JOURNAL OF APPROXIMATION THEORY 5, 137-148 (1972)

Second-Degree Iterative Methods for the Solution of Large Linear Systems*

DAVID M. YOUNG

Center for Numerical Analysis,
University of Texas, Austin, Texas 78712

Communicated by Oved Shisha

Received October 9, 1970

DEDICATED TO PROFESSOR J. L. WALSH ON THE OCCASION OF HIS 75TH BIRTHDAY

1. INTRODUCTION

In this paper we consider the convergence properties of various iterative methods for solving the linear system

$$Au = b, \tag{1.1}$$

where A is a given real nonsingular $N \times N$ matrix, b is a given real $N \times 1$ column matrix, and u is an unknown $N \times 1$ column matrix. We consider methods derived from the linear stationary method of first degree defined by

$$u^{(n+1)} = Gu^{(n)} + k, \tag{1.2}$$

where G is a real $N \times N$ matrix such that $I - G$ is nonsingular and

$$k = (I - G) A^{-1}b. \tag{1.3}$$

The iterative method (1.2) is completely consistent with the system (1.1) in the sense that the solution of (1.1) is the same as the solution of the related equation

$$u = Gu + k \tag{1.4}$$

(see [1]). Moreover, if for some $u^{(0)}$ the sequence defined by (1.2) converges,

* Work on this paper was supported in part by the U. S. Army Research Office (Durham) through Grant DA-ARO(D)-31-124-G1050 and by the National Science Foundation through Grant GP-8442 to The University of Texas at Austin. The author wishes to acknowledge the helpful suggestions of David R. Kincaid in the preparation of this paper.

it converges to the solution of (1.1). It is well known that the sequence defined by (1.2) converges for all $u^{(0)}$ to a limit independent of $u^{(0)}$ if and only if $S(G)$, the spectral radius of G, is less than unity. However, we do not make this assumption.

The convergence properties of an iterative method can often be improved by the use of a semi-iterative method based on the given method (see [2–5]). To define a semi-iterative method one chooses constants $\alpha_{n,k}$, $k = 0, 1,..., n$, $n = 0, 1, 2,...$ such that

$$\sum_{k=0}^{n} \alpha_{n,k} = 1, \qquad n = 0, 1, 2,..., \tag{1.5}$$

and one lets

$$v^{(n)} = \sum_{k=0}^{n} \alpha_{n,k} u^{(k)}, \qquad n = 0, 1, 2,... . \tag{1.6}$$

If $\bar{u}$ is the exact solution of (1.1), then $\bar{u}$ satisfies (1.4) and we have

$$v^{(n)} - \bar{u} = P_n(G)(u^{(0)} - \bar{u}), \tag{1.7}$$

where, in general,

$$P_n(x) = \sum_{k=0}^{n} \alpha_{n,k} x^k. \tag{1.8}$$

If the eigenvalues μ of G are real and lie in the interval

$$\alpha \leqslant \mu \leqslant \beta < 1 \tag{1.9}$$

then the choice of the $\alpha_{n,k}$ given by

$$\sum_{k=0}^{n} \alpha_{n,k} x^k = \frac{T_n\left(\dfrac{2x - (\beta + \alpha)}{\beta - \alpha}\right)}{T_n(z)} = \bar{P}_n(x), \tag{1.10}$$

where

$$z = [2 - (\beta + \alpha)]/(\beta - \alpha)$$

is optimal in the sense of minimizing the virtual spectral radius $\bar{S}(\bar{P}_n(G))$. Here, for any polynomial $P_n(x)$, we let

$$\bar{S}(P_n(G)) = \max_{\alpha \leqslant \mu \leqslant \beta} | P_n(\mu)|. \tag{1.11}$$

The $T_n(x)$ are the Chebyshev polynomials of degree n defined by

$$T_n(x) = \cos(n \cos^{-1} x) = \tfrac{1}{2}\{[x + \sqrt{(x^2 - 1)}]^n + [x + \sqrt{(x^2 - 1)}]^{-n}\}. \tag{1.12}$$

Moreover, we have

$$\bar{S}(\bar{P}_n(G)) = 2r^{n/2}/(1 + r^n), \tag{1.13}$$

where

$$r = \hat{\omega}_b - 1 = \{\sigma/[1 + \sqrt{(1 - \sigma^2)}]\}^2, \tag{1.14}$$

$$\hat{\omega}_b = 2/[1 + \sqrt{(1 - \sigma^2)}], \tag{1.15}$$

and

$$\sigma = 1/z. \tag{1.16}$$

By virtue of the relations

$$\begin{aligned} T_0(x) &= 1, \qquad T_1(x) = x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x), \end{aligned} \tag{1.17}$$

one can derive the following three-term relation involving $v^{(n+1)}$, $v^{(n)}$, and $v^{(n-1)}$:

$$\begin{aligned} v^{(n+1)} = \frac{1}{z}\,\omega_{n+1}\left[\frac{2}{\beta - \alpha}\,G - \frac{\beta + \alpha}{\beta - \alpha}\,I\right] v^{(n)} \\ + (1 - \omega_{n+1})\,v^{(n-1)} + \frac{2\omega_{n+1}}{z(\beta - \alpha)}\,k, \end{aligned} \tag{1.18}$$

where

$$\begin{aligned} \omega_1 &= 1, \qquad \omega_2 = 1\Big/\left(1 - \frac{1}{2z^2}\right), \\ \omega_{n+1} &= 1\Big/\left(1 - \frac{\omega_n}{4z^2}\right), \qquad n = 2, 3, \ldots . \end{aligned} \tag{1.19}$$

We remark that $\omega_n \to \omega_b$ as $n \to \infty$. (See [5].)

We now compare the convergence of the semi-iterative method with that of the method

$$\begin{aligned} u^{(n+1)} &= \bar{P}_1(G)\,u^{(n)} + \frac{2}{2 - (\beta + \alpha)}\,k \\ &= \frac{1}{2 - (\beta + \alpha)}\,[2G - (\beta + \alpha)I]\,u^{(n)} + \frac{2}{2 - (\beta + \alpha)}\,k. \end{aligned} \tag{1.20}$$

As a measure of the rapidity of the convergence we take the asymptotic average rate of convergence defined by

$$R_{\mathrm{SI}} = \lim_{n\to\infty}\left(-\frac{1}{n}\log \bar{S}(\bar{P}_n(G))\right) = \lim_{n\to\infty}\left[-\frac{1}{n}\log\frac{2r^{n/2}}{1 + r^n}\right] = -\frac{1}{2}\log r. \tag{1.21}$$

For the method (1.20) we have

$$R_1 = -\log \bar{S}(\bar{P}_1(G)) = -\log \sigma. \tag{1.22}$$

It is easy to show that for σ close to unity we have

$$R_{SI}/\sqrt{R_1} \sim \sqrt{2}. \tag{1.23}$$

Thus, in a sense, there is an order-of-magnitude improvement in the convergence rate of the semi-iterative method as compared with the method (1.20).

Evidently, (1.18) can be written in the form

$$v^{(n+1)} = v^{(n)} + d_{n+1}(v^{(n)} - v^{(n-1)}) + e_{n+1}(Gv^{(n)} + k - v^{(n)}), \tag{1.24}$$

where

$$d_{n+1} = \omega_{n+1} - 1, \qquad e_{n+1} = 2\omega_{n+1}/z(\beta - \alpha). \tag{1.25}$$

The method (1.24) is said to be a *nonstationary method* of second degree. One of the objects of the present paper is to consider *stationary* second-degree methods based on (1.2) of the form

$$u^{(n+1)} = u^{(n)} + d(u^{(n)} - u^{(n-1)}) + e(Gu^{(n)} + k - u^{(n)}), \qquad n = 1, 2, \ldots . \tag{1.26}$$

Here $u^{(0)}$ is arbitrary and $u^{(1)}$ is determined by a special procedure such as (1.20). It is shown that by a suitable choice of d and e one can achieve a convergence rate nearly, though not quite, as good as that of the optimum semi-iterative method. This result is known for the case of the Jacobi method; the associated second-degree method is referred to as the "second-order Richardson method"; it has been studied by Frankel [6], Riley [7], Golub [4], Golub and Varga [5], and others. However, it does not seem to be generally recognized that second-degree methods can be effectively applied to other methods as well.

A second object of the present paper is to apply the above results to the symmetric successive overrelaxation method (SSOR method) for linear systems arising from the five-point discrete analog of the Dirichlet problem. We give a formula for a relaxation factor whose use, together with semi-iteration or the corresponding second-degree method, results in a reciprocal rate of convergence of $O(h^{-1/2})$ where h is the mesh size. For the ordinary successive overrelaxation method with the optimum relaxation factor the reciprocal rate of convergence is known to be $O(h^{-1})$ [8]. Thus the semi-iterative method and the second-degree method based on the SSOR method are better than the SOR method by an order-of-magnitude. This is true for nonrectangular as well as rectangular regions. Moreover, an explicit procedure is given for the choice of the iteration parameters corresponding to each value of the mesh size h.

It appears that for most of the methods considered in the literature for obtaining a reciprocal convergence rate less than $O(h^{-1})$ either the improvement in the convergence cannot be shown to hold for nonrectangular regions or else an explicit procedure for choosing the iteration parameters is not available. Thus the Peaceman–Rachford alternating-direction implicit method [9] can be shown to have a reciprocal convergence rate much less than $O(h^{-1/2})$ for rectangular regions. However, while numerical evidence (see, for instance, [10, 11]) indicates that the improvement holds for other regions as well, no proof has as yet been given. For convex regions Guilinger [12] has shown that the reciprocal rate of convergence can be as small as $O(1)$ provided certain assumptions are made as to the choice of the starting vector $u^{(0)}$.

Habetler and Wachspress [13] have proved the existence of a relaxation factor whose use with the SSOR method and semi-iteration leads to a reciprocal convergence rate of $O(h^{-1/2})$. However, in their analysis the determination of the relaxation factor involves the solution of a highly implicit equation.

2. Second-degree Methods

Let $\bar{u}$ be the exact solution of (1.1) and let

$$\epsilon^{(n)} = u^{(n)} - \bar{u}, \tag{2.1}$$

where $u^{(2)}, u^{(3)},\ldots,$ are determined by the second-degree method (1.26) with $u^{(0)}$ and $u^{(1)}$ arbitrary. Since $G\bar{u} + k - \bar{u} = 0$ we have

$$\epsilon^{(n+1)} = \epsilon^{(n)} + d(\epsilon^{(n)} - \epsilon^{(n-1)}) + e(G\epsilon^{(n)} - \epsilon^{(n)}), \qquad n = 1, 2,\ldots . \tag{2.2}$$

To study the convergence of (2.2) we observe that

$$w^{(n)} = \Gamma w^{(n-1)} = \Gamma^n w^{(0)}, \qquad n = 1, 2,\ldots ,$$

where

$$\Gamma = \begin{pmatrix} 0 & I \\ -dI & (1 + d - e) I + eG \end{pmatrix}, \qquad w^{(n)} = \begin{pmatrix} \epsilon^{(n)} \\ \epsilon^{(n+1)} \end{pmatrix}.$$

We seek to choose d and e so as to minimize $\bar{S}(\Gamma)$. To do this we observe that

$$\Gamma \begin{pmatrix} S \\ t \end{pmatrix} = \lambda \begin{pmatrix} S \\ t \end{pmatrix}$$

for some vectors s and t if and only if $t = \lambda s$ and

$$\{\lambda[eG + (1 + d - e)I] - dI\}s = \lambda^2 s. \tag{2.3}$$

Unless both s and t vanish we must have

$$\det(\lambda^2 I - \lambda(eG + (1 + d - e)I) + dI) = 0$$

and

$$\lambda^2 - \lambda(e\mu + (1 + d - e)) + d = 0 \tag{2.4}$$

for some eigenvalue μ of G. Thus

$$\bar{S}(\Gamma) = \rho = \max_i (\max(|\lambda_i^+|, |\lambda_i^-|), \tag{2.5}$$

where, for each eigenvalue $\mu_1, \mu_2, \dots, \mu_N$ of G, λ_i^+ and λ_i^- are the roots of (2.4) with $\mu = \mu_i$.

It follows from the analysis of Frankel [6] that if μ varies over the range $\alpha \leqslant \mu \leqslant \beta < 1$, then the choice of d and e which minimizes ρ is given by†

$$d = \hat{\omega}_b - 1, \qquad e = 2\hat{\omega}_b/[2 - (\beta + \alpha)], \tag{2.6}$$

where

$$\hat{\omega}_b = 2/[1 + \sqrt{(1 - \sigma^2)}], \qquad \sigma = 1/z = (\beta - \alpha)/[2 - (\beta + \alpha)]. \tag{2.7}$$

The corresponding value of ρ is

$$\rho = \sqrt{(\hat{\omega}_b - 1)} = \sigma/[1 + \sqrt{(1 - \sigma^2)}] = r^{1/2}. \tag{2.8}$$

Thus with this choice of d and e, (1.26) becomes

$$\begin{aligned} u^{(n+1)} &= u^{(n)} + (\hat{\omega}_b - 1)(u^{(n)} - u^{(n-1)}) + \frac{2\hat{\omega}_b}{2 - (\beta + \alpha)} (Gu^{(n)} + k - u^{(n)}) \\ &= \frac{\hat{\omega}_b}{z} \left[\frac{2}{\beta - \alpha} G - \frac{(\beta + \alpha)}{\beta - \alpha} I\right] u^{(n)} + (1 - \hat{\omega}_b)\, u^{(n-1)} + \frac{2\hat{\omega}_b}{z(\beta - \alpha)} k. \end{aligned} \tag{2.9}$$

A more precise assessment of the convergence rate can be made if we specify the choice of $u^{(1)}$. It seems reasonable to let $u^{(1)}$ be the same as for the corresponding semi-iterative method. Thus from (1.20) we have

$$u^{(1)} = [1/z(\beta - \alpha)][2G - (\beta + \alpha)I]\, u^{(0)} + [2/z(\beta - \alpha)]k. \tag{2.10}$$

Let us consider the sequence of polynomials defined by

$$\begin{gathered} \hat{Q}_0(G') = I, \qquad \hat{Q}_1(G') = G', \\ \hat{Q}_{n+1}(G') = \hat{\omega}_b G' \hat{Q}_n(G') + (1 - \hat{\omega}_b)\, \hat{Q}_{n-1}(G'), \quad n \geqslant 1, \end{gathered} \tag{2.11}$$

† *Note added in proof.* For details see D. Kincaid, Report CNA-23, Center for Numerical Analysis, University of Texas, Austin, Tex., 1971.

where

$$G' = \{1/[2 - (\beta + \alpha)]\}[2G - (\beta + \alpha)I]. \tag{2.12}$$

Evidently by (2.9) and (2.10) we have

$$\epsilon^{(n)} = \hat{Q}_n(G')\,\epsilon^{(0)}. \tag{2.13}$$

Corresponding to the polynomial $\bar{P}_n(x)$ defined by (1.10) let us define

$$\hat{P}_n(y) = \bar{P}_n\{[(2 - (\beta + \alpha))\,y + (\beta + \alpha)]/2\} = T_n(zy)/T_n(z). \tag{2.14}$$

From (1.17) and (1.19) we have

$$\begin{gathered} \hat{P}_0(G') = I, \qquad \hat{P}_1(G') = G', \\ \hat{P}_{n+1}(G') = \omega_{n+1}G'\hat{P}_n(G') + (1 - \omega_{n+1})\,\hat{P}_{n-1}(G'), \qquad n = 1, 2, \ldots \end{gathered} \tag{2.15}$$

which is the same as (2.11) except that $\hat{\omega}_b$ is replaced by ω_{n+1}. Because of the similarity between (2.11) and (2.15) it seems reasonable to expect that the polynomials $\hat{Q}_n(y)$ will be good approximations to the $\hat{P}_n(y)$.

Let us now determine $\bar{S}(\hat{Q}_n(G'))$. Golub [4] has shown that

$$\max_{-\sigma \leqslant y \leqslant \sigma} |\hat{Q}_n(y)| = \hat{Q}_n(\sigma) \tag{2.16}$$

(see also Young and Kincaid [14]). It is easy to verify that

$$\hat{Q}_n(\sigma) = \frac{2r^{n/2}}{1 + r}\left[1 + \left(\frac{n-1}{2}\right)(1 - r)\right]; \tag{2.17}$$

hence we have

$$\bar{S}(\hat{Q}_n(G')) = \frac{2r^{n/2}}{1 + r}\left[1 + \left(\frac{n-1}{2}\right)(1 - r)\right] = r^{n/2}\left[1 + n\left(\frac{1-r}{1+r}\right)\right]. \tag{2.18}$$

It can be shown that $\bar{S}(\hat{Q}_n(G')) \geqslant \bar{S}(\hat{P}_n(G'))$ (see [5, 14]). On the other hand, the asymptotic average rate of convergence R_{SD} is given by

$$R_{\mathrm{SD}} = \lim_{n\to\infty}\left[-\frac{1}{n}\log r^{n/2}\left(1 + n\left(\frac{1-r}{1+r}\right)\right)\right] = -\frac{1}{2}\log r \tag{2.19}$$

as for the semi-iterative method. Thus the second-degree method, like the semi-iterative method, is better than (1.20) by an order-of-magnitude.

As an example, let us consider the case where $\alpha = -0.95$, $\beta = 0.95$. For (1.20), the semi-iterative method, and the second-degree method we seek the smallest integer n for which the spectral radius associated with it is less than 10^{-6}. Thus for (1.20) we solve

$$\sigma^n = 10^{-6},$$

where

$$\sigma = 1/z = (\beta - \alpha)/[2 - (\beta + \alpha)] = 0.95,$$

and obtain

$$n_B \doteq 269.$$

Evidently, by (2.7) and (2.8) we have

$$\hat{\omega}_b \doteq 1.524, \qquad r = \hat{\omega}_b - 1 \doteq 0.524.$$

For the semi-iterative method, we solve

$$2r^{n/2}/(1 + r^n) = 10^{-6}$$

obtaining

$$n_{\mathrm{SI}} \doteq 45.$$

For the second-degree method we solve

$$r^{n/2}\left[1 + n\left(\frac{1 - r}{1 + r}\right)\right] = 10^{-6},$$

obtaining

$$n_{\mathrm{SD}} \doteq 51$$

which is only slightly larger than the corresponding number for the semi-iterative method. Both the semi-iterative method and the second-degree method are better than the basic method by a factor greater than five. This factor of improvement increases as σ increases.

3. The Symmetric Successive Overrelaxation Method (SSOR Method)

We now consider the application of the above results to the case of the SSOR method. For simplicity we assume that A is a positive definite matrix with unit diagonal elements and we let

$$A = I - L - U, \tag{3.1}$$

where L and U are strictly lower and strictly upper triangular matrices, respectively. Since A is symmetric we have

$$L^T = U. \tag{3.2}$$

The SSOR method is defined by

$$u^{(n+1)} = \mathscr{S}_\omega u^{(n)} + \omega(2-\omega)(I-\omega U)^{-1}(I-\omega L)^{-1}b, \tag{3.3}$$

where

$$\mathscr{S}_\omega = I - \omega(2-\omega)(I-\omega U)^{-1}(I-\omega L)^{-1}A. \tag{3.4}$$

Sheldon [15] considered the use of semi-iterative methods to accelerate the convergence of the SSOR method. Subsequent work was done by Habetler and Wachspress [13] and by Ehrlich [16, 17]. Habetler and Wachspress proved the existence of a unique value of ω in the range $0 < \omega < 2$ which minimizes $S(\mathscr{S}_\omega)$. However, as mentioned earlier, the determination of this value of ω involves the solution of a highly implicit equation. For our purposes it is sufficient to give a "good" value of ω for which a bound on $S(\mathscr{S}_\omega)$ can be found for a special case. We prove

THEOREM 3.1.† *Let A be a positive definite matrix with unit diagonal elements such that $\bar{\mu} < 1$ and*

$$S(LU) \leqslant 1/4, \tag{3.5}$$

where L and U are, respectively, a strictly lower and a strictly upper triangular matrix such that (3.1) *holds. Then*

$$S(\mathscr{S}_{\omega_1}) \leqslant [1-\sqrt{(1-\bar{\mu})/2}]/[1+\sqrt{(1-\bar{\mu})/2}], \tag{3.6}$$

where

$$\bar{\mu} = S(L+U) \tag{3.7}$$

and

$$\omega_1 = 2/[1+\sqrt{2(1-\bar{\mu})}]. \tag{3.8}$$

Proof. For $0 < \omega < 2$ we have by (3.4)

$$\mathscr{S}_\omega = I - R(\omega)^{-1}, \tag{3.9}$$

where

$$R(\omega) = [1/\omega(2-\omega)]\, A^{-1}(I-\omega L)(I-\omega U). \tag{3.10}$$

If λ is an eigenvalue of $R(\omega)$ and if v is an associated eigenvector, then we have

$$\lambda = (1-\omega\xi+\omega^2\eta)/\omega(2-\omega)(1-\xi), \tag{3.11}$$

where

$$\xi = (v, (L+U)v), \qquad \eta = (v, LUv). \tag{3.12}$$

† *Note added in proof.* In more recent work it is shown that (3.5) implies that $\bar{\mu} \leqslant 1$; moreover, $\bar{\mu}$ can be replaced in (3.8) and (3.6) by β, the largest eigenvalue of $L + U$.

Here we define the inner product (u, v) of two vectors by

$$(u, v) = \sum_{i=1}^{N} \bar{u}_i v_i \,.$$

We assume that $(v, v) = 1$. Since $\eta \leqslant \| LU \| = S(LU) \leqslant 1/4$ and since

$$1 - \omega\xi + \omega^2\eta = (v, (I - \omega L)(I - \omega U)v) = ((I - \omega U)v, (I - \omega U)v) \geqslant 0, \tag{3.13}$$

it follows that

$$\lambda \leqslant (1 - \omega\xi + \tfrac{1}{4}\omega^2)/\omega(2 - \omega)(1 - \xi). \tag{3.14}$$

Moreover,

$$\frac{d\lambda}{d\xi} = (1 - \omega\xi + \tfrac{1}{4}\omega^2)/\omega(2 - \omega)(1 - \xi) = \frac{2 - \omega}{4\omega(1 - \xi)^2} > 0.$$

Since $\xi \leqslant \| L + U \| = S(L + U) = \bar{\mu}$, we have

$$\lambda \leqslant [1/\omega(2 - \omega)][(1 - \omega\bar{\mu} + \tfrac{1}{4}\omega^2)/(1 - \bar{\mu})]. \tag{3.15}$$

The derivative of the right member of the above equation with respect to ω vanishes when

$$\omega^2(\bar{\mu} - \tfrac{1}{2}) = 2(\omega - 1). \tag{3.16}$$

The root of (3.16) in the interval $0 < \omega < 2$ is clearly ω_1 as given by (3.8). From (3.16) and (3.15) we obtain

$$\lambda \leqslant \left(1 - \frac{\omega_1\bar{\mu}}{2}\right)\Big/\omega_1(1 - \bar{\mu}). \tag{3.17}$$

The result (3.6) follows from (3.9) and (3.8).

For any $h > 0$ let Ω_h be a set of points (ih, jh), where i and j are integers. Two points (ih, jh) and $(i'h, j'h)$ are *adjacent* if $| i - i' | + | j - j' | = 1$. Let R_h be any finite subset of Ω_h and let S_h be the set of all points of Ω_h which are not in R_h but are adjacent to points of R_h. We define the discrete analog of the Dirichlet problem as that of finding a function $u(x, y)$ defined on $R_h + S_h$ which assumes prescribed values on S_h and satisfies on R_h the difference equation

$$u(x, y) - \tfrac{1}{4}u(x + h, y) - \tfrac{1}{4}u(x - h, y) - \tfrac{1}{4}(x, y + h) - \tfrac{1}{4}u(x, y - h) = 0. \tag{3.18}$$

If we label the points of R_h in their "natural order" with (x, y) following (x', y') if $y > y'$ or if $y = y'$ and $x > x'$, then for the associated matrix A

we have (3.5). For each of the matrices L and U has at most two nonzero elements, namely 1/4, in any row. Hence $\| L \|_\infty \leqslant \frac{1}{2}$, $\| U_\infty \| \leqslant \frac{1}{2}$ and

$$S(LU) \leqslant \| LU \|_\infty \leqslant \| L \|_\infty \| U \|_\infty \leqslant \tfrac{1}{4}. \tag{3.19}$$

Here for any $N \times N$ matrix A we let

$$\| A \|_\infty = \max_{1 \leqslant i \leqslant N} \sum_{j=1}^{N} | a_{i,j} |.$$

Suppose that the points of R_h and S_h belong to the unit square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. It is easy to show that if S_h is the set of all points of Ω_h on the boundary of the square then

$$\bar{\mu} = \bar{\mu}_s = \cos \pi h \tag{3.20}$$

and in general

$$\bar{\mu} \leqslant \cos \pi h. \tag{3.21}$$

Since $\mathscr{S}_\omega$ has real nonnegative eigenvalues [15] we can apply the semi-iterative method or the second-degree method using

$$\alpha = 0, \beta = \frac{1 - \sqrt{\dfrac{1 - \bar{\mu}_s}{2}}}{1 + \sqrt{\dfrac{1 - \bar{\mu}_s}{2}}} = \frac{1 - \sin \dfrac{\pi h}{2}}{1 + \sin \dfrac{\pi h}{2}} = 1 - \pi h + O(h^2). \tag{3.22}$$

We let

$$\omega_1 = 2/[1 + \sqrt{2(1 - \bar{\mu}_s)}] = 2/[1 + 2 \sin(\pi h/2)]. \tag{3.23}$$

Evidently

$$\bar{S}(\bar{P}_1(\mathscr{S}_{\omega_1})) = \frac{\beta}{2 - \beta} = \frac{1 - \sin \dfrac{\pi h}{2}}{1 + 3 \sin \dfrac{\pi h}{2}} = 1 - 2\pi h + O(h^2) \tag{3.24}$$

and

$$R_1 = 2\pi h + O(h^2). \tag{3.25}$$

Consequently, by (1.23) it follows that

$$R_{\mathrm{SI}} \sim 2 \sqrt{\pi h}.$$

Similarly, for the second-degree method we have

$$R_{\mathrm{SD}} \sim 2 \sqrt{\pi h}.$$

Thus the reciprocal rate of convergence for the semi-iterative method and for the second-degree method is $O(h^{-1/2})$.

References

1. David M. Young, On the consistency of linear stationary iterative methods, to appear in *SIAM J. Numer. Anal.*
2. Richard S. Varga, A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials, *J. Soc. Indus. Appl. Math.* **5** (1957), 39–46.
3. A. Blair, N. Metropolis, J. von Neumann, A. H. Taub, and M. Tsingori, A study of a numerical solution of a two-dimensional hydrodynamical problem, *Math. Tables Aids Comput.* **13** (1959), 145–184.
4. G. H. Golub, "The use of Chebyshev matrix polynomials in the iterative solution of linear systems compared with the method of successive overrelaxation," doctoral thesis, University of Illinois, Urbana, IL, 1959.
5. G. H. Golub and R. S. Varga, Chebyshev semi-iterative methods, successive over-relaxation iterative methods, and second-order Richardson iterative methods, I, II, *Numer. Math.* **3** (1961), 147–168.
6. S. P. Frankel, Convergence rates of iterative treatments of partial differential equations, *Math. Tables Aids Comput.* **4** (1950), 65–75.
7. James D. Riley, Iteration procedures for the Dirichlet difference problem, *Math. Tables Aids Comput.* **8** (1954), 125–131.
8. David M. Young, Iterative methods for solving partial difference equations of elliptic type, *Trans. Amer. Math. Soc.* **76** (1954), 92–111.
9. D. W. Peaceman and H. H. Rachford, Jr., The numerical solution of parabolic and elliptic differential equations, *J. Soc. Indus. Appl. Math.* **3** (1955), 28–41.
10. David M. Young and L. Ehrlich, Some numerical studies of iterative methods for solving elliptic difference equations, *in* "Boundary Problems in Differential Equations" (R. E. Langer, Ed.), pp. 143–162, University of Wisconsin Press, Madison, WI, 1960.
11. Garrett Birkhoff, Richard S. Varga, and David Young, "Alternating Direction Implicit Methods," Advances in Computers Series, (F. Alt and M. Rubinoff, Eds.), Vol. 3, pp. 189–273, Academic Press, New York, 1962.
12. Willis H. Guilinger, Jr., The Peaceman–Rachford method for small mesh increments, *J. Math. Anal. Appl.* **11** (1965), 261–277.
13. G. J. Habetler and E. L. Wachspress, Symmetric successive overrelaxation in solving diffusion difference equations, *Math. Comp.* **15** (1961), 356–362.
14. David M. Young and David R. Kincaid, "Norms of the Successive Overrelaxation Method and Related Methods," TNN-94, Computation Center, University of Texas, Austin, Tex., 1969.
15. J. Sheldon, On the numerical solution of elliptic difference equations, *Math. Tables Aids Comput.* **9** (1955), 101–112.
16. Louis W. Ehrlich, "The Block Symmetric Successive Overrelaxation Method," doctoral thesis, University of Texas, Austin, Tex., 1963.
17. Louis W. Ehrlich, The block symmetric successive overrelaxation method, *J. Soc. Indus. Appl. Math.* **12** (1964), 807–826.