Numer. Math. 48, 543–560 (1986)

Numerische Mathematik

The Rate of Convergence of Conjugate Gradients

A. van der Sluis[1] and H.A. van der Vorst[2]

[1] Mathematical Institute, University of Utrecht, Budapestlaan 6, Utrecht, The Netherlands
[2] Department of Mathematics, University of Technology, Julianalaan 132, Delft, The Netherlands

Summary. It has been observed that the rate of convergence of Conjugate Gradients increases when one or more of the extreme Ritz values have sufficiently converged to the corresponding eigenvalues (the "superlinear convergence" of CG). In this paper this will be proved and made quantitative. It will be shown that a very modest degree of convergence of an extreme Ritz value already suffices for an increased rate of convergence to occur.

Subject Classifications: AMS (MOS): 65F10; CR: G.1.3.

1. Introduction and Outline

It is the purpose of this paper to study the convergence behaviour of the Conjugate Gradients (CG) method for solving a linear system $Ax=b$, with A a positive definite symmetric matrix.

In particular we will explain the so-called superlinear convergence behaviour as observed in practice.

We will not consider the influence of rounding errors, i.e., we assume exact arithmetic.

Although the CG method, mathematically speaking, is *finite*, it is quite popular as an *iterative* method for many problems where A is large and sparse in some sense (e.g., A may be a preconditioned matrix that arises in discretized elliptic PDE's), since in those situations it tends to produce a sufficiently accurate approximation to x in far fewer steps than required for exact termination. In those situations it has also been observed that, during the process, the rate of convergence has the tendency to increase. This phenomenon is generally referred to as "superlinear convergence", although, mathematically speaking, this notion is not applicable on account of the finiteness of the process.

Already in the early fifties proofs for a true superlinear convergence behaviour of CG have been given for certain linear operators on a Hilbert space (where, on account of the infinite dimensionality, the mathematical notion of superlinear convergence does make sense: cf. [3, 7]; also [14]). However, these theories do not lead to useful estimates for the convergence behaviour during the small number of iterations that we are interested in.

In [1], p. 320, a link is made between the increasing rate of convergence as observed in practice, and the convergence of the so-called Ritz values. It is first observed that the rate of convergence of CG may be bounded in terms of the spectral condition number $\lambda_{\max}/\lambda_{\min}$, where $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalue of A for which a corresponding eigenvector is a component of the initial error vector (see also [5, 8]). Then they continue: "the extremal eigenvalues are approximated especially well (by the Ritz values) as CG proceeds, the iteration then behaving as if the corresponding eigenvectors are not present" (thus leading to a smaller "effective" condition number which might then explain the faster convergence).

Since the vanishing of eigenvectors from the error vector corresponds to the convergence of the corresponding Ritz *vectors*, it would thus seem that the convergence of the Ritz vectors is involved. Ritz vectors, however, converge much more slowly than Ritz values, and so, in this line of thought, one would expect the faster convergence of CG to occur only once the Ritz values have converged very well. Numerical experiments indicate otherwise, however. Anyhow, until now no proof nor a quantification has been given. It is the purpose of this paper to do this.

We will compare the convergence behaviour of the given CG process from a certain stage on (stage i, say) with the convergence behaviour of a comparison CG process for $Ax=b$, which has as its initial error vector the i-th error vector of the original process from which one or more extremal eigenvector components have been deleted (and which therefore really has a smaller condition number).

We will:

- prove that if i is such that one or more extremal Ritz values have arrived in their final intervals, and we remove the corresponding eigenvector components from the initial error vector of the comparison process, that then the error reduction in the next j steps of the original CG process is at most a fixed factor worse than the error reduction in the first j steps of the comparison process;
- prove that this factor is already close to 1 when the Ritz values in question have achieved a very modest degree of convergence (and hence from that moment on the rate of convergence of the CG process may indeed be bounded by the reduced condition number);
- explain why this factor may even be moderate if the Ritz values in question have barely entered their final intervals – or not even that.

For information about the convergence of Ritz values we may refer to [12] and [13]. The convergence of the Ritz vectors plays no role at all in our analysis.

At the end we will give some numerical illustrations.

2. Definitions and Basic Properties

Since the properties of CG and Ritz values which we need are so dispersed in the literature, we collect them in this section. It will be indicated how they follow from the basic projection property of CG. The defining formulae of CG will play no role, nor will tridiagonal matrices.

2.1. Characterizing Properties of the Conjugate Gradients Method

Let a linear system $Ax=b$ be given, A a symmetric positive definite matrix. Let x_0 denote an (arbitrary) starting vector and $r_0 := b - Ax_0$ the corresponding initial residual.

Then the standard Conjugate Gradients method (cf. [4] or [2], p. 362) for solving $Ax=b$ generates a sequence $x_1, x_2, \ldots$ with the following characterizing properties:

$$x_i \in x_0 + K_i \tag{2.1}$$

with $K_i = \operatorname{span}(r_0, Ar_0, \ldots, A^{i-1} r_0)$ (Krylov subspace)

$$\|x - x_i\|_A = \min_{u \in x_0 + K_i} \|x - u\|_A \tag{2.2}$$

where $\|v\|_A := (Av, v)^{\frac{1}{2}}$ for any v, or, equivalently,

$$x - x_i \perp_A K_i \tag{2.3}$$

where $v \perp_A w$ means $(Av, w) = 0$.

We note that (as is easily verified) for any vector u

$$u \in x_0 + K_i \Leftrightarrow x - u = q(A)(x - x_0) \quad \text{with } q \in \Pi_i^{ct1} \tag{2.4}$$

where Π_i^{ct1} denotes the class of polynomials of degree at most i and constant term 1, and then (2.2) implies

$$\|x - x_i\|_A = \min_{q \in \Pi_i^{ct1}} \|q(A)(x - x_0)\|_A. \tag{2.5}$$

2.2. The Active Eigenvalues and Eigenvectors

Corresponding to r_0 there are uniquely determined eigenvalues $\lambda_1 < \lambda_2 < \ldots < \lambda_m$ and normalized eigenvectors $z_1, \ldots, z_m$ of A such that $r_0 = \sum_1^m \mu_j z_j$, $\mu_j > 0$ for all j. These eigenvalues and eigenvectors are the *active* ones; in view of (2.1) the (possibly) other eigenvalues and eigendirections do not participate in the CG process.

Obviously $\dim K_i = i$ for $i \leq m$.

2.3. The Ritz Values and Vectors

For any $i \leq m$ the *Ritz values* $\theta_1^{(i)} < \theta_2^{(i)} < \ldots < \theta_i^{(i)}$ of A with respect to K_i are defined as the eigenvalues of the mapping $A_i := \pi_i A|_{K_i}$, π_i denoting the orthogonal projection upon K_i.

For any fixed k, $\theta_k^{(i)}$ decreases and $\theta_{i-k}^{(i)}$ increases as a function of i (as follows from Property 2.8 (below) since the polynomials R_i satisfy the orthogonality relation $(R_i(A)r_0, R_j(A)r_0)=0$ for $i \neq j$ on account of (2.3) and the roots of successive orthogonal polynomials separate each other; actually, we will only need that $\theta_k^{(i)}$ is non-increasing and $\theta_{i-k}^{(i)}$ non-decreasing, which already follows from a simpler minimax argument).

As a consequence, if $\theta_k^{(i)} \in [\lambda_k, \lambda_{k+1}]$ for a certain value of i, then also for all larger values of i. For this reason we shall refer to $[\lambda_k, \lambda_{k+1}]$ as the *final interval* of the k-th Ritz value. Similarly for the k-th Ritz value from above.

Any two Ritz values $\theta_k^{(i)}$ and $\theta_{k+1}^{(i)}$ are separated by at least one eigenvalue (since polynomials orthogonal with respect to a discrete innerproduct have the property that there is always at least one node between two zeros; or, cf. e.g. [10], theorem (10-4-6)).

The Ritz vectors $y_1^{(i)}, \ldots, y_i^{(i)}$ are normalized eigenvectors of A_i corresponding to $\theta_1^{(i)}, \ldots, \theta_i^{(i)}$.

Note that

$$A y_j^{(i)} - \theta_j^{(i)} y_j^{(i)} \perp K_i. \tag{2.6}$$

Moreover

$$(y_j^{(i)}, r_0) \neq 0 \qquad \text{for all } j. \tag{2.7}$$

In order to prove (2.7) we note that for $k \leqq i$: $(y_j^{(i)}, A^k r_0) = (A y_j^{(i)}, A^{k-1} r_0) = \theta_j^{(i)} (y_j^{(i)}, A^{k-1} r_0)$ (cf. (2.6)) $= \ldots = \theta_j^{(i)k} (y_j^{(i)}, r_0)$ and hence $(y_j^{(i)}, r_0) = 0$ leads to $y_j^{(i)} \perp K_i$, which is contradictory (cf. [6], Lemma 3 for an early occurrence of this property).

2.4. *A fundamental Relation Between CG Iterates and Ritz Values*

There exists the following fundamental property:

Property 2.8. *The CG iterates and the Ritz values are related by*

$$x - x_i = R_i(A)(x - x_0) \tag{2.9}$$

with $R_i(t) = \dfrac{(\theta_1^{(i)} - t) \ldots (\theta_i^{(i)} - t)}{\theta_1^{(i)} \ldots \theta_i^{(i)}}$.

Thus the minimum in (2.5) *is attained for* $q = R_i$.

Proof. Write $x - x_i = q(A)(x - x_0)$ (cf. (2.4)). Then $q(A)(x - x_0) \perp_A K_i$ (cf. (2.3)), implying $q(A) r_0 \perp K_i$. Hence $q(A_i) r_0 = q(\pi_i A) r_0 = \pi_i q(A) r_0 = 0$. Therefore $q(\theta_j^{(i)}) = 0$ for all j since $(r_0, y_j^{(i)}) \neq 0$ for all j (cf. (2.7)). □

This property is implicit in [4], it is explicit in [3], formula (3.7). We gave a more explicit proof.

We finally note that the property is equivalent to [10], Corollary (12-3-7), as is seen by observing that the polynomial μ considered there satisfies the same orthogonality relation as our R_i.

3. Results

3.1. Main Theorem, First Version

We will now present a first version of our main result. As has already been announced in the introduction, we will compare the error reduction of the CG process from a certain stage on (stage i, say) with the error reduction of a comparison CG process starting with an initial error vector obtained from the i-th error vector of the original process by deleting an appropriate eigenvector component.

Theorem 3.1. *Let x_i be the i-th iterate of the* CG *process for $Ax=b$* (cf. 2.1). *For a given integer i let $\bar{x}_j$ denote the j-th iterate of the comparison* CG *process for this equation, starting with $\bar{x}_0$ such that $x-\bar{x}_0$ is the projection of $x-x_i$ on* span$(z_2, \ldots, z_m)$. *Then for any j there holds*

$$\|x-x_{i+j}\|_A \leqq F_i \|x-\bar{x}_j\|_A \leqq F_i \frac{\|x-\bar{x}_j\|_A}{\|x-\bar{x}_0\|_A} \|x-x_i\|_A \tag{3.2}$$

with

$$F_i := \frac{\theta_1^{(i)}}{\lambda_1} \max_{k \geqq 2} \left| \frac{\lambda_k - \lambda_1}{\lambda_k - \theta_1^{(i)}} \right|. \tag{3.3}$$

Proof. Writing

$$x - x_0 = \sum_1^m \gamma_k z_k \tag{3.4}$$

we have $x-x_i = \sum_1^m R_i(\lambda_k)\gamma_k z_k$ (cf. (2.9)), and hence

$$x - \bar{x}_0 = \sum_2^m R_i(\lambda_k)\gamma_k z_k. \tag{3.5}$$

Hence (cf. (2.4)):

$$x - \bar{x}_j = \sum_2^m \bar{q}(\lambda_k) R_i(\lambda_k)\gamma_k z_k \tag{3.6}$$

for some $\bar{q} \in \Pi_j^{ct1}$.

Now define a polynomial $q \in \Pi_i^{ct1}$ as follows:

$$q(t) = \frac{(\lambda_1 - t)(\theta_2^{(i)} - t)\ldots(\theta_i^{(i)} - t)}{\lambda_1 \theta_2^{(i)} \ldots \theta_i^{(i)}} = \frac{\theta_1^{(i)}(\lambda_1 - t)}{\lambda_1(\theta_1^{(i)} - t)} R_i(t). \tag{3.7}$$

Then $\bar{q}q \in \Pi_{i+j}^{ct1}$ and hence, on account of (2.5)

$$\|x-x_{i+j}\|_A^2 \leqq \|\bar{q}(A)\,q(A)(x-x_0)\|_A^2 = \sum_2^m \lambda_k \bar{q}(\lambda_k)^2 q(\lambda_k)^2 \gamma_k^2. \tag{3.8}$$

Since $|q(\lambda_k)| \leqq F_i |R_i(\lambda_k)|$ for $k \geqq 2$ we get

$$\begin{aligned} \|x-x_{i+j}\|_A^2 &\leqq F_i^2 \sum_2^m \lambda_k \bar{q}(\lambda_k)^2 R_i(\lambda_k)^2 \gamma_k^2 \\ &= F_i^2 \|x-\bar{x}_j\|_A^2 \leqq F_i^2 \frac{\|x-\bar{x}_j\|_A^2}{\|x-\bar{x}_0\|_A^2} \|x-x_i\|_A^2. \quad \square \end{aligned} \tag{3.9}$$

Note. The idea of the theorem and its proof is the following. If you take an $\tilde{x}_0$ in x_0+K_i as starting vector for still another CG process with iterates $\tilde{x}_1, \tilde{x}_2, \ldots$ then $\tilde{x}_j \in x_0+K_{i+j}$, and hence $\|x-x_{i+j}\|_A \leqq \|x-\tilde{x}_j\|_A$. Thus, if you take in addition $\tilde{x}_0$ such that $x-\tilde{x}_0$ has no z_1-component then $\|x-x_{i+j}\|_A$ is bounded by the error $\|x-\tilde{x}_j\|_A$ of a process without z_1-component (which we wanted in the first place). Naturally, one will want $\tilde{x}_0$ to be close to x_i if $\theta_1^{(i)}$ is close to λ_1. This led to the choice of q (and $x-\tilde{x}_0=q(A)(x-x_0)$). Then, since $\tilde{x}_0 \approx \bar{x}_0$, bounding $\|x-\tilde{x}_j\|_A$ in terms of $\|x-\bar{x}_j\|_A$ seems reasonable.

3.2. Variants

Other choices for q and $\bar{q}$ in (3.8) give other results. We now give two such results which, under circumstances, give better insight in what happens than Theorem 3.1.

In the first variant we use the smallest Ritz value of the comparison process, which we denote by $\bar{\theta}_2^{(j)}$ since λ_2 is the smallest active eigenvalue for this process.

Property 3.10. *In* (3.2) F_i *may be replaced by*

$$\bar{F}_{ij} := \frac{\theta_1^{(i)} \bar{\theta}_2^{(j)}}{\lambda_1 \lambda_2} \max_{k \geqq 3} \left| \frac{\lambda_k - \lambda_1}{\lambda_k - \theta_1^{(i)}} \frac{\lambda_k - \lambda_2}{\lambda_k - \bar{\theta}_2^{(j)}} \right| \tag{3.11}$$

Proof. In (3.8) replace $\bar{q}$ by $\bar{\bar{q}} \in \Pi_j^{ct1}$, defined by

$$\bar{\bar{q}}(t) = \frac{\bar{\theta}_2^{(j)}}{\lambda_2} \frac{\lambda_2 - t}{\bar{\theta}_2^{(j)} - t} \bar{q}(t). \quad \square$$

Property 3.12. *The following inequality holds*

$$\|x-x_{i+j}\|_A \leqq \bar{F}_i \|x-\bar{x}_{j-p}\|_A \leqq \bar{F}_i \frac{\|x-\bar{x}_{j-p}\|_A}{\|x-\bar{x}_0\|_A} \|x-x_i\|_A \tag{3.13}$$

with

$$\bar{F}_i := \frac{\theta_1^{(i)}}{\lambda_1} \max_{k \geqq 3} \left| \frac{\lambda_k - \lambda_1}{\lambda_k - \theta_1^{(i)}} \right| \tag{3.14}$$

and

$$p = \left\lceil \frac{\pi}{4 \arcsin \sqrt{\frac{\lambda_2}{\lambda_m}}} \right\rceil \leqq \left\lceil \frac{\pi}{4} \sqrt{\frac{\lambda_m}{\lambda_2}} \right\rceil \tag{3.15}$$

where $\lceil \ldots \rceil$ *denotes rounding up.*

Proof. Let T_p be the Chebyshev polynomial of degree p, and $\mu := -\cos \frac{\pi}{2p}$ its left-most zero. Then $\cos \frac{\pi}{2p} \geqq 1 - 2\frac{\lambda_2}{\lambda_m}$ on account of $\frac{\pi}{2p} \leqq 2 \arcsin \sqrt{\frac{\lambda_2}{\lambda_m}}$. Transform the interval $[\mu, 1]$ linearly to $[\lambda_2, \lambda_m]$. Then -1 is mapped into $\frac{2\lambda_2-(1+\mu)\lambda_m}{1-\mu} \geqq 0$. Thus, if $\tilde{T}_p$ denotes the correspondingly "shifted" Chebyshev polynomial, we have $|\tilde{T}_p(0)| \geqq 1$ and hence $|\tilde{T}_p(\lambda_k)/\tilde{T}_p(0)| \leqq 1$ for $k \geqq 3$ and $\tilde{T}_p/\tilde{T}_p(0) \in \Pi_p^{ct1}$. Now replace in (3.8) $\bar{q}$ by $\bar{q}_{j-p}\tilde{T}_p/\tilde{T}_p(0) \in \Pi_j^{ct1}$, $\bar{q}_{j-p}$ denoting the polynomial $\bar{q}$ one has in (3.6) when j is replaced by $j-p$. □

Note. As is clear from the proof we might have defined

$$\bar{F}_i := \max_{k \geqq 3} \frac{\theta_1^{(i)}}{\lambda_1} \left| \frac{\lambda_k-\lambda_1}{\lambda_k-\theta_1^{(i)}} \cdot \frac{\tilde{T}_p(\lambda_k)}{\tilde{T}_p(0)} \right|. \tag{3.16}$$

We note that

$$\left| \frac{\tilde{T}_p(\lambda_k)}{\tilde{T}_p(0)} \right| \leqq \frac{2}{\lambda_m} \frac{p}{\sin \frac{\pi}{2p}} (\lambda_k-\lambda_2) \approx \frac{\pi}{4} \cdot \frac{\lambda_k-\lambda_2}{\lambda_2}$$

which gives a marked improvement if, e.g., $\theta_1^{(i)} \approx \lambda_2$ and λ_3 is close to λ_2.

4. Discussion

The theorem shows that from any stage i on for which $\theta_1^{(i)}$ does not happen to coincide with an eigenvalue λ_k, the error reduction in the next j steps is at most the fixed factor F_i worse than the error reduction in the first j steps of the comparison process in which the error vector has no z_1-component. We now discuss the magnitude of this factor in some cases, as well as alternative approaches when it is large.

4.1. The First Ritz Value Rather Close to λ_1

When $\lambda_1 < \theta_1^{(i)} < \lambda_2$ we have

$$F_i = \frac{\theta_1^{(i)}}{\lambda_1} \cdot \frac{\lambda_2-\lambda_1}{\lambda_2-\theta_1^{(i)}} \tag{4.1}$$

and we note that this is a kind of relative convergence measure for $\theta_1^{(i)}$, viz. relative to the quantities λ_1 and $\lambda_2-\lambda_1$. Obviously, F_i will already be close to 1 if $\theta_1^{(i)}$ has only achieved a very modest degree of convergence. E.g. $\frac{\theta_1^{(i)}-\lambda_1}{\lambda_1} < 0.1$ and $\frac{\theta_1^{(i)}-\lambda_1}{\lambda_2-\lambda_1} < 0.1$ is already enough to have $F_i < 1.25$. Hence, already for this modest degree of convergence of $\theta_1^{(i)}$, the process virtually converges as well as the comparison process, i.e. it virtually converges as well as if the z_1-component were not present. This proves and quantifies the observation in [1] mentioned in our introduction.

Numerical experiments (cf. §6) suggest that for F_i close to 1 the inequality (3.2) is reasonably sharp.

4.2. The First Ritz Value not Close to any Eigenvalue

(a) As we see from (4.1), if $\theta_1^{(i)} \in (\lambda_1, \lambda_2)$ but is not close to λ_2, and $\frac{\lambda_2}{\lambda_1}$ is not large, then F_i will be quite moderate.

Likewise, if $\theta_1^{(i)}$ is somewhere in the middle between λ_2 and λ_3 and both $\frac{\lambda_3}{\lambda_1}$ and $\frac{\lambda_3-\lambda_1}{\lambda_3-\lambda_2}$ are moderate, then F_i is moderate.

Hence, the convergence behaviour of a CG process may already be quite comparable to that of the comparison process (without the z_1-component) – in the sense that the error lags at most a fixed and moderate factor behind the error of the comparison process – well before $\theta_1^{(i)}$ has even entered its final interval.

(b) Numerical experiments (cf. § 6) suggest that if $\theta_1^{(i)}$ is not close to λ_1, (3.2) is not very sharp and that, indeed, F_i had better be replaced by a quantity rather close to $\frac{\theta_1^{(i)}}{\lambda_1}$. It is not hard to guess why this might be so. Actually, replacing $q(\lambda_k)$ by $F_i R_i(\lambda_k)$, as we did in going from (3.8) to (3.9), is reasonable only for those terms in (3.8) that have λ_k close to $\theta_1^{(i)}$, whereas for the other terms we have $q(\lambda_k) \approx \frac{\theta_1^{(i)}}{\lambda_1} R_i(\lambda_k)$; for the dense spectra that interest us the latter terms form by far the majority, and there is no evidence that the former terms are particularly big.

Unfortunately, since the distribution of the values of the polynomial R_i (cf. 2.4) in dependence on the γ_k is an open problem, we cannot make this reasoning quantitative.

(c) The mishap reported in (b) can, to some extent, be remedied by using Property 3.10. If $\theta_1^{(i)} < \lambda_3$ and $\bar{\theta}_2^{(j)} < \lambda_3$ then

$$\bar{F}_{ij} = \frac{\theta_1^{(i)}}{\lambda_1} \cdot \frac{\lambda_3-\lambda_1}{\lambda_3-\theta_1^{(i)}} \cdot \frac{\bar{\theta}_2^{(j)}}{\lambda_2} \cdot \frac{\lambda_3-\lambda_2}{\lambda_3-\bar{\theta}_2^{(j)}}. \tag{4.4}$$

Hence, for the values of j large enough in order that the factors containing $\bar{\theta}_2^{(j)}$ be close to 1, (3.2) holds with F_i replaced by $\bar{F}_{ij} \approx \bar{F}_j := \frac{\theta_1^{(i)}}{\lambda_1} \frac{\lambda_3-\lambda_1}{\lambda_3-\theta_1^{(i)}}$; in words: eventually the CG error is again majorized by the error of the comparison process, apart from the factor $\bar{F}_i$, which may be a good deal smaller than F_i if $\theta_1^{(i)}$ is below or not much above λ_2.

In order that this property be interesting in practical situations, $\bar{\theta}_2^{(j)}$ should already be close to λ_2 for moderate values of j. Unfortunately this will only happen if the z_2-component of $x-\bar{x}_0$ is not too small, and this requires that $\theta_1^{(i)}$ be not too close to λ_2. If, in particular, $\theta_1^{(i)} = \lambda_2$ then $x-\bar{x}_0$ has no z_2-component at all, and $\bar{\theta}_2^{(j)}$ will now approach λ_3 as j increases, making $\bar{F}_{ij}$ very large.

A less speculative, but cruder, result will be given in 4.3.

(d) Anyway, if $\theta_1^{(i)}$ is not close to λ_1, then we can say nothing better than that the error lags only a moderate factor behind the error of the comparison process, and that is worse than what we could say in 4.1. However, the reasoning in 4.1 is applicable only from larger values of i on than that in (a) or (c). Now, if $\theta_1^{(i)}$ goes to λ_1 rather rapidly, one may get a better description of the convergence behaviour of the CG process by waiting a few iterations and apply 4.1. If, however, $\theta_1^{(i)}$ goes to λ_1 very slowly (as it may when γ_1 is very small; cf. e.g. [11, 13]) then (a) or (c) will be more informative.

4.3. The First Ritz Value Close to a Wrong Eigenvalue

If $\theta_1^{(i)}$ is close to or equal to λ_2, say, we get very large values of F_i, or even ∞, and now the reasonings in 4.2(a) or (c) are not informative anymore. Moreover, in view of the arguments in 4.2(b), the value of F_i may now be very unrealistic.

If $\theta_1^{(i)}$ is close to λ_2 only in passing, then one need only apply one of the previous reasonings for a somewhat larger or smaller value of i. It may happen, however, that $\theta_1^{(i)}$ remains close to λ_2 for quite a few values of i (e.g. if γ_1 is small, and the process has, so to speak, not yet discovered that there still is an eigenvalue below λ_2; see again [11, 13]).

Here Property 3.12 helps. If $\theta_1^{(i)} < \lambda_3$ we have again $\bar{F}_i = \frac{\theta_1^{(i)}}{\lambda_1} \cdot \frac{\lambda_3 - \lambda_1}{\lambda_3 - \theta_1^{(i)}}$, and the property now says that the error lags at most a factor $\bar{F}_i$ and a fixed number of steps behind the error of the comparison process. This is rather surprising, in particular if $\theta_1^{(i)} = \lambda_2$, since then in the comparison process not only the z_1-component but also the z_2-component is missing.

In a somewhat speculative way we may translate the fixed number of steps mentioned above into an additional factor we "loose" when comparing the original process with the comparison process. Assuming that the "convergence factor" of the comparison process is reasonably given by $\frac{\sqrt{c_3}-1}{\sqrt{c_3}+1}$ with $c_3 = \lambda_m/\lambda_3$, then $p \approx \frac{\pi}{4}\sqrt{\frac{\lambda_m}{\lambda_2}}$ steps of this process correspond to a factor $\approx e^{\frac{\pi}{2}\sqrt{\lambda_3/\lambda_2}}$.

5. Main Theorem, General Version

Theorem 3.1 is a special case of the following more general theorem, in which an arbitrary number of eigenvector components in the error vector are replaced by 0. The proof requires only a different, but obvious, choice of q.

Theorem 5.1. *Let x_j be the j-th iterate of the* CG *process for $Ax=b$* (cf. § 2). *For given integers i, l and r let $\bar{x}_j$ denote the j-th iterate of the comparison* CG *process for this equation, starting with $\bar{x}_0$ such that $x-\bar{x}_0$ is the projection of $x-x_i$ on* $\operatorname{span}(z_{l+1}, \ldots, z_{m-r})$. *Then for any j there holds*

$$\|x - x_{i+j}\|_A \leqq F_{i,l,r} \|x - \bar{x}_j\|_A \leqq F_{i,l,r} \frac{\|x - \bar{x}_j\|_A}{\|x - \bar{x}_0\|_A} \|x - x_i\|_A \tag{5.2}$$

with

$$F_{i,l,r} = \max_{\substack{l' \geq l \\ r' \geqq r}} \prod_{j=1}^{l} \frac{\theta_j^{(i)}(\lambda_{l'} - \lambda_j)}{\lambda_j |\lambda_{l'} - \theta_j^{(i)}|} \prod_{j=1}^{r} \frac{\theta_{i+1-j}^{(i)}(\lambda_{m+1-j} - \lambda_{m-r'})}{\lambda_{m+1-j} |\theta_{i+1-j}^{(i)} - \lambda_{m-r'}|}. \tag{5.3}$$

Note 1. Obviously, F_i in (3.3) is identical with $F_{i,1,0}$.

Note 2. If the l left-most and r right-most Ritz values are in their final intervals, i.e.

$$\lambda_j < \theta_j^{(i)} < \lambda_{j+1} \quad \text{for } j \leqq l; \qquad \lambda_{m-j} < \theta_{i+1-j}^{(i)} < \lambda_{m+1-j} \quad \text{for } j \leqq r \tag{5.4}$$

then

$$F_{i,l,r} = \prod_{j=1}^{l} \frac{\theta_j^{(i)}(\lambda_{l+1} - \lambda_j)}{\lambda_j(\lambda_{l+1} - \theta_j^{(i)})} \prod_{j=1}^{r} \frac{\theta_{i+1-j}^{(i)}(\lambda_{m+1-j} - \lambda_{m-r})}{\lambda_{m+1-j}(\theta_{i+1-j}^{(i)} - \lambda_{m-r})}. \tag{5.5}$$

Note 3. Variants as in 3.2 may be given. Comments as in §4 may be made.

6. Numerical Experiments

6.1. Description of the Experiments

In this section we report some numerical experiments illustrating our theoretical results.

Since our present work was motivated by our research at preconditioned discretized differential operators we centered our experiments around matrices with spectra similar to those of the matrix P of the 5 point finite difference discretized Poisson operator over the unit square with stepsizes $\frac{1}{31}$ and Dirichlet boundary conditions (leading to a 900×900 matrix), using the ICCG(0) preconditioning (cf. [9]).

Actually, let $\tau_1 < \ldots < \tau_{900}$ denote the eigenvalues of P, $\tau_1, \ldots, \tau_7$ to 5 decimal places being 0.03420, 0.08179, 0.08293, 0.12707, 0.15508, 0.15537, 0.19110 and $\tau_{900} = 1.20455$. Let σ denote the set consisting of $\tau_1, \ldots, \tau_{21}$ and $\tau_{892}, \ldots, \tau_{900}$ and the 870 numbers equidistantly situated between τ_{21} and τ_{892} (the reason for this seemingly odd choice being that we used the same program to do some other experiments as well, and that we had already experienced that the eigenvalue distribution in the interior of the spectrum is of very little consequence for the convergence behaviour of CG). Then we may describe the matrices A we used as follows:

Experiment I. Since narrow pairs of eigenvalues near the endpoints of the spectrum tend to complicate the convergence behaviour of CG, we took as matrix A in our first experiment the 898×898 diagonal matrix with the elements of σ on its diagonal, leaving out τ_3 and τ_6 (thus avoiding two narrow pairs at the beginning).

Experiment II. Here we took as matrix A the 900×900 diagonal matrix with all the elements of σ its diagonal.

Experiment III. Here A differed from the matrix in Experiment I to the extent that we replaced its second eigenvalue (0.08179) by 0.0771268, with the effect that $\theta_1^{(6)} = \lambda_2$.

In all experiments we took $b = (1, 1, \ldots, 1)^T$ and $x_0 = 0$, implying $\mu_j = \lambda_j \gamma_j = 1$ (see 2.2 and formula (3.4)).

In the experiments we compared the error reduction from various stages on with the error reduction in the comparison process (as defined in Sects. 3.1 and 5) starting with one or more error components 0.

The experiments were done on a CDC Cyber 855 computer with 48 bits in the mantissa. In order to get some confidence that rounding errors did not spoil the results, we monitored the Ritz values. During our experiments we found no multiplet Ritz values, which is an indication that there was no serious loss of orthogonality.

6.2. *Convergence Behaviour of* CG *in Experiments I and II*

In order to illustrate the convergence behaviour of CG in experiments I and II we have plotted the quotients ("local convergence factors") $\dfrac{\|x - x_i\|_A}{\|x - x_{i-1}\|_A}$ in Fig. 1. We note that, on the whole, these quotients decrease, but that "locally" the behaviour may be rather bizarre, in particular for Experiment II.

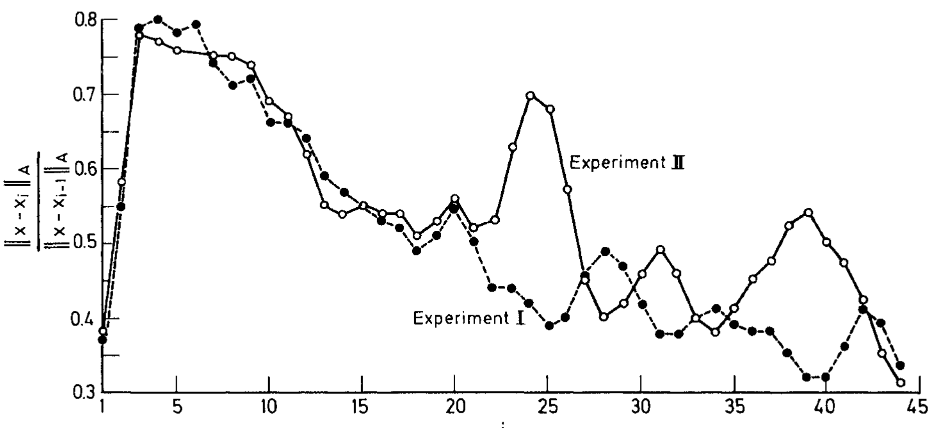

Fig. 1

Regarding the Ritz values we found that even after 44 iteration steps the largest Ritz value had not moved to a position between the largest and one but largest eigenvalue of A. Therefore we will restrict ourselves to considering only effects at the lower end of the spectrum (see 6.8, however).

6.3. *Experiment I*

(a) In Table 1 we have listed a few Ritz values before and after they entered their final intervals, and the corresponding F-values (an asterisk indicating that the Ritz value is not yet in its final interval). We note that the F-values decrease to 1 rather rapidly as i increases.

(b) Tables 2 and 3 give the following information:

- F gives the value of $F_{i,l,0}$ for the specified values of l (asterisks to be understood as in (a)).
- $\dfrac{\theta_1}{\lambda_1}$ gives the value of $\theta_1^{(i)}/\lambda_1$; likewise $\prod_1^4 \dfrac{\theta_k}{\lambda_k}$;
- the columns marked *orig* list the error reduction $\|x-x_{i+j}\|_A/\|x-x_i\|_A$ of the original process;
- the columns marked *comp* list the error reduction $\|x-\bar{x}_j\|_A/\|x-\bar{x}_0\|_A$ of the comparison process starting with initial error vector $x-x_i$ after its first l coordinates have been replaced by 0;
- the columns marked *cond* list the quantity $2\left(\dfrac{\sqrt{c_{l+1}}-1}{\sqrt{c_{l+1}}+1}\right)^j$ where $c_k=\lambda_m/\lambda_k$.

Table 1. Experiment I

i	$\theta_1^{(i)}$	$F_{i,1,0}$	i	$\theta_2^{(i)}$	$F_{i,2,0}$	i	$\theta_3^{(i)}$	$F_{i,3,0}$	i	$\theta_4^{(i)}$	$F_{i,4,0}$
5	0.11026	*17.8	10	0.13094	*22.0	14	0.16244	*5.83	18	0.19168	*943
6	0.07824	30.7	11	0.11403	5.33	15	0.15160	10.5	19	0.18365	6.42
7	0.05802	3.40	12	0.10245	2.40	16	0.14373	2.90	20	0.17762	3.34
8	0.04763	1.93	13	0.09357	1.56	17	0.13813	1.83	21	0.17028	2.00
9	0.04213	1.48	14	0.08827	1.27	18	0.13419	1.43	22	0.16281	1.37
10	0.03844	1.23	15	0.08520	1.12	19	0.13189	1.26	23	0.15857	1.14
11	0.03648	1.12	16	0.08346	1.06	20	0.13046	1.17	24	0.15652	1.05
$\lambda_1=0.03420$			$\lambda_2=0.08179$			$\lambda_3=0.12707$			$\lambda_4=0.15508$		

Table 2. Experiment I

	$i=5$, $l=1$; $F=$*17.8; $\frac{\theta_1}{\lambda_1}=3.22$		$i=6$, $l=1$; $F=30.7$; $\frac{\theta_1}{\lambda_1}=2.29$		$i=9$, $l=1$; $F=1.48$; $\frac{\theta_1}{\lambda_1}=1.23$		$l=1$
j	orig	comp	orig	comp	orig	comp	cond
2	0.58	0.33	0.53	0.35	0.43	0.51	0.69
4	0.30	0.14	0.25	0.091	0.16	0.20	0.24
6	0.13	0.053	0.11	0.029	0.052	0.053	0.081
8	0.048	0.016	0.035	0.0_289	0.014	0.019	0.028
10	0.015	0.0_244	0.010	0.0_233	0.0_236	0.0_244	0.0_296
12	0.0_243	0.0_213	0.0_227	0.0_213	0.0_399	0.0_211	0.0_233
14	0.0_211	0.0_333	0.0_375	0.0_343	0.0_319	0.0_320	0.0_211

Table 3. Experiment I

	$i=18,\ l=4$		$i=20,\ l=4$		$l=4$
	$F=*943$		$F=3.34$		
	$\prod_1^4 \frac{\theta_k}{\lambda_k}=1.32$		$\prod_1^4 \frac{\theta_k}{\lambda_k}=1.18$		
j	orig	comp	orig	comp	cond
2	0.28	0.20	0.22	0.26	0.37
4	0.062	0.035	0.041	0.053	0.069
6	0.011	0.0_269	0.0_264	0.0_285	0.013
8	0.0_218	0.0_214	0.0_214	0.0_215	0.0_224
10	0.0_340	0.0_324	0.0_328	0.0_326	0.0_344
12	0.0_478	0.0_434	0.0_439	0.0_459	0.0_481
14	0.0_411	0.0_552	0.0_564	0.0_599	0.0_415

(c) Looking at Table 2 we see that the numbers under *orig* are less than F times the corresponding ones under *comp*, as they should (cf. (3.2)). We also see that the inequality is rather sharp for the moderate value of F which occurs when $i=9$, but very crude for the large values of F when $i=5$ and 6; in the latter case $\frac{\theta_1}{\lambda_1}$ as multiplier instead of F does indeed seem more realistic (cf. 4.2(b)).

(d) For the quantity $\bar{F}_i$ occurring in (3.13) we find for $i=5$ that $\bar{F}_i=F_i$, and hence in (3.11) $\bar{F}_{ij}>F_i$, so now the "variants" don't help (the reason is that $\theta_1^{(i)}$ is too close to λ_3).

For $i=6$ we find $\bar{F}_i=4.35$, which means a considerable improvement. Hence Property 3.10 will now do pretty well if you iterate long enough. In this respect (see also 4.2(c)) we mention that $\bar{\theta}_2^{(j)}$ (for $i=6$) has the values 0.104 and 0.0882 for $j=12$ and 14, respectively, the corresponding factor $\frac{\lambda_k-\lambda_2}{\lambda_k-\bar{\theta}_2^{(j)}}$ in (3.11) then being 1.96 and 1.16, respectively. Hence, iterating long enough means in this case iterating until $j\approx 14$.

Property 3.12 says that you may not only loose the factor $\bar{F}$ but also p steps, and $p=3$ in experiment I. Since the local convergence factor of the comparison process is about 0.55 this means loss of another factor 6, making this result crude.

(e) Table 3 offers much the same picture as Table 2, and the same was the case for comparisons using $l=2$ and $l=3$ (which we made, but do not report here).

6.4. Experiment II

(a) For the explanation of Table 4, see 6.3(a). We note again that the F-values decrease rather rapidly to 1 as i increases, except for $F_{i,2,0}$, which moves rather slowly for $i=19,\ldots,25$ due to the stagnation of $\theta_2^{(i)}$. Note also that $\theta_3^{(i)}$ and $\theta_4^{(i)}$

Table 4. Experiment II (For the interpretation of the table cf. 6.3(a)

i	$\theta_1^{(i)}$	$F_{i,1,0}$	i	$\theta_2^{(i)}$	$F_{i,2,0}$	i	$\theta_3^{(i)}$	$F_{i,3,0}$	i	$\theta_4^{(i)}$	$F_{i,4,0}$
6	0.08226	*244	16	0.08325	*3.63	22	0.12711	*1960	22	0.15743	*67.3
7	0.06361	4.87	17	0.08279	8.22	23	0.12580	53.7	23	0.15387	106
8	0.05254	2.51	19	0.08243	2.30	24	0.12445	25.7	24	0.15000	23.3
9	0.04575	1.77	21	0.08235	1.98	25	0.12043	9.81	25	0.13974	6.13
10	0.04067	1.38	23	0.08232	1.88	26	0.10198	2.20	26	0.12934	1.88
11	0.03768	1.19	25	0.08230	1.82	27	0.08765	1.21	27	0.12758	1.17
12	0.03584	1.08	28	0.08210	1.38	28	0.08378	1.04	28	0.12721	1.03
$\lambda_1=0.03420$			$\lambda_2=0.08179$			$\lambda_1=0.08293$			$\lambda_4=0.12707$		

Table 5. Experiment II (For the interpretation of the table (cf. 6.3 (b))

	$i=12,\ l=1$ $F=1.08$ $\frac{\theta_1}{\lambda_1}=1.05$		$i=23,\ l=4$ $F=106$ $\prod_1^4 \frac{\theta_k}{\lambda_k}=1.85$		$i=28,\ l=4$ $F=1.03$ $\prod_1^4 \frac{\theta_k}{\lambda_k}=1.01$	
j	orig	comp	orig	comp	orig	comp
2	0.30	0.45	0.47	0.21	0.19	0.24
4	0.087	0.17	0.12	0.042	0.044	0.054
6	0.024	0.065	0.020	0.011	0.0_267	0.010
8	0.0_272	0.023	0.0_246	0.0_225	0.0_212	0.0_223
10	0.0_219	0.0_250	0.0_385	0.0_361	0.0_331	0.0_364
12	0.0_387	0.0_218	0.0_313	0.0_313	0.0_482	0.0_314
14	0.0_334	0.0_340	0.0_428	0.0_425	0.0_416	0.0_425

enter their final intervals simultaneously at $i=23$. The corresponding $F_{i,3,0}$ and $F_{i,4,0}$ values then decrease about equally fast. Indeed, these effects may be expected for the situation of a pair of close eigenvalues, as is shown in [13]. In 6.6 and 6.7 we will further comment on the observed convergence behaviour in view of the fact that λ_2 and λ_3 are close.

(b) For the explanation of Table 5, see 6.3(b). For $i=12$, $l=1$ and $i=28$, $l=4$ we have F-values close to 1 and we note again that the bounds are surprisingly sharp. We also note that for $i=23$, $l=4$ the large F-value is again unrealistic, and that $\prod_1^4 \frac{\theta_k}{\lambda_k}$ is much more realistic.

6.5. *Experiment III*

Experiment III was so designed that $\theta_1^{(6)}=\lambda_2$, hence $F_6=\infty$. In Table 6 we display what happens. In Property 3.12 we now have $\bar{F}_6=4.91$ and $p=3$. We note that the factor $\bar{F}_6$ alone is not able to bridge the gap between the two columns, but the 3 steps are, and we note that actually the error reduction of the original process lags 3 steps behind the error reduction of the comparison process to a remarkable degree of accuracy.

Table 6. Experiment III (For the interpretation of the table cf. 6.3 (a), (b))

	$i=6,\ l=1$ $\bar{F}=4.91$ $\frac{\theta_1}{\lambda_1}=2.39$	
j	orig	comp
3	0.38	0.17
6	0.11	0.027
9	0.022	$0.0_2 32$
12	$0.0_2 29$	$0.0_3 43$
15	$0.0_3 41$	$0.0_4 36$
18	$0.0_4 33$	$0.0_5 23$

6.6. Experimental Observations About the Relation Between CG *and Ritz Values*

Until now we have just compared the convergence behaviour of the original CG process with that of a comparison process. This does not explain, of course, the bizarre behaviour displayed in Fig. 1, nor can we explain it. We do want to report, however, some striking correspondences between the local convergence factors $\|x-x_i\|_A/\|x-x_{i-1}\|_A$ and the local behaviour of the Ritz values, in the hope that this may open ways for learning to understand this behaviour.

(a) Up to iteration 21 the Ritz values of processes I and II are quite similar. In particular, the first 3 Ritz values of process II seem to be heading for λ_1, $\lambda_2 \approx \lambda_3$ and λ_4 (as is the case for process I, of course). Process II has, so to speak, not yet discovered that $\lambda_2 \neq \lambda_3$ and $\lambda_5 \neq \lambda_6$ (a matter which can be explained using techniques as in [13]). Thus the corresponding polynomials R_i (cf. 2.4) are about the same, which makes it understandable that the two CG processes behave about the same, as we see in Fig. 1. Note, however, that as long as process II does not recognize that $\lambda_2 \neq \lambda_3$, it should consider $\lambda_2 \approx \lambda_3$ as one eigenvalue with weight $\sqrt{2}$. This does not seem to disturb the picture, however. This is understandable if the corresponding terms do not contribute greatly to $\|x-x_i\|_A$, and that is, indeed, what we have noticed during our computations.

(b) As i increases to 22, the terms in $\|x-x_i\|_A$ corresponding to λ_2 and λ_3 are rapidly becoming more and more important until, at $i=22$, these terms contribute 85% of the value of $\|x-x_i\|_A$. At this time we observe that θ_3 changes its course, and starts heading for λ_3 (at $i=23$ it has moved past λ_4). Moreover, at the same time we note that the local convergence factor of the CG process increases considerably (see Fig. 1), i.e. the process now converges as if a smaller number of extremal eigenvalues is reasonably approximated.

(c) The bulge for process II around $i=40$ coincides with process II finding out that $\lambda_5 \neq \lambda_6$.

(d) Virtually each "irregularity" in Fig. 1, for process I as well as for process II, could be traced back to "irregular" behaviour of some Ritz value at the same stage. By this we mean that this Ritz value changes course and starts heading for a limit different from the one it seemed to be heading for first, a sign of which is that its rate of change increases instead of decreases.

6.7. *The Effect of Almost Double Eigenvalues*

(a) As we see in Fig. 1, the presence of almost double eigenvalues may have rather dramatic effects on the rate of convergence of CG. Just how dramatic may this be?

(b) In Fig. 2 we have plotted the values of $\|x-x_i^{\mathrm{II}}\|_A/\|x-x_i^{\mathrm{I}}\|_A$ and we note that all these quotients are larger than 1 and may have rather large values.

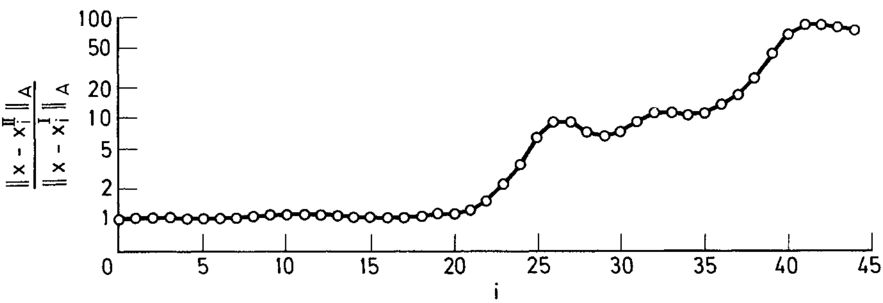

Fig. 2

We also note that the decelerating effect of the bulges in Fig. 1 is clearly visible.

(c) We get a somewhat different picture, however, if we look at the number of steps that process II lags behind process I, defined as $p_i = i - k_i$ where k_i is such that $\|x-x_{k_i}^{\mathrm{I}}\|_A < \|x-x_i^{\mathrm{II}}\|_A \leqq \|x-x_{k_i-1}^{\mathrm{I}}\|_A$. These numbers are plotted in Fig. 3 and we note that they are not larger than 5.

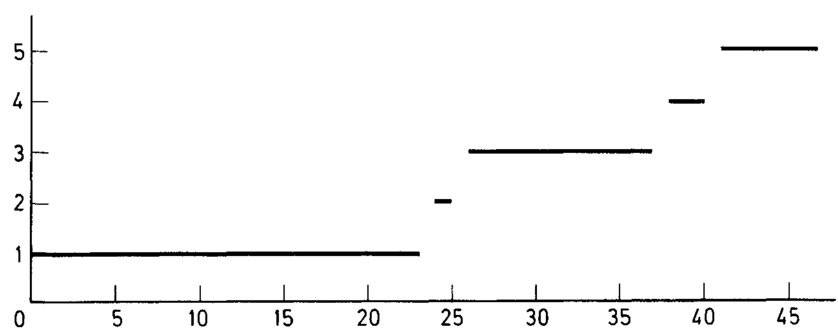

Fig. 3

(d) We can actually *prove* our findings in (b) and (c) and show the remarkable fact that the number of steps that process II is lagging behind does not depend on how close the almost double eigenvalues are.

Let p and p' denote the quantities defined by (3.15) if we replace λ_2 by λ_3 or λ_6 respectively. Then, as we shall prove, we have

$$\|x - x_i^{\mathrm{I}}\|_A \leqq \|x - x_i^{\mathrm{II}}\|_A \leqq \|x - x_{i-p-p'}^{\mathrm{I}}\|_A. \tag{6.1}$$

In our case we have $p = p' = 3$ and hence process II will never be more than 6 steps behind, and this number depends only on λ_m/λ_3 and λ_m/λ_6, and not on how close $\lambda_3(\lambda_6)$ is to $\lambda_2(\lambda_5)$.

In order to prove (6.1), we define $R_{\mathrm{I},i}$ and $R_{\mathrm{II},i}$ as the polynomials corresponding to process I and II according to 2.4. Let $\tilde{T}_p$ and $\tilde{T}_{p'}$ denote the polynomials as defined in the proof of Property 3.12 such that $\tilde{T}_p(\lambda_3) = \tilde{T}_{p'}(\lambda_6) = 0$. Then

$$\begin{aligned}\|x - x_i^{\mathrm{II}}\|_A^2 &\leqq \sum_{\mathrm{II}} \lambda_k \gamma_k^2 \left[\frac{\tilde{T}_p(\lambda_k)}{\tilde{T}_p(0)} \cdot \frac{\tilde{T}_{p'}(\lambda_k)}{\tilde{T}_{p'}(0)} R_{\mathrm{I},i-p-p'}(\lambda_k)\right]^2 \\ &\leqq \sum_{\mathrm{I}} \lambda_k \gamma_k^2 R_{\mathrm{I},i-p-p'}(\lambda_k)^2 = \|x - x_{i-p-p'}^{\mathrm{I}}\|_A^2\end{aligned} \tag{6.2}$$

where $\sum_{\mathrm{II}}$ and $\sum_{\mathrm{I}}$ denote summation over the active eigenvalues in the processes in question. This proves one of the inequalities. The other one follows from

$$\|x - x_i^{\mathrm{I}}\|_A^2 \leqq \sum_{\mathrm{I}} \lambda_k \gamma_k^2 R_{\mathrm{II},i}(\lambda_k)^2 \leqq \sum_{\mathrm{II}} \lambda_k \gamma_k^2 R_{\mathrm{II},i}(\lambda_k)^2 = \|x - x_i^{\mathrm{II}}\|_A^2. \tag{6.3}$$

6.8. *The Effects of Convergence at the Upper End of the Spectrum*

So far our experiments dealt only with the effects on the convergence of CG by the convergence of Ritz values at the lower end of the spectrum.

Our theory, of course, applies to the upper end as well. But in practice, due to the presence of rounding errors, the situation there is much more cumbersome. For, suppose that convergence of a Ritz value to λ_m would lead to the same sort of faster convergence as we have seen in, e.g., Experiment I, then, taking condition numbers as a measure, we should have $\lambda_{m-1} \approx \frac{0.034}{0.082} \lambda_m \approx 0.4\,\lambda_m$. This then leads to a gap ratio $\frac{\lambda_m - \lambda_{m-1}}{\lambda_{m-1} - \lambda_1} \approx 1.5$, which causes very rapid convergence of the Ritz value (cf. [10], p. 242–244), and consequently very rapid loss of orthogonality ([10], p. 270). This spoils any convergence pattern that we might observe in exact computation. Therefore, convergence of Ritz values at the upper end of the spectrum will rarely lead to impressive increases of the rate of convergence of the CG process. This was confirmed by experiments.

Acknowledgements. We thank Prof. Olof Widlund and our student Ton Peerdeman for pointing out to us references [3] and [14], respectively, and our colleague Gerard Sleypen for stimulating discussions.

References

1. Concus, P., Golub, G.H., O'Leary, D.P.: A generalized Conjugate Gradient method for the numerical solution of elliptic partial differential equations. In: Sparse Matrix Computations (J.R. Bunch and D.J. Rose, eds.), pp. 309–332. New York: Academic Press 1976
2. Golub, G.H., Van Loan, C.F.: Matrix Computations. Oxford: North Oxford Academic 1983
3. Hayes, R.M.: Iterative methods of solving linear problems on Hilbert space. National Bureau of Standards, Appl. Math. Ser., **39**, 71–103 (1954)
4. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. **49**, 409–436 (1952)
5. Kaniel, S.: Estimates for some computational techniques in linear algebra. Math. Comput. **20**, 369–378 (1966)
6. Karush, W.: An iterative method for finding characteristic vectors of a symmetric matrix. Pacific J. Math. **1**, 233–248 (1950)
7. Karush, W.: Convergence of a method of solving linear problems. Proc. Amer. Math. Soc. **3**, 839–851 (1952)
8. Meinardus, G.: Über eine Verallgemeinerung einer Ungleichung von L.V. Kantorovitch. Numer. Math. **5**, 14–23 (1963)
9. Meyerink, J.A., van der Vorst, H.A.: An iterative solution method for linear systems of which the coefficient matrix is a symmetric *M*-matrix. Math. Comput. **31**, 148–162 (1977)
10. Parlett, B.N.: The Symmetric Eigenvalue Problem. Englewood Cliffs: Prentice Hall 1980
11. Parlett, B.N., Simon, H., Stringer, L.M.: On estimating the largest eigenvalue with the Lanczos Algorithm. Math. Comput. **38**, 153–165 (1982)
12. Saad, Y.: On the rates of convergence of the Lanczos and Block Lanczos methods. SIAM J. Numer. Anal. **17**, 687–706 (1980)
13. van der Sluis, A., van der Vorst, H.A.: The convergence behaviour of Ritz values in the presence of close eigenvalues. (To be published)
14. Winther, R.: Some superlinear convergence results for the conjugate gradient methods. SIAM J. Numer. Anal. **17**, 14–17 (1980)

Received November 2, 1984 / August 16, 1985

Note Added in Proof

In relation to 6.6 we note that we have also run experiment I with $(\tau_2+\tau_3)/2$ and $(\tau_5+\tau_6)/2$ instead of τ_2 and τ_5 and weights $\sqrt{2}$ instead of 1. This produced a picture which was indistinguishable from the one given in Fig. 1.