Journal of Computational Mathematics, Vol.22, No.6, 2004, 817–832.

CONVERGENCE DOMAINS OF AOR TYPE ITERATIVE MATRICES FOR SOLVING NON-HERMITIAN LINEAR SYSTEMS *1)

Li Wang
(*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)
(*School of Mathematics and Computer Sciences, Nanjing Normal University, Nanjing 210097, China*)

Abstract

We discuss AOR type iterative methods for solving non-Hermitian linear systems based on Hermitian splitting and skew-Hermitian splitting. Convergence domains of iterative matrices are given and optimal parameters are investigated for skew-Hermitian splitting. Numerical examples are presented to compare the effectiveness of the iterative methods in different points in the domain. In addition, a model problem of three-dimensional convection-diffusion equation is used to illustrated the application of our results.

Mathematics subject classification: 65F10.
Key words: Convergence, AOR type iterative methods, Skew-Hermitian splitting, Hermitian splitting, Linear systems.

1. Introduction

Given a nonsingular system of linear equations

$$Ax = b, \quad A \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n, \tag{1.1}$$

where the coefficient matrix A is non-Hermitian, we assume that $D = diag(A)$ 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

$$A = I - B, \quad \text{where} \quad diag(B) = 0. \tag{1.2}$$

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

$$M = I - M_B, \quad \text{and} \quad N = B - M_B. \tag{1.3}$$

Then, with AOR type iteration matrix [4]

$$T_{\omega,\gamma} = (I - \gamma M_B)^{-1}\{(1-\omega)I + (\omega-\gamma)M_B + \omega N\}, \tag{1.4}$$

and $c_{\omega,\gamma} = \omega(I - \gamma M_B)^{-1}b$, we have the associated AOR type iterative method [1, 3]

$$x^{(i+1)} = T_{\omega,\gamma}x^{(i)} + c_{\omega,\gamma} \quad i = 1, 2, \ldots.$$

* Received October 29, 2002.
1) Subsidized by the Special Funds For Major State Basic Research Projects G1999032803, China.

Letting

$$F = \frac{B + B^*}{2} \quad \text{and} \quad G = \frac{B - B^*}{2}$$

denote, respectively, the Hermitian and skew-Hermitian parts of B, then the Hermitian splitting of A is defined by [4]

$$A = M^h - N^h \quad \text{with} \quad M^h = I - F \quad \text{and} \quad N^h = G \tag{1.5}$$

where we assume that M^h is invertible, which is, for instance, guaranteed if the Hermitian part M^h of A is positive definite. The associated skew-Hermitian splitting of A is given by [4]

$$A = M^s - N^s \quad \text{with} \quad M^s = I - G \quad \text{and} \quad N^s = F. \tag{1.6}$$

In this way, the specific splitting defined in (1.5) and (1.6) generates the following two AOR type iterative methods

$$x^{(m)} = T^h_{\omega,\gamma} x^{(m-1)} + c^h_{\omega,\gamma} \quad (m = 1, 2, \cdots), \tag{1.7}$$

where

$$T^h_{\omega,\gamma} = (I - \gamma F)^{-1}\{(1-\omega)I + (\omega - \gamma)F + \omega G\}, \quad c^h_{\omega,\gamma} = \omega(I - \gamma F)^{-1} b,$$

and

$$x^{(m)} = T^s_{\omega,\gamma} x^{(m-1)} + c^s_{\omega,\gamma} \quad (m = 1, 2, \cdots), \tag{1.8}$$

where

$$T^s_{\omega,\gamma} = (I - \gamma G)^{-1}\{(1-\omega)I + (\omega - \gamma)G + \omega F\}, \quad c^s_{\omega,\gamma} = \omega(I - \gamma G)^{-1} b,$$

Each of these methods depends on two parameters γ and ω.

The last fourty years have produced many methods for solving linear systems. Much is known in the literature [7, 8] about basic ones. AOR type method which was proposed by A Hadjidimos in [3] in 1978 is a accelerated overrelaxation method. Using Hermitian and skew-Hermitian matrix splitting and combining with krylov subspace iterative methods, many methods have been developed [2, 5]. In [1] Bai has also given the convergence domain of the matrix multisplitting relaxation methods. Based on the technique in [4], further discuss of the convergence of AOR type methods will be given in our paper.

The organization of this paper is as follows. In section 2, we study the convergence properties of AOR type iterative methods for Hermitian splitting and skew-Hermitian splitting and give the near optimal parameters for skew-Hermitian splitting. In section 3, the three-dimensional convection-diffusion equation is employed as a model problem to illustrate the application of our results. Numerical experiments are presented in section 4 to compare the effectiveness of our methods in different points of convergence domains.

2. Convergence of AOR type Iterative Methods

Lemma 2.1. *If $I-\gamma M_B$ is nonsingular and if τ is a eigenvalue of $T_{\omega,\gamma}$ of (1.4) with eigenvector v, normalized by $v^*v = 1$, then*

$$\tau = \frac{1 - \omega + (\omega - \gamma)m + \omega\eta}{1 - \gamma m}, \quad \textit{where} \quad \eta = v^* N v \quad \textit{and} \quad m = v^* M_B v. \tag{2.1}$$

Proof. Using (1.4), we obtain from $T_{\omega,\gamma}v = \tau v$ that $\{(1-\omega)I + (\omega-\gamma)M_B + \omega N\}v = \tau(I-\gamma M_B)v$, so that $v^*\{(1-\omega)I + (\omega-\gamma)M_B + \omega N\}v = \tau v^*(I-\gamma M_B)v$. Thus, with the definitions of η and m, $1-\omega+(\omega-\gamma)m+\omega\eta = \tau(1-\gamma m)$, which gives (2.1).

The general complex numbers m and η appearing in (2.1) 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 [6]) 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 (2.1) can be directly estimated in these special cases. This is done below. We assume that the Hermitian part, namely, $M = I - F$, of A is positive definite. If $\{f_j\}_{j=1}^n$ denotes the eigenvalues of F, with $\alpha = f_1 \leq f_2 \cdots \leq f_n = \beta$, then as $diag(B) = 0$ implies $diag(F) = 0$, it follows that $\alpha \leq 0 \leq \beta$, and since $I - F$ is assumed to be positive definite, then $\beta < 1$, i.e.,

$$\alpha \leq 0 \leq \beta < 1, \tag{2.2}$$

Now we can get the following results.

Theorem 2.1. *If the Hermitian part $\frac{A+A^*}{2}$ of A is positive definite, then $T^h_{\omega,\gamma}$ of (1.7) is convergent for*

$$\begin{cases} \frac{1}{\alpha} + \frac{(\alpha-1)^2+\rho^2(G)}{2\alpha(\alpha-1)}\omega & < \gamma < \frac{1}{\beta} + \frac{(\beta-1)^2+\rho^2(G)}{2\beta(\beta-1)}\omega \\ 0 & < \omega < \frac{2(1-\alpha)(1-\beta)}{(1-\alpha)(1-\beta)+\rho^2(G)(1-\alpha-\beta)}. \end{cases} \quad \textit{where} \quad F \neq 0$$
$$\textit{or} \tag{2.3}$$
$$\begin{cases} 0 < \omega < \frac{2}{1+\rho^2(G)} \\ 0 < \gamma < \frac{1}{\beta} \end{cases} \quad \textit{where} \qquad F = 0$$

Proof. Let τ be any eigenvalue of $T^h_{\omega,\gamma}$. It follows from (2.1) that

$$\tau = \frac{1-\omega+(\omega-\gamma)m+i\omega\eta}{1-\gamma m}, \quad \text{where} \quad i\eta = v^*Gv, \quad m = v^*Fv. \tag{2.4}$$

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

$$\alpha \leq m \leq \beta \quad \text{and} \quad -\rho(G) \leq \eta \leq \rho(G). \tag{2.5}$$

Let us assume that

$$0 < \gamma < \frac{1}{\beta}, \quad \textit{or} \quad \frac{1}{\alpha} < \gamma < 0,$$

then τ in (2.4) is defined, since, by (2.5), we have $\alpha \leq m \leq \beta$. Now we have to show that there exists an area such that for all possible values m, η and γ, ω with (γ,ω) in the area $T^h_{\omega,\gamma}$ of (1.7) is convergent.

The inequality $|\tau| < 1$ is guaranteed if

$$[1-\gamma m+\omega(m-1)]^2 + \omega^2\eta^2 < (1-\gamma m)^2,$$

which is equivalent to

$$f(m) + \omega^2\eta^2 < 0.$$

where

$$f(m) = 2\omega(1-\gamma)(m-1) + (\omega^2 - 2\gamma\omega)(m-1)^2.$$

It is easy to know that $m_1 = 1, \quad m_2 = \frac{\omega-2}{\omega-2\gamma}$ satisfy $f(m) = 0$. By analysing the geometric properties of $f(m)$ and combining with $f(m) < 0$ we investigate the convergent domains as follows:

I If $F \neq 0$, we have $\alpha \neq 0$ and $\beta \neq 0$.

1. If $0 < \omega < 2\gamma$ and $0 \leq \gamma \leq 1$, it holds that $m_2 \geq 1$. The value of $f(m) + \omega^2\eta^2$ becomes maximal, when $\eta = \rho(G)$ and $m = \beta$. (See Figure 2.1(1)). Thus, the inequality $|\tau| < 1$ is guaranteed if $f(\beta) + \omega^2\rho^2(G) < 0$, which is equivalent to

$$\gamma < \frac{1}{\beta} + \frac{(\beta-1)^2 + \rho^2(G)}{2\beta(\beta-1)}\omega.$$

So τ satisfies $|\tau| < 1$ for all (ω, γ) with

$$\begin{cases} 0 < \frac{\omega}{2} < \gamma < \frac{1}{\beta} + \frac{(\beta-1)^2+\rho^2(G)}{2\beta(\beta-1)}\omega, \\ \gamma \leq 1, \\ 0 < \omega < \frac{2(1-\beta)}{\rho^2(G)+(1-\beta)}; \end{cases}$$

2. If $0 < \omega < 2\gamma$ and $\gamma > 1$, it holds that $m_2 < 1$. By $\beta < 1$ and $f(m) < 0$, see Figure 2.1(2), we have $\beta < m_2$, which means

$$\gamma < \frac{\beta\omega - \omega + 2}{2\beta},$$

so that the value of $f(m) + \omega^2\eta^2$ becomes maximal, when $\eta = \rho(G)$ and $m = \beta$. Thus, the inequality $|\tau| < 1$ is guaranteed if $f(\beta) + \omega^2\rho^2(G) < 0$, which is equivalent to

$$\gamma < \frac{1}{\beta} + \frac{(\beta-1)^2 + \rho^2(G)}{2\beta(\beta-1)}\omega.$$

So τ satisfies $|\tau| < 1$ for all (ω, γ) with

$$\begin{cases} \max\{\frac{\omega}{2}, 1\} < \gamma < \frac{1}{\beta} + \frac{(\beta-1)^2+\rho^2(G)}{2\beta(\beta-1)}\omega, \\ 0 < \omega < \frac{2(\beta-1)^2}{(\beta-1)^2+\rho^2(G)}; \end{cases}$$

3. If $\omega > 2\gamma$ and $0 \leq \gamma < 1$, it holds that $m_2 < 1$. By $\beta < 1$ and $f(m) < 0$, see Figure 2.1(3), we have $m_2 < \alpha$, which means

$$\gamma > \frac{\alpha\omega - \omega + 2}{2\alpha}.$$

Thus, the inequality $|\tau| < 1$ is guaranteed if

$$\begin{cases} f(\alpha) + \omega^2\rho^2(G) & < & 0, \\ f(\beta) + \omega^2\rho^2(G) & < & 0, \end{cases}$$

which is equivalent to

$$\frac{1}{\alpha} + \frac{(\alpha-1)^2+\rho^2(G)}{2\alpha(\alpha-1)}\omega < \gamma < \frac{1}{\beta} + \frac{(\beta-1)^2+\rho^2(G)}{2\beta(\beta-1)}\omega.$$

So τ satisfies $|\tau|<1$ for all (ω,γ) with

$$\begin{cases} \frac{1}{\alpha} + \frac{(\alpha-1)^2+\rho^2(G)}{2\alpha(\alpha-1)}\omega < \gamma < \min\{\frac{1}{\beta} + \frac{(\beta-1)^2+\rho^2(G)}{2\beta(\beta-1)}\omega, \frac{\omega}{2}, 1\}, \\ \gamma \geq 0, \\ 0<\omega<\min\{\frac{2(1-\alpha)}{(1-\alpha)+\rho^2(G)}, \quad \frac{2(1-\beta)}{(1-\beta)^2+\rho^2(G)}, \quad \frac{2(1-\alpha)(1-\beta)}{(1-\alpha)(1-\beta)+\rho^2(G)(1-\alpha-\beta)}\}; \end{cases}$$

4. If $\omega > 2\gamma$ and $\gamma \geq 1$, it holds that $m_2 \geq 1$. With $\alpha < 0$ it is easy to know that there is no (ω,γ) for $|\tau|<1$. (See Figure 2.1(4)).
5. If $\omega = 2\gamma$ and $0<\gamma<1$, it holds that $f(m) = 4\gamma(1-\gamma)(m-1)$. By $\beta < 1$ and $f(m) < 0$, the value of $f(m)+\omega^2\eta^2$ becomes maximal, when $m=\beta$ and $\eta = \rho(G)$. (See Figure 2.1(5)). Thus, the inequality is guaranteed if

$$0 < \frac{\omega}{2} = \gamma < \frac{1-\beta}{1-\beta+\rho^2(G)};$$

6. If $\omega = 2\gamma$ and $\gamma \geq 1$, it is easy to see that there is no (ω,γ) which satisfies $|\tau|<1$. (See Figure 2.1(6)).
7. If $\omega > 0$ and $\frac{1}{\alpha} < \gamma < 0$, it holds that $m_2 < 1$. By $\beta < 1$ and $f(m) < 0$, we have $m_2 < \alpha$ which means $\gamma > \frac{1}{\alpha} + \frac{\alpha-1}{2\alpha}\omega$. Thus, with analogous technique used in 3, and see Figure 2.1(3), the inequality $|\tau|<1$ is guaranteed if

$$\begin{cases} f(\alpha)+\omega^2\rho^2(G) & < & 0, \\ f(\beta)+\omega^2\rho^2(G) & < & 0. \end{cases}$$

So, τ satisfies $|\tau|<1$ for all (ω,γ) with

$$\begin{cases} \frac{1}{\alpha} + \frac{(\alpha-1)^2+\rho^2(G)}{2\alpha(\alpha-1)}\omega < \gamma < \min\{\frac{1}{\beta} + \frac{(\beta-1)^2+\rho^2(G)}{2\beta(\beta-1)}\omega, 0\}, \\ 0<\omega<\min\{\frac{2(1-\alpha)}{(1-\alpha)^2+\rho^2(G)}, \quad \frac{2(1-\alpha)(1-\beta)}{(1-\alpha)(1-\beta)+\rho^2(G)(1-\alpha-\beta)}\}; \end{cases}$$

8. With similar technique used above we can prove that there is no solution for $|\tau|<1$ if $\omega<0$ and $\gamma>0$, or $\omega<0$ and $\gamma<0$.

II If $F=0$, we have $\alpha=0$ and $\beta=0$, which means $m\equiv 0$. It is easy to obtain that τ satisfies $|\tau|<1$ for all (ω,γ) with

$$0< \quad \omega \quad < \tfrac{2}{1+\rho^2(G)}.$$

With above results, (2.3) is proved.

We now consider skew-Hermitian splitting (1.6). 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 (1.8) and (2.1), for an arbitrary eigenvalue τ of $T^s_{\omega,\gamma}$ with normalized eigenvector v, we have

$$\tau = \frac{1-\omega+\omega m+(\omega-\gamma)\eta i}{1-\gamma\eta i}, \quad \text{where} \quad m = v^*Fv \in \mathbb{R} \quad \text{and} \quad i\eta = v^*Gv \quad (\eta\in\mathbb{R}). \tag{2.6}$$

Figure 2.1(1)

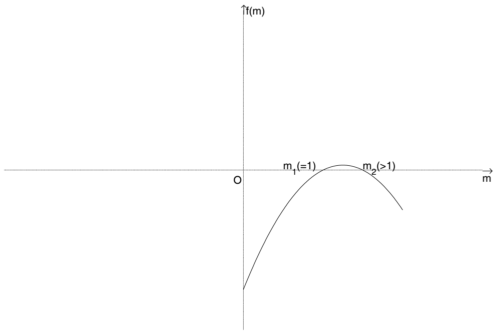


Figure 2.1(2)

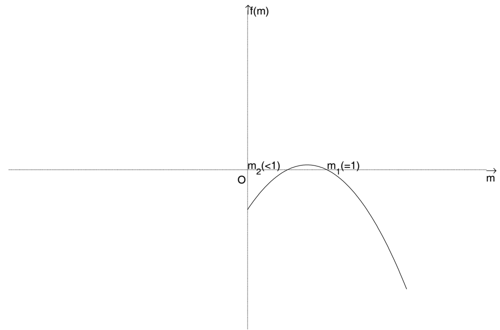


Figure 2.1(3)

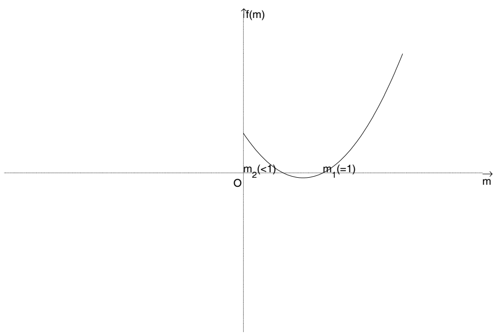
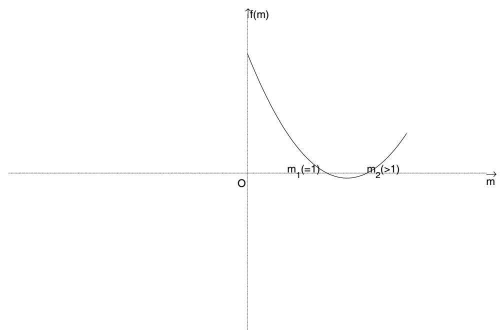


Figure 2.1(4)



For γ and η, again (2.5) holds: $\alpha \leq \gamma \leq \beta < 1, \quad -\rho(G) \leq \eta \leq \rho(G)$, where α and β satisfy (2.2). In this case, we have the following theorem.

Theorem 2.2.

Let the Hermitian part $\frac{A+A^*}{2} = I - F$ of A be positive definite, and let the eigenvalues of F satisfy $\alpha = f_1 \leq \ldots \leq f_n = \beta$. Then, $T^s_{\omega,\gamma}$ is convergent for

$$\begin{cases} 0 < \omega < \frac{2}{1-\alpha} \\ \frac{\omega}{2} < \gamma, \end{cases}$$

Figure 2.1(5)

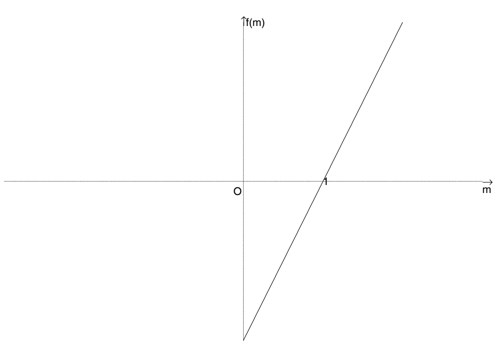
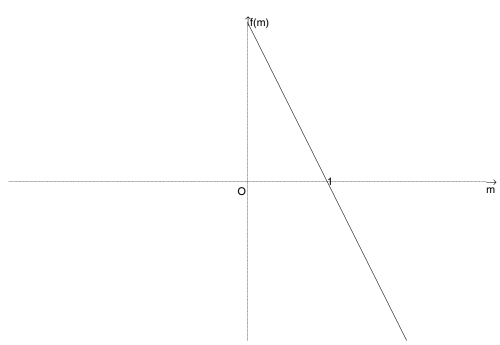


Figure 2.1(6)



The near optimal parameters ω and γ are

$$\begin{cases} \omega_0 = \frac{2}{2-(\beta+\alpha)}, \\ \frac{1}{1-\alpha} \leq \gamma \leq \frac{1}{1-\beta}, \end{cases}$$

with

$$\rho(T^s_{\omega,\gamma}) = \frac{\beta-\alpha}{2-(\beta+\alpha)}.$$

Proof. Noticing that τ in 2.6 is of the form

$$\tau = 1 - \frac{\omega}{\gamma} + \frac{s}{1+it},$$

where $s = \frac{\omega}{\gamma} - \omega + \omega m$ and $t = -\gamma\eta$. For $\omega > 0$, it holds that $s_\alpha \leq s \leq s_\beta$ where $s_\alpha = \frac{\omega}{\gamma} - \omega + \omega\alpha$ and $s_\beta = \frac{\omega}{\gamma} - \omega + \omega\beta$.

Now, $\frac{s}{1+it}$ is on the circle with center $\frac{s}{2}$ and radius $\frac{|s|}{2}$, then $1 - \frac{\omega}{\gamma} + \frac{s}{1+it}$ is on the circle with center $\frac{s}{2} + 1 - \frac{\omega}{\gamma}$ and radius $\frac{|s|}{2}$. It does not traverse the full boundary of this circle, but only moves through those values for which $0 \leq t^2 \leq \gamma^2\rho^2(G)$. All those 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 $(1 - \frac{\omega}{\gamma}, 0)$ where

$$c_1 = \frac{\frac{\omega}{\gamma} - \omega + \omega\beta}{2} + 1 - \frac{\omega}{\gamma}$$
$$c_2 = \frac{\frac{\omega}{\gamma} - \omega + \omega\alpha}{2} + 1 - \frac{\omega}{\gamma}$$

where $c_2 < c_1$, and it is easy to know that $T^s_{\omega,\gamma}$ is convergent for

$$\begin{cases} 0 < \omega < \frac{2}{1-\alpha}, \\ \frac{\omega}{2} < \gamma. \end{cases}$$

For $0 < \frac{\omega}{\gamma} \leq 1$,

1. Increasing ω from 0 to $\frac{1}{1-\alpha}$ with $0 < \omega \leq \gamma \leq \frac{1}{1-\alpha}$, which guarantee $s_\beta \geq 0$ and $s_\alpha \geq 0$, means that the maximal distance $1 - \omega + \omega\beta$ of D_1 from the origin is decreased, while $1 - \omega + \omega\alpha$, the maximal distance of D_2 from the origin, is decreased too, but smaller than the former. (Figure 2.2(1));

2. Increasing ω from 0 to $\frac{1}{1-\alpha}$ with $\frac{1}{1-\alpha} < \gamma < \frac{1}{1-\beta}$, which guarantee $s_\beta > 0$, $\quad s_\alpha < 0$, and $1 - \omega + \omega\alpha > 0$, $\quad 1 - \omega + \omega\beta > 0$, means that the maximal distance $1 - \omega + \omega\beta$ of D_1 from the origin is decreased, while the maximal distance of D_2 from the origin, (Figure 2.2(2)) is $1 - \frac{\omega}{\gamma}$ which is smaller than the former.

3. Increasing ω from 0 to $\frac{1}{1-\alpha}$ with $\gamma \geq \frac{1}{1-\beta}$, which guarantee $s_\beta < 0$, $\quad s_\alpha < 0$, $\quad 1 - \omega + \omega\alpha > 0$ and $1 - \omega + \omega\beta > 0$, means that the maximal distance $1 - \frac{\omega}{\gamma}$ of D_1 and D_2 gets the minimal value $\frac{\beta-\alpha}{1-\alpha}$ when $\omega = \frac{1}{1-\alpha}$ and $\gamma = \frac{1}{1-\beta}$. (Figure 2.2(3));

4. Increasing ω from $\frac{1}{1-\alpha}$ to $\frac{2\gamma}{1+\gamma-\gamma\alpha}(< \frac{1}{1-\beta})$ with $\frac{1}{1-\alpha} < \gamma < \frac{1}{1-\beta}$, which guarantee $1-\omega+\omega\alpha < 0$, $\quad 1-\omega+\omega\beta > 0$, $\quad s_\alpha < 0$, $\quad s_\beta > 0$ and $\frac{s_\alpha}{2} + 1 - \frac{\omega}{\gamma} > 0$, means that the maximal distance $1-\omega+\omega\beta$ of D_1 from the origin is decreased, while OB_1, the maximal distance of D_2 from the origin is smaller than the former. (Figure 2.2(4)). In this case, the minimum of $1-\omega+\omega\beta$ is $\frac{\beta-\alpha}{2-(\beta+\alpha)}$ when $\omega = \frac{2\gamma}{1+\gamma-\gamma\alpha}$ and $\gamma = \frac{1}{1-\beta}$;

5. Increasing ω from $(\frac{1}{1-\alpha} <)\frac{2\gamma}{1+\gamma-\gamma\alpha}$ to $\min\{\frac{1}{1-\beta}, \frac{2}{1-\alpha}\}$ with $\frac{1}{1-\alpha} < \gamma < \frac{1}{1-\beta}$, which guarantee $1-\omega+\omega\beta > 0$, $\quad 1-\omega+\omega\alpha < 0$, $\quad \frac{s_\alpha}{2}+1-\frac{\omega}{\gamma} < 0$, $\quad s_\alpha < 0$ and $s_\beta > 0$, means that the maximal distance $1-\omega+\omega\beta$ of D_1 from the origin is decreased, while $|1-\omega+\omega\alpha|$, the maximal distance of D_2 from the origin is increased. (Figure 2.2(5));

6. Increasing ω from $\frac{1}{1-\alpha}$ to $\frac{2\gamma}{1+\gamma-\gamma\alpha}(< \frac{2}{1-\alpha})$ with $\gamma \geq \frac{1}{1-\beta}$ which guarantee $1-\omega+\omega\alpha \leq 1-\omega+\omega\beta \leq 1-\frac{\omega}{\gamma}$, $\quad 1-\omega+\omega\alpha < 0$, $\quad \frac{s_\alpha}{2}+1-\frac{\omega}{\gamma} > 0$, $\quad s_\alpha < 0$ and $s_\beta < 0$, implies that the maximal distance $1-\frac{\omega}{\gamma}$ of D_1 and D_2 get minimal value $\frac{\beta-\alpha}{2-\alpha-\beta}$ when $\gamma = \frac{1}{1-\beta}$ and $\omega = \frac{2\gamma}{1+\gamma-\gamma\alpha}$. (Figure 2.2(6));

7. Increasing ω from $(\frac{1}{1-\alpha} <)\frac{2\gamma}{1+\gamma-\gamma\alpha}$ to $\frac{2}{1-\alpha}$ with $\gamma \geq \frac{1}{1-\beta}$ which guarantee $1-\omega+\omega\alpha < 0$, $\quad 1-\omega+\omega\alpha \leq 1-\omega+\omega\beta \leq 1-\frac{\omega}{\gamma}$, $\quad 1-\omega+\omega\alpha < 0$, $\quad s_\alpha < 0$, $\quad s_\beta < 0$, and $\frac{s_\alpha}{2}+1-\frac{\omega}{\gamma} < 0$, implies that the maximal distance $\omega-1-\omega\alpha$ of D_1 and D_2 get minimal value $\frac{\beta-\alpha}{2-\alpha-\beta}$ when $\gamma = \frac{1}{1-\beta}$ and $\omega = \frac{2\gamma}{1+\gamma-\gamma\alpha}$. (Figure 2.2(7)).

Figure 2.2(1)

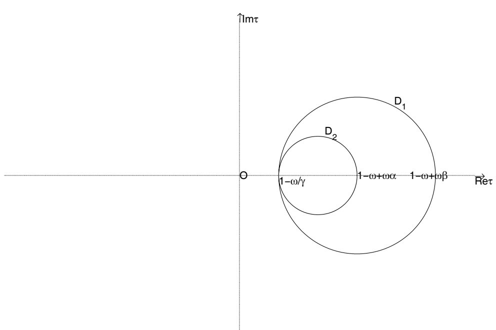


Figure 2.2(2)

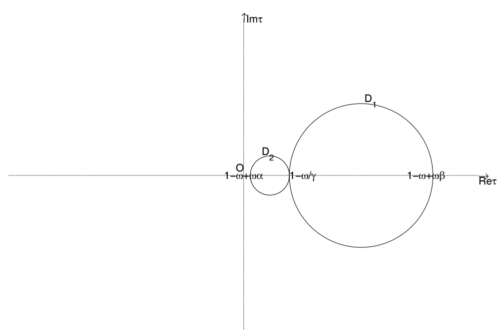


Figure 2.2(3)

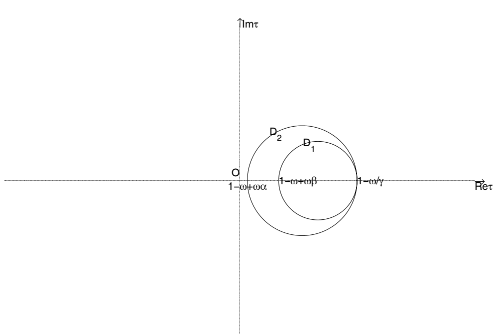
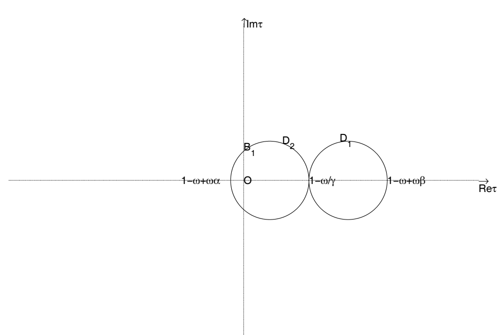


Figure 2.2(4)



For $1-\frac{\omega}{\gamma} \leq 0$, we can get analogous results when ω decreases. The detail is as follows.

Figure 2.2(5)

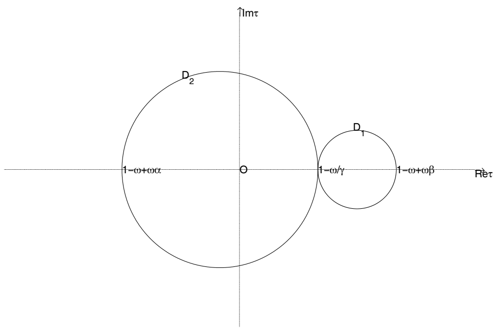


Figure 2.2(6)

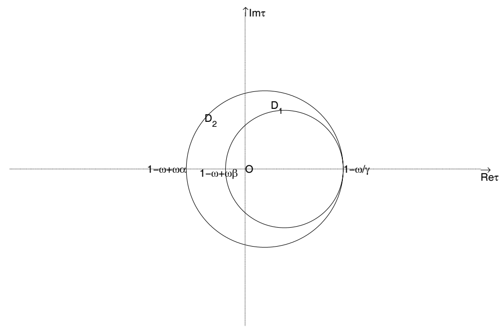


Figure 2.2(7)

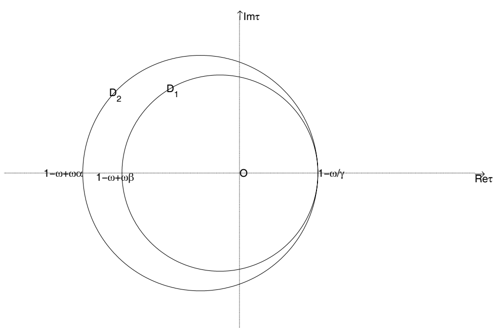
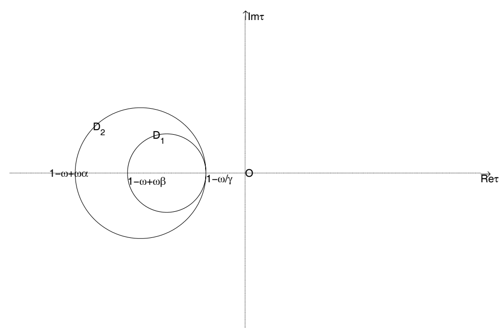


Figure 2.2(8).



1. Decreasing ω from $\max\{\frac{2}{1-\alpha}, \frac{1}{1-\beta}\}$ to $\frac{1}{1-\beta}$ with $\frac{1}{1-\beta} \le \gamma \le \omega$, which guarantee $s_\alpha \le 0, \quad s_\beta \le 0$, means that the maximal distance $|1-\omega+\omega\alpha|$ of D_2 from the origin is decreased, while $|1-\omega+\omega\beta|$, the maximal distance of D_1 from the origin is decreased too, but smaller than the former. (Figure 2.2(8));

2. Decreasing ω from $\max\{\frac{2}{1-\alpha}, \frac{1}{1-\beta}\}$ to $\frac{1}{1-\beta}$ with $\frac{1}{1-\alpha} \le \gamma \le \frac{1}{1-\beta}$, which guarantee $s_\alpha \le 0, \quad s_\beta \ge 0$, and $1-\omega+\omega\beta \le 0$, means that the maximal distance $|1-\omega+\omega\alpha|$ of D_2 from the origin is decreased, while the maximal distance of D_1 from the origin is $|1-\frac{\omega}{\gamma}|$ which is smaller than the former. (Figure 2.2(9));

3. Decreasing ω from $\max\{\frac{2}{1-\alpha}, \frac{1}{1-\beta}\}$ to $\frac{1}{1-\beta}$ with $\gamma < \frac{1}{1-\alpha}$, which guarantee $s_\alpha > 0, \quad s_\beta >$

Figure 2.2(9)

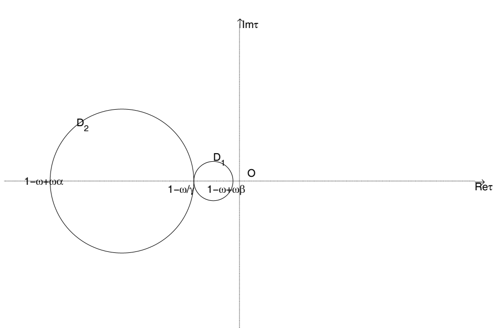


Figure 2.2(10)

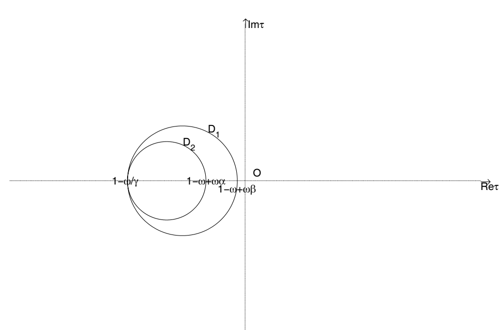


Figure 2.2(11)

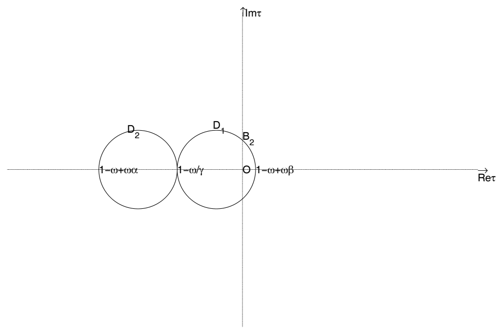


Figure 2.2(12)

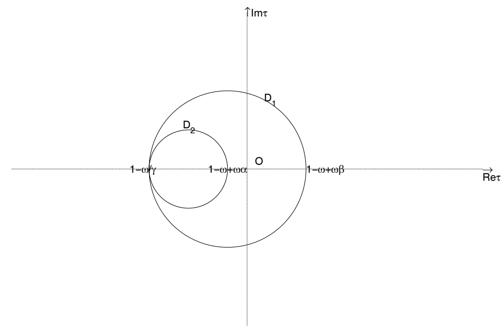


Figure 2.2(13)

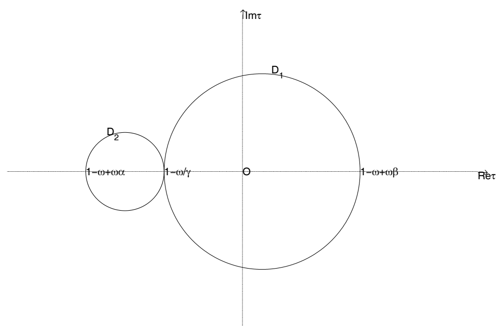
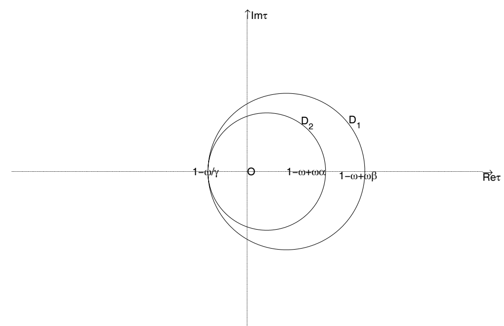


Figure 2.2(14)



0, $\quad 1-\omega+\omega\alpha \le 0$ and $1-\omega+\omega\beta \le 0$, implies that the maximal distance $|1-\frac{\omega}{\gamma}|$ of D_1 and D_2 get minimal value $\frac{\beta-\alpha}{1-\beta}$ when $\omega = \frac{1}{1-\beta}, \quad \gamma = \frac{1}{1-\alpha}$. (Figure 2.2(10));

4. Decreasing ω from $\frac{1}{1-\beta}$ to $\frac{2\gamma}{1+\gamma-\gamma\beta}(> \frac{1}{1-\alpha})$ with $\frac{1}{1-\alpha} \le \gamma \le \omega$, which guarantee $s_\alpha \le 0, \quad s_\beta \ge 0, \quad \frac{s_\beta}{2}+1-\frac{\omega}{\gamma} \le 0$, and $1-\omega+\omega\beta > 0$, means that the maximal distance $|1-\omega+\omega\alpha|$ of D_2 from the origin is decreased, while OB_2, the maximal distance of D_1 from the origin is smaller than it. (Figure 2.2(11));

5. Decreasing ω from $\frac{1}{1-\beta}$ to $\frac{2\gamma}{1+\gamma-\gamma\beta}(< \frac{1}{1-\beta})$ with $\gamma < \frac{1}{1-\alpha}$, which guarantee $s_\alpha \ge 0, \quad s_\beta \ge 0, \quad 1-\omega+\omega\beta \ge 0, \quad \frac{s_\beta}{2}+1-\frac{\omega}{\gamma} < 0$, implies that $\frac{\omega}{\gamma}-1$, the maximal distance of D_1 and D_2 get minimal value $\frac{\beta-\alpha}{2-\alpha-\beta}$ when $\omega = \frac{2\gamma}{1+\gamma-\gamma\beta}$ and $\gamma = \frac{1}{1-\alpha}$. (Figure 2.2(12));

6. Decreasing ω from $\frac{2\gamma}{1+\gamma-\gamma\beta}$ to $\frac{1}{1-\alpha}$ with $\frac{1}{1-\alpha} \le \gamma \le \frac{1}{1-\beta}$, which guarantee $s_\alpha \le 0, \quad s_\beta \ge 0$, and $\frac{s_\beta}{2}+1-\frac{\omega}{\gamma} \ge 0$, means that the maximal distance $|1-\omega+\omega\alpha|$ of D_2 from the origin is decreased, while $1-\omega+\omega\beta$, the maximal distance of D_1 from the origin is increased. (Figure 2.2(13));

7. Decreasing ω from $\frac{2\gamma}{1+\gamma-\gamma\beta}$ to $\frac{1}{1-\alpha}$ with $\gamma < \frac{1}{1-\alpha}$, which guarantee $s_\alpha > 0, \quad s_\beta > 0$, and $\frac{s_\beta}{2}+1-\frac{\omega}{\gamma} > 0$, means that the maximal distance $1-\omega+\omega\beta$ get minimal value $\frac{\beta-\alpha}{2-\alpha-\beta}$ when $\omega = \frac{2\gamma}{1+\gamma-\gamma\beta}$ and $\gamma = \frac{1}{1-\alpha}$. (Figure 2.2(14));

8. Decreasing ω from $\frac{1}{1-\alpha}$ to 0, with $0 < \gamma \le \omega \le \frac{1}{1-\alpha}$, which guarantee $s_\alpha \ge 0, \quad s_\beta \ge 0, \quad 1-\omega+\omega\beta > 0, \quad 1-\omega+\omega\alpha \ge 0$, which means that the maximal distance of D_1 reach minimum $\frac{\beta-\alpha}{1-\alpha}$ when $\omega = \frac{1}{1-\alpha}$ which is larger than $\frac{\beta-\alpha}{2-(\alpha+\beta)}$. (See Figure 2.2(14)).

Thus, an optimal value ω_0 arises from the condition that the disks D_1 and D_2 have equal distance from the origin. This means $-(1-\omega+\omega\alpha) = 1-\omega+\omega\beta$, and this gives $\omega_0 = \frac{2}{2-(\alpha+\beta)}$ with $\frac{1}{1-\alpha} \le \gamma \le \frac{1}{1-\beta}$ and $\rho(T^s_{\omega,\gamma}) = 1-\omega_0+\omega_0\beta = \frac{\beta-\alpha}{2-(\alpha+\beta)}$.

Remark. The near optimal parameters are not unique. The eigenvalues of iterative matrix $T^s_{\omega,\gamma}$ do not traverse the full boundaries of extremal circles D_1 and D_2, so the exact optimal value of the spectral radius is not larger than $\frac{\beta-\alpha}{2-(\beta+\alpha)}$.

3. Application to the Model Convection-diffusion Equation

For the three-dimensional(3D) convection-diffusion equation [2]

$$-(u_{xx}+u_{yy}+u_{zz})+q(u_x+u_y+u_z) = f(x,y,z)$$

on the unit cube $\Omega = [0,1]\times[0,1]\times[0,1]$, with constant coefficient q and subject to Dirichlet-type boundary conditions, we use the seven-point finite difference discretization with the centered differences to all the terms. Then we get the system of linear equations (1.1) with the coefficient matrix which has been transformed to the form (1.2):

$$A = T_x \otimes I \otimes I + I \otimes T_y \otimes I + I \otimes I \otimes T_z, \tag{3.1}$$

where the equidistant step-size $h = \frac{1}{n+1}$ is used in the discretization on all the three directions and the natural lexicographic ordering is employed to the unknowns. In addition, $\otimes$ denotes the Kroneck product, T_x, T_y and T_z are tridiagonal matrices given by

$$T_x = tridiag(t_2,t_1,t_3), \quad T_y = tridiag(t_2,0,t_3), \quad \text{and} \quad T_z = tridiag(t_2,0,t_3),$$

with

$$t_1 = 1, \quad t_2 = (-1-r)/6, \quad t_3 = (-1+r)/6.$$

Here, $r = \frac{qh}{2}$ is the mesh Reynolds number. It is easy to know that the Hermitian part H and the skew-Hermitian part S of the matrix A are

$$H = H_x \otimes I \otimes I + I \otimes H_y \otimes I + I \otimes I \otimes H_z$$

and

$$S = S_x \otimes I \otimes I + I \otimes S_y \otimes I + I \otimes I \otimes S_z,$$

where

$$H_x = tridiag(\tfrac{t_2+t_3}{2}, t_1, \tfrac{t_2+t_3}{2}), \quad H_y = H_z = tridiag(\tfrac{t_2+t_3}{2}, 0, \tfrac{t_2+t_3}{2}),$$
$$S_\xi = tridiag(\tfrac{t_2-t_3}{2}, 0, -\tfrac{t_2-t_3}{2}), \quad \xi \in \{x,y,z\}.$$

Then, we know that

$$F = F_x \otimes I \otimes I + I \otimes F_y \otimes I + I \otimes I \otimes F_z$$

and

$$G = G_x \otimes I \otimes I + I \otimes G_y \otimes I + I \otimes I \otimes G_z,$$

where

$$F_x = tridiag(-\tfrac{t_2+t_3}{2}, 0, -\tfrac{t_2+t_3}{2}), \quad F_y = F_z = tridiag(-\tfrac{t_2+t_3}{2}, 0, -\tfrac{t_2+t_3}{2}),$$
$$G_\xi = tridiag(-\tfrac{t_2-t_3}{2}, 0, \tfrac{t_2-t_3}{2}), \quad \xi \in \{x,y,z\}.$$

From Lemma 7.1 in Appendix in [2] and by direct computations, we get the concrete forms of Theorem 2.1 and Theorem 2.2.

Theorem 3.1. *For the system of linear equations (1.1) with the coefficient matrix (3.1) arising from the centered difference scheme, the iteration (1.7) converges for any initial guess to the unique solution when* ω, γ *satisfy:*

$$\begin{cases} 0<\omega<\frac{8\sin^2(\pi/(n+1))}{4\sin^2(\pi/(n+1))+(q\cos(\pi/(n+1))/(n+1))^2}, \\ u=-\frac{1}{cos(\pi/(n+1))}+\frac{(cos(\pi/(n+1))+1)^2+(q/(2(n+1))cos(\pi/(n+1)))^2}{2cos(\pi/(n+1))(1+cos(\pi/(n+1)))}\omega, \\ v=\frac{1}{cos(\pi/(n+1))}+\frac{(1-cos(\pi/(n+1))^2+(q/(2(n+1))cos(\pi/(n+1)))^2}{2cos(\pi/(n+1))(cos(\pi/(n+1))-1)}\omega, \\ u<\gamma<v. \end{cases}$$

Theorem 3.2. *For the system of linear equations (1.1) with the coefficient matrix (3.1) arising from the centered difference scheme, the iteration (1.8) converges for any initial guess to the unique solution when* ω, γ *satisfy:*

$$0<\omega<\frac{2}{1+cos(\pi/(n+1))}.$$

The near optimal parameters ω *and* γ *are*

$$\begin{cases} \omega_0=1, \\ \frac{1}{1+cos(\pi/(n+1))}\le\gamma\le\frac{1}{1-cos(\pi/(n+1))} \quad (n>2) \end{cases},$$

with

$$\rho(T_{\omega,\gamma}{}^s)=cos(\pi/(n+1)).$$

4. Numerical Example

In this section, we give a numerical example to illustrate the effectiveness of AOR type iterative methods for Hermitian splitting and skew-Hermitian splitting.

4.1 Spectral radius

In this subsection, we have computed and compared spectral radius of various points in the convergence domain. For Hermitian splitting, it is found that spectral radius decreases when the point moves from the origin to the boundary along straight line of the domain. So, we need only to compare the points in boundary. As the result of comparison, Table 4.1.1 shows that the optimal spectral radius for the model problem in section 3 can be got in the point which near $\omega=1$ and $\gamma=1$. For skew-Hermitian splitting, if we choose $n=10$, Theorem 3.2 shows that the near optimal radius is about 0.9595. The optimal radius should be not larger than this number. By computing, we have found that the optimal point is $\omega=1$, $\gamma=1$ and the optimal spectral radius is much smaller than the near one for the numerical example when q became larger. We can see these numerical results from Tables 4.1.2-4.1.3.

Table 4.1.1 Spectral radius for the model problem in section 3 ($n = 10$, $q = 1$)

ω	γ	$\rho(T^h_{\omega,\gamma})$	$\rho(T^s_{\omega,\gamma})$
1.6766	0.6707	0.9939	2.2852
1.9103	0.9551	0.9166	2.7418
1.6144	0.9686	0.6144	2.1621
1.3979	0.9785	0.3979	1.7379
1.2325	0.9860	0.2325	1.4141
1.1022	0.9920	0.1292	1.1588
0.9968	0.9968	0.1493	0.9587
1	1	0.1548	0.9586
0.9098	1.0008	0.1639	0.9623
0.8367	1.0041	0.1752	0.9653
0.7745	1.0069	0.2255	0.9679
0.7210	1.0094	0.2790	0.9701
0.6743	1.0115	0.3257	0.9721
0.6333	1.0134	0.3667	0.9738
0.5971	1.0150	0.4029	0.9753
0.5647	1.0165	0.4353	0.9766
0.5155	-0.5155	0.9987	0.9786
0.5424	-0.4881	0.9986	0.9775
0.6427	-0.3856	0.9985	0.9733
0.6849	-0.3425	0.9985	0.9716
0.7331	-0.2932	0.9984	0.9696
0.7886	-0.2366	0.9984	0.9673
0.8531	-0.1706	0.9985	0.9646
0.9292	-0.0929	0.9982	0.9614
1.0202	5.66e-17	0.9980	0.9980
1.1309	0.1131	0.9977	1.2152
1.2685	0.2537	0.9973	1.4853
1.4443	0.4333	0.9965	1.8300

Table 4.1.2 Spectral radius for the model problem in section 3 ($n = 10$, $q = 1$)

ω	γ	$\rho(T^h_{\omega,\gamma})$	$\rho(T^s_{\omega,\gamma})$
0.1	92	1.2801	0.9931
0.2	92	1.5603	0.9863
0.3	92	1.8404	0.9794
0.4	92	2.1205	0.9725
0.5	92	2.4007	0.9656
0.6	92	2.6808	0.9588
0.7	92	2.9610	0.9519
0.8	92	3.2411	0.9450
0.9	92	3.5212	0.9382
1	92	3.8014	0.9313
1.02	92	3.8574	0.9299
1	1	0.1548	0.9586

Table 4.1.3 Spectral radius for the model problem in section 3 ($n = 10$, $q = 3$)

ω	γ	$\rho(T^h_{\omega,\gamma})$	$\rho(T^s_{\omega,\gamma})$
0.1	42	1.2689	0.9945
0.2	42	1.5378	0.9890
0.3	42	1.8067	0.9835
0.4	42	2.0756	0.9780
0.5	42	2.3446	0.9725
0.6	42	2.6135	0.9670
0.7	42	2.8824	0.9614
0.8	42	3.1513	0.9560
0.9	42	3.4202	0.9504
1	42	3.6891	0.9449
1.02	42	3.7429	0.9438
1	1	0.4644	0.9514

Table 4.1.4 Spectral radius for the model problem in section 3 ($n = 10$, $q = 10$)

ω	γ	$\rho(T^h_{\omega,\gamma})$	$\rho(T^s_{\omega,\gamma})$
0.9	0.9	0.8032	0.8898
0.9	1	1.3968	0.8915
0.91	0.9	0.8068	0.8885
0.91	0.91	0.8348	0.8887
0.92	0.9	0.8105	0.8873
0.92	0.92	0.8707	0.8877
0.92	0.93	0.9062	0.8879
0.93	0.9	0.8142	0.8861
0.93	0.93	0.9119	0.8867
0.93	0.94	0.9527	0.8868
0.94	0.9	0.8181	0.8849
0.94	0.94	0.9593	0.8856
0.95	0.9	0.8221	0.8836
0.95	0.95	1.0148	0.8846
0.95	0.96	1.0717	0.8848
0.95	1	1.4715	0.8855
0.96	0.9	0.8262	0.8824
0.96	0.96	1.0804	0.8836
0.96	1	1.4867	0.8843
0.97	0.9	0.8304	0.8812
0.97	0.97	1.1595	0.8825
0.97	1	1.5019	0.8831
0.98	0.9	0.8347	0.8800
0.98	0.98	1.2572	0.8815
0.98	1	1.5172	0.8819
0.98	1.01	1.7304	0.8820
0.99	0.9	0.8392	0.8881
0.99	0.99	1.3818	0.8805
0.99	1	1.5326	0.8807
0.99	1.01	1.7488	0.8809
1	0.9	0.8437	0.9072
1	1	1.5480	0.8795
1	1.01	1.7673	0.8797

4.2 Iterative Effectiveness

The second test is for the three-dimensional convection-diffusion equation

$$-(u_{xx}+u_{yy}+u_{zz})+qexp(x+y+z)(xu_x+yu_y+zu_z)=f(x,y,z)$$

on the unit cube $\Omega=[0,1]\times[0,1]\times[0,1]$, with the homogeneous Dirichlet boundary conditions. In Tables 4.3.1-4.3.3 we list the numerical results for the centered difference schemes when $q=1,\quad n=10$. The results also show that for Hermitian splitting, the optimal value can not be got in $\omega=1$, $\gamma=1$. That is to say, SOR type iterative method is not optimal. For skew-Hermitian splitting, it is found that the larger q is, the faster the convergence speed is. When q is large enough the iterative methods for Hermitian splitting can hardly convergent. In the Tables, the symbol "t_i" $(i=1,2)$, represent CPU times, "R_i" $(i=1,2)$, denote spectral radius, "IT_i" $(i=1,2)$, mean iteration numbers, while $i=1,2$ are of Hermitian splitting and skew-Hermitian splitting respectively.

Table 4.2.1 $(n=10,\ q=1)$

ω	γ	t_1	IT_1	R_1	t_2	IT_2	R_2
0.1	1	12.406	137	0.9035	90.845	>1000	0.9966
0.2	1	6.700	66	0.8157	105.300	>1000	0.9932
0.3	1	4.942	45	0.7396	109.384	>1000	0.9899
0.4	1	4.733	35	0.6792	135.698	>1000	0.9865
0.5	1	3.846	30	0.6390	112.077	880	0.9831
0.6	1	2.858	29	0.6229	70.799	732	0.9797
0.7	1	2.947	30	0.6327	61.586	626	0.9764
0.8	1	3.288	33	0.6673	54.512	547	0.9730
0.9	1	7.577	42	0.7232	84.487	486	0.9696
1	0.9	4.555	44	0.7061	45.194	436	0.9689
1	1	6.301	59	0.7958	65.160	436	0.9662
1	1.1	NaN	NaN	5.9836	44.245	437	0.9663
1.1	0.8	7.739	79	0.8247	103.896	>1000	1.1684
1.1	0.9	3.728	39	0.6768	101.211	>1000	1.1658
1.1	1	11.900	105	0.8811	122.312	>1000	1.1628

Table 4.2.2 $(n=10,\ q=1)$

ω	γ	t_1	IT_1	R_1	t_2	IT_2	R_2
1	0.9	4.561	44	0.7061	45.549	436	0.968917
1	0.91	5.021	40	0.6789	79.224	436	0.968662
1	0.92	3.507	35	0.6461	44.977	436	0.968403
1	0.93	3.004	31	0.6056	43.795	436	0.968140
1	0.94	5.869	27	0.5547	95.876	436	0.967874
1	0.95	2.159	22	0.5433	43.796	436	0.967605
1	0.96	1.915	20	0.5704	42.238	436	0.967333
1	0.97	2.792	22	0.6039	55.393	436	0.967058
1	0.98	2.728	26	0.6481	44.353	436	0.966780
1	0.99	3.839	37	0.7102	47.774	436	0.966499
1	1	10.631	59	0.7958	81.192	436	0.966216
1	1.01	12.449	131	0.8955	41.733	436	0.966221
1	1.02	NaN	NaN	2.6246	42.052	436	0.966226
1	1.03	NaN	NaN	8.0909	55.647	436	0.966231
1	1.04	NaN	NaN	3.1230	41.971	437	0.966236
1	1.05	NaN	NaN	2.3553	44.982	437	0.966241
1	1.06	NaN	NaN	6.7334	44.559	437	0.966245

Table 4.2.3 ($n = 10$, $q = 10$)

ω	γ	t_1	IT_1	R_1	t_2	IT_2	R_2
0.9	0.8	NaN	NaN	8.0652	26.039	192	0.9175
0.9	0.9	NaN	NaN	14.5706	39.737	195	0.9185
1	0.9	NaN	NaN	16.1432	19.372	201	0.9358
1	0.92	NaN	NaN	27.7726	18.099	186	0.9305
1	0.94	NaN	NaN	9.1147	16.416	178	0.9253
1	0.95	NaN	NaN	9.0488	19.236	177	0.9228
1	0.96	NaN	NaN	13.8846	17.661	176	0.9203
1	1	NaN	NaN	8.9481	17.629	177	0.9106
1	1.1	NaN	NaN	11.6271	17.523	179	0.9117

Table 4.2.4 ($n = 10$, $q = 100$)

ω	γ	t_1	IT_1	R_1	t_2	IT_2	R_2
0.98	0.98	NaN	NaN	35.7764	5.094	51	0.6846
0.98	1	NaN	NaN	171.7967	5.176	50	0.6877
0.99	0.98	NaN	NaN	36.1517	7.658	53	0.6814
0.99	0.99	NaN	NaN	32.9507	4.942	52	0.6830
0.99	1	NaN	NaN	173.5395	5.010	51	0.6845
1	0.99	NaN	NaN	33.2834	8.625	55	0.6895
1	1	NaN	NaN	175.2823	5.587	53	0.6813
1	1.01	NaN	NaN	32.3496	5.521	52	0.6829

Acknowledgement. The author thanks for her supervisor Professor Zhong-Zhi Bai for his many suggestions and discussions.

References

[1] Z.Z. Bai, On the convergence domain of the matrix multisplitting relaxation methods for linear systems, *Appl. Math.-JCU*, **13B** (1998), 45-52.

[2] Z.Z. Bai, G.H.Golub, and Michael K.Ng,Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *SIAM J. Matrix Anal. Appl.,* to appear, 2002.

[3] A Hadjidimos, Accelerated overrelaxation method, *Math. Comput.*, **32** (1978), 149-157.

[4] M. Eiermann, W. Niethammer, and R.S. Varga, Acceleration of relaxation methods for non-Hermitian linear systems, *SIAM J. Matrix Anal. Appl.*, **13**:3 (1992), 979-991.

[5] O. Axelsson, Z.Z. Bai and S.X. Qiu, A class of nested iteration schemes for linear systems with a coefficient matrix with a dominant positive definite symmetric part, *Appl. Numer. Math.,* to appear, 2002.

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

[7] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1962.

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