Electronic Transactions on Numerical Analysis.
Volume 2, pp. 57-75, March 1994.

ISSN 1068-9613.

MINIMIZATION PROPERTIES AND SHORT RECURRENCES FOR KRYLOV SUBSPACE METHODS*

RÜDIGER WEISS †

Dedicated to Wilhelm Niethammer on the occasion of his 60th birthday.

Abstract. It is well known that generalized conjugate gradient (cg) methods, fulfilling a minimization property in the whole spanned Krylov space, cannot be formulated with short recurrences for nonsymmetric system matrices. Here, Krylov subspace methods are proposed that do fulfill a minimization property and can be implemented as short recurrence method at the same time. These properties are achieved by a generalization of the cg concept. The convergence and the geometric behavior of these methods are investigated.

Practical applications show that first realizations of these methods are already competitive with commonly used techniques such as smoothed biconjugate gradients or QMR. Better results seem to be possible by further improvements of the techniques. However, the purpose of this paper is not to propagate a special method, but to stimulate research and development of new iterative linear solvers.

Key words. conjugate gradients, convergence, linear systems, Krylov methods.

AMS subject classifications. 65F10, 65F50, 40A05.

1. Introduction.

We are interested in the solution of the linear system

$$Ax = b. \tag{1.1}$$

The matrix A is a real square matrix of dimension n, i. e. $A \in \mathbb{R}^{n\times n}$, and $x, b \in \mathbb{R}^n$. In general the matrix A is nonsymmetric and not positive definite. Let us assume A to be nonsingular.

We use the following notations: Let Z be a symmetric, positive definite matrix, then the norm $\|y\|_Z$ of any vector $y \in \mathbb{R}^n$ is defined by $\|y\|_Z = \sqrt{y^T Z y}$. If Z is nonsymmetric and not positive definite, then $\|y\|_Z^2$ is a mnemonic abbreviation for $y^T Z y$. $\|y\|_I$ is the Euclidean norm $\|y\|$. Let $K_k(B, y) = span(y, By, \ldots, B^k y)$ be the Krylov space spanned by the matrix $B \in \mathbb{R}^{n\times n}$ and the vector $y \in \mathbb{R}^n$. The symmetric part of the matrix $B \in \mathbb{R}^{n\times n}$ is $\frac{1}{2}(B + B^T)$ and the skew-symmetric part is $\frac{1}{2}(B - B^T)$.

For large and sparse linear systems arising from the discretization and linearization of systems of partial differential equations, iterative solution techniques have become a powerful solution tool, because they are limited in their storage requirements and generally need less computing time than direct solvers if a limited accuracy is required.

Usually the linear system (1.1) is preconditioned in order to accelerate the convergence. We apply preconditioning from the right-hand side and consider the linear system

$$APy = b \tag{1.2}$$

---

* Received January 19, 1994. Accepted for publication June 6, 1994. Communicated by L. Reichel.

† Numerikforschung für Supercomputer, Rechenzentrum der Universität Karlsruhe, Postfach 6980, D-76128 Karlsruhe, Germany, e-mail: weiss@rz.uni-karlsruhe.de.

instead of (1.1), where $P \in I\!R^{n\times n}$ is a nonsingular preconditioning matrix. The solution of (1.1) is obtained from the solution of (1.2) by

$$x = Py. \tag{1.3}$$

Any iterative method applied to (1.2) with approximations y_k can be reformulated so that an iterative method for the original system (1.1) is induced by

$$x_k = Py_k. \tag{1.4}$$

We will always reformulate preconditioned iterative methods accordingly in the following definitions.

Among iterative methods generalized conjugate gradient (cg) methods converge quickly in many cases. These methods are vectorizable, parallelizable, parameter-free and therefore widely used. The technique is as follows:

Choose a right-hand preconditioning matrix P and an initial guess x_0. Calculate approximations x_k and residuals $r_k = Ax_k - b$ for $k \geq 1$ so that

$$x_k \in x_0 + K_{k-1}(PA, Pr_0), \tag{1.5}$$

with

$$r_k^T Z r_{k-i} = 0 \tag{1.6}$$

for $i = 1, \ldots, \sigma_k$, where Z is an auxiliary, nonsingular matrix. The method is called exact if $\sigma_k = k$, restarted if $\sigma_k = (k-1) \bmod \sigma_{res} + 1$ with σ_{res} fixed, truncated if $\sigma_k = \min(k, \sigma_{max})$ with σ_{max} fixed, combined if the truncated method is restarted.

Convergence properties are well-known for exact generalized cg methods [16]. If APZ^{-1} is positive real, i.e. the symmetric part of APZ^{-1} is positive definite, then

$$\|r_k\|_{ZP^{-1}A^{-1}} \leq \sqrt{1 + \frac{\rho^2(R)}{\mu_m^2}} \min_{\Pi_k} \|\Pi_k(AP)r_0\|_{ZP^{-1}A^{-1}} \tag{1.7}$$

is valid for exact generalized cg methods. Π_k is a polynomial of degree k with $\Pi_k(0) = 1$. $\rho(R)$ is the spectral radius of the skew-symmetric part R of $ZP^{-1}A^{-1}$. μ_m is the minimum eigenvalue of M, the symmetric part of $ZP^{-1}A^{-1}$. In particular if $Z = AP$, then

$$\|r_k\| = \min_{\Pi_k} \|\Pi_k(AP)r_0\|. \tag{1.8}$$

If we apply an exact method to systems with arbitrary matrices, then in general the required storage and the computational work increase with the iteration. The reason is that for the calculation of the new approximation all preceding approximations are required to fulfill equation (1.6). Thus we get long recurrences except for certain favorable cases. Therefore, exact methods are not feasible for large systems. In this section we will survey well-known techniques to obtain short recurrences.

Faber and Manteuffel [2] give conditions for short recurrences of exact ORTHODIR implementations. Joubert and Young [8] prove similar results for the simplification of ORTHOMIN and ORTHORES implementations. An exact ORTHORES method can be formulated as a three-term recurrence, if

$$P^T A^T Z = ZAP. \tag{1.9}$$

Condition (1.9) is valid if AP is symmetric and $Z = I$, but in most practical cases AP is nonsymmetric. Jea and Young [7] show that a matrix Z fulfilling (1.9) for fixed AP always exists, but the determination of such a matrix is usually impossible for systems arising from practical applications.

The choice $Z = I$ and $P = A^T$ satisfies (1.9) for an arbitrary matrix A (Craig's method [1]), but then the iteration matrix is AA^T resulting in slow convergence for systems where the eigenvalues of AA^T, the singular values of A, are scattered, see inequality (1.7).

For the biconjugate gradients (BCG) [3, 9], the double system $\hat{A}\hat{x} = \hat{b}$, i. e.

$$\begin{pmatrix} A & 0 \\ 0 & A^T \end{pmatrix} \begin{pmatrix} x \\ x^* \end{pmatrix} = \begin{pmatrix} b \\ b^* \end{pmatrix},$$

is considered. b^* is arbitrary. The residuals have the form $\hat{r} = \begin{pmatrix} r \\ r^* \end{pmatrix}$ and $Z = Z_B = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$. P is the unit matrix. As $\hat{A}^T Z_B = Z_B \hat{A}$, condition (1.9) is valid, thus we obtain a short recurrence. Property (1.7) becomes

$$\left(r_k^*\right)^T A^{-1} r_k = \min_{\Pi_k} \left(r_k^*\right)^T A^{-1} \Pi_k(A) r_0, \tag{1.10}$$

where we cannot easily estimate the Euclidean norm of the residual r_k.

Sonnefeld's CGS [14] is a method using a short recurrence that minimizes the same expression as the biconjugate gradients, equation (1.10), but uses as residual polynomial

$$r_k = \Pi_k^2(A) r_0.$$

As for the biconjugate gradients the Euclidean norm of the residual is hard to determine from equation (1.10).

Freund's and Nachtigal's QMR method [5] and the biconjugate gradients smoothed by Schönauer's minimal residual algorithm [11, 12] fulfill the minimization property, see also [17],

$$\begin{aligned} \|r_k\| &\leq \sqrt{k+1} \min_{z_1,\ldots,z_k} \left\| \, \|r_0\| e_1 - H_k z \, \right\|, \\ &\leq \sqrt{k+1} \, \|V_n^{-1}\| \, \min_{\Pi_k} \|\Pi_k(A) r_0\| \end{aligned} \tag{1.11}$$

where e_1 is the first unit vector, $z = (z_1, \ldots, z_k)^T \in I\!R^k$, $V_k = (\frac{r_0}{\|r_0\|}, \ldots, \frac{r_{k-1}}{\|r_{k-1}\|})$ and

$$AV_k = V_{k+1} H_k. \tag{1.12}$$

$H_k \in I\!R^{(k+1)\times k}$ is the tridiagonal matrix resulting from the nonsymmetric Lanczos process, or a block tridiagonal matrix resulting from a look-ahead Lanczos process to avoid breakdowns. Similar results are valid for Freund's TFQMR [4]. The minimization property is quite more complex than (1.7) and the right-hand side of (1.11) is growing - however small it may be - with the iteration.

In order to obtain short recurrences for nonsymmetric matrices A there are various other possibilities, among them

- restarted or truncated versions,
- CGSTAB approaches [6, 13, 15] introduced by van der Vorst.

However, all techniques mentioned above produce short recurrences, but do not fulfill the convergence properties (1.7) or (1.8). We will show in the following that we can enforce automatic termination of the sequence by allowing the matrix Z to be dependent on the iteration step and to maintain at the same time convergence property (1.7).

2. Conjugate Krylov Subspace Methods.

We start by further generalizing the generalized cg methods and by showing some fundamental convergence properties. The difference from cg methods in the following definition is that the matrix Z is substituted by step-depending matrices Z_k.

DEFINITION 1. *Let x_0 be any initial guess, $r_0 = Ax_0 - b$ the starting residual. The following recurrence is called a conjugate Krylov subspace (CKS) method. Choose a right-hand preconditioning matrix P and calculate for $k \geq 1$ the residuals r_k and approximations x_k so that*

$$x_k \in x_0 + K_{k-1}(PA, Pr_0), \tag{2.1}$$

with

$$r_k^T Z_k r_{k-i} = 0 \tag{2.2}$$

for $i = 1, \ldots, k$, where Z_k are auxiliary, nonsingular matrices.

If $Z_k = Z = const$, then definition 1 describes exact generalized cg methods as special case.

It is quite easy to verify that the approximations x_k, the residuals r_k and the errors $e_k = x_k - x$ of CKS methods can be represented as follows:

$$\begin{aligned} x_k &= \sum_{i=1}^{k} \nu_{i,k} P(AP)^{i-1} r_0 + x_0 \\ &= \sum_{i=1}^{k} \mu_{i,k} P r_{k-i} + x_0 \\ &= \beta_{0,k} P r_{k-1} + \sum_{i=1}^{k} \beta_{i,k} x_{k-i}, \end{aligned} \tag{2.3}$$

$$r_k = \sum_{i=1}^{k} \nu_{i,k} (AP)^i r_0 + r_0 \tag{2.4}$$

$$
\begin{aligned}
&= \sum_{i=1}^{k} \mu_{i,k} APr_{k-i} + r_0 \\
&= \beta_{0,k} APr_{k-1} + \sum_{i=1}^{k} \beta_{i,k} r_{k-i},
\end{aligned}
$$

$$
\begin{aligned}
e_k &= \sum_{i=1}^{k} \nu_{i,k} (PA)^i e_0 + e_0 \\
&= \sum_{i=1}^{k} \mu_{i,k} PAe_{k-i} + e_0 \\
&= \beta_{0,k} PAe_{k-1} + \sum_{i=1}^{k} \beta_{i,k} e_{k-i},
\end{aligned} \tag{2.5}
$$

with

$$\sum_{i=1}^{k} \beta_{i,k} = 1. \tag{2.6}$$

We can prove the following theorems in analogy to generalized cg methods. The next lemma is the basis for all convergence analysis and it is equivalent to the same statements for generalized cg methods [16].

LEMMA 2. *For CKS methods*

$$r_k^T Z_k P^{-1} A^{-1} r_k = r_k^T Z_k P^{-1} A^{-1} \Pi_k(AP) r_0 \tag{2.7}$$

is satisfied for all matrix polynomials $\Pi_k(AP) = \sum_{i=1}^{k} \theta_i (AP)^i + I$ *(i. e.* $\theta_1, \ldots, \theta_k$ *are arbitrary). In particular*

$$r_k^T Z_k (AP)^i r_0 = 0 \tag{2.8}$$

for $i = 0, \ldots, k-1$.

Proof. By analogy with lemma 3.5 in [16]. □

The next theorem shows the geometric behavior of the approximations of CKS methods and generalizes the result for exact generalized cg methods.

THEOREM 3. *The residuals* r_k *and the errors* e_k *of CKS methods satisfy*

$$\left\| r_k - \frac{\tilde{r}_j}{2} \right\|^2_{Z_k P^{-1} A^{-1}} = \frac{\|\tilde{r}_j\|^2_{Z_k P^{-1} A^{-1}}}{4}, \tag{2.9}$$

$$\left\| e_k - \frac{\tilde{e}_j}{2} \right\|^2_{A^T Z_k P^{-1}} = \frac{\|\tilde{e}_j\|^2_{A^T Z_k P^{-1}}}{4} \tag{2.10}$$

for $j = 0, \ldots, k$ *with*

$$\tilde{r}_j = 2 \left(Z_k P^{-1} A^{-1} + (Z_k P^{-1} A^{-1})^T \right)^{-1} Z_k P^{-1} A^{-1} r_j, \tag{2.11}$$

$$\tilde{e}_j = 2 \left(Z_k P^{-1} + (Z_k P^{-1} A^{-1})^T A \right)^{-1} Z_k P^{-1} e_j. \tag{2.12}$$

In particular if $A^T Z_k P^{-1}$ *is symmetric, then*

$$\tilde{r}_j = r_j, \tag{2.13}$$

$$\tilde{e}_j = e_j. \tag{2.14}$$

Proof. By analogy with theorem 3 in [18]. □

Theorem 3 describes geometric figures, in general hyperellipsoids, see [18] for a classification and an explanation. The next theorem analyzes the convergence behavior of the approximations of CKS methods. We obtain the same results as for exact generalized cg methods.

Theorem 4. *If* APZ_k^{-1} *is positive real, i.e. the symmetric part of* APZ_k^{-1} *is positive definite, then*

$$\|r_k\|_{Z_k P^{-1} A^{-1}} \leq \sqrt{1 + \frac{\rho^2(R)}{\mu_m^2}} \min_{\Pi_k} \|\Pi_k(AP) r_0\|_{Z_k P^{-1} A^{-1}} \tag{2.15}$$

$$\|e_k\|_{A^T Z_k P^{-1}} \leq \sqrt{1 + \frac{\rho^2(R)}{\mu_m^2}} \min_{\Pi_k} \|\Pi_k(PA) e_0\|_{A^T Z_k P^{-1}} \tag{2.16}$$

holds for CKS methods. Π_k *is a polynomial of degree* k *with* $\Pi_k(0) = 1$. $\rho(R)$ *is the spectral radius of the skew-symmetric part* R *of* $Z_k P^{-1} A^{-1}$. μ_m *is the minimum eigenvalue of* M*, the symmetric part of* $Z_k P^{-1} A^{-1}$.

Proof. By analogy with theorem 3.9 in [16]. □

We have seen by the theorems 3 and 4 that CKS methods have a similar convergence behavior as generalized cg methods, see [16, 18]. In the next section we will show how to construct short recurrences.

3. Short Recurrences.

The next lemma is the key to construct the matrices Z_k of a CKS method so that we will obtain short recurrences.

Lemma 5. *If* $B \in \mathbb{R}^{k \times k}$, $y, b \in \mathbb{R}^k$, *and*

$$B = \begin{pmatrix} B_{1,1} & 0 \\ B_{2,1} & B_{2,2} \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, b = \begin{pmatrix} 0 \\ b_2 \end{pmatrix}, \tag{3.1}$$

with $B_{1,1} \in \mathbb{R}^{k-j \times k-j}$ *nonsingular,* $y_1 \in \mathbb{R}^{k-j}$, $B_{2,1} \in \mathbb{R}^{j \times k-j}$, $B_{2,2} \in \mathbb{R}^{j \times j}$ *nonsingular,* $y_2, b_2 \in \mathbb{R}^j$, *then the solution of the system*

$$By = b \tag{3.2}$$

is

$$y_1 = 0, \tag{3.3}$$

$$y_2 = B_{2,2}^{-1} b_2. \tag{3.4}$$

Proof. The proof is trivial. □

Let us apply lemma 5 to CKS methods. If $\beta_{0,k} \neq 0$, then equation (2.2) is equivalent to

$$(APr_{k-1} + \sum_{i=1}^{k} \alpha_{i,k} r_{k-i})^T Z_k r_{k-j} = 0 \tag{3.5}$$

for $j = 1, \dots, k$, following from (2.4), where

$$\alpha_{i,k} = \frac{\beta_{i,k}}{\beta_{0,k}} \tag{3.6}$$

for $i = 1, \dots, k$. The $\alpha_{i,k}$ can be determined by the solution of the linear system

$$\sum_{i=1}^{k} \alpha_{i,k} r_{k-i}^T Z_k r_{k-j} = -r_{k-1}^T P^T A^T Z_k r_{k-j} \tag{3.7}$$

for $j = 1, \dots, k$. Let

$$R_k = (r_0, \dots, r_{k-1}), \tag{3.8}$$

$$\alpha_k = (\alpha_{k,k}, \dots, \alpha_{1,k})^T, \tag{3.9}$$

then (3.7) can be written in the short form

$$R_k^T Z_k^T R_k \alpha_k = -R_k^T Z_k^T APr_{k-1}. \tag{3.10}$$

From lemma 5 follows directly that $\alpha_{i,k} = 0$ for $i = 3, \dots, k$, if

$$r_{k-1}^T Z_k r_{k-i} = 0, \tag{3.11}$$

$$r_{k-2}^T Z_k r_{k-i} = 0, \tag{3.12}$$

$$r_{k-1}^T P^T A^T Z_k r_{k-i} = 0 \tag{3.13}$$

for $i = 3, \dots, k$.

THEOREM 6. *CKS methods can be formulated as three-term recurrences, if*

$$r_{k-1}^T Z_k = r_{k-1}^T Z_{k-1}, \tag{3.14}$$

$$r_{k-2}^T Z_k = r_{k-2}^T Z_{k-1}, \tag{3.15}$$

$$r_{k-1}^T P^T A^T Z_k = r_{k-1}^T Z_{k-1} AP, \tag{3.16}$$

for $k \geq 3$ in the following way:

$$r_k = \phi_k (APr_{k-1} + \alpha_{1,k} r_{k-1} + \alpha_{2,k} r_{k-2}), \tag{3.17}$$

$$x_k = \phi_k (Pr_{k-1} + \alpha_{1,k} x_{k-1} + \alpha_{2,k} x_{k-2}), \text{ where} \tag{3.18}$$

$$\alpha_{2,k} = -\frac{r_{k-2}^T Z_k^T APr_{k-1}}{r_{k-2}^T Z_k r_{k-2}}, \tag{3.19}$$

$$\alpha_{1,k} = -\frac{1}{r_{k-1}^T Z_k r_{k-1}} \left(r_{k-1}^T Z_k^T APr_{k-1} + \alpha_{2,k} r_{k-2}^T Z_k r_{k-1} \right), \tag{3.20}$$

$$\phi_k = \frac{1}{\alpha_{1,k} + \alpha_{2,k}}. \tag{3.21}$$

Proof. From (3.14) follows

$$r_{k-1}^T Z_k r_{k-i} = r_{k-1}^T Z_{k-1} r_{k-i} = 0 \tag{3.22}$$

for $i = 2, \ldots, k$ by (2.2) and from (3.15) follows

$$r_{k-2}^T Z_k r_{k-i} = r_{k-2}^T Z_{k-1} r_{k-i} = r_{k-2}^T Z_{k-2} r_{k-i} = 0 \tag{3.23}$$

for $i = 3, \ldots, k$ by (3.14) and (2.2). Thus (3.11) and (3.12) are fulfilled. From (3.16) follows

$$\begin{aligned} r_{k-1}^T P^T A^T Z_k r_{k-i} &= r_{k-1}^T Z_{k-1} A P r_{k-i} \\ &= r_{k-1}^T Z_{k-1} A P \left(\sum_{j=1}^{k-i} \nu_{j,k-i} (AP)^j r_0 + r_0 \right) \\ & \quad \text{by (2.4)} \\ &= r_{k-1}^T Z_{k-1} \left(\sum_{j=1}^{k-i} \nu_{j,k-i} (AP)^{j+1} r_0 + APr_0 \right) = 0 \end{aligned}$$

because of (2.8). Thus (3.13) is fulfilled and the sequence terminates automatically. From lemma 5 and (3.7) follows that $\alpha_{1,k}$ and $\alpha_{2,k}$ can be calculated by

$$\begin{aligned} \alpha_{1,k} r_{k-1}^T Z_k r_{k-1} + \alpha_{2,k} r_{k-2}^T Z_k r_{k-1} &= -r_{k-1}^T P^T A^T Z_k r_{k-1}, \\ \alpha_{1,k} r_{k-1}^T Z_k r_{k-2} + \alpha_{2,k} r_{k-2}^T Z_k r_{k-2} &= -r_{k-1}^T P^T A^T Z_k r_{k-2}. \end{aligned}$$

Following from (3.22) the second equation is equivalent to

$$\alpha_{2,k} r_{k-2}^T Z_k r_{k-2} = -r_{k-1}^T P^T A^T Z_k r_{k-2}.$$

(3.17) - (3.20) follow from simple calculations and from (3.6) and (2.4), (2.5), respectively. □

Theorem 6 describes an ORTHORES-like implementation. The method breaks down if $\alpha_{1,k} + \alpha_{2,k} = 0$. We will assume in the following that the method does not break down.

If $Z_k = Z = \mathit{const}$, then Condition (3.16) follows from condition (1.9) and the conditions (3.14) and (3.15) are always fulfilled. Thus we have got a generalization where the global condition (1.9) is substituted by local conditions. It is easy to verify that (3.14) - (3.16) are equivalent to

$$S_k^T Z_k = Y_k^T, \tag{3.24}$$

where

$$S_k = (r_{k-1}, r_{k-2}, APr_{k-1}) \in I\!R^{n \times 3}, \tag{3.25}$$

$$Y_k = (Z_{k-1}^T r_{k-1}, Z_{k-1}^T r_{k-2}, P^T A^T Z_{k-1}^T r_{k-1}) \in I\!R^{n \times 3}. \tag{3.26}$$

Thus

$$S_k^T Z_k = \begin{pmatrix} 0 \\ 0 \\ r_{k-1}^T (Z_{k-1} AP - P^T A^T Z_{k-1}) \end{pmatrix} + S_k^T Z_{k-1}. \tag{3.27}$$

Equation (3.27) shows that the change of Z_k with respect to Z_{k-1} is caused by the magnitude of $Z_{k-1}AP - P^T A^T Z_{k-1}$. For $Z_{k-1} = I$ the change of Z_k is caused by the nonsymmetric part of AP.

4. Rank-Three Updates. In this section we propose a method how to construct the matrices Z_k that fulfill the assumptions of theorem 6. (3.14) - (3.16) are vector equations that can be fulfilled by choosing

$$Z_k = Z + a_k b_k^T + c_k d_k^T + e_k f_k^T \tag{4.1}$$

as rank-three update with $a_k, b_k, c_k, d_k, e_k, f_k \in I\!R^n$. It is obvious that the following equations have to be satisfied so that (3.14) - (3.16) are valid:

$$\begin{aligned}
(4.2)\quad & r_{k-1}^T a_k b_k^T + r_{k-1}^T c_k d_k^T + r_{k-1}^T e_k f_k^T \\
&\quad = r_{k-1}^T a_{k-1} b_{k-1}^T + r_{k-1}^T c_{k-1} d_{k-1}^T + r_{k-1}^T e_{k-1} f_{k-1}^T, \\
(4.3)\quad & r_{k-2}^T a_k b_k^T + r_{k-2}^T c_k d_k^T + r_{k-2}^T e_k f_k^T \\
&\quad = r_{k-2}^T a_{k-1} b_{k-1}^T + r_{k-2}^T c_{k-1} d_{k-1}^T + r_{k-2}^T e_{k-1} f_{k-1}^T, \\
(4.4)\quad & r_{k-1}^T P^T A^T Z + r_{k-1}^T P^T A^T a_k b_k^T + r_{k-1}^T P^T A^T c_k d_k^T + r_{k-1}^T P^T A^T e_k f_k^T \\
&\quad = r_{k-1}^T ZAP + r_{k-1}^T a_{k-1} b_{k-1}^T AP \\
&\qquad + r_{k-1}^T c_{k-1} d_{k-1}^T AP + r_{k-1}^T e_{k-1} f_{k-1}^T AP,
\end{aligned}$$

or in matrix form

$$\Psi_k \begin{pmatrix} b_k^T \\ d_k^T \\ f_k^T \end{pmatrix} = \theta_k \tag{4.5}$$

with

$$\begin{aligned}
(4.6)\quad \Psi_k &= \begin{pmatrix} r_{k-1}^T a_k & r_{k-1}^T c_k & r_{k-1}^T e_k \\ r_{k-2}^T a_k & r_{k-2}^T c_k & r_{k-2}^T e_k \\ r_{k-1}^T P^T A^T a_k & r_{k-1}^T P^T A^T c_k & r_{k-1}^T P^T A^T e_k \end{pmatrix} \\
&= \begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (a_k, c_k, e_k)
\end{aligned}$$

and

$$\theta_k = \begin{pmatrix} 0 \\ 0 \\ r_{k-1}^T (ZAP - P^T A^T Z) \end{pmatrix} \tag{4.7}$$

$$
\begin{aligned}
&+\begin{pmatrix} r_{k-1}^T a_{k-1} b_{k-1}^T + r_{k-1}^T c_{k-1} d_{k-1}^T + r_{k-1}^T e_{k-1} f_{k-1}^T \\ r_{k-2}^T a_{k-1} b_{k-1}^T + r_{k-2}^T c_{k-1} d_{k-1}^T + r_{k-2}^T e_{k-1} f_{k-1}^T \\ r_{k-1}^T a_{k-1} b_{k-1}^T AP + r_{k-1}^T c_{k-1} d_{k-1}^T AP + r_{k-1}^T e_{k-1} f_{k-1}^T AP \end{pmatrix} \\
= &\begin{pmatrix} 0 \\ 0 \\ r_{k-1}^T (Z_{k-1} AP - P^T A^T Z_{k-1}) \end{pmatrix} \\
&+\begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (Z_{k-1} - Z).
\end{aligned}
$$

Note that θ_k depends only on approximations of previous steps. If $ZAP = P^T A^T Z$, then (4.5) is satisfied by $b_k = d_k = f_k = 0$ for all k following from (4.7), thus coinciding with $Z_k = Z = \mathit{const}$ and (1.9). If $ZAP \neq P^T A^T Z$, then choose $a_0 = b_0 = c_0 = d_0 = e_0 = f_0 = 0$ and a_k, c_k, e_k so that Ψ_k can be inverted for $k \geq 1$ and calculate

$$
\begin{pmatrix} b_k^T \\ d_k^T \\ f_k^T \end{pmatrix} = \Psi_k^{-1} \theta_k. \tag{4.8}
$$

In each iteration step two matrix-vector multiplications with the matrices AP, $P^T A^T$, respectively, have to be performed:

$$
APr_{k-1}
$$

and

$$
P^T A^T (Z^T r_{k-1} + r_{k-1}^T a_{k-1} b_{k-1} + r_{k-1}^T c_{k-1} d_{k-1} + r_{k-1}^T e_{k-1} f_{k-1})
$$

for the determination of θ_k. The work counted in matrix-vector multiplications as essential operations corresponds therefore to the work of the biconjugate gradients.

Z_k is of the form

$$
Z_k = Z + (a_k, c_k, e_k) \left[\begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (a_k, c_k, e_k) \right]^{-1} \theta_k. \tag{4.9}
$$

The vectors a_k, c_k, e_k still have to be determined. A natural choice would be to minimize

$$
\|Z_k - Z\| = \left\| (a_k, c_k, e_k) \left[\begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (a_k, c_k, e_k) \right]^{-1} \theta_k \right\|, \tag{4.10}
$$

so that Z_k is close to Z. If the Frobenius norm is admissible, then the update

$$
Z_k = Z + S_k \left(S_k^T S_k \right)^{-1} (Y_k^T - S_k^T Z) \tag{4.11}
$$

minimizes (4.10) and fulfills (3.24). Numerical tests indicate that this choice of Z_k is not optimal. However for the Euclidean norm, the determination of Z_k from (4.10) is a problem because of the complex structure of Z_k. The determination of a_k, c_k, e_k is still an unsolved problem. At least we can formulate the following equivalence:

LEMMA 7. *Let*

$$Z_k = Z + (a_k, c_k, d_k) \begin{pmatrix} b_k^T \\ d_k^T \\ f_k^T \end{pmatrix}, \tag{4.12}$$

$$\tilde{Z}_k = Z + (\tilde{a}_k, \tilde{c}_k, \tilde{d}_k) \begin{pmatrix} \tilde{b}_k^T \\ \tilde{d}_k^T \\ \tilde{f}_k^T \end{pmatrix}. \tag{4.13}$$

If

$$(a_k, c_k, d_k)\ C = (\tilde{a}_k, \tilde{c}_k, \tilde{d}_k) \tag{4.14}$$

with $C \in I\!\!R^{3\times 3}$*, nonsingular, and* Z_k *and* $\tilde{Z}_k$ *satisfy (3.24), then*

$$Z_k = \tilde{Z}_k. \tag{4.15}$$

Proof. From (4.9) follows

$$\begin{aligned} Z_k &= Z + (a_k, c_k, e_k) \left[\begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (a_k, c_k, e_k) \right]^{-1} \theta_k \\ &= Z + (a_k, c_k, e_k)\ C\ C^{-1} \left[\begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (a_k, c_k, e_k) \right]^{-1} \theta_k \\ &= Z + (a_k, c_k, e_k)\ C \left[\begin{pmatrix} r_{k-1}^T \\ r_{k-2}^T \\ r_{k-1}^T P^T A^T \end{pmatrix} (a_k, c_k, e_k)\ \ C \right]^{-1} \theta_k \\ &= \tilde{Z}_k. \end{aligned}$$

□

For $Z = I$ we get the following result for the nonsingularity of rank-j updates.

LEMMA 8. *Let* $D, E \in I\!\!R^{n\times j}$*, let be* I *the unit matrix in* $I\!\!R^{n\times n}$ *and* I_j *the unit matrix in* $I\!\!R^{j\times j}$*, then* $I + DE^T$ *is invertible if* $I_j + E^T D$ *is invertible and*

$$[I + DE^T]^{-1} = I - D[I_j + E^T D]^{-1} E^T. \tag{4.16}$$

Proof.

$$\begin{aligned} &(I + DE^T)(I - D[I_j + E^T D]^{-1} E^T) \\ &\qquad = I + DE^T - D[I_j + E^T D]^{-1} E^T - DE^T D[I_j + E^T D]^{-1} E^T \\ &\qquad = I + D\left(I_j - [I_j + E^T D]^{-1} - E^T D[I_j + E^T D]^{-1}\right) E^T \\ &\qquad = I + D\left(I_j - [I_j + E^T D][I_j + E^T D]^{-1}\right) E^T = I. \end{aligned}$$

□

5. Algorithmic Considerations. The Euclidean norm of the residuals of generalized cg methods may oscillate heavily. The same is true for CKS methods. Therefore we apply Schönauer's minimal residual smoothing [11] to the original CKS sequence. The sequence

$$s_0 = r_0 \ , z_0 = x_0 \ , \tag{5.1}$$

$$s_k = s_{k-1} + \gamma_k (r_k - s_{k-1}) \ , \tag{5.2}$$

$$z_k = z_{k-1} + \gamma_k (x_k - z_{k-1}) \tag{5.3}$$

is a corresponding *smoothed method* delivering approximations z_k and residuals s_k. γ_k is determined from $\|s_k\| = \min$:

$$\gamma_k = -\frac{s_{k-1}^T (r_k - s_{k-1})}{\|r_k - s_{k-1}\|^2}. \tag{5.4}$$

The technique guarantees a monotonous decrease of the residuals

$$\|s_k\| \le \|r_k\|, \tag{5.5}$$

$$\|s_k\| \le \|s_{k-1}\|. \tag{5.6}$$

For a theoretical investigation see [16, 17]. The implementation according to (5.2) and (5.3) can give deceptive results in practice because the updates for s_k and z_k are decoupled. Alternative implementations were proposed by Zhou and Walker [19] that perform better in some circumstances.

We can formulate the following rank-three update CKS method by means of the preceding reflections.

Algorithm I

Choose x_0 *as initial guess for the solution of the system* $Ax = b$*, let* $r_0 = Ax_0 - b$ *be the starting residual, set* $r_{-1} = 0$*,* $a_0 = b_0 = c_0 = d_0 = e_0 = f_0 = 0$*,* $a_1 = b_1 = c_1 = d_1 = e_1 = f_1 = 0$ *and* $Z_0 = Z_1 = Z = I$*. Set the initial values for the smoothed sequence* $z_0 = x_0$ *and* $s_0 = r_0$*.*
For $k > 2$ *calculate*

$$\beta_1 = -\frac{r_{k-1}^T A P r_{k-1}}{\|r_{k-1}\|^2}, \tag{5.7}$$

$$\beta_2 = -\frac{r_{k-2}^T A P r_{k-1}}{\|r_{k-2}\|^2}. \tag{5.8}$$

If $\beta_1 + \beta_2 = 0$*, then set* $\beta_1 = \beta_2 = 0$ *and* $\beta_0 = 1$*, else set* $\beta_0 = \frac{1}{\beta_1 + \beta_2}$*.*

$$\tilde{e}_k = \beta_0 (APr_{k-1} + \beta_1 r_{k-1} + \beta_2 r_{k-2}), \tag{5.9}$$

$$a_k = \frac{APr_{k-2}}{r_{k-1}^T APr_{k-2}}, \tag{5.10}$$

$$c_k = \frac{APr_{k-3}}{r_{k-2}^T APr_{k-3}}, \tag{5.11}$$

$$e_k = \frac{\tilde{e}_k}{\tilde{e}_k^T APr_{k-1}}, \tag{5.12}$$

If Ψ_k *in (4.6) is singular, then restart from the smoothed sequence, else determine* b_k, d_k, f_k *from (4.8). Set*

$$Z_k = I + a_k b_k^T + c_k d_k^T + e_k f_k^T. \tag{5.13}$$

For $k \geq 1$ *calculate*

$$\alpha_{2,k} = -\frac{r_{k-2}^T Z_k^T APr_{k-1}}{r_{k-2}^T Z_k r_{k-2}}, \tag{5.14}$$

$$\alpha_{1,k} = -\frac{1}{r_{k-1}^T Z_k r_{k-1}} \left(r_{k-1}^T Z_k^T APr_{k-1} + \alpha_{2,k} r_{k-2}^T Z_k r_{k-1} \right), \tag{5.15}$$

If $\alpha_{1,k} + \alpha_{2,k} = 0$*, then restart from the smoothed sequence, else*

$$\phi_k = \frac{1}{\alpha_{1,k} + \alpha_{2,k}}, \tag{5.16}$$

$$r_k = \phi_k (APr_{k-1} + \alpha_{1,k} r_{k-1} + \alpha_{2,k} r_{k-2}), \tag{5.17}$$

$$x_k = \phi_k (Pr_{k-1} + \alpha_{1,k} x_{k-1} + \alpha_{2,k} x_{k-2}). \tag{5.18}$$

Calculate the smoothed quantities z_k *and* s_k *from* x_k *and* r_k*.*

$$\tilde{r}_k \quad = \quad Z_k r_k, \tag{5.19}$$

Determine the approximation $\tilde{x}_k$ *corresponding to the residual* $\tilde{r}_k = A\tilde{x}_k - b$ *from*

$$\begin{aligned} \tilde{x}_k \quad = \quad & \frac{1}{1 + \frac{f_k^T r_k}{\tilde{e}_k^T APr_{k-1}}} \left(x_k + b_k^T r_k \frac{Pr_{k-2}}{r_{k-1}^T APr_{k-2}} + d_k^T r_k \frac{Pr_{k-3}}{r_{k-1}^T APr_{k-2}} \right. \\ & \left. + f_k^T r_k \frac{\beta_0}{\tilde{e}_k^T APr_{k-1}} (Pr_{k-1} + \beta_1 x_{k-1} + \beta_2 x_{k-2}) \right). \end{aligned} \tag{5.20}$$

Calculate the smoothed quantities z_k *and* s_k *again from* $\tilde{x}_k$ *and* $\tilde{r}_k$*. If at least five steps have been performed without restart and*

$$\frac{\|s_k - s_{k-1}\|}{\|s_k\|} \leq 10^{-3}, \tag{5.21}$$

then restart from the smoothed sequence.

By this choice of a_k, c_k, e_k the main diagonal of Ψ_k is equal to 1. We got very bad convergence for the choice

$$a_k = \frac{r_{k-1}}{\|r_{k-1}\|^2}, \tag{5.22}$$

$$c_k = \frac{r_{k-2}}{\|r_{k-2}\|^2} \tag{5.23}$$

instead of (5.10) and (5.11) showing that the methods are sensitive with respect to these vectors. From lemma 7 it follows that the choice according to (5.22), (5.23) and (5.12) is equivalent to

$$(a_k, c_k, e_k) = S_k.$$

Thus Z_k fulfills (4.11) and the Frobenius norm in (4.10) is minimized.

Condition (5.21) should prevent that the smoothed residuals stagnate. The value was optimized by numerical tests. In many cases the method restarts once in the first iteration steps and then proceeds without restart. This can be considered as an adaptive search for a better initial guess.

The second calculation of the smoothed quantities z_k and s_k from $\tilde{x}_k$ and $\tilde{r}_k$ is implemented in order to exploit as much information as possible without essential work. Note that (5.19) and (5.20) consist only of dot products and triadic operations. Thus we put the information of r_k with the corresponding x_k and the information of $Z_k r_k$ with the corresponding approximation $\tilde{x}_k$ into the smoothing algorithm. Our tests indicate that the convergence is accelerated by the second smoothing, whereas the first smoothing could be omitted.

6. Numerical Experiments. Let us consider the rough model of the 3–dimensional Navier–Stokes equations

$$\Delta v + v + \rho (v^T \nabla) v = h, \tag{6.1}$$

with $v = (v_1, v_2, v_3)^T$, $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})^T$, $\Delta = \nabla^T \nabla$. The calculations have been performed on a $20 \times 20 \times 20$ grid with Dirichlet boundary conditions on the unit cube. The right-hand side is determined so that the exact solution of equation (6.1) consists of trigonometric functions. The linear system arises from a finite difference discretization with consistency order 2 and from the linearization in the first Newton step. The matrix is normalized, i.e. every row is divided by the sum of the absolute entries in that row and all diagonal entries have a positive sign. The parameter ρ simulates a Reynolds number. For increasing ρ the skew-symmetric part of the system matrix increases.

We compare a CKS method according to algorithm I, denoted by R3-CKS, with the biconjugate gradients (BCG) [3, 9], smoothed by Schönauer's residual smoothing [11], with QMR [5] and with GMRES [10] introduced by Saad and Schultz. GMRES minimizes the residuals in the Euclidean norm in the whole spanned Krylov space. In general it is not feasible because the storage requirements and the computational work increase with the iteration. It is used as reference for the best possible reduction of the residuals in the spanned space.

We present the tests for $\rho = 10$, $\rho = 50$, $\rho = 100$ and $\rho = 1000$. We always count the matrix-vector multiplications as essential operations instead of the iteration steps. One iteration is equivalent to one matrix-vector multiplication for GMRES, while for BCG, QMR and R3-CGS two matrix-vector multiplications have to be performed in each step. For all Reynolds numbers the R3-CKS algorithm restarts because of condition (5.21) monitoring the convergence, see table 6.1. The relation between the original R3-CKS residuals and the smoothed residuals according to the strategy of algorithm I for $\rho = 50$ is depicted in figure 6.1. As already mentioned, the norm of the original residuals oscillates heavily, as it does in general for related cg methods, see also [16]. In figure 6.2 the non-smoothed R3-CKS residuals are shown in comparison with the non-smoothed original BCG residuals. The qualitative behavior is the same. The investigation shows that smoothing is very advisable. Therefore, only the smoothed residuals are presented in the following tests.

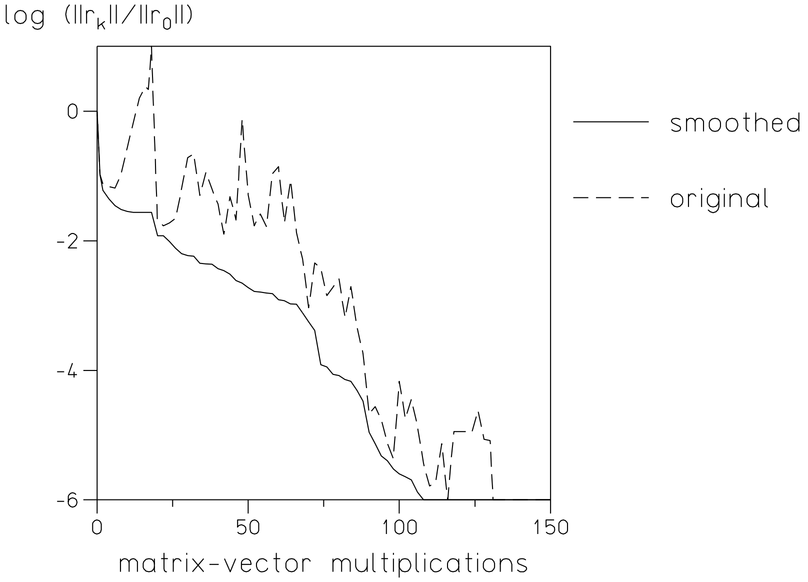

FIG. 6.1. *Relative residuals of R3-CKS for* $\rho = 50$

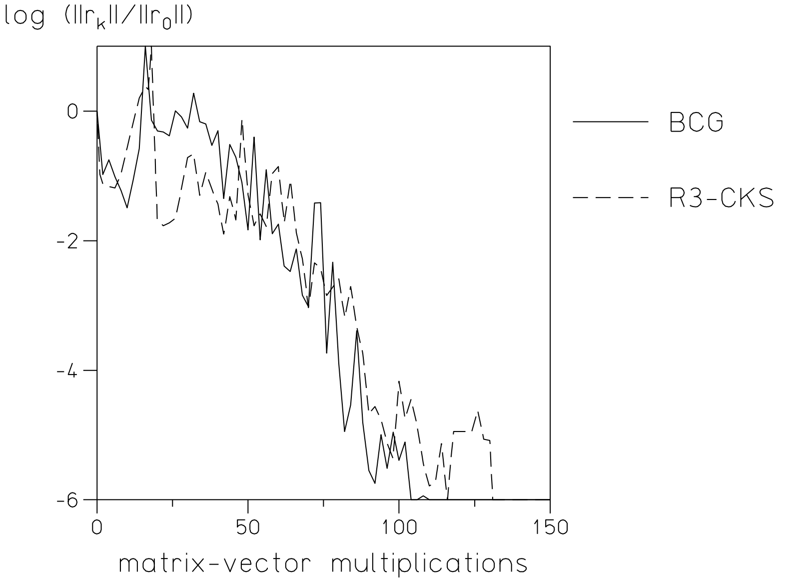

FIG. 6.2. *Relative, non-smoothed residuals of BCG and R3-CKS for* $\rho = 50$

ρ	restart at matrix-vector multiplication (mvm)
10	46
50	16
100	18
1000	16 and then every 10 mvm

TABLE 6.1
Restart because of slow convergence

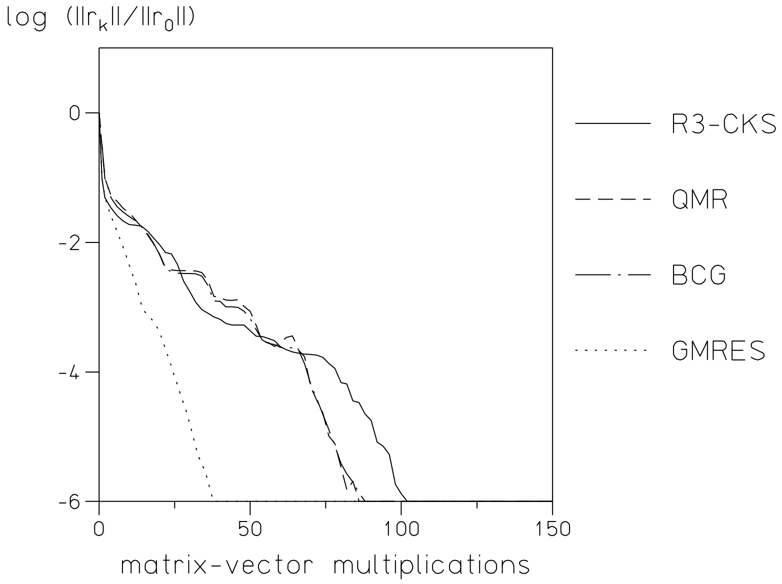

FIG. 6.3. *Relative residuals for $\rho = 10$*

For $\rho = 10$, $\rho = 50$ and $\rho = 100$, R3-CKS is competitive with BCG and QMR, see figures 6.3, 6.4 and 6.5. Up to a reduction of the relative residual of 10^{-4} for $\rho = 10$, of 10^{-3} for $\rho = 50$ and of 10^{-2} for $\rho = 100$ R3-CKS is mostly better than BCG and QMR. For higher accuracies R3-CKS becomes worse. As the linear system comes from a Newton linearization high accuracies are in general not required, so that the method becomes attractive for practical applications. Moreover, for a small reduction of the relative residual R3-CKS seems to be very close to GMRES if we count the iteration steps. Note that one R3-CKS iteration needs two matrix-vector multiplications, while GMRES needs one.

For $\rho = 1000$ R3-CKS fails while BCG and QMR still slowly converge, see figure 6.6. This may be due to the fact that the method restarts too often, see table 6.1.

However, R3-CKS has been proven to be competitive with the commonly used BCG and QMR algorithms. We think that further improvements of R3-CKS are possible by changing the restart philosophy and in particular by another choice of the rank-three update vectors.

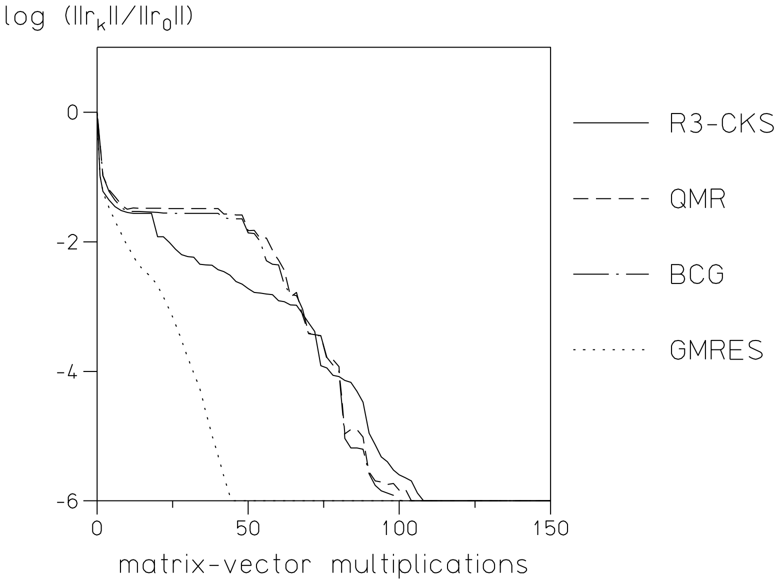

FIG. 6.4. *Relative residuals for* $\rho = 50$

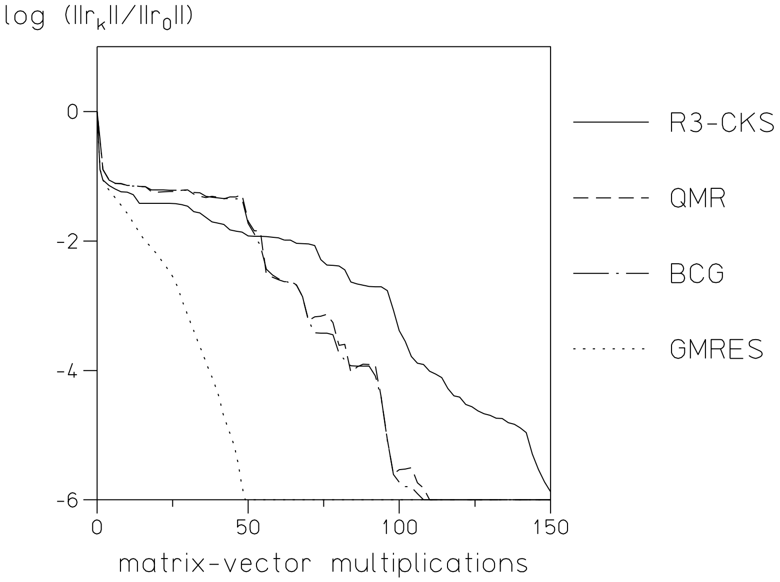

FIG. 6.5. *Relative residuals for* $\rho = 100$

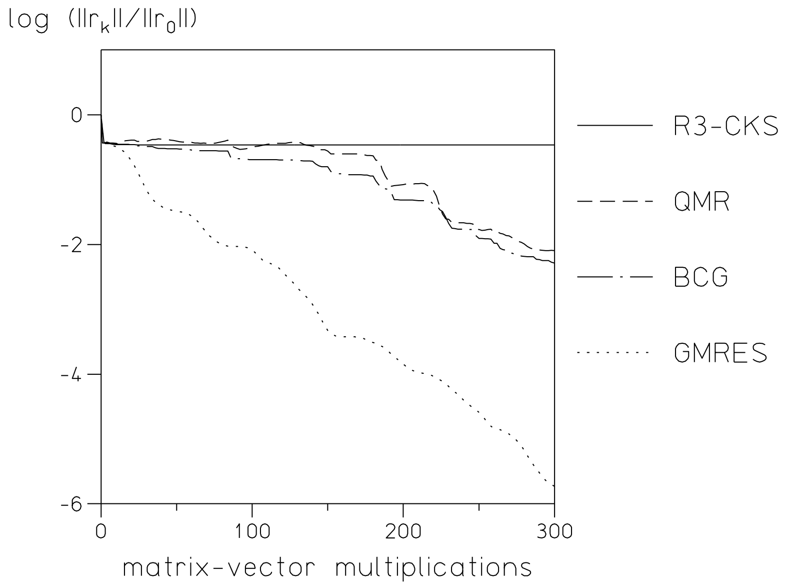

FIG. 6.6. *Relative residuals for* $\rho = 1000$

7. Outlook. CKS methods minimize a norm depending on Z_k, thus depending on the iteration step. Moreover, the properties of Z_k are up to now unknown. By choosing the preconditioning matrix

$$P_k = Z_k, \tag{7.1}$$

also depending on the iteration step, we obtain from (2.16)

$$\|e_k\|_A \quad \leq \quad \sqrt{1 + \frac{\rho^2(R)}{\mu_m^2}} \min_{\Pi_k} \|\Pi_k(Z_k A)e_0\|_A, \tag{7.2}$$

where $\rho(R)$ is the spectral radius of the skew-symmetric part of A^{-1}. μ_m is the minimum eigenvalue of the symmetric part of A^{-1}. Thus we get a Krylov subspace method that minimizes the energy norm of the error and can be formulated as a short recurrence at the same time. Of course the determination of the rank-three update vectors becomes more complex. This will be the subject of further research.

8. Conclusion. CKS methods have been proposed generalizing the cg concept. Convergence and geometric properties have been shown that are similar to well known cg results. It has been proven that CKS methods, minimizing residuals in the whole spanned space, can be implemented as short recurrences. A first realization of these methods based on a rank-three update is competitive with smoothed biconjugate gradients and QMR. There still are many open questions for the choice of the here introduced matrices Z_k. The potential of CKS methods is not yet exhausted and further research seems to be promising.

Our proper aim has been to stimulate research and development for new iterative approaches. The goal is to improve CKS methods so that they become more robust and efficient. If it is possible to get a convergence close to GMRES, then for certain cases very elaborate preconditioning techniques may be unnecessary. As nearly all robust preconditioners are recursive by nature it is difficult to implement them on vector and parallel computers, whereas CKS methods are optimally suited for advanced computer architectures.

REFERENCES

[1] E.J. Craig, *The n-step iteration procedures*, Math. Phys. 34 (1955), pp. 64–73.

[2] V. Faber and T. Manteuffel, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal. 21(1984), pp. 352–362.

[3] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Proc. Dundee Biennial Conf. on Num. Anal., G.A. Watson, ed., Springer-Verlag, 1975, pp. 73–89.

[4] R.W. Freund, *A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems*, SIAM J. Sci. Comput., 14 (1993), pp. 470–482.

[5] R.W. Freund and N.M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[6] M.H. Gutknecht, *Variants of BICGSTAB for matrices with complex spectrum*, Research Report 91-14, Eidgenössische Technische Hochschule Zürich, Interdisziplinäres Projektzentrum für Supercomputing, ETH-Zentrum, CH-8092 Zürich, August 1991.

[7] K.C. Jea and D.M. Young *On the simplification of generalized conjugate gradient methods for nonsymmetrizable linear systems.* Lin. Alg. Appl., 52/53 (1983), pp. 399–417.

[8] W.D. Joubert and D. Young, *Necessary and sufficient conditions for the simplification of generalized conjugate-gradient algorithms*, Lin. Alg. Appl., 88/89 (1987) pp. 449–485.

[9] C. Lanczos, *Solution of systems of linear equations by minimized iterations*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 33–53.

[10] Y. Saad and M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 7 (1986), pp. 856–869.

[11] W. Schönauer, *Scientific Computing on Vector Computers*, North-Holland, Amsterdam, 1987.

[12] W. Schönauer, H. Müller, and E. Schnepf, *Pseudo-residual type methods for the iterative solution of large linear systems on vector computers*, in Parallel Computing 85, M. Feilmeier, J. Joubert, U. Schendel eds., Elsevier Science Publishers B.V., North-Holland, 1986, pp. 193–198.

[13] G.L.G. Sleijpen and D.R. Fokkema, *BiCGStab(l) for linear equations involving unsymmetric matrices with complex spectrum*, Electron. Trans. Numer. Anal., 1 (1993), pp. 11–32.

[14] P. Sonneveld, *CGS, a fast Lanczos-type solver for nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 10 (1989), pp. 36–52.

[15] H.A. van der Vorst, *BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 13 (1992), pp. 631–644.

[16] R. Weiss, *Convergence behavior of generalized conjugate gradient methods*, Internal report 43/90, University of Karlsruhe, Computing Center, 1990.

[17] R. Weiss, *Relations between smoothing and QMR*, submitted to Numer. Math., 1993.

[18] R. Weiss, *Error-minimizing Krylov subspace methods*, SIAM J. Sci. Comput., 15 (1994), pp 511–527.

[19] L. Zhou and H.F. Walker, *Residual smoothing techniques for iterative methods*, SIAM J. Sci. Comput, 15 (1994), pp 297–312.