MATHEMATICS OF COMPUTATION
VOLUME 58, NUMBER 197
JANUARY 1992, PAGES 191–212

KRYLOV APPROXIMATIONS FOR MATRIX SQUARE ROOTS IN STIFF BOUNDARY VALUE PROBLEMS

BERNHARD A. SCHMITT

ABSTRACT. Recently, we proposed an algebraic difference scheme, with extended stability properties, for linear boundary value problems involving stiff differential equations of first order. Here, an efficient approximation scheme is presented for matrix square roots, which provides the stabilization of that scheme in case of stiffness. It combines the use of low-rank matrix approximations from projections onto Krylov subspaces with an accelerated sign iteration for the matrix square root. The Krylov approximation, being accurate in eigenspaces with large eigenvalues, preserves the stability of the scheme, and the $O(n^3)$ square root computation need be performed only in lower dimension. Operation counts and numerical results show that the effort for the numerical scheme is essentially proportional to the number of stiff components, but not to the norm of the coefficient matrix. Approximation properties of low-rank Krylov matrices, which may be of independent interest, are analyzed.

1. THE SQRT ONE-STEP DIFFERENCE SCHEME

A standard approximation for the differential equation in the linear boundary value problem (BVP)

$$\begin{aligned} &u'(x) = A(x)u(x) + g(x), \qquad x \in [0, 1],\ u(x) \in \mathbf{R}^n, \\ &B_0u(0) + B_1u(1) = \beta, \end{aligned} \tag{1.1}$$

is the trapezoidal rule on a suitable grid, $0 = x_0 < x_1 < \cdots < x_N = 1$,

$$[I - \tfrac{1}{2}h_kA_{k+1}]y_{k+1} - [I + \tfrac{1}{2}h_kA_k]y_k = h_k\tilde{g}_{k+1/2}, \tag{1.2}$$

$A_j := A(x_j)$, $h_k := x_{k+1} - x_k$, $\tilde{g}_{k+1/2} := \frac{1}{2}[g(x_k) + g(x_{k+1})]$. Because of stability reasons it is usually necessary to restrict the stepsize,

$$h_k < 2/\|A_{k+j}\|, \qquad j = 0, 1, \tag{1.3}$$

to assure regularity of the matrices $(2/h_k)I \pm A$. If the coefficient in the differential equation in (1.1) is "large," $\|A(x)\| \gg 1$, this restriction leads to unacceptable small stepsizes in regions where the solution is smooth. This situation is usually referred to as "stiffness." In certain cases, however, a relaxation

Received March 27, 1990; revised December 5, 1990.
1991 *Mathematics Subject Classification.* Primary 65L10, 65L20, 65F30.
Key words and phrases. Stiff boundary value problems, matrix square roots, Krylov subspace methods.

0025-5718/92 $1.00 + $.25 per page

of (1.3) is possible (cf. [2]). In order to avoid stepsize restrictions, the following SQuare Root Trapezoidal scheme (SQRT-scheme), with a parameter ω, $0 < \omega \leq 2$, has been proposed [17, 18],

$$\varphi_\omega(-h_k A_{k+1})y_{k+1} - \varphi_\omega(h_k A_k)y_k = h_k \hat{g}_{k+1/2}, \tag{1.4a}$$

$$\varphi_\omega(z) := 1 + \frac{z}{2} - \frac{\omega}{2} + \frac{1}{2}\sqrt[+]{\omega^2 + z^2}, \qquad z \in \omega D, \tag{1.4b}$$

$$D := \{z \in \mathbf{C}\colon \operatorname{Re} z \neq 0 \text{ or } |\operatorname{Im} z| < 1\}. \tag{1.4c}$$

The square root denotes the unique branch with positive real part, i.e., the matrix $W := (\omega^2 I + h^2 A^2)^{1/2}$ is the solution of $W^2 = \omega^2 I + h^2 A^2$ with $\operatorname{Re} \mu > 0$ for every eigenvalue μ of W. This solution exists uniquely under the assumption that A has no purely imaginary eigenvalues λ with absolute value $|\lambda| \geq \omega/h$, i.e., $h\lambda \notin \omega D$. The scheme (1.4) is a stabilized version of the trapezoidal rule, which corresponds to the limit $\omega \to \infty$, and is unconditionally stable (in the absence of purely imaginary eigenvalues), since its coefficient matrices $\varphi_\omega(\pm hA)$ possess no eigenvalues with nonpositive real part. As a consequence, the stability function $f(z) := \varphi_\omega(z)/\varphi_\omega(-z)$ of this scheme is bounded by one from above (below) in the left (right) complex halfplane and has no zeros or poles in the finite complex plane. In the usual notation, the scheme is both symmetric, $f(z) \cdot f(-z) \equiv 1$, and stiffly A-stable, $f(z) \to 0$ $(\operatorname{Re} z \to -\infty)$, $f(z) \to \infty$ $(\operatorname{Re} z \to \infty)$, which is impossible for rational schemes. The convergence is of second order; for singularly perturbed equations $\varepsilon u' = Au + g$, $0 < \varepsilon \ll 1$, the global error has the form $O(\min\{h^2, \varepsilon h\})$. More explicitly, for nonstiff eigencomponents we have an h^2-scheme; for stiff components, however, the order reduces to one. This situation is ameliorated by the presence of the factor ε.

This description states the background for our discussion of the implementation of the scheme. In §2 we discuss the full computation of the matrix square root in (1.4) through an optimally accelerated sign iteration. In spite of very fast convergence, this procedure, however, is probably too costly for use in a difference scheme. In §3, replacement of the matrix

$$X := [\alpha^2 I + A^2]^{1/2} - \alpha I, \qquad \alpha := \omega/h, \tag{1.5}$$

from (1.4) by a cheaper approximation is considered. This matrix X provides the stabilization of the trapezoidal rule for large eigenvalues. But the change for small eigenvalues is only minimal. Thus, a convenient approximation should contain large eigenvalues of A with relatively high precision, and small eigenvalues may be approximated by zero. This leads to replacement of X by a low-rank approximation $Y \simeq X$, with the aim of keeping all coefficient matrices of the approximate scheme

$$\frac{1}{h} I + \frac{1}{2}(Y \pm A) \tag{1.6}$$

regular. A construction based on Krylov subspaces of A is developed in §4, and several of its approximation properties are presented in §5. The computational effort for these approximations is essentially proportional to their rank. This leads to a stable scheme, with a computational overhead to the trapezoidal rule only proportional to the local stiffness, i.e., the number of eigenvalues with absolute value exceeding the reciprocal $1/h$ of the local stepsize. This could

make this scheme competitive with others suitable for stiff BVP's, e.g., [12, 11, 3, 7]. Numerical examples for some turning point BVP's on realistic grids are presented in §6.

For any real square matrix A we denote its spectrum, spectral radius, and symmetric part by $\sigma(A)$, $\rho(A)$, and $\operatorname{Re} A = (A + A^T)/2$, respectively. Unless stated otherwise, the Euclidean vector norm and the spectral matrix norm $\|A\|_2$ are used.

2. An accelerated sign iteration

The square root $B^{1/2}$ of a matrix $B := \alpha^2 I + A^2$ is closely related to the sign function of this matrix [14]. The real sign function is easily extended to the complex plane by defining $\operatorname{sign}(z) := \operatorname{sign}(\operatorname{Re} z)$, $z \in \mathbf{C} \setminus i\mathbf{R}$. The matrix $S := \operatorname{sign}(B)$ is diagonalizable, commutes with B, and has an eigenvalue $+1$ (-1), whenever B has an eigenvalue with positive (negative) real part. For the computation of $B^{1/2}$ we consider the iteration [14]

$$\begin{aligned} &P_0 := B, && R_0 := I, \\ &P_{k+1} := \alpha_k P_k + \beta_k R_k^{-1}, && R_{k+1} := \alpha_k R_k + \beta_k P_k^{-1}, \end{aligned} \tag{2.1}$$

$k = 0, 1, 2, \ldots$, with parameters α_k, $\beta_k > 0$ to be chosen later. The simplest version uses

$$\alpha_k = \beta_k \equiv \tfrac{1}{2}. \tag{2.2}$$

Convergence properties of this iteration are identical to those of the sign iteration,

$$\begin{aligned} &S_0 := M, \\ &S_{k+1} := \alpha_k S_k + \beta_k S_k^{-1}, \qquad k = 0, 1, 2, \ldots. \end{aligned} \tag{2.3}$$

With (2.2), convergence takes place, $S_k \to S := \operatorname{sign}(M)$, $k \to \infty$, if M has no purely imaginary eigenvalues. With the iterates P_k from (2.1), the matrices $S_k := P_k W^{-1} = W R_k$ also satisfy (2.3) [17], where W is any square root of B. Thus, the P_k converge to the positive square root of B, with $\operatorname{sign}(B^{1/2}) = I$, if B has no nonpositive real eigenvalues. The iteration (2.1) has been considered in [14] with nonoptimal parameters. The optimally accelerated Newton method $W_k W_{k+1} + W_{k+1} W_k = a_k(B + W_k^2)$ for real spectrum was seen by Albrecht [1] to be related to the Wachspress parameters in the ADI-iteration (cf. [22]). The construction follows from the observation that the map $w(s) := (s + 1/s)/2$ folds the positive real axis around the point 1 such that $w(1/s) = w(s)$. By scaling the spectrum after every step to an interval of the form $[1/r, r]$, $r \geq 1$, convergence is improved.

Before restating this result, we note the relationship

$$[A^2]^{1/2} = \operatorname{sign}(A) \cdot A, \tag{2.4}$$

which has the practical consequence that for $\alpha \simeq 0$ the matrix X (cf. (1.5)) may be computed by (2.3) instead of (2.1), saving half the computational effort, since only one matrix inversion per step is needed. The relevant formulation from [1], for both iterations (2.1), (2.3), is

Lemma 2.1. *Let* $\sigma(B) \in [a^2, b^2]$, $\sigma(A) \in [a, b]$, $a > 0$, *and* $W := B^{1/2}$. *Define*

$$\begin{aligned} \alpha_0 &:= 1 \Big/ \sqrt{2(a+b)\sqrt{ab}}, & \beta_0 &:= \alpha_0 ab, \\ \alpha_1 &:= 1 \Big/ \sqrt{4\sqrt{\alpha_0\beta_0} + 1/\sqrt{\alpha_0\beta_0}}, & \beta_k &:= \alpha_k \quad (k \geq 1), \\ \alpha_{k+1} &:= 1 \Big/ \sqrt{4\alpha_k + 1/\alpha_k}, & k &= 1, 2, \ldots . \end{aligned} \tag{2.5}$$

Then $\alpha_k \to 1/2$, $k \to \infty$, *and the iterates of* (2.1), (2.3) *converge,* $P_k \to W$, $S_k \to I$, *with*

$$\sigma(P_{k+1}W^{-1}), \sigma(S_{k+1}) \in \left[2\alpha_k, \frac{1}{2\alpha_k}\right], \qquad k = 1, 2, \ldots . \tag{2.6}$$

The initial range of the spectrum $a = \rho(S_0^{-1})^{-1}$, $b = \rho(S_0)$ may be estimated by standard means, e.g., weighted norms, from the first iterate and its inverse, which is computed anyway. These estimates need not be very accurate, since the iteration (2.5), which is closely related to the Gaussian arithmetic-geometric mean iteration, shows a very rapid global convergence to $1/2$ in $\log(\log \varepsilon) + \log(\log(b/a))$ steps for a relative error criterion ε. This follows from the local quadratic convergence and the relations $1 \geq 2\alpha_{k+1} \geq \sqrt{2\alpha_k} \geq (2\alpha_1)^{2^{-k}}$, $k \geq 0$. Thus, the numerical iteration stops in a fixed number of steps for any realistic floating point range. In fact, this number only depends on the condition b/a of A, resp. $B^{1/2}$, since $1/(\alpha_0\beta_0) = 2(\sqrt{b/a} + \sqrt{a/b})$. As an example, the following table shows the (theoretical) number of iterations (2.3) necessary to satisfy the moderate stopping criterion $1 - 2\alpha_k \leq 10^{-4}$ (cf. (2.6)):

$b/a =$	10^2	$10^3 \ldots 10^5$	$10^6 \cdots 10^{12}$	$10^{13} \ldots 10^{25}$
$k+1 =$	4	5	6	7

This accuracy of the square root is sufficient to provide the stabilizing effect of the difference scheme. Thus, we will use as computation count

$$\text{one sign computation} = c \cdot n^3\text{FLOPS}, \qquad c \simeq 6, \tag{2.7}$$

since one matrix inversion needs n^3 FLOPS. One square root computation has the double cost, $\simeq 12n^3$, which still is competitive with the Schur method [5], since the average for one Schur decomposition alone is estimated to cost $15n^3$ FLOPS [10]. However, for a difference scheme, all this may be far too much overhead. In the following sections we will show that it often is possible to perform the costly sign computation in lower-dimensional spaces $\mathbf{R}^m$, $m < n$.

The acceleration was designed for real spectrum only. The global convergence (in $\mathbf{C}$) for the original iteration with parameters (2.2) follows from the well-known identity

$$(S_k - I)(S_k + I)^{-1} = [(S_0 - I)(S_0 + I)^{-1}]^{2^k}. \tag{2.8}$$

It may be deduced from, e.g., [22, §17.5] that the accelerated iteration (2.3) produces a rational Chebyshev approximation of degree 2^k to the zero function.

It has the form

$$(S_k - I)(S_k + I)^{-1} = \prod_{j=1}^{2^k} (S_0 - \gamma_j I)(S_0 + \gamma_j I)^{-1}, \tag{2.9}$$

with $\gamma_j \in (a, b)$. The global approximation properties of such functions are much better than those of polynomials. Every factor in (2.9) is smaller than 1 in the right complex halfplane and greater than 1 in the left. Hence, the iteration still converges if S_0 has no purely imaginary eigenvalues. In fact, convergence deteriorates only slowly if eigenvalues move away from the real axis. In the numerical experiments (Example 6.3), after ten unsuccessful sign iterations, the existence of nearly imaginary eigenvalues is presumed and the stepsize is reduced near the grid point involved.

3. Low-rank approximations for the matrix square root

The matrix X from (1.5) has the sole purpose of providing regularity of the coefficient matrices $\varphi_\omega(\pm hA)$ of the scheme (1.4). The matrix X has the following properties. If A is small compared to $\alpha = \omega/h$, $\|A\| \ll \alpha$, X is very small, too, $X = O(\|A\|^2/\alpha) \simeq 0$. But if A has only large eigenvalues, e.g., $\|A^{-1}\| < \alpha^{-1}$, we see that $X \simeq [A^2]^{1/2} = \operatorname{sign}(A) \cdot A$ (cf. (2.4)). It is our aim to replace X by a "cheaper" matrix with similar properties.

In this section, we discuss the approximation of X under the assumption that an approximate block Schur form of the matrix A is known, i.e., a unitary similarity transformation to (2×2)-block (almost) upper triangular structure. The construction of such a decomposition will be considered in the next section. Let $U := (Q, P) \in \mathbf{R}^{nn}$ be a partitioned unitary matrix. In practice, we will assume explicit knowledge of the part $Q \in \mathbf{R}^{nm}$ only. This introduces a partition of A,

$$\hat{A} := U^{\mathrm{T}} A U = \begin{pmatrix} H & B \\ C & D \end{pmatrix}, \qquad U = (Q, P). \tag{3.1}$$

For the following discussion we assume that the eigenvalues of A with large modulus are essentially contained in the principle submatrix H, and the matrices in the second row are relatively small. The case of a vanishing second row is particularly simple. In this case, a square root still exists, since zero is a nondefective eigenvalue.

Lemma 3.1. *Let $C, D = 0$ in (3.1), and let H have no eigenvalues with zero real part. Then, $Y = (A^2)^{1/2} := \lim_{\alpha \to 0} (\alpha^2 I + A^2)^{1/2}$ is given by*

$$Y = \tilde{S} A = U \hat{Y} U^{\mathrm{T}}, \qquad \tilde{S} := Q S Q^{\mathrm{T}},$$

where

$$\hat{Y} := \begin{pmatrix} SH & SB \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} H & B \\ 0 & 0 \end{pmatrix}, \qquad S := \operatorname{sign}(H). \tag{3.2}$$

Proof. With (3.1) and $K := (\alpha^2 I + H^2)^{1/2}$, a simple computation gives

$$(\alpha^2 I + A^2)^{1/2} = \begin{pmatrix} K & (\alpha I + K)^{-1} H B \\ 0 & \alpha I \end{pmatrix}. \tag{3.3}$$

But $K - SH = \alpha^2(K + SH)^{-1}$. By assumption and Theorem 1.4 of [17] there follows $\|K - SH\| \leq \alpha^2/(2\mu)$ in a suitable norm, where $\mu := \min\{|\operatorname{Re}\lambda| : \lambda \in \sigma(H)\} > 0$. Finally, $(\alpha I + K)^{-1}HB = (SH)^{-1}HB + O(\alpha) = SB + O(\alpha)$ completes the description of the limit $\alpha \to 0$ in (3.3). □

Remark. The matrix QSQ^{T} in Lemma 3.1 is not the full sign matrix of A.

The approximate square root of rank $m < n$ in the general case will be constructed according to Lemma 3.1 by neglecting the (C, D) part of the matrix A. The practical construction depends on the specification of Q alone, through

$$Y := QSQ^{\mathrm{T}}A \in \mathbf{R}^{nn}, \quad \widetilde{S} := QSQ^{\mathrm{T}}, \qquad Q = (q_1, \ldots, q_m) \in \mathbf{R}^{nm}, \tag{3.4}$$

where $S \in \mathbf{R}^{mm}$ is the sign matrix in m-space,

$$S := \operatorname{sign}(H), \qquad H := Q^{\mathrm{T}}AQ. \tag{3.5}$$

In the approximate SQRT-scheme we attach to all matrices in (3.4), (3.5) the subscript k of the matrix $A_k = A(x_k)$ from which they are constructed. A criterion for the choice of their rank will be presented later on. It will depend on the local stepsize h_k. Since every matrix A_k appears in two steps of the scheme (1.4) with general stepsizes $h_{k-1} \neq h_k$, we must allow for approximations with different ranks in the steps through $[x_{k-1}, x_k]$ and $[x_k, x_{k+1}]$. Within one interval, however, matrices must change smoothly in order to maintain consistency, which, in general, precludes a rank change. Thus, at every grid point we have to work with two square root approximations Y_{k-}, Y_{k+} derived from A_k by (3.4), where

$$\operatorname{rank}(Q_{k+}) = \operatorname{rank}(Q_{k+1-}) =: m_k$$

and the two matrices Q_k agree in their first columns, $Q_{k+} = (Q_{k-}, \ldots)$ if $m_k > m_{k-1}$, resp. $Q_{k-} = (Q_{k+}, \ldots)$ if $m_k < m_{k-1}$.

Thus, the approximate SQRT-scheme takes the form

$$\left[\frac{2}{h_k}I + Y_{k+1-} - A_{k+1}\right] y_{k+1} - \left[\frac{2}{h_k}I + Y_{k+} + A_k\right] y_k = \overline{g}_{k+1/2}, \tag{3.6}$$

where

$$\overline{g}_{k+1/2} := \tfrac{1}{2}[I + \widetilde{S}_{k+}]g(x_k) + \tfrac{1}{2}[I - \widetilde{S}_{k+1-}]g(x_{k+1}).$$

Looking at the eigencomponents and assuming that the eigenstructure of the Y and A matrices are sufficiently similar, we see that (3.6) reduces to the trapezoidal rule for nonstiff components. For stiff ones, it corresponds to the implicit or explicit Euler scheme according to the sign of the eigenvalue. Thus, this scheme looks similar to that of Kreiss, Nichols, and Brown [11]. However, we want to stress the fact that (3.6) does not work with transformed coordinates and the need to keep transformations smooth, which may be difficult to achieve in practice (see also Remark 3 in §4). It is more akin to automatic partitioning methods in stiff initial value problems [8, 4, 9].

We now consider only one single matrix and drop all subscripts with respect to the grid numbering. The matrix $Q \in \mathbf{R}^{nm}$ will be constructed by induction on m as explained in the next section. One of the most important aspects of this procedure is the choice of the rank m. Rather than discussing the approximation properties of Y with respect to the full square root X, we will consider two important properties of the original SQRT-scheme (1.4) directly.

These are the regularity of the matrices $\varphi(\pm hA)$ and one-sided stability estimates. We analyze their counterparts for the approximation scheme (3.6). The next lemma is merely a motivation for the rank criterion. It is formulated under the simplifying assumption that H in (3.1) is diagonalizable.

Lemma 3.2. *With A from* (3.1), *consider one of the coefficient matrices $F := \alpha I + Y \pm A$ of the scheme* (3.6). *The projector onto the invariant subspace of H corresponding to eigenvalues of positive, resp. negative, sign is $T_\pm := \frac{1}{2}[\operatorname{sign}(H) \pm I]$. Let k be its rank, and $T_\pm := VW^{\mathrm{T}}$, V, $W \in \mathbf{R}^{nk}$, $W^{\mathrm{T}}V = I$ such that $W^{\mathrm{T}}HV$ is diagonal. Let*

$$\mu := \min\{|\operatorname{Re}\lambda| : \lambda \in \sigma(H)\} > 0, \tag{3.7a}$$

$$b := \|W^{\mathrm{T}}B\|_2, \quad c := \|CV\|_2, \quad d := \|D\|_2. \tag{3.7b}$$

Then the matrix F is regular if

$$\alpha > d + \frac{2bc}{2\mu + d}. \tag{3.8}$$

Proof. We drop the subscript on $T = \frac{1}{2}[S \pm I]$ (cf. (3.5)). With $v \in \mathbf{R}^m$, $w \in \mathbf{R}^{n-m}$, $y^{\mathrm{T}} := (v^{\mathrm{T}}, w^{\mathrm{T}})$, the homogeneous system $Fy = 0$ reads

$$\begin{aligned}(\alpha I \pm 2TH)v \pm 2TBw &= 0,\\ \pm Cv + (\alpha I \pm D)w &= 0.\end{aligned}$$

From the first equation there follows $v = Tv$. By (3.8), $\alpha I \pm D$ is regular, and by elimination of w we arrive at

$$\alpha Tv \pm 2THTv - 2TB(\alpha I \pm D)^{-1}CTv = 0.$$

Using $v = Tv = Vz$, $z \in \mathbf{R}^k$, and $W^{\mathrm{T}}V = I$, we now have the equation

$$\alpha z \pm 2W^{\mathrm{T}}HVz - 2W^{\mathrm{T}}B(\alpha I \pm D)^{-1}CVz = 0. \tag{3.9}$$

If H is diagonalized by the matrix Z, $Z^{-1}HZ = \operatorname{diag}(\lambda_j)$, V and W may be assumed to be the columns of Z^{-1}, resp. Z^{T}, corresponding to eigenvalues with appropriate sign. Since $\pm z^{\mathrm{T}}WHVz \geq \mu$ by (3.7a), multiplication of (3.9) by z^{T} leads to

$$0 \leq [\alpha + 2\mu - 2\|W^{\mathrm{T}}B\|\|CV\|/(\alpha - \|D\|)]\|z\|^2.$$

By assumption (3.8), the term in square brackets is positive, whence $z = 0$, $v = 0$, and $y = 0$. □

If the matrix A is not extremely nonnormal, the off-diagonal block B in (3.1) should not exceed the main diagonal H in size. Thus, in (3.8), $b/\mu \leq 1$ may be expected and the regularity criterion (3.8) simplifies to $\alpha > c + d$ or $\alpha^2 > c^2 + d^2$, where $\alpha = 2/h$. This is the motivation for the use of the following assumption on the second row of (3.1):

$$\|(C, D)\|_2 = \|P^{\mathrm{T}}A\|_2 = \|(I - QQ^{\mathrm{T}})A\|_2 \leq \delta. \tag{3.10}$$

The use of the part $Q^{\mathrm{T}}A$ in approximating the matrix A and the rank criterion (3.10) were already discussed in [8]. However, in §4, we will describe an implementation which only requires computation of the main part $Q^{\mathrm{T}}A$ of A.

The part $(I - QQ^T)A = PP^TA$ will be called the residual of the low-rank approximation of the matrix A. In view of (3.6), (3.8) we will use $\delta = 1/h < 2/h$ in (3.10) to determine the rank of Q. A rigorous analysis of the effect of this rank criterion on the stability of the approximation scheme (3.6) will be given now.

The norm estimates for the transition matrix $f(hA)$ in Theorem 1.4 of [17] played a crucial role in the stability analysis there. The function $f(z) = z + \sqrt{1+z^2}$ was the stability function of the SQRT-scheme with $\omega = 1$. We reprove this theorem for the approximate case. But before that, a short explanation is necessary. For the nonconstant coefficient case, the scheme (3.6) has the form $T(h_k, -A_{k+1})y_{k+1} - T(h_k, A_k)y_k = r_k$, where the matrices $T(h, \pm A)$ are derived from different coefficients $A(x)$. This makes the analysis more difficult. However, representing y_{k+1} by terms with smaller index $j < k$, we get

$$T(h_k, -A_{k+1})y_{k+1} = F_kF_{k-1}\cdots F_{j+1}T(h_j, A_j)y_j + \cdots,$$

where the subexpressions $F_i := T(h_i, A_i)T(h_{i-1}, -A_i)^{-1}$ contain only one coefficient matrix A_i. In order to show uniform boundedness of, e.g., y_{k+1} in terms of y_j in the case $\operatorname{Re} A < 0$, coarse bounds suffice for $T(h_j, A_j)$ and $T(h_k, -A_{k+1})^{-1}$, but tight bounds of size $1 + O(h)$ are needed for $\|F_i\|$. This is the subject of the next theorem. For ease of formulation we will consider equal stepsizes $h_i = h_{i-1}$ only.

Theorem 3.3. *Let the matrix A have the form* (3.1) *and satisfy the bound* (3.10) *with $0 \le \delta \le h^{-1}$, $h > 0$. Let the matrix Y be computed by* (3.4), (3.5), *and define the transition matrix by*

$$F := \left[I + \frac{h}{2}(Y + A)\right]\left[I + \frac{h}{2}(Y - A)\right]^{-1}. \tag{3.11}$$

Then

(a) $\operatorname{Re} A \le \mu I$, $\mu < 0$ *implies* $\|F\|_2 \le \frac{2+h\delta}{2+h\delta-2h\mu} < 1$;

(b) $\operatorname{Re} A \ge \nu I$, $\nu > 0$ *implies* $\|F^{-1}\|_2 \le \frac{2+h\delta}{2+h\delta+2h\nu} < 1$.

Proof. Since all estimates are invariant under unitary transformations, we may assume $U = I$. From the assumption in part (a) there follows $\operatorname{Re} H \le \mu I$, which gives $S = \operatorname{sign}(H) = -I$. Thus, the matrices $Y \pm A$ have the form

$$Y - A = \begin{pmatrix} -2H & -2B \\ -C & -D \end{pmatrix}, \qquad Y + A = \begin{pmatrix} 0 & 0 \\ C & D \end{pmatrix}.$$

Now, for arbitrary $v \in \mathbf{R}^n$, $v \ne 0$, we introduce

$$\begin{pmatrix} x \\ y \end{pmatrix} := u := \left[I + \frac{h}{2}(Y - A)\right]^{-1} v, \qquad \begin{pmatrix} a \\ b \end{pmatrix} := Au = \begin{pmatrix} Hx + By \\ Cx + Dy \end{pmatrix},$$

with $x, a \in \mathbf{R}^m$, $y, b \in \mathbf{R}^{n-m}$. Then,

$$(Y - A)u = \begin{pmatrix} -2a \\ -b \end{pmatrix}, \qquad (Y + A)u = \begin{pmatrix} 0 \\ b \end{pmatrix},$$

and

$$\begin{aligned}\frac{\|Fv\|^2}{\|v\|^2} &= \frac{\|u+(h/2)(Y+A)u\|^2}{\|u+(h/2)(Y-A)u\|^2}\\ &= \frac{\|u\|^2+hy^{\mathrm{T}}b+h^2\|b\|^2/4}{\|u\|^2-2hx^{\mathrm{T}}a-hy^{\mathrm{T}}b+h^2\|a\|^2+h^2\|b\|^2/4}\\ &= \frac{1+\beta+\eta^2/4}{1-2\alpha-\beta+\xi^2+\eta^2/4} =: \psi(\alpha,\beta,\xi,\eta),\end{aligned} \tag{3.12}$$

where $\alpha := hx^{\mathrm{T}}a/\|u\|^2$, $\beta := hy^{\mathrm{T}}b/\|u\|^2$, $\xi := h\|a\|/\|u\|$, and $\eta := h\|b\|/\|u\|$. The following relations among these variables hold:

$$|\alpha| \le \xi, \quad |\beta| \le \eta \le h\delta, \quad \text{and} \quad \alpha+\beta \le h\mu < 0. \tag{3.13}$$

The first two estimates, $|\alpha| \le \xi$, $|\beta| \le \eta$, follow from the Cauchy-Schwarz inequality, the third, $\eta \le h\delta$, is a consequence of (3.10), since $\|b\|/\|u\| \le \|(C, D)\|_2 \le \delta$, and the last estimate is equivalent to the assumption $\operatorname{Re} A \le \mu I$, since

$$\alpha+\beta = h(x^{\mathrm{T}}a+y^{\mathrm{T}}b)/\|u\|^2 = hu^{\mathrm{T}}Au/\|u\|^2 = hu^{\mathrm{T}}(\operatorname{Re} A)u/\|u\|^2.$$

Evidently, ψ is decreasing in ξ. Thus, the norm of F can be bounded by

$$\|F\|^2 \le \max\{\psi(\alpha,\beta,\alpha,\eta): |\beta| \le \eta \le h\delta,\ \alpha+\beta \le h\mu < 0\}.$$

By tedious computations we will show in the technical Lemma 3.4 below that the last expression satisfies the bound given in statement (a) of the theorem. Part (b) may be reduced to (a) by changing the sign of the matrix with $\mu := -\nu$. □

Remarks. (1) The estimates of Theorem 3.3 are similar to those of the original SQRT-scheme. Even for the extremal case that the "perturbation" (C, D), which is neglected in the construction of the square root approximation Y, has size $\delta = 1/h$, both bounds do not exceed one and still have the form $\|F\| \le 1 + O(h)$ $(h \to 0)$, resp. $\|F\| \le 1/O(h\mu)$ $(h\mu \to -\infty)$.

(2) We have to note that the theorem does not cover the important case of a rank change $\operatorname{rank}(Y_{k-}) \ne \operatorname{rank}(Y_{k+})$. This possibility was explained in the discussion preceding the definition (3.6) of the approximation scheme and is due to the use of (3.10) as the practical rank criterion (cf. §4) with $\delta_- = 1/h_{k-1}$ and $\delta_+ = 1/h_k$. The capability to adapt the rank to the local situation, especially the stepsize, is a prerequisite for an efficient implementation. □

Proof of the final estimate in Theorem 3.3 was deferred to

Lemma 3.4. *For $\lambda \le 0$ and $e \le 1$, the function ψ from* (3.12) *satisfies*

$$\max\{\psi(\alpha,\beta,\alpha,\eta)^{1/2}: |\beta| \le \eta \le e,\ \alpha+\beta \le \lambda\} \le \frac{2+e}{2+e-2\lambda} \le \frac{1}{1-2\lambda/3}.$$

Proof. First, we observe that ψ is an increasing function of η^2 for $\alpha+\beta \le 0$ and attains its maximum at $\eta^2 = e^2$. Second, the function

$$\psi(\alpha,\beta,\alpha,e) = \frac{1+\beta+e^2/4}{(1-\alpha)^2-\beta+e^2/4}$$

is increasing in α, since $\alpha \le \lambda - \beta \le e \le 1$. Thus, the overall maximum occurs on the line $\alpha = \lambda - \beta$, $-e \le \beta \le e$. Now we show

$$\psi(\lambda-\beta, \beta, \lambda-\beta, e) = \frac{1+\beta+e^2/4}{(1-\lambda)^2+(1-2\lambda)\beta+\beta^2+e^2/4} \le \left(\frac{1+e/2}{1+e/2-\lambda}\right)^2, \qquad |\beta| \le e.$$

Crossmultiplication of the denominators and the scaling $\beta = xe$, $|x| \le 1$, lead to the equivalent condition

$$(3.14)\quad e(1+e/2)^2x^2 - \lambda(\lambda+e+e^2/2)x - \lambda(1-\lambda-e^2/4) \ge 0, \qquad |x| \le 1.$$

This inequality holds, because sufficient conditions for the polynomial $p(x) := ax^2+bx+c$ to be nonnegative on the interval $[-1, 1]$ are

$$(3.15)\qquad a \ge 0, \quad c \ge 0, \quad c^2+2ac-b^2 \ge 0.$$

This can be seen from the implication

$$\begin{aligned} & c \ge \sqrt{a^2+b^2} - a \\ & \Rightarrow p(x) \ge \sqrt{a^2+b^2} - a(1-x^2) + bx \\ & \ge \sqrt{a^2+b^2}\left[1 - \sqrt{(1-x^2)^2+x^2}\right] \ge 0 \qquad (x^2 \le 1), \end{aligned}$$

where the Cauchy-Schwarz inequality was used in the main step. Now, for the polynomial in (3.14) the conditions (3.15) read

$$e(1+e^2/2)^2 \ge 0, \qquad \lambda^2 - \lambda(1-e^2/4) \ge \lambda^2 - 3\lambda/4 \ge 0,$$

$$c^2+2ab-b^2 = -\lambda^3\left(2+2e+\frac{1}{2}e^2\right) + \lambda^2\left(1+\frac{e}{2}\right)^2\left(1+e-\frac{3}{4}e^2\right) - 2\lambda e\left(1+\frac{e}{2}\right)^2\left(1-\frac{e^2}{4}\right) \ge 0.$$

Here, the nonnegativity follows from the assumptions $\lambda \le 0$ and $0 \le e \le 1$. □

The motivation for the specific form of the rank criterion (3.10), $\delta = 1/h$, was essentially heuristic. Theorem 3.3 gives it a sounder theoretical basis, at least with respect to stability questions, which is also supported by the numerical experiments in §6.

4. Krylov subspaces

Recently, the application of the Arnoldi iterative method has been discussed in the solution of implicit equations in stiff initial value problems (cf. [9, 6]). The salient feature of this iteration is that convergence is fastest in subspaces corresponding to eigenvalues of the coefficient matrix with large modulus. This is due to the fact that the Arnoldi method solves the orthogonal projection of the linear system in a Krylov subspace. Thus the stabilization from the implicit method is achieved early in the iteration while the accuracy of the scheme is already provided by its explicit part. As a consequence, Brown and Hindmarsh report surprisingly low dimensions of the Krylov subspaces in their implementation.

The situation for the square root scheme (1.4) contains similar features. The square root X (cf. (1.5)) provides a stabilization of the trapezoidal rule for large eigenvalues, although this increases its consistency error. Thus, an approach based on Krylov subspaces promises a sensible approximation. But there are also differences to stiff initial value problems. For the solution of a linear system $Ax = b$, the inverse A^{-1} need only be approximated at one point, $A^{-1}b$. By choosing the right-hand side b as the starting vector in the Krylov process, a possible degeneration of this vector with respect to the eigensystem of A is not crucial, since the same degeneration will occur in the solution x. In the SQRT-scheme (1.4), the algebraic matrix function $\varphi(hA)$ multiplies the unknown vector y. In an iterative environment, e.g., iterative mesh refinement, approximations of y might be available in all but the first steps, but these might not be reliable for rapidly changing solutions. Hence, since we do not (implicitly) know the eigencomponent structure of the vector being mapped by $\varphi(hA)$, we need an approximation of the full matrix X in (1.5), i.e., one satisfying a matrix norm estimate like (3.10), not only a pointwise estimate.

The Arnoldi process generates the column vectors q_k of the orthogonal matrix Q in (3.1) as orthonormal basis vectors of the Krylov spaces

$$\begin{aligned} K_k &= \operatorname{span}\{z_1, \ldots, z_k\} \\ &= \operatorname{span}\{q_1, \ldots, q_k\}, \qquad z_k := A^{k-1}z_1, \ k \le m. \end{aligned} \tag{4.1}$$

In consequence of (4.1) the matrix H in (3.1) is upper Hessenberg. The matrix Q may be obtained from the matrix $(z_1, \ldots, z_m)$ by a QR factorization. This can either be based on Gram-Schmidt orthogonalization or on Householder reflections. Both versions have been discussed in the literature (cf. [16, 21]), where only the matrices Q and H are actually computed in accordance with (3.4), (3.5). The Arnoldi method, described by Saad [16], is based on a Gram-Schmidt process. However, this algorithm may suffer a breakdown if there is a rank deficit in the Krylov subspace, $\operatorname{rank}(q_1, Aq_1, \ldots, A^{j-1}q_1) = \operatorname{rank}(q_1, Aq_1, \ldots, A^{j}q_1)$. In cg-iterations, this degeneration is a "lucky breakdown" [16], since the corresponding cg-iterate is the exact solution already. In our context, however, a restart with some (unknown) $q_{j+1} \perp K_j$ would be necessary. Walker [21] proposed an alternative construction of Q using Householder transformations. This process additionally provides a basis of the orthogonal complement of K_j [10] and no breakdown can occur. Since it is only marginally more expensive than Arnoldi's method, this decisive advantage led us to choosing the Householder implementation. The full algorithm will be described soon.

Both Krylov processes construct parts of the full transformed matrix $U^{\mathrm{T}}AU$. An inspection of the Walker implementation shows that, after the first step $(j = 1)$, it is analytically equivalent to the usual Householder reduction to Hessenberg form (of $A_1 = R_1AR_1$). The latter was used by Enright and Kamel [8]. However, for small ranks it is more expensive than the Walker approach, since whole matrices are transformed. Its only advantage would be the possibility to check the rank criterion (3.10) or more complicated ones.

But a slight modification of this criterion allows an implementation with only knowledge of A and Q. It relies on the unitary invariance and additivity of

the Frobenius norm of a matrix,

$$\|A\|_F^2 := \operatorname{trace}(A^T A) = \sum_{i,j} a_{ij}^2 .$$

Since Q in (3.1) is unitary, we are able to satisfy (3.10) by requiring the residual bound

$$\|(I - QQ^T)A\|_2^2 \le \|(I - QQ^T)A\|_F^2 = \|A\|_F^2 - \|Q^T A\|_F^2 \le \delta^2 . \tag{4.2}$$

The left products $q_j^T A$ forming $Q^T A$ are needed for the computation of the matrix Y as well (cf. (3.4)). Thus, this rank criterion (4.2) represents a minor computational effort only. It should be expected that the value of the norm $\|A\|_F$, which has to be computed in the first step, is very large in stiff BVP's. Thus, in practice, it is advisable to scale the coefficient matrix in order to minimize this norm. This should also improve the performance of the Krylov approximation. The EISPACK subroutine BALANC [19] may be easily modified to compute a diagonal similarity transformation of A with nearly minimal Frobenius norm.

We now give the formal description of the construction of one approximate square root. There is a notational difficulty in the trivial case $m = 0$, i.e., $\|A\|_F \le \delta$, where it is understood that the algorithms produce $Y = 0$ with H and Q undefined.

Algorithm 4.1. Computation of the Krylov-Approximation Y, (3.4), (4.1) with rank criterion (4.2), using Householder reflections $R := I - 2uu^T$, $\|u\|_2 = 1$.

1. Choose q_1, $\|q_1\|_2 = 1$, R_1 such that $R_1 q_1 = e_1$.
   Compute $r_0 := \|A\|_F^2$, let $j := 0$.
2. If $r_j \le \delta^2$ then go to Step 5.
3. [21] Let $j := j + 1$, $w_j := R_j \cdots R_1 A q_j$.
   Choose R_{j+1} such that $(R_{j+1} w_j)^T e_k = 0$, $k > j + 1$,
   i.e., $R_{j+1} \cdots R_1 (q_1, Aq_1, \ldots, Aq_j)$ is upper triangular,
   $q_{j+1} := R_1 \cdots R_{j+1} e_{j+1}$.
4. $r_j := r_{j-1} - \|q_j^T A\|^2$.
   Repeat with Step 2.
5. Let $m := j$, $U := R_1 \cdots R_m$, $Q := (q_1, \ldots, q_j)$, $He_j := R_{j+1} w_j$, $j = 1, \ldots, m$. $S := \operatorname{sign}(H)$, $Y := QS(Q^T A)$.

Remarks. (1) The algorithm uses the matrix A only via products Aq and $q^T A$. Thus, sparseness properties of the coefficient matrix may be easily exploited.

(2) The computation of the left products $q^T A$ is not necessary for an almost symmetric matrix A. In this case the principal submatrix $H = Q^T A Q$ contains the essential information, and in place of Y from (3.4) the simpler square root approximation $X \simeq Q(SH)Q^T$, $SH = (H^2)^{1/2}$ may be used. The corresponding rank criterion for the algorithm is $\|A\|_F^2 - \|Q^T A Q\|_F^2 \le \delta^2$. For general matrices, however, this criterion leads to much too large ranks (usually $m = n$), since the submatrix $B = Q^T A P$ often is of the same magnitude as H.

(3) Considering neighboring points on the grid, the square root approximation $Y = Y(x)$ must depend smoothly on x, at least within one subinterval, in order to preserve the consistency of the approximation scheme (3.6). This could be expected from Gram-Schmidt, since it is a deterministic algorithm

without branching. All terms depend continuously on A and the initial Krylov vector q_1 in the case of nondegeneracy. But the sign decision involved in stable Householder reductions may introduce discontinuities into H and Q. However, different choices correspond to different similarity transformations of H with unitary diagonal matrices. But these transformations are again (implicitly) applied in Step 5, $Y = Q \cdot \operatorname{sign}(H) \cdot Q^T A$ of Algorithm 4.1, providing smooth dependence of Y on A and q_1.

In estimating the computational cost of Algorithm 4.1 we rely on the results of §2 and count the sign computation in m-space with $6m^3$ FLOPS. In Step 5 of the algorithm, we can exploit the fact that the rows of $Q^T A$ have already been computed in Step 4. For dense matrices the highest-order terms of the computation count are $mn^2 + 2m^2n - 2m^3/3$ FLOPS in Step 3 and mn^2 in Step 4. The two matrix multiplications $Q(S(Q^T A))$ in Step 5 need $mn^2 + m^2n$ FLOPS. Thus, if Algorithm 4.1 produces a rank-m approximation $H = Q^T AQ$ which satisfies the assumptions of §2, the computation count for the Krylov approximation Y is $3mn^2 + 3m^2n + \frac{16}{3}m^3$ FLOPS. The Arnoldi process [16] would be slightly cheaper in Step 3, with an overall count of $3mn^2 + 2m^2n + 6m^3$ FLOPS. But since it runs the risk of a breakdown (resp. numerical instability) in the case of a (nearly) degenerate Krylov space, we opted for the Householder version, Algorithm 4.1, which is used in the numerical experiments.

Now, we are able to make a realistic assessment of the approximate SQRT-scheme. The numerical solution of a two-point boundary value problem requires the solution of a linear system with staircase matrix. If Gaussian elimination is used, the computational effort is $\frac{4}{3}n^3$ [20] per grid point. Thus, on any subinterval of the grid for, e.g., $m \leq 0.3 \cdot n$, the approximate square root is cheaper than the elimination of the corresponding part of the linear system. In this case the SQRT-scheme is more efficient than the trapezoidal rule if the latter needs a subdivision into two or more subintervals for stability reasons. Further, it should be kept in mind that the rank m decreases with the local stepsize. On realistic grids, which concentrate points in layers, the mean value of the matrix ranks on the grid is fairly low (see §6). But this overhead provides a scheme with the original stability properties of (1.4), which allows for reliable adaptive grid placement procedures [18].

5. Convergence properties of Krylov approximations

The algorithm of the last section chooses the rank m such that the residual norm of the matrix A satisfies the rank criterion (3.10). Thus, the regularity requirements of Theorem 3.3 are met by construction. Still, it is interesting to relate the performance of the Krylov approximation to the eigenstructure of the matrix A. From [15] it follows that any particular dominant eigenvalue is well represented in H for increasing rank m. Thus, the approximate SQRT-scheme has no stability problems. However, for eigencomponents contained in Y, the scheme reduces to the implicit or explicit Euler scheme with order one (see §3). Thus, it is important to ensure that no small eigenvalues are contained in Y, since this would produce large errors in the solution of the differential equation. For this reason we now consider estimates for the residual norm $\|(I - QQ^T)A\|_2$ under the assumption that there exists a gap between the absolute values of the largest m and the remaining eigenvalues. If this gap is sufficiently large, under

suitable regularity assumptions on the Krylov basis, then Theorem 5.1 shows that the criterion (3.10), (4.2) produces the correct rank m if the upper limit δ lies somewhere in this gap. The objective of this analysis is similar to that of Björck [4], however we do not use asymptotic estimates for the QR iteration, but concentrate on the initial situation after reduction to Hessenberg form.

To be explicit, we consider the matrix A to be in block Schur form,

$$A = \begin{pmatrix} L_1 & N \\ 0 & L_2 \end{pmatrix}, \tag{5.1}$$

with large eigenvalues contained in the upper $(m \times m)$-block L_1 and small ones in L_2. The gap between both sets of eigenvalues is assumed via the inequality $\|L_1^{-1}\|\|L_2\| < 1$. Then A is transformed to block diagonal form by the matrix

$$\begin{pmatrix} I & M \\ 0 & I \end{pmatrix}, \qquad L_1 M - M L_2 = N \tag{5.2}$$

[10]. It is shown in the next theorem that the residual of the Krylov approximation of the same rank m exceeds its minimum, $\|L_2\|$, by a multiple of a certain condition number κ_m. This number will be estimated in a later lemma, using the explicit eigenstructure of L_1. There is a slight convenience in assuming $z_1 = Az_0$. We denote the space of polynomials with degree k or less by π_k.

Theorem 5.1. *Let the matrix A be in block Schur form* (5.1) *with blocks $L_1 \in \mathbf{R}^{m\times m}$, $L_2 \in \mathbf{R}^{(n-m)\times(n-m)}$ satisfying*

$$\|L_1^{-1}\| \le \gamma_1^{-1}, \quad \|L_2\| \le \gamma_2, \quad \|L_1^{-1}N\| \le \nu, \tag{5.3}$$

and assume $\gamma_2/\gamma_1 < 1$. Let $Q \in \mathbf{R}^{nm}$ be constructed by (4.1) *from the Krylov vectors $z_k := A^k z_0$, $k := 1, \dots, m$, which may be partitioned in the same way as A, $z_k^{\mathrm{T}} = (x_k^{\mathrm{T}}, y_k^{\mathrm{T}})$. Assume that the minimal polynomial of the leading bock L_1 has full degree m, and that the vector $x_0^* := x_0 + My_0$, with M from* (5.2), *is nondegenerate, i.e., has a nonzero component in every eigenvector and principal vector of L_1. Then the number*

$$\kappa_m(A, z_0) := \max(\|p(L_2)y_0\|/\|p(L_1)x_0^*\| : p \in \pi_{m-1}) \tag{5.4}$$

is finite and, with constant $c^2 := 1 + \nu^2/(1 - \gamma_2/\gamma_1)^2$, the residual satisfies

$$\|(I - QQ^{\mathrm{T}})A\| \le \|L_2\|[1 + c^2\kappa_m(A, z_0)]. \tag{5.5}$$

Proof. By assumption, no polynomial $p \in \pi_{m-1}$ satisfies $p(L_1)x_0^* = 0$. Hence, κ_m in (5.4) exists. A standard argument for Krylov approximations is the identity $Q\mathbf{R}^m = K_m = \{p(A)z_1 : p \in \pi_{m-1}\}$. Now, for arbitrary $z \in \mathbf{R}^n$, $\|z\| = 1$, the matrix residual satisfies

$$\begin{aligned} \|(I - QQ^{\mathrm{T}})Az\|^2 &= \min\{\|Az - u\|^2 : u \in K_m\} \\ &= \min\{\|A[z - p(A)z_0]\|^2 : p \in \pi_{m-1}\}. \end{aligned} \tag{5.6}$$

By assumption (5.3) the matrix M exists and can be bounded by

$$\|M\| \le \nu/(1 - \gamma_2/\gamma_1), \tag{5.7}$$

since M is the fixed point of the contractive map $X \to L_1^{-1}(N + XL_2)$. Now it is easy to see that the following representations are valid:

$$y_k = L_2^k y_0, \quad x_k = L_1^k x_0^* - ML_2^k y_0, \qquad k = 1, \dots, m. \tag{5.8}$$

Thus, a linear combination $\sum a_k z_k = p(A)Az_0$, $p \in \pi_{m-1}$, of the Krylov vectors has the two components

$$\sum_{k=1}^{m} a_k y_k = L_2 p(L_2) y_0, \qquad \sum_{k=1}^{m} a_k x_k = L_1 p(L_1) x_0^* - M L_2 p(L_2) y_0.$$

Now we set $z^T = (x^T, y^T)$ and construct an upper bound for the minimum in (5.6) by the choice of a special polynomial. Let $q \in \pi_{m-1}$ satisfy $q(L_1)x_0^* = r := x + L_1^{-1}Ny$. This polynomial exists, since κ_m in (5.4) is finite. Since a polynomial is a linear expression of its coefficients, we have

$$\begin{aligned} &\{\|q(L_2)y_0\| : q(L_1)x_0^* = r\} \\ &\quad \le \max\{\|p(L_2)y_0\| : \|p(L_1)x_0^*\| = 1,\ p \in \pi_{m-1}\}\|r\| = \kappa_m \|r\|. \end{aligned} \tag{5.9}$$

For the polynomial q, the norm in (5.6) is

$$\begin{aligned} \|A[z - q(A)z_0]\|^2 &= \|ML_2 q(L_2)y_0\|^2 + \|L_2[y - q(L_2)y_0]\|^2 \\ &\le [\|L_2 y\| + (1 + \|M\|^2)^{1/2}\|L_2 q(L_2)y_0\|]^2 \\ &\le \|L_2\|^2(\|y\| + c\kappa_m\|r\|)^2, \end{aligned}$$

with the constant c from (5.5), by (5.7) and (5.9). The estimate is completed with $\|r\| \le (1 + \nu^2)^{1/2} \le c$, since $\|z\| = 1$. □

Remarks. (1) The term κ_m in (5.4) describes the conditioning of the Krylov basis with respect to a uniform approximation of the matrix A. It is important to note that this condition depends linearly on the ratio $\|y_0\|/\|x_0 + My_0\|$ of the two components of the initial vector z_0. Thus, for nondegenerate x_0 and $\|y_0\| \to 0$, $\|z_0\| = 1$, the Krylov process produces the exact block Schur form of A.

(2) The norm ratio $\|y_0\|/\|x_0^*\|$ can be improved in practice by performing preiterations, $z_0 := A^j z'/\|A^j z'\|$, $j > 0$. Every such iteration decreases the condition κ_m by a factor $\gamma_2/\gamma_1 < 1$, since, e.g.,

$$x_1^* = x_1 + My_1 = (L_1 x_0 + L_1 M y_0 - M L_2 y_0) + M L_2 y_0 = L_1 x_0^*$$

(see (5.2)). In fact, by using $z_1 = Az_0$ as the first Krylov vector, we already eliminated the impact of very small eigenvalues (e.g., zero). In the numerical implementation, this single preiteration will be used, at the least.

The following lemma contains an estimate of the global Krylov condition number κ_m under the simplifying assumption that the leading block L_1 has only simple eigenvalues. However, we note that κ_m may still be finite for multiple eigenvalues if their geometric multiplicity is one.

Lemma 5.2. *Under the assumptions of Theorem* 5.1, *let the matrix* L_1 *have the diagonal form* $L_1 = V^{-1}\operatorname{diag}(\lambda_j)V$ *with simple eigenvalues, and let* $\xi = (\xi_1, \dots, \xi_m)^T := Vx_0^*$ *be such that*

$$d := \min\left\{ |\xi_i| \prod_{j \ne i} |\lambda_i/\lambda_j - 1| : i = 1, \dots, m \right\} > 0. \tag{5.10}$$

Then

$$\kappa_m(A, z_0) \le d^{-1}\|y_0\|\sqrt{m}\|V\|(1 + \gamma_2/\gamma_1)^{m-1}.$$

Proof. Recalling (5.9), with $t \in \mathbf{R}^m$, $\|t\| = 1$, we consider the interpolation problem $t = p(L_1)x_0^* = V^{-1}\operatorname{diag}(p(\lambda_j))\xi$, $p \in \pi_{m-1}$, i.e.,

$$p(\lambda_i) = r_i/\xi_i, \qquad i = 1, \dots, m, \quad \tau = (\tau_1, \dots, \tau_m) := Vt.$$

Denoting by $b_i(x)$, $i = 1, \dots, m$, the Lagrangian basis polynomials associated with the points $\lambda_1, \dots, \lambda_m$, we get the following expression for the numerator in (5.4):

$$\begin{aligned}\|p(L_2)y_0\| &= \left\| \sum_{i=1}^{m} b_i(L_2)y_0\tau_i/\xi_i \right\| \\ &\le \max\{\|b_i(L_2)\|/|\xi_i| : i = 1, \dots, m\}\|y_0\|\|\tau\|\sqrt{m}.\end{aligned}$$

The explicit representation of the Lagrangian polynomials now leads to

$$\begin{aligned}\|b_i(L_2)\|/|\xi_i| &\le \prod_{j\neq i} \|L_2 - \lambda_j\| \Big/ \left(|\xi_i| \prod_{j\neq i} |\lambda_i - \lambda_j| \right) \\ &\le (\|L_2\|/|\lambda_{\min}| + 1)^{m-1}/d.\end{aligned}$$

The assertion finally follows from the observation $\|\tau\| \le \|V\|$, $\|t\| = 1$. □

The Krylov condition number introduced in Theorem 5.1 allows an easy description of several other properties of the Krylov approximation of A. One is the angle, resp. the distance, between the dominant invariant subspace of A and the Krylov space K_m.

Lemma 5.3. *Under the assumptions of Theorem* 5.1, *denote by* $\Omega_1 := E_1\mathbf{R}^m$, $E_1^{\mathrm{T}} := (I_m, 0)^{\mathrm{T}}$, *the dominant invariant subspace of the matrix* A. *Then the angle* ϑ *between* Ω_1 *and* K_m *satisfies*

$$\sin\vartheta = \operatorname{dist}(\Omega_1, K_m) = \|E_1E_1^{\mathrm{T}} - QQ^{\mathrm{T}}\| \quad \textit{and} \quad \cos\vartheta = \|(E_1^{\mathrm{T}}Q)^{-1}\|^{-1}$$

(*cf.* [10, §2.4]), *where*

$$\tan\vartheta \le \frac{\kappa_m(A, z_1)}{1 - c\kappa_m(A, z_1)}, \tag{5.11}$$

if $1 - c\kappa_m(A, z_1) > 0$, *with* $c := \nu/(1 - \gamma_2/\gamma_1)$. *Note that* $\kappa_m(A, z_1) \le (\gamma_2/\gamma_1)\kappa_m(A, z_0)$.

Proof. With $E_2 := (0, I_{n-m})^{\mathrm{T}}$, the matrix Q has the components $Q_j := E_j^{\mathrm{T}}Q$, $j = 1, 2$, where Q_1 is square and $\|Q_2\| = \sin\vartheta$ [10]. If $|\sin\vartheta| < 1$, by virtue of $I = Q^{\mathrm{T}}Q = Q_1^{\mathrm{T}}Q_1 + Q_2^{\mathrm{T}}Q_2$ there follows

$$\begin{aligned}\|Q_2Q_1^{-1}\|^2 &= \rho(Q_2(I - Q_2^{\mathrm{T}}Q_2)^{-1}Q_2^{\mathrm{T}}) \\ &= \rho((I - Q_2^{\mathrm{T}}Q_2)^{-1}Q_2^{\mathrm{T}}Q_2) = \frac{\sin^2\vartheta}{1 - \sin^2\vartheta},\end{aligned}$$

i.e., $\|Q_2Q_1^{-1}\| = \tan\vartheta$. Since Q is the unitary factor in the QR-decomposition of the Krylov matrix,

$$Z := (z_1, \dots, z_m)\begin{pmatrix} Z_1 - MZ_2 \\ Z_2 \end{pmatrix} = Q \cdot R, \qquad R \in \mathbf{R}^{mm},$$

where $Z_1 = (x_1^*, \dots, L_1^{m-1}x_1^*)$, $Z_2 = (y_1, \dots, L_2^{m-1}y_1)$ (cf. (5.8)), $x_1^* := x_1 + My_1$, we have

$$Q_2Q_1^{-1} = Z_2(Z_1 - MZ_2)^{-1} = Z_2Z_1^{-1}(I - MZ_2Z_1^{-1})^{-1}.$$

With the estimate (5.7), $\|M\| \le c$, the assertion follows from

$$\max\left\{\frac{\|Z_2Z_1^{-1}x\|}{\|x\|} : x \in \mathbf{R}^m\right\} = \max\left\{\frac{\|Z_2v\|}{\|Z_1v\|} : v \in \mathbf{R}^m\right\} = \kappa_m(A, z_1).$$

The relation between $\kappa_m(A, z_1)$ and $\kappa_m(A, z_0)$ has been discussed in Remark (2) after Theorem 5.1. □

The following corollary shows that the leading block H in (3.5) approximates L_1 (up to unitary transformations) and, hence, is nonsingular for small enough ϑ. This has the practical consequence that the sign computation needs fewer iterations for $\operatorname{sign}(H)$ than for $\operatorname{sign}(A)$, since $\operatorname{cond}(H) < \operatorname{cond}(A)$ (see §2).

Corollary 5.4. *Under the assumptions of Theorem* 5.1, *there exists a unitary* $(m \times m)$*-matrix* V *such that* $H = Q^TAQ$ *satisfies*

$$\|H - V^TL_1V\| \le \sin\vartheta(\|N\| + 2\|L_1\|\tan(\vartheta/2) + \|L_2\|\sin\vartheta).$$

The angle ϑ *can be estimated by* (5.11).

Proof. With the notation of Lemma 5.3 we may write $A = E_1(L_1E_1^T + NE_2^T) + E_2L_2E_2^T$. This leads to

$$\begin{aligned} Q^TAQ - V^TL_1V &= Q^TE_1L_1E_1^TQ - V^TL_1V + (Q^TE_1N + Q^TE_2L_2)E_2^TQ \\ &= (Q^TE_1 - V^T)L_1E_1^TQ + V^TL_1(E_1^TQ - V) \\ &\quad + (Q^TE_1N + Q^TE_2L_2)E_2^TQ. \end{aligned}$$

Finally, we use $\|Q^TE_2\| = \sin\vartheta$ and $\|Q^TE_1\| \le 1$, and choose V^T as the orthogonal factor in the polar decomposition of Q^TE_1. Then VQ^TE_1 is symmetric, nonnegative definite and

$$\|I - VQ^TE_1\| = 1 - \lambda_{\min}(VQ^TE_1) = 1 - \cos\vartheta = \sin\vartheta \cdot \tan(\vartheta/2)$$

(cf. Lemma 5.3). □

In Lemma 5.2 we saw, as was to be expected, that the accuracy of the Krylov approximation with fixed rank m depends crucially on the eigenvalue separation in the leading block L_1 and the presence of corresponding eigencomponents in the starting vector z_0. If dominant eigenvalues are clustered, or z_0 is degenerate with respect to one of these eigenvalues, the procedure of §4 will end with a rank (much) larger than m, even if $\gamma_2 \le \delta \le \gamma_1$. While the clustering of eigenvalues is beyond our scope, we briefly discuss the choice of z_0, resp. z_1.

A "good" starting vector z_1 should contain only small components corresponding to small eigenvalues, while dominant eigencomponents should be present with equal magnitude. In the original approach of Enright and Kamel [8] a global approximation of the matrix A was also constructed. An improvement was achieved by an a priori pivoting procedure, which permutes the dominant rows or columns of the matrix to the first position. In the Walker formulation (Algorithm 4.1) a similar effect is possible by the choice $z_1 = q_1 := e_k$,

where e_k is the unit vector with the index of the dominant row. Other a priori constructions of the first vector z_1 are obvious. The simplest examples are

$$z_1 := Ae_k, \quad \text{with } \|Ae_k\| = \max\{\|Ae_j\| : j = 1, \dots, n\}, \tag{5.12a}$$

$$e_i^T z_1 := \|e_i^T A\|, \quad i = 1, \dots, n. \tag{5.12b}$$

However, we have to recall that the vector z_1 must depend smoothly on the point x in the interval $[0, 1]$ (see the discussion before (3.6)). Evidently, this is not the case with (5.12a). The rule (5.12b) appears to be a better choice, especially if the l_2-norm is used. Unfortunately, it leads to a degenerate starting vector in some simple cases. For instance, at turning points, coefficient matrices often (see §6) contain a submatrix of the form $\left(\begin{smallmatrix} 0 & \lambda \\ \lambda & 0 \end{smallmatrix}\right)$, $|\lambda| \gg 1$, with the eigenvalue pair $\pm\lambda$. Here, (5.12b) produces the exact eigenvector corresponding to the eigenvalue $+\lambda$, which gives a degenerate Krylov space of dimension 1. On the other hand, (5.12a) gives $z_1 = \lambda e_1$ or $z_1 = \lambda e_2$ which, being sums, resp. differences, of the dominant eigenvectors, are optimal choices. Hence, we are not able to offer a satisfactory rule, or heuristics, for the construction of the starting vector z_1. In the experiments of §6 we used a constant vector z' having no obvious degenerations (e.g., constancy, zero components, symmetry), namely $(e_j, z') := j(n + 2 - j)$, $j = 1, \dots, n$, with one or two preiterations $z_1 := A^j z'$, $j = 1, 2$. Evidently, the question of a careful choice of the starting vector, resp. the number of preiterations, needs further investigation.

6. Numerical examples

Since this paper is concerned with the implementation of the SQRT-scheme for higher dimensions, we discuss three examples with dimensions $n = 16, 4$, and 8. A discussion of the computational expense of the scheme only makes sense under realistic circumstances, which means on a practical grid being coarse in those parts of the interval where the solution is smooth and concentrating points in boundary and interior layers. For this reason we rely on the procedure described in [18] for the mesh generation. The mesh is constructed adaptively from an initial, equidistant, grid by inserting additional points in all subintervals where a scaled estimate of the local error exceeds a prescribed bound. Stepsize ratios are restricted to negative powers of 2.

The first example contains no serious difficulties, but its dimension and explicit construction allow us to study the dependence of the rank of the Krylov approximation on the number of stiff eigenvalues in the coefficient matrix. The subsequent two examples are more difficult, since they have turning points.

Example 6.1. The Hadamard matrices V_n, $n = 2^k$, defined recursively by

$$V_{2m} = 2^{-1/2} \begin{pmatrix} V_m & V_m \\ V_m & -V_m \end{pmatrix}, \qquad V_1 = 1,$$

are symmetric and unitary. Then, $T(x) := \cosh(x)I + \sinh(x)V_n$, $x \in \mathbf{R}$, satisfies $T(x)^{-1} = T(-x)$ and $T'(x)T(x)^{-1} = T(x)^{-1}T'(x) = V_n$. Substitution of the new variable $u(x) := T(x)y(x)$ in the simple equation $y' = Dy + f$ with a real, constant, diagonal matrix D, yields the equation

$$u' = A(x)u + g, \qquad A(x) = T(x)DT(x)^{-1} + V_n, \quad g = T(x)f, \tag{6.1}$$

with explicit solution $u = T(x)\exp(xD)a + \cdots$, $a \in \mathbf{R}^n$. Since A is similar to the symmetric matrix $D + V_n$, the eigenvalues of A and D differ by 1 at most. We take $n = 16$ and let the entries of D depend on two parameters $r \in \{0, \ldots, n\}$ and $\varepsilon > 0$,

$$D = D(\varepsilon, r) = \operatorname{diag}(d_i),$$
$$d_i := \begin{cases} (2i-n)^3, & |2i-n| \le n-r, \\ (2i/n-1)^3/\varepsilon, & |2i-n| > n-r, \end{cases} \qquad i = 1, \ldots, n. \tag{6.2}$$

Thus, the largest kinematic eigenvalue is always $1/\varepsilon$ $(r > 0)$, but the number of stiff eigenvalues is only r, if r is odd. Boundary conditions and inhomogeneity are transformed from $u_i(0) = y_i(0) = 1$ $(i \le 8)$, $y_i(1) = 1$ $(i > 8)$, and $f_7 := d_7 \cdot \exp(-x)$, $f_i := 0$ $(i \ne 7)$. Thus, there is one boundary layer at $x = 1$ for $r = 1$ and two at both ends for $r > 1$. The results of runs with $\varepsilon = 10^{-6}$ and the sequence $r = 1, 3, \ldots, 15$ are shown in Table 1. In this example, two preiterations of the initial Krylov vector were necessary, and the sign computation needed seven iterations at most. The mean value of ranks on the final grid is denoted by $\overline{m}$; N_0 is the number of intervals where rank zero (i.e., the trapezoidal rule) was used. The final mesh is characterized by its size N, and the number NL of points in the layers, i.e., within 0.01 distance of $x = 0$ or $x = 1$. The error tolerance was 0.01, the exact errors lie below $0.93 \cdot 10^{-2}$. The numbers clearly show that the rank of the Krylov approximation is indeed essentially proportional to the number of stiff eigenvalues of the coefficient matrix A, and the number of intervals with rank zero plays only a minor role.

TABLE 1
Example 6.1, $\varepsilon = 10^{-6}$, *mean rank and mesh parameters*

r	1	3	5	7	9	11	13	15
$\overline{m}$	1.4	2.5	3.9	4.7	5.2	6.0	6.2	8.0
N_0	17	33	35	47	57	47	87	0
N	114	156	154	176	172	202	244	196
NL	43	90	92	102	129	161	205	167

Example 6.2. Dimension $n = 4$, $\varepsilon > 0$, interval $[-1, 1]$. The equations

$$\begin{aligned} &\varepsilon y'' + xy' - y = -(\varepsilon\pi^2 + 1)\cos \pi x - \pi x \sin \pi x, \\ &v'' - 2v' - 3v = y - xy' \end{aligned} \tag{6.3}$$

are transformed to the variables $u(x) = (y - xv, y' - xv', v, v')^{\mathrm{T}}$. Boundary conditions are $u_1(\pm 1) = u_3(\pm 1) = 0$, and the resulting coefficient matrix is

$$A(x) := \begin{pmatrix} 0 & 1 & -1 & 0 \\ 1/\varepsilon - x & -x/\varepsilon + x^2 & x/\varepsilon - x^2 - 3x & -x^2/\varepsilon - 1 - 2x + x^3 \\ 0 & 0 & 0 & 1 \\ 1 & -x & x+3 & 2 - x^2 \end{pmatrix}.$$

It has two large and two small eigenvalues. A turning point exists at $x = 0$, where the solution $y(x)$ shows a cusp and $y'(x)$ a jump. The explicit solution is known. Table 2 contains the results of runs, where the stiffness parameter ε varies between 10^{-2} and 10^{-9}. The notation is the same as in Table 1, where NL now denotes the number of meshpoints near $x = 0$. The error criterion, 10^{-2}, produced errors below $1.03 \cdot 10^{-2}$. Again, seven sign iterations were necessary, at most. The scheme was successful with one preiteration in the Krylov process, except in the case $\varepsilon = 10^{-8}$, where two were needed. The numbers in Table 2 show that the mean rank $\overline{m}$ increases very slowly with $1/\varepsilon$ and stays well below two, the number of stiff eigenvalues.

TABLE 2
Example 6.2, mean rank and mesh parameters

ε	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
$\overline{m}$	0.72	0.78	0.76	0.76	1.07	0.98	1.21	1.28
N_0	16	14	22	30	28	26	30	46
N	56	64	92	116	152	162	236	272
NL	1	1	9	30	50	50	72	94

Example 6.3. The Orr-Sommerfeld equations describe the temporal stability of plane parallel flow in a channel. It is a complex-valued eigenvalue problem for the wave velocity c and the factor φ in the stream function $\varphi(x) \cdot \exp(i\alpha(y - ct))$, $i^2 = -1$; for Poiseuille flow it reads

$$
\begin{aligned}
&M := -d^2/dx^2 + \alpha^2, \\
&M^2\varphi + i\alpha R\{(1 - x^2 - c)M\varphi - 2\varphi\} = 0, \\
&\varphi'(0) = 0, \quad \varphi'''(0) = 0, \quad \varphi(1) = 0, \quad \varphi'(1) = 0.
\end{aligned} \tag{6.4}
$$

By introduction of $\psi := M\varphi$, $\varphi := u_1 + iu_2$, $\psi := u_3 + iu_4$, $u_{j+4} := u_j'$, $j = 1, \ldots, 4$, (6.4) is reduced to a first-order eigenvalue problem of dimension 8. If an eigenvalue c is known, a BVP (1.1) may be obtained by replacing one of the homogeneous boundary conditions (e.g., $u_1'(0) = 0$) by a norming condition (e.g., $u_1(0) = 1$). We use the values $R = 10^6$, $\alpha = 1$, with the eigenvalue $c = 0.066592523 - 0.013983266 \cdot i$ from [13]. The coefficient matrix has two groups of eigenvalues with four-fold symmetry $\pm\lambda_j$, $\pm\bar{\lambda}_j$, $j = 1, 2$, which move near the diagonals $(1 \pm i)\mathbf{R}$, the smaller group coming very close to the imaginary axis. There is a layer near $x = 1$.

A run with an error criterion of one percent ended with a mesh of 390 intervals and an error estimate of 0.37 percent. The neglected boundary condition was almost satisfied with $u_5(0) \simeq -10^{-5}$. One preiteration was sufficient. The final mesh consisted of 55 subintervals on $[0, 0.9]$ with a Krylov approximation of rank 4. The refinement in this part was triggered by the device explained at the end of §2 to avoid almost imaginary eigenvalues. After that, the number of sign iterations was again seven. On [0.9,1], the rest of 335 intervals was

concentrated with Krylov rank 0, only 41 points within 10^{-2} of $x = 1$. This concentration might be due to the fact that the solution shows several oscillations in this interval, which may be difficult to approximate with a low-order scheme.

In this paper we did not only discuss the computational expense of the partitioned, approximate SQRT-scheme (3.6), but also its stability. In a second kind of test we investigate the stability and accuracy of the Krylov approximations. The validity of the heuristic argument following Lemma 3.2, which led to our rank criterion (3.10), is checked by computing the eigenvalues λ_i of the coefficient matrices in (3.6) with minimal real part, namely

$$RM := \frac{h}{2} \min\{\operatorname{Re} \lambda_i(Y + A), \operatorname{Re} \lambda_i(Y - A) : i = 1, \dots, n\}. \tag{6.5}$$

If RM is greater than -1 everywhere, all matrices of the approximation scheme are regular. However, it is also of great interest to see how far the eigenstructure of the coefficient matrices is deformed by the Krylov approximation. It is difficult to assess this deformation, but a comparison of the eigenvalues of the matrices $(Y + A)/2$, $(Y - A)/2$ and A might be an indication. Since $(A + Y)/2$ and $(A - Y)/2$ add up to A, this should also be the case for their spectra, if Y is a good approximation to X and (almost) commutes with A. This might be checked by the relative difference

$$DM := \max \left\{ \frac{|\lambda_i(A) - \frac{1}{2}[\lambda_{j_i}(A + Y) + \lambda_{k_i}(A - Y)]|}{1 + |\lambda_i(A)|} : i = 1, \dots, n \right\} \tag{6.6}$$

for suitable permutations (j_i), (k_i).

These eigenvalue bounds were computed for the turning point problems 6.2 and 6.3. But it is not very instructive to display detailed numbers, since there seems to be no regular behavior. We just note that for the runs of Table 2 the overall minimum for RM was -0.5082 on all meshes. This shows that the approximate scheme is stable. The eigenvalue difference had a smaller bound on the final grids, $DM \leq 0.207$, than on the intermediate meshes, where $DM = 0.402$ was reached. An eigenvalue perturbation of 20 percent could be an indication of unacceptable deformations of solutions; in the presence of an eigenvalue spread of 10^9, however, these numbers may not be too bad. Still, there is probably a need to develop a better estimation of the quality of Krylov approximations. The results for the Orr-Sommerfeld equation, Example 6.3, are better. Since a brute-force search for optimal permutations in (6.6) is very time-consuming, the numbers are given for the final grid only. Here, RM was never smaller than -0.11 and DM did not exceed 0.033.

Computations were performed on the IBM 4381 of the University Computing Center at Marburg in double precision.

Conclusion

We have shown that it may be possible to approximate the SQRT-scheme through a partitioned scheme, which preserves many of its original stability properties. The cost of this implementation is not determined by the norm of the coefficient matrix in the differential equation, but by the number of stiff components only. There are open questions concerning the choice of the Krylov starting vector or the number of preiterations. This question is closely related

to an estimation of the accuracy of Krylov matrices with respect to the eigenstructure of the original matrix.

BIBLIOGRAPHY

1. J. Albrecht, *Quadratisch konvergente Iterationsverfahren zur Berechnung von* $A^{1/2}$ *und* A^{-1}, Internat. Ser. Numer. Math., vol. 32, Birkhäuser, Basel-Stuttgart, 1976, pp. 9–15.
2. U. Ascher and R. Weiss, *Collocation for singular perturbation problems* II: *Linear first order systems without turning points*, Math. Comp. **43** (1984), 157–187.
3. I. Babuška and V. Majer, *The factorization method for the numerical solution of two point boundary value problems for linear ODE's*, SIAM J. Numer. Anal. **24** (1987), 1301–1334.
4. A. Björck, *A block QR algorithm for partitioning stiff differential systems*, BIT **23** (1983), 329–345.
5. A. Björck and S. Hammarling, *A Schur method for the square root of a matrix*, Linear Algebra Appl. **52–53** (1983), 127–140.
6. P. N. Brown and A. C. Hindmarsh, *Matrix-free methods for stiff systems of ODE's*, SIAM J. Numer. Anal. **23** (1986), 610–638.
7. L. Dieci, M. R. Osborne, and R. D. Russell, *A Riccati transformation method for solving linear BVPs.* II: *Computational aspects*, SIAM J. Numer. Anal. **25** (1988), 1074–1092.
8. W. H. Enright and M. S. Kamel, *Automatic partitioning of stiff systems and exploiting the resulting structure*, ACM Trans. Math. Software **4** (1979), 127–136.
9. C. W. Gear and Y. Saad, *Iterative solution of linear equations in ODE codes*, SIAM J. Sci. Statist. Comput. **4** (1983), 583–601.
10. G. H. Golub and C. F. van Loan, *Matrix computations*, North Oxford Academic, London, 1986.
11. H. O. Kreiss, N. K. Nichols, and D. L. Brown, *Numerical methods for stiff two-point boundary value problems*, SIAM J. Numer. Anal. **23** (1986), 325–368.
12. R. M. M. Mattheij and G. W. Staarink, *An efficient algorithm for solving general linear two-point BVPs*, SIAM J. Sci. Statist. Comput. **5** (1984), 745–763.
13. G. H. Meyer, *Continuous orthonormalization for boundary value problems*, J. Comput. Phys. **62** (1986), 248–262.
14. J. D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control **32** (1980), 677–687.
15. Y. Saad, *Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices*, Linear Algebra Appl. **34** (1980), 269–295.
16. ____, *Krylov subspace methods for solving large unsymmetric linear systems*, Math. Comp. **37** (1981), 105–126.
17. B. A. Schmitt, *An algebraic approximation for the matrix exponential in singularly perturbed boundary value problems*, SIAM J. Numer. Anal. **27** (1990), 51–66.
18. B. A. Schmitt and K. H. Schild, *Error estimation and mesh adaptation for an algebraic difference scheme in stiff BVPs*, Numerical Treatment of Differential Equations, NUMDIFF-5 (K. Strehmel, ed.), Teubner-Texte Math., Band 121, Teubner, Stuttgart, Leipzig, 1991.
19. B. T. Smith et al., *Matrix eigensystem routines—EISPACK Guide*, Lecture Notes in Comput. Sci., Springer, Berlin-New York, 1976.
20. J. M. Varah, *Alternate row and column elimination for solving linear systems*, SIAM J. Numer. Anal. **13** (1976), 71–75.
21. H. F. Walker, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Statist. Comput. **9** (1988), 152–163.
22. D. M. Young, *Iterative solution of large linear systems*, Academic Press, 1971.

FACHBEREICH MATHEMATIK, UNIVERSITÄT MARBURG, LAHNBERGE, D-3550 MARBURG, GERMANY

E-mail address: schmitt@dmrhrz11.bitnet