Results in Mathematics
Vol. 16 (1989)

0378-6218/89/040308-13$1.50+0.20/0

Relaxation Methods for Non–Hermitian Linear Systems

W. Niethammer and R. S. Varga*

Abstract

For the matrix equation $A\mathbf{x} = \mathbf{b}$, we consider here two splittings $A = M_1 - N_1 = M_2 - N_2$ of the matrix A, where $M_1 := (A + A^*)/2$ is the Hermitian part of A, and $M_2 := I + (A - A^*)/2$ is the identity plus the skew–Hermitian part of A. To these two splittings of A, we apply an extrapolation, with extrapolation factor ω, and we find associated regions for ω, in the complex plane, for which these extrapolated splittings yield convergent iterative methods. From this, further applications to semiiterative methods are indicated.

Dedicated to Professor Karl Zeller on the occasion of his 65th birthday.

*The research of this author was supported by the Alexander von Humboldt Foundation.

1 Introduction

Given a nonsingular system of linear equations

$$(1.1) \qquad A\mathbf{x} = \mathbf{b} \qquad (A \in \mathbb{C}^{n,n}),$$

let $A = M - N$ be any *splitting* of A, i.e., M and N are in $\mathbb{C}^{n,n}$ with M nonsingular. From this splitting, we have

$$(1.2) \qquad M\mathbf{x} = N\mathbf{x} + \mathbf{b},$$

or equivalently, on setting $T := M^{-1}N$ and $\mathbf{c} := M^{-1}\mathbf{b}$,

$$(1.3) \qquad \mathbf{x} = T\mathbf{x} + \mathbf{c}.$$

This last equation induces the iterative method

$$(1.4) \qquad \mathbf{x}_{m+1} := T\mathbf{x}_m + \mathbf{c} \qquad (m = 0, 1, \ldots).$$

It is well–known that the vector iterates $\{\mathbf{x}_m\}_{m=0}^{\infty}$ of (1.4) converge (for an arbitrary starting vector $\mathbf{x}_0$) to the unique solution of (1.1), iff $\rho(T) < 1$ (where, as usual, $\rho(T) := \max\{|\lambda_j| : \lambda_j \text{ is an eigenvalue of } T\}$).

To introduce an *extrapolation* in (1.4), consider the splittings of $A = M_\omega - N_\omega$ where

$$(1.5) \qquad M_\omega := \left(\frac{1}{\omega} - 1\right) I + M, \text{ and } N_\omega := \left(\frac{1}{\omega} - 1\right) I + N,$$

for any $\omega \neq 0$ for which M_ω is nonsingular. Then, with

$$(1.6) \qquad \mathcal{T}_\omega := M_\omega^{-1} N_\omega \text{ and } \mathbf{c}_\omega := M_\omega^{-1}\mathbf{b},$$

we have, in analogy with (1.4), the associated *extrapolated* iterative method

$$(1.7) \qquad \mathbf{y}_{m+1} := \mathcal{T}_\omega \mathbf{y}_m + \mathbf{c}_\omega \qquad (m = 0, 1, \ldots).$$

Since the iterative method of (1.7) reduces to that of (1.4) when $\omega = 1$, the idea is to select appropriate values of ω which either *produce* convergence in (1.7) (when (1.4) is divergent), or produce *faster* convergence in (1.7) than that of (1.4).

Returning to (1.2) and (1.5), it is tacitly assumed here that equations of the form $M\mathbf{y} = \mathbf{g}$ or $M_\omega \tilde{\mathbf{y}} = \tilde{\mathbf{g}}$ can be *directly* solved when $\mathbf{g}$ or $\tilde{\mathbf{g}}$ is given. This obviously places restrictions on the matrices M and M_ω. If we represent the original matrix A of (1.1) as

$$(1.8) \qquad A = D - L - U,$$

where D is a nonsingular (block) diagonal matrix and L and U are, respectively, strictly lower and strictly upper triangular matrices, then the choice $M := D$ and $N := L + U$ in (1.2) defines the (block) Jacobi iterative method, while the choice of $M := D - L$ and $N := U$ in (1.2) defines the (block) Gauss–Seidel iterative method. Similarly the (block) Jacobi overrelaxation method (called

JOR) and the (block) succesive overrelaxation method (called SOR) have, as is well–known (cf. Young [15]), the associated iteration matrices:

$$\mathcal{J}_\omega := D^{-1}\{(1-\omega)D + \omega(L+U)\} \quad (JOR), \tag{1.9}$$

and

$$\mathcal{L}_\omega := (D-\omega L)^{-1}\{(1-\omega)D + \omega U\} \quad (SOR). \tag{1.10}$$

Much is known in the literature (cf. Varga [14] and Young [15]) about the influence of the extrapolation (or relaxation) parameter ω on the convergence behavior of the matrices $\mathcal{J}_\omega$ and $\mathcal{L}_\omega$ of (1.9) and (1.10), especially in the case when A is an Hermitian positive definite matrix.

The last twenty years have produced new fast direct methods for solving particular matrix equations such as the model problem (i.e., the finite difference approximation to the Dirichlet problem in the unit square), and these fast direct methods can be used to *define* the matrix M in the splitting $A = M - N$. The same is true of the powerful preconditioned conjugate gradient method (pcg), i.e., one 'peels off' from the matrix A that part, M, to which such pcg methods can be used to rapidly solve equations of the form $M\mathbf{x} = \mathbf{g}$. Such uses of fast direct solvers and the pcg method are generally applied to Hermitian positive definite matrix systems. Our interest here is in applications to *non–Hermitian non–positive definite* matrices, since such matrices arise naturally in the numerical solution of convection–diffusion equations.

A brief outline of our paper follows. In §2 we discuss some lesser known results on JOR and SOR when applied to general complex linear systems. In §§3 and 4, two new methods are examined which arise from a splitting $A = M - N$ where M is either the Hermitian part of A, or the identity plus the skew–Hermitian part of A. We study the influence of the relaxation parameter ω on convergence rates, and we describe regions in the complex plane which contain the spectrum of the corresponding iteration operators. The knowledge of these eigenvalue regions is useful for further applications of semiiterative methods.

2 Remarks on JOR and SOR

In the general framework (1.7) of extrapolation applied to arbitrary *complex* matrices, there is no reason to confine attention to *real* values for the extrapolation factor ω. Thus, we assume that ω is a complex parameter in what follows.

One of the most useful results in the general SOR theory for the convergence of the matrix $\mathcal{L}_\omega$ of (1.10) (in terms of complex values of ω) is the following well–known result of W. Kahan:

Lemma 2.1 (Kahan [9]). *For $A = D - L - U$, assume that D is nonsingular with $D^{-1}L$ and $D^{-1}U$ respectivly strictly lower and strictly upper triangular matrices. Then, $\mathcal{L}_\omega$ of* (1.10) *satisfies*

$$\rho(\mathcal{L}_\omega) \geq |\omega - 1|, \tag{2.1}$$

with equality only if all eigenvalues of $\mathcal{L}_\omega$ are in modulus $|\omega - 1|$. Thus, a necessary condition for $\mathcal{L}_\omega$ to be convergent is that ω be contained in the open disk with center 1 *and radius* 1.

We next discuss some lesser known results of Buoni and Varga [2,3] on necessary and sufficient conditions for the convergence of $\mathcal{J}_\omega$. Writing again $A = D - L - U$, we now assume only that D is nonsingular, i.e., we do not assume that D is block–diagonal, nor that $D^{-1}L$ and $D^{-1}U$ are triangular. With $\sigma(E) := \{\lambda : \det(\lambda I - E) = 0\}$ denoting the spectrum of an arbitrary matrix E in $\mathbb{C}^{n,n}$, then we set

$$\mathcal{K}(D^{-1}A) := \text{ closed convex hull of } \sigma(D^{-1}A), \tag{2.2}$$

and we let $\mathcal{K}^\circ(D^{-1}A)$ denote the interior of $\mathcal{K}(D^{-1}A)$. With this notation, we state the following result:

Lemma 2.2 (Buoni and Varga [2]). *$\mathcal{J}_\omega$ of (1.9) is divergent for all complex ω iff*

$$0 \in \mathcal{K}(D^{-1}A). \tag{2.3}$$

Moreover, if $0 \in \mathcal{K}^\circ(D^{-1}A)$, then there is a $r_0 > 0$ such that $\mathcal{L}_\omega$ of (1.10) also diverges for all complex ω with $0 \le |\omega| < r_0$.

This geometrical use in Lemma 2.2 of the convex set $\mathcal{K}(D^{-1}A)$, to determine convergence and divergence criteria for both $\mathcal{J}_\omega$ and $\mathcal{L}_\omega$, also gives geometrical conditions for the *simultaneous* convergence and divergence of $\mathcal{J}_\omega$ and $\mathcal{L}_\omega$, in the spirit of the classical Stein–Rosenberg Theorem (cf. [14] and [15]). With the notation of

$$\begin{cases} \Omega_{\mathcal{L}} := \{\omega \in \mathbb{C} : \rho(\mathcal{L}_\omega) < 1\}; & \mathcal{D}_{\mathcal{L}} := \{\omega \in \mathbb{C} : \rho(\mathcal{L}_\omega) > 1\}, \\ \Omega_{\mathcal{J}} := \{\omega \in \mathbb{C} : \rho(\mathcal{J}_\omega) < 1\}; & \mathcal{D}_{\mathcal{J}} := \{\omega \in \mathbb{C} : \rho(\mathcal{J}_\omega) > 1\}, \end{cases} \tag{2.4}$$

then the second part of Lemma 2.2 gives

$$\mathcal{D}_{\mathcal{J}} \cap \mathcal{D}_{\mathcal{L}} \supseteq \{\omega \in \mathbb{C} : |\omega| < r_0 \text{ and } \omega \neq 1\} \text{ for some } r_0,$$ $\wedge\,|\vartheta|$

which is a simultaneous *divergence* result. For simultaneous *convergence*, we similarly have

Lemma 2.3 (Buoni and Varga [2]). *With $A = D - L - U$ where D is nonsingular, then*

$$\Omega_{\mathcal{L}} \cap \Omega_{\mathcal{J}} \neq \emptyset \quad \textit{iff} \quad 0 \notin \mathcal{K}(D^{-1}A). \tag{2.5}$$

More precisely, if $0 \notin \mathcal{K}(D^{-1}A)$, there are real values of $\hat{\theta}$ such that

$$\min \operatorname{Re}\left\{e^{i\hat{\theta}}\xi : \xi \in \sigma(D^{-1}A)\right\} > 0, \tag{2.6}$$

and

$$\begin{cases} \rho(\mathcal{J}_{re^{i\hat{\theta}}}) < 1 & \textit{for all } r > 0 \textit{ sufficiently small, and} \\ \rho(\mathcal{L}_{re^{i\hat{\theta}}}) < 1 & \textit{for all } r > 0 \textit{ sufficiently small.} \end{cases} \tag{2.7}$$

In the case when $D^{-1}A$ is *strongly stable*, i.e., $\operatorname{Re} \xi > 0$ for all $\xi \in \sigma(D^{-1}A)$, we see that (2.6) of Lemma 2.3 is satisfied for $\hat{\theta} = 0$. This gives us the useful result of the

Corollary 2.4 (Buoni and Varga [2]). *With $A = D - L - U$ where D is nonsingular, assume that $D^{-1}A$ is strongly stable. Then $\mathcal{L}_\omega$ and $\mathcal{J}_\omega$ are simultaneously convergent for all $\omega > 0$ sufficiently small.*

Assuming $0 \notin \mathcal{K}(D^{-1}A)$, then from (2.7), there are real values of $\hat{\theta}$ such that *both* $\mathcal{T}_{re^{i\hat{\theta}}}$ and $\mathcal{L}_{re^{i\hat{\theta}}}$ are convergent for all $r > 0$ sufficiently small. Clearly, this implies, for each fixed $r > 0$ sufficiently small, that there is a *best* real $\hat{\theta}_1(r)$ which minimizes $\rho(\mathcal{T}_{re^{i\theta}})$ as a function of θ, and a *best* real $\hat{\theta}_2(r)$ which minimizes $\rho(\mathcal{L}_{re^{i\theta}})$ as a function of θ. It turns out (cf. Buoni and Varga [3]) that there is a *unique* path $re^{i\hat{\theta}(r)}$ in the complex plane (for all r sufficiently small) which minimizes $\rho(\mathcal{T}_{re^{i\theta}})$, and this unique path in the complex plane of $re^{i\hat{\theta}_1(r)}$ is determined solely for $r > 0$ sufficiently small, by the convex set $\mathcal{K}(D^{-1}A)$. (Similar results for $\mathcal{L}_{re^{i\theta}}$ are given in [3].)

If the Hermitian part $(A + A^*)/2$ of A is positive definite, then A is evidently strongly stable as can be deduced from Bendixson's Theorem (cf. [13, p. 395]). Thus, from the Corollary 2.4, it follows that there exists an $\omega_g > 0$ such that the matrix $\mathcal{L}_\omega$ of (1.10) is convergent for all ω with $0 < \omega < \omega_g$; in Niethammer [10,11], ω_g is determined as a function of the extremal eigenvalues of some matrices related to A. If A is moreover Hermitian and positive definite, then the upper bound on ω for convergence becomes $\omega_g = 2$, as is well-known from the Theorem of Ostrowski [14, p. 77]; from this theorem, in addition, it follows that if, for a linear system with an Hermitian matrix A, $\mathcal{L}_\omega$ is convergent for all ω with $0 < \omega < 2$, then A is necessarily positive definite.

Now, let us assume that SOR, applied to a linear system with a non-Hermitian matrix A, converges for $0 < \omega < \omega_g$ with $0 < \omega_g \leq 2$; one may ask, analogous to Ostrowski's Theorem, if it follows that $(A + A^*)/2$ is positive definite? In general, this is not true, as can be seen from the following example: Let

$$A := \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}.$$

Then SOR converges for $0 < \omega < 2$ and arbitrary complex a, but $(A + A^*)/2$ is positive definite only for $|a| < 2$.

3 Hermitian Splittings

Starting with the matrix equation of (1.1), assume that $D := \operatorname{diag}(A)$ is nonsingular. Since both splittings $A = M - N$ and $D^{-1}A = D^{-1}M - D^{-1}N$ lead to the same iteration operator, we may assume, without loss of generality, that

$$A = I - B, \text{ where } \operatorname{diag}(B) = O. \tag{3.1}$$

It is convenient to regard any splitting $M - N$ of $A = I - B$ as having the identity incorporated into M, and we thus write

$$M := I - M_B, \text{ and } N := B - M_B. \tag{3.2}$$

(For example, on writing $B = L + U$ where L and U are strictly lower and strictly upper triangular matrices, then the choice $M_B = L$ gives the standard

Gauss–Seidel iteration matrix.) The corresponding extrapolated form of the splitting of (3.2) is, from (1.5),

$$M_\omega = \frac{1}{\omega} I - M_B; \text{ and } N_\omega = \left(\frac{1}{\omega} - 1\right) I + N, \tag{3.3}$$

whose associated iteration matrix is

$$\mathcal{T}_\omega = (I - \omega M_B)^{-1}\{(1-\omega)I + \omega N\}. \tag{3.4}$$

The following useful result is essentially well–known in the literature, but we include its proof, as it is quite short.

Lemma 3.1. *If $I - \omega M_B$ is nonsingular and if τ is an eigenvalue of $\mathcal{T}_\omega$ of (3.4) with eigenvector* $\mathbf{v}$, *normalized by $\mathbf{v}^*\mathbf{v} = 1$, then*

$$\tau = \frac{(1-\omega) + \omega\eta}{1 - \omega m}, \quad \textit{where } \eta := \mathbf{v}^* N \mathbf{v} \textit{ and } m := \mathbf{v}^* M_B \mathbf{v}. \tag{3.5}$$

Proof. Using (3.4), we obtain from $\mathcal{T}_\omega \mathbf{v} = \tau\mathbf{v}$ that $((1-\omega)I + \omega N)\mathbf{v} = \tau(I - \omega M_B)\mathbf{v}$, so that $\mathbf{v}^*\{(1-\omega)I + \omega N\}\mathbf{v} = \tau \mathbf{v}^*(I - \omega M_B)\mathbf{v}$. Thus, with the definitions of η and m, then $1 - \omega + \omega\eta = \tau(1 - \omega m)$, which gives (3.5). □

We remark that Albrecht [1] has applied the representation (3.5), in the case of the matrix $\mathcal{L}_\omega$, to obtain convergence results for the SOR iterative method for Hermitian positive definite matrices A for $0 < \omega < 2$. (This is of course related to Ostrowski's Theorem (cf. Varga [14, p. 77]).) Similarly, Niethammer [10,11] examined systems $A\mathbf{x} = \mathbf{b}$ where $(A + A^*)/2$ was positive definite, and he obtained, using (3.5), an interval $(0, \omega_g)$ for which the SOR iterative method is convergent for all $0 < \omega < \omega_g$. (It should be remarked here, that the first author did his thesis [11] with Professor Karl Zeller, who is being honored with this volume).

The (generally complex) numbers m and η appearing in (3.5) are all elements of the *field of values* for the matrices M and N, respectively. It is, of course, well–known (see, e.g., Stoer–Bulirsch [13, p. 85]) that the field of values of a matrix Q is the convex hull of the eigenvalues of Q when Q is a *normal* matrix. Because Hermitian and skew–Hermitian matrices are particular normal matrices, the numbers m and η of (3.5) can be directly estimated in these special cases. This is done below.

We next remark that Concus and Golub [5] describe a generalized conjugate gradient method for non–symmetric systems of linear equations. They use the splitting $A = \tilde{M} - \tilde{N}$ with $\tilde{M} := (A + A^*)/2$, and they assume that $\tilde{M}$ is positive definite. Then, $\tilde{M}$ is used as a preconditioner which means that in each step a system $\tilde{M}\mathbf{y} = \mathbf{g}$ has to be solved. This was in fact the motivation for this section.

From $A = I - B$ in (3.1), we write

$$\begin{aligned} &A = I - B = I - (F + G) \\ &\text{where } F := (B + B^*)/2 \text{ and } G := (B - B^*)/2, \end{aligned} \tag{3.6}$$

and we assume that the Hermitian part, namely $M := I - F$, of A is *positive definite*. This 'Hermitian' splitting leads, by (3.4), to the iteration operator

$$\mathcal{T}_\omega^h := (I - \omega F)^{-1}\{(1-\omega)I + \omega G\}. \tag{3.7}$$

If $\{f_j\}_{j=1}^n$ denotes the eigenvalues of F, with $\alpha := f_1 \leq f_2 \ldots \leq f_n =: \beta$, then as $\operatorname{diag}(B) = O$ implies $\operatorname{diag}(F) = O$, it follows that $\alpha \leq 0 \leq \beta$, and since $I - F$ is assumed to be positive definite, then $\beta < 1$, i.e.,

$$\alpha \leq 0 \leq \beta < 1. \tag{3.8}$$

Our next result has its analogue in [10, Theorem 1].

Theorem 3.2. *If the Hermitian part $(A + A^*)/2$ of A of (3.1) is positive definite, then $\mathcal{T}_\omega^h$ of (3.7) is convergent for*

$$0 < \omega < \omega_g := \frac{2(1-\beta)}{1 + \rho^2(G) - \beta^2}. \tag{3.9}$$

Proof. Let τ be any eigenvalue of $\mathcal{T}_\omega^h$. It follows from (3.5) of Lemma 3.1 that

$$\tau = \frac{(1-\omega) + i\omega\eta}{1 - \omega\gamma}, \quad \text{where } i\eta := \mathbf{v}^* G\mathbf{v} \text{ and } \gamma := \mathbf{v}^* F\mathbf{v}. \tag{3.10}$$

Since the eigenvalues $\{f_j\}_{j=1}^n$, of F lie in the interval $[\alpha, \beta]$, then η and γ necessarily satisfy

$$\alpha \leq \gamma \leq \beta, \text{ and } -\rho(G) \leq \eta \leq \rho(G). \tag{3.11}$$

Let us assume that

$$0 < \omega < 1/\beta. \tag{3.12}$$

Then, τ in (3.10) is defined since, by (3.11), we have $\gamma \leq \beta$. Now we have to show that there exists an ω_g such that, in (3.10), $|\tau| < 1$ for all possible values of η, γ and ω_g with $0 < \omega < \omega_g$. The absolute value of τ in (3.10) becomes maximal, when $\eta = \rho(G)$ and $\gamma = \beta$. Thus, the inequality $|\tau| < 1$ is guaranteed if (with $\rho = \rho(G)$)

$$\begin{aligned}
(1-\omega\beta)^2 &> (1-\omega)^2 + \omega^2\rho^2 \\
1 - 2\omega\beta + \omega^2\beta^2 &> 1 - 2\omega + \omega^2 + \omega^2\rho^2 \\
\omega[2(1-\beta) + \omega\beta^2 - \omega(1+\rho^2)] &> 0 \\
0 < \omega < \omega_g &:= \frac{2(1-\beta)}{1+\rho^2-\beta^2}
\end{aligned}$$

A direct calculation shows that $\omega_g \leq 1/\beta$ holds, i.e., τ satisfies $|\tau| < 1$ for all ω with $0 < \omega < \omega_g$. □

To decide whether it is worthwile to apply a semiiterative method (SIM) with respect to $\mathcal{T}_\omega^h$ (see [8]), we need first to determine a region Ω_ω in the complex plane which contains the spectrum of $\mathcal{T}_\omega^h$. We can directly see from (3.10) that Ω_ω can be taken to be the following rectangle in the complex plane:

$$\Omega_\omega = \{z \in \mathbf{C} : c \leq \operatorname{Re} z \leq d,\ |\operatorname{Im} z| \leq f\} \tag{3.13}$$
$$\text{with } c := \frac{1-\omega}{1-\omega\alpha},\ d := \frac{1-\omega}{1-\omega\beta} \text{ and } f := \frac{\omega\rho}{1-\omega\beta}.$$

This is indicated in Figure 1 below.

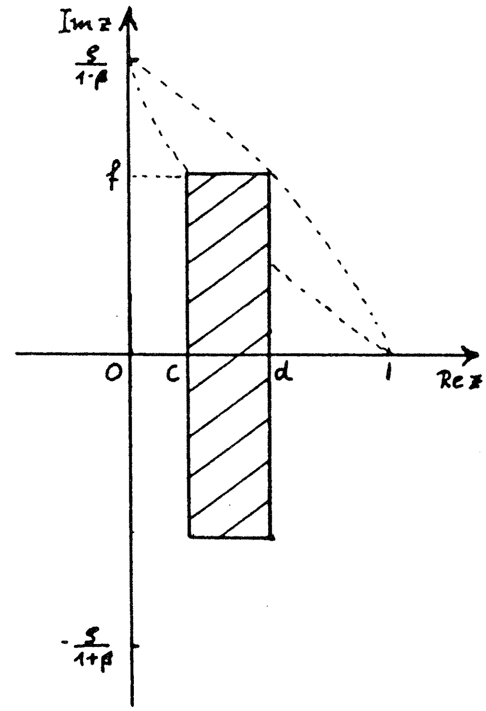

Figure 1: The set Ω_ω for $\mathcal{T}_\omega^h$

Let us first consider the case $\omega = 1$; here, $\Omega_\omega = \Omega_1$ of (3.13) degenerates to the interval

$$\Omega_1 := \{z = it :\ -\tilde{\rho} \leq t \leq \tilde{\rho} \text{ where } \tilde{\rho} := \rho/(1-\beta)\}. \tag{3.14}$$

(This was observed, too, in Concus and Golub [5]). A two-step method, which is an *asymptotically optimal semiiterative method* (AOSIM) with respect to Ω_1, is described in Niethammer and Varga [12] and has the *asymptotic convergence factor* (see [12], formula (7.3) with $\alpha := -i\tilde{\rho}$, $\beta := i\tilde{\rho}$)

$$\kappa(\Omega_1) = \frac{\tilde{\rho}}{1+\sqrt{1+\tilde{\rho}^2}}. \tag{3.15}$$

Now, for the case when $\omega \neq 1$, Ω_ω is the rectangle indicated in Figure 1. An AOSIM with respect to Ω_ω exists, but such a method requires storing all preceeding iterates of the associated iterative method. However a four-step method which approximates the AOSIM very well, is studied by Eiermann ([6]). A decision about the optimal choice of ω, i.e., about the optimal choice of Ω_ω, can be made using the hybrid results in Eiermann, Li and Varga ([7]).

4 A Skew–Hermitian Splitting

Let us again begin with the splitting (3.6) of a non–Hermitian matrix A:

$$A = I - B = I - F - G, \quad F = (B + B^*)/2, \quad G = (B - B^*)/2,$$

where the Hermitian part $I - F$ is assumed to be positive definite. We now consider the 'skew–Hermitian' splitting

$$A = M - N \quad \text{with} \quad M := I - G,\ N := F. \tag{4.1}$$

As this splitting requires the solution of a system with matrix $I - G$ in each iteration step, we will assume that such systems can be *easily* solved.

From (3.4), the resulting iteration operator is

$$\mathcal{T}_\omega^s := (I - \omega G)^{-1}\{(1-\omega)I + \omega F\}. \tag{4.2}$$

From Lemma 3.1, for an arbitrary eigenvalue τ of $\mathcal{T}_\omega^s$ with normalized eigenvector $\mathbf{v}$, we have

$$\tau = \frac{(1-\omega) + \omega\gamma}{1 - i\omega\eta}, \quad \text{where } \gamma := \mathbf{v}^* F \mathbf{v} \in \mathbb{R} \text{ and } i\eta := \mathbf{v}^* G \mathbf{v} \ \ (\eta \in \mathbb{R}). \tag{4.3}$$

For γ and η, again (3.11) holds: $\alpha \le \gamma \le \beta < 1$, $-\rho(G) \le \eta \le \rho(G)$ where α and β satisfy (3.8).

In order to describe a region Ω_ω in the complex plane containing the eigenvalues of $\mathcal{T}_\omega^s$, we remark that τ in (4.3) is of the form

$$\tau = s/(1 + it) \tag{4.4}$$

where $1 - \omega - \omega\alpha \le s \le 1 - \omega + \omega\beta$, and $-\omega\rho(G) \le t \le \omega\rho(G)$. Now, $1/(1+it)$ is on the circle with center $1/2$ and radius $1/2$; thus for fixed s, the complex number τ moves on a circle with center $s/2$ and radius $|s|/2$. It does not traverse the full boundary of this circle, but only moves through those values for which $0 \le t^2 \le \omega^2\rho^2(G)$. All these circles are contained in the two 'extremal' disks D_1 and D_2, having centers c_1 and c_2, with both disks touching at the origin, where

$$c_1 := (1 - \omega + \omega\beta)/2, \quad c_2 := (1 - \omega + \omega\alpha)/2. \tag{4.5}$$

This is indicated in Figure 2.

We remark that Chin and Manteuffel ([4]) found a similar region (cf. Figure 2) which contains the eigenvalues of the block SOR operator when applied to a discretized convection–diffusion equation. Because of its shape, they call this region a *bow-tie-region*. The slopes of the straight–line portions of these regions increase when the spectral radius $\rho(G)$ increases, but, independent of the size of $\rho(G)$, the eigenvalues of $\mathcal{T}_\omega^s$ are *always* contained in the two disks D_1 and D_2.

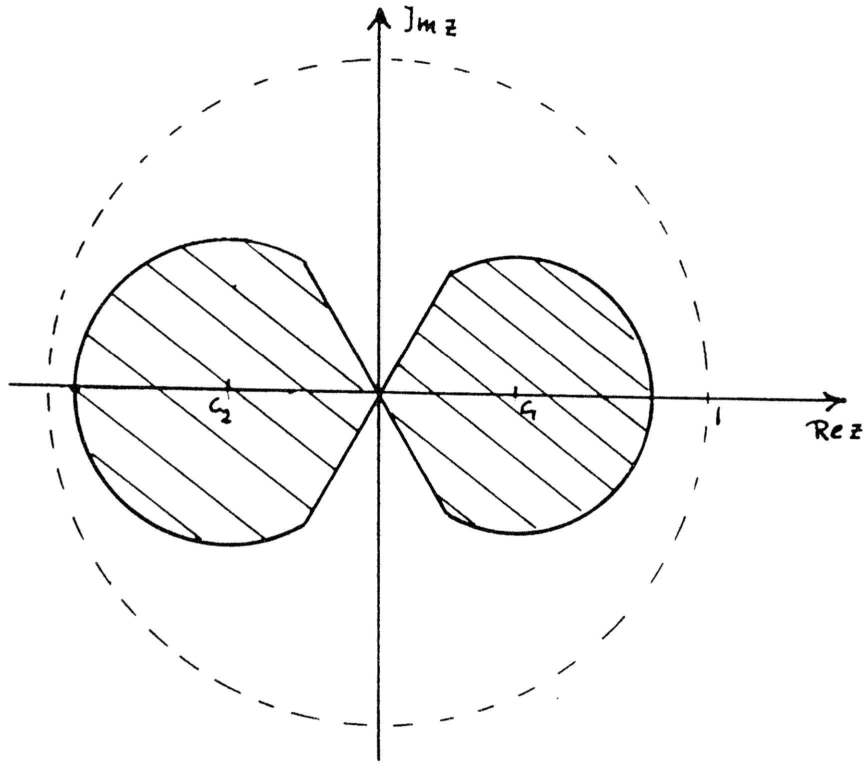

Figure 2: The set Ω_ω for $\mathcal{T}_\omega^s$

From Figure 2, it is clear that $\mathcal{T}_\omega^s$ is convergent if $c_2 > -1/2$, which yields $\omega < 2/(1-\alpha)$.

Increasing ω means that the disk D_2 is *increased*, while the disk D_1 is *decreased.* Thus, an optimal value ω_0 arises from the condition that the disks D_1 and D_2 have *equal radii.* This means $c_2 = -c_1$, and this gives

$$(4.6) \qquad \omega_0 = \frac{2}{2-(\alpha+\beta)}, \quad \rho(\mathcal{T}_\omega^s) = (1-\omega_0) + \omega_0\beta = \frac{\beta-\alpha}{2-(\beta-\alpha)}.$$

Thus, we have established

Theorem 4.1. *Let the Hermitian part* $(A+A^*)/2 = I - F$ *of* A *be positive definite, and let the eigenvalues of* F *satisfy* $\alpha := f_1 \leq \cdots \leq f_n =: \beta$. *Then,* $\mathcal{T}_\omega^s$ *is convergent for*

$$(4.7) \qquad 0 < \omega < \omega_g := 2/(1-\alpha).$$

A near optimal relaxation factor ω_0 *is*

$$(4.8) \qquad \omega_0 = \frac{2}{2-(\beta-\alpha)}$$

with

$$(4.9) \qquad \rho(\mathcal{T}_\omega^s) = \frac{\beta-\alpha}{2-(\beta-\alpha)}.$$

Remark: For $\alpha = -\beta$ we get $\omega_0 = 1$ and $\rho(\mathcal{T}_1^s) = \beta$, i.e., relaxation does not yield any acceleration of convergence. On the other hand, ω_g and ω_0 are *independent* of $\rho(G)$, i.e., the non–Hermitian part of A has no influence on ω_g and ω_0.

For the application of semiiterative methods, we can use the above information on the spectrum $\sigma(\mathcal{T}_\omega^s)$ of $\mathcal{T}_\omega^s$. Assuming $\omega = \omega_0$, as in (4.8), we may assume that, for arbitrary G, $\sigma(\mathcal{T}_\omega^s) \subset D_1 \cup D_2$ (see Figure 2) where the disks D_1 and D_2 have equal radii. To apply a one– or a two–step semiiterative method is equivalent (see [8]) to enclosing the set $D_1 \cup D_2$ by disks or ellipses, respectively. Since such an enveloping disk would have its center at the origin, no acceleration could be expected from the corresponding one–step semiiterative method. Even inclosing these disks by an ellipse (a two–step method) would yield only to a minor improvement in convergence rates. But, *hybrid methods*, introduced in [7], can be successfully applied here. This will be examined in a forthcoming paper.

5 Final Remarks

Many iterative methods for the solution of systems of linear equations $A\mathbf{x} = \mathbf{b}$ arise from a splitting $A = M - N$; the requirement that only such splittings are useful when the system $M\mathbf{y} = \mathbf{g}$ can be easily solved, resulted traditionally in choosing, for M, either diagonal or triangular matrices (or their corresponding block forms). Now, because of the development of very fast direct methods for the solution of special systems, and because of the construction of new (parallel) computer architectures, it is in order to think of new splittings.

We proposed two possibilities. The advantage of both may be that it is very easy to determine the Hermitian and the non–Hermitian part of A; a disadvantage may be (as it is indicated by the example at the end of §2) that the domains in the complex plane for the eigenvalues of the Hermitian and the skew–Hermitian parts of A, can be much larger than those of A, resp. the Jacobi matrix B.

Our aim here in this paper was to show how Lemma 3.1 can be used to describe regions in the complex plane which contain the spectrum of the corresponding iteration operators. It should be remarked that Lemma 3.1 can be used for other splittings in the same way. In a second step (which will be done in a forthcoming paper), this information concerning the spectrum will be used to define appropriate semiiterative methods through the use of hybrid methods.

Acknowledgment. We thank Dr. Michael Eiermann for his careful reading of the manusript and for his useful comments.

References

[1] J. Albrecht. Fehlerabschätzungen bei Relaxationsverfahren zur numerischen Auflösung linearer Gleichungssysteme. *Numer. Math.* 3, 188–201, 1961.

[2] J. J. Buoni and R. S. Varga. Theorems of Stein–Rosenberg type. In R. Ansorge, K. Glashoff, and B. Werner, editors, *Numerical Mathematics. ISNM 49*, p. 65–75, Birkhäuser Verlag, Basel, 1979.

[3] J. J. Buoni and R. S. Varga. Theorems of Stein–Rosenberg type II, Optimal paths of relaxation in the complex domain. In M. H. Schultz, editor, *Elliptic Problem Solvers*, p. 231–240, Academic Press, New York, 1981.

[4] R. C. Y. Chin and T. A. Manteuffel. An analysis of block succesive overrelaxation for a class of matrices with complex spectra. *SIAM J. Numer. Anal.* 25, 564–585, 1988.

[5] P. Concus and G. H. Golub. A generalized conjugate gradient method for nonsymmetric systems of linear equations. In R. Glowinski and J. R. Lions, editors, *Lect. Notes in Econ. and Math. Systems 134*, p. 56–65, Springer Verlag, New York, 1976.

[6] M. Eiermann. On semiiterative methods generated by Faber polynomials. (to appear in Numer. Math.).

[7] M. Eiermann, X. Li, and R. S. Varga. On hybrid semi-iterative methods. *SIAM J. Numer. Anal.* 26, 152–168, 1989.

[8] M. Eiermann, W. Niethammer, and R. S. Varga. A study of semiiterative methods for nonsymmetric systems of linear equations. *Numer. Math.* 47, 505–533, 1985.

[9] W. Kahan. *Gauss-Seidel methods of solving large systems of linear equations.* PhD thesis, University of Toronto, 1958.

[10] W. Niethammer. Relaxation bei nichtsymmetrischen Matrizen. *Math. Z.* 85, 319–327, 1964.

[11] W. Niethammer. *Überrelaxation bei linearen Gleichungssystemen mit schiefsymmetrischer Koeffizientenmatrix.* Dissertation, Universität Tübingen, 1964.

[12] W. Niethammer and R. S. Varga. The analysis of k-step iterative methods for linear systems from summability theory. *Numer. Math.* 41, 177–206, 1983.

[13] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer-Verlag, New York, Berlin, Heidelberg, 1980.

[14] R. S. Varga. *Matrix Iterative Analysis.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1962.

[15] D. M. Young. *Iterative Solution of Large Linear Systems.* Academic Press, New York, London, 1971.

Institut für Praktische Mathematik
Universität Karlsruhe
7500 Karlsruhe 1
Federal Republic of Germany

Institute for Computational Mathematics
Kent State University
Kent, Ohio 44242
U.S.A.

Eingegangen am 4. September 1989