has several advantages over Chebyshev 
acceleration. First, it can be shown that the convergence of conjugate gra-
dient acceleration, measured in a certain norm, is at least as fast as that of 
Chebyshev acceleration. For a proof, see, e.g., Hageman and Young [1981], 
Chapter 7. A second advantage of conjugate gradient acceleration is that 
no parameter estimates are required unless, of course, the basic iterative 
method involves a parameter, as in the case of the SSOR method. It should 
be noted, on the other hand, that conjugate gradient acceleration involves
the computation of inner products for each iteration; hence, the work required per iteration is somewhat greater than for Chebyshev acceleration.

Generalized conjugate gradient methods can be used to accelerate the convergence of basic iterative methods which are not symmetrizable, see the discussion in Section 10.
7. Alternating Direction Implicit Methods

Alternating direction implicit methods were introduced in two companion papers which appeared in the mid 1950s. One paper by Peaceman and Rachford [1955] was concerned with an alternating direction implicit method for solving linear systems arising from the finite difference solution of elliptic partial differential equations. The other paper by Douglas [1955] was concerned with the solution of parabolic partial differential equations.

Let us assume that the coefficient matrix $A$ of (1.1) is SPD and can be represented as a sum of two easily invertible SPD matrices, $H$ and $V$. If the linear system is derived from an elliptic five-point finite difference equation, $H$ could correspond to terms involving partial derivatives in $x$ and would be tri-diagonal, whereas $V$ would correspond to terms involving partial derivatives in $y$ and would be permutationally similar to a tridiagonal matrix. Thus both $H$ and $V$ would be easily invertible. In the simplest case, we define the "Peaceman-Rachford method" or the "ADI method" by

$$(H + \rho I)u^{(n+\frac{1}{2})} = b - (V - \rho I)u^{(n)}$$

(7.1)

$$(V + \rho I)u^{(n+1)} = b - (H - \rho I)u^{(n+\frac{1}{2})}$$

Here $\rho > 0$ is a parameter. Frequently, the parameter $\rho$ is allowed to vary from iteration to iteration. A similar method, which can be defined in 3-dimensions as well as in 2-dimensions, was given by Douglas and Rachford [1956].

28
In certain cases the ADI method converge very rapidly. For the model problem, for example, with a suitable choice of parameters, the number of iterations required for convergence is on the order of \( \log h^{-1} \), as compared with the order of \( h^{-1} \) for the SOR method. As shown by Birkhoff and Varga [1959], see also Wachspress and Habetler [1960] and Birkhoff, Varga and Young [1962], the key condition is the commutativity condition \( HV = VH \). This condition holds for certain separable differential equations defined over rectangles, see Birkhoff, Varga and Young [1962]. Procedures for choosing the parameters, given the commutativity condition and also bounds for the eigenvalues of \( H \) and of \( V \), are given by Peaceman and Rachford [1955], Wachspress [1957]; see also Birkhoff, Varga and Young [1962] and Young [1971], Chapter 17.

Unfortunately, attempts to extend the theory beyond the commutative case have been largely futile; see however, Widlund [1966]. An example of Price and Varga [1962] shows that the method can fail in some cases. At present, for the general (noncommutative) case there does not appear to be any way to predict in advance how the method will perform. Also it is my understanding that alternating direction implicit methods are not widely used at the present time. However, evidence to the contrary would be welcomed.
8. The SSOR Method

The symmetric SOR method (SSOR method) was introduced by Sheldon [1955]. It can easily be shown that if the coefficient matrix $A$ of (1.1) is SPD then the matrix, $S_{\omega}$, corresponding to the SSOR method is similar to an SPD matrix and hence the eigenvalues of $S_{\omega}$ are real, positive and less than one. It is thus possible to accelerate the convergence of the SSOR method by an order-of-magnitude using Chebyshev acceleration or conjugate gradient acceleration. Thus, for example, it can be shown, see, e.g., Young [1971, 1972] that for the model problem with the natural ordering of the mesh points, the number of iterations required for convergence using the unaccelerated SSOR method is $O(h^{-1})$, the same as for the SOR method. Actually, the SSOR method, with optimum extrapolation, converges about as fast as the SOR method and the optimum values of omega are nearly the same for both methods as well. However, when the SSOR method is accelerated, the number of iterations required for convergence is $O(h^{-\frac{3}{2}})$. It was shown by Young [1977] that this result also holds for more general elliptic problems provided that the coefficients are sufficiently "smooth." Unfortunately, it appears that the SSOR method, like the ADI method, is quite effective for many "well-behaved" problems but is much less effective for other problems.

The analysis of the SSOR method is complicated by the fact that to date it has not been possible to find a relation between the eigenvalues of the
matrix $S_\omega$ and the eigenvalues of the matrix $B$ for the Jacobi method except in the special case where the coefficient matrix $A$ of (1.1) is red-black. In the red-black case, it can be shown, see, e.g., Young [1971], that the optimum value of $\omega$ is one and the SSOR method reduces to the symmetric Gauss-Seidel method which is much less effective than the SOR method. Indeed for the model problem with the red-black ordering the symmetric Gauss-Seidel method requires $O(h^{-2})$ iterations. When accelerated it requires $O(h^{-1})$ iterations and is no more effective than the SOR method.

The convergence properties of the SSOR method have been studied by a number of researchers including Habetler and Wachspress [1961], Ehrlich [1963, 1964], Evans and Forrington [1963], Young [1971, 1972], Axelsson [1974], and by Benokraitis [1974]. Habetler and Wachspress showed that the key quantities for the SSOR method are the largest eigenvalue, $M(B)$, of the matrix $B$ for the Jacobi method, and the spectral radius, $S(LU)$ of the matrix $LU$, where $L$ and $U$ are strictly lower triangular and strictly upper triangular matrices, respectively, such that $B = L + U$. If $S(LU) \leq \frac{1}{2} + \epsilon$ where $\epsilon$ is "small," then the SSOR method is effective. For the model problem with the natural ordering $S(LU) \leq \frac{1}{4}$. However, if, as in the case of the model problem with the red-black ordering, $S(LU)$ is close to $\frac{1}{2}$, the SSOR method, even with acceleration, is no more effective than the SOR method.

Adaptive procedures for determining $\omega$ for the SSOR method were de-
veloped by Benokraitis [1974] and by Hayes and Young [1977]. The procedure of Hayes and Young has been incorporated into the ITPACK 2C and the ITPACK 3A software packages; see Kincaid et al [1982], and Young and Mai [1984].
9. Algorithms and Software

Several mathematical software projects were undertaken in the mid 1970s including the LINPACK project for solving linear systems and the ELLPACK project for solving elliptic partial differential equations. Around that time Garrett Birkhoff suggested that the time was ripe to start developing software based on iterative methods for solving large sparse linear systems.

One of the reasons for the lack of iterative software prior to that time was the lack of any general procedure for determining iteration parameters such as omega for the SOR method. For a survey of the various methods which have been used to choose iteration parameters see Young and Mai [1988]. Another question which previously had received little attention, and which Birkhoff suggested should be considered, involved the determination of the accuracy of an approximate solution of the system (1.1) to the true solution. Accurate estimates are needed to develop reasonable tests for deciding when to terminate the iterative process.

Over the next several years there was intensive activity on the development of adaptive parameter determination procedures as well as stopping procedures. Algorithms based on these procedures are described in the book by Hageman and Young [1981]. The research-oriented software package, ITPACK 2C, was developed based on these algorithms, see Kincaid et al [1982]. The ITPACK 2C package includes the Jacobi, SSOR and RS
methods*, with both Chebyshev acceleration and conjugate gradient acceleration as well as the SOR method. Adaptive procedures for determining any necessary iteration parameters are included. Also, stopping procedures are included which are based on error estimation procedures which have been found to be accurate over a wide range of cases. The ITPACK 2C package was modified and incorporated into the ELLPACK package; see Rice and Boisvert [1985].

The ITPACK 2C package was intended primarily for linear systems where the coefficient matrix, is SPD, or nearly so. Other packages have since been developed which include the capability of handling nonsymmetric systems. These include the “new Yale square matrix package” described by Eisenstat, Elman, Schultz and Sherman [1984], as well as several ITPACK packages including ITPACK 3A, ITPACK 3B, and NSPCG. (For a review of the ITPACK packages see Kincaid and Young [1988].

---

* The RS method is defined for a red-black system and is essentially Richardson's method applied to the reduced system obtained by eliminating the “red unknowns.”
10. **Nonsymmetric Systems**

The problem of solving the linear system (1.1), where the coefficient matrix $A$ is not necessarily SPD or even symmetric, is equivalent to the problem of solving the normal equations.

\[ A^T A u = A^T b \]

Unfortunately, although the matrix $A^T A$ is SPD, the condition number of $A^T A$ is often very much greater than that of $A$, frequently so much greater that the use of the normal equations is impractical.*

**SOR Method**

Considerable work has been done on the choice of the optimum $\omega$ for the case where $A$ is consistently ordered but not SPD and where some of the eigenvalues of the matrix $B$ for the Jacobi method are complex. Several programs are available for choosing the optimum $\omega$ if all of the eigenvalues of $B$ are known or if a convex region containing them is known; see e.g., Young and Eidson [1970], Young and Huang [1983] and Huang [1983].

Ehrlich [1981] and others have considered the use of an "ad hoc SOR" procedure for solving linear systems corresponding to partial difference equations. Here $\omega$ is allowed to vary from point to point. The method works well in many cases, in spite of the lack of any rigorous theory

*However, Danquarra et al [1981] considered the use of generalized normal equations which are the normal equations for the preconditioned system corresponding to a basic iterative method.
Chebyshev Acceleration

Chebyshev acceleration can be applied to a basic iterative method whose iteration matrix $G$ has complex eigenvalues. An early paper was written by Wrigley [1963]. As in the case where $G$ has real eigenvalues, one must estimate two parameters which define a key ellipse containing all of the eigenvalues. Manteuffel [1977,1978] described procedure for choosing the two parameters, given the location of all of the eigenvalues of $G$, or at least the location of certain key eigenvalues. Another program for doing this was developed by Huang [1983]. Manteuffel also described an adaptive procedure for finding the key eigenvalues based on results obtained from several iterations which were carried out with estimated values of the parameters.

Generalized Conjugate Gradient Acceleration


Young and Jea [1980] defined three generalized conjugate gradient acceleration methods. These methods were referred to as ORTHODIR, ORTHOMIN, and ORTHORES. It was shown that under fairly general conditions these methods converge, in the absence of rounding errors, in a finite number of iterations. Numerical experiments indicate that in many
cases the methods converge as fast or faster than Chebyshev acceleration procedures.

Unfortunately, the generalized conjugate gradient procedures require, in general, at each stage, the storage and use of at least one vector from every previous iteration. (However, in the symmetrizable case, information from only one or two previous iterations is required.) To overcome this difficulty, "truncated" forms of the methods are often used wherein information from all but a fixed number of the most recent iterations are discarded. Another procedure is to use nontruncated methods but to "restart" the procedure periodically.

Further information concerning generalized conjugate gradient methods can be found in papers by Eisenstat, Elman and Schultz [1979], Young and Jea [1980], Axelsson [1980], Elman [1982], Jea [1982], Saad and Schultz [1983], and Jea and Young [1983], to mention only a few.

Lanczos Method

Fletcher [1976] developed a method called the "biconjugate gradient method" which can be used to accelerate the convergence of a nonsymmetrizable basic iterative method. Moreover the method requires only the storage and use of vectors from one previous iteration. The method is, however, at least in theory, subject to the possibility of breakdown.

Jea and Young [1983] derived the biconjugate gradient method as well as other forms of the Lanczos method by considering generalized conju-
gate gradient acceleration methods corresponding to certain linear systems derived from the system (1.1).

Future Work

If the eigenvalues of the iteration matrix $G$ for a basic iteration method all lie in a circle in the complex-plane and are densely distributed on the boundary of the circle, then no acceleration of the convergence is possible. An example is the SOR method for a consistently ordered SPD matrix with the optimum value of $\omega$. All of the eigenvalues lie in the boundary of the circle $|z| = \omega - 1$. This illustrates the fact that for a given basic iterative method little or no acceleration may be possible. On the other hand, for a given linear system there exists a preconditioner, namely, $A$ itself which, if practical, would lead to a basic iterative method which would converge in one iteration. It seems reasonable to expect that, at least in most cases, there should exist an easily invertable matrix $Q$ which is "close" to $A$ and which yields an effective basic iterative matrix.

To date, most of the effort which has been spent on nonsymmetric systems has been devoted to the accelerator and very little has been spent on preconditioners. The above considerations suggest that much more effort should be spent in finding ways to construct effective preconditioners.
11. A Short Guide to the Literature

In this section, we first mention some early papers, each of which has had an important impact on the field. Next we mention some other historical reviews. Finally we list some of the books on iterative methods which have appeared over the years.

Some Important Early Papers

We briefly discuss four papers which have stimulated the development of iterative methods. First, the paper of Shortley and Weller [1938] showed how the convergence of certain iterative methods could be analyzed in terras of the eigenvalues and the Jordan canonical forms of certain matrices. Second, the paper of Geiringer [1949] treated iterative methods from a more general point of view — not tied directly to elliptic difference equations. Third, the paper of Forsythe [1953] was not only very interesting and full of good ideas but it was also important in that it helped to establish the respectability of the field. This was no mean feat since to most pure mathematicians the solution of a linear system with a square nonsingular matrix is a completely trivial problem which could be solved using Cramer's rule. Finally, we mention the paper of Flanders and Shortley [1950] which introduced the idea of using Chebyshev polynomials to solve problems in numerical linear algebra.
Some Other Historical Surveys

Books on Iterative Methods

We list below several books which are devoted, entirely or at least in substantial part, to iterative methods.

Forsythe and Wasow [1960]
Varga [1962]
Wachspress [1966]
Young [1971]
Birkhoff [1971]
Hageman and Young [1981]
Birkhoff and Lynch [1984]
Axelsson and Barker [1984]
Ortega [1988]
References


die Methode der Kleinsten Quadrate führt, sowie lineare Gleichungen
überhaupt, durch successive Annäherung aufzulösen, Abh. Math-


[82] Shortley, G.H. [1953]. "Use of Tschebyscheff Polynomial Operators
Phys., 24, 392–396.


Laplace's and Poisson's Equation with Applications to Photoclasticity
and Torsion, Ohio State University, Engineering Experiment Station,

[85] Stone, H.L. [1968]. "Iterative Solutions of Implicit Approximations of
Anal., 5, 530–538.

[86] Southwell, R.V. [1946]. Relaxation Methods in Theoretical Physics,

[87] Swartztrauber, P.M. [1974]. "A Direct Method for the Discrete Solu-
tion of Separable Elliptic Equations," SIAM J. Numer. Anal., 11,
1136–1150.


[90] Varga, R.S. [1957]. "A Comparison of the Successive Overrelaxation
Method and Semi-iterative Methods Using Chebyshev Polynomials;"
J. SIAM, 5, 39–46.

[91] Varga, R.S. [1957]. "p-cyclic Matrices: A Generalization of the Young-
Frankel Successive Overrelaxation Scheme," Pacific J. Math. 9, 617–
628.


