ANALYSIS OF SOME KRYLOV SUBSPACE APPROXIMATIONS TO THE MATRIX EXPONENTIAL OPERATOR

Y. SAAD *

Abstract. In this note we present a theoretical analysis of some Krylov subspace approximations to the matrix exponential operation $exp(A)v$ and establish a priori and a posteriori error estimates. Several such approximations are considered. The main idea of these techniques is to approximately project the exponential operator onto a small Krylov subspace and carry out the resulting small exponential matrix computation accurately. This general approach, which has been used with success in several applications, provides a systematic way of defining high order explicit-type schemes for solving systems of ordinary differential equations or time-dependent Partial Differential Equations.

1. Introduction. The problem of approximating the operation $exp(A)v$ for a given vector v and a matrix A is of considerable importance in many applications. For example, this basic operation is at the core of many methods for solving systems of ordinary differential equations (ODE's) or time-dependent partial differential equations (PDE's). Recently, the use of Krylov subspace techniques in this context has been actively investigated in the literature [2, 3, 4, 9, 10]. Friesner et al. [2] and Gallopoulos and Saad [3] introduced a few different ways of applying this approximation to the solution of systems of ordinary differential equations. The paper [3] presents some analysis on the quality of the Krylov approximation and on the ODE integration schemes derived from it. In this note we make the following contributions.

1. We introduce and justify a few new approximation schemes (Section 2);
2. We analyze the Krylov subspace approach from an approximation theory viewpoint. In particular we establish that the Krylov methods are equivalent to interpolating the exponential function on the associated Ritz values (Section 3);
3. We generalize some of the a priori error bounds proved in [3] (Section 4);
4. Finally, we present *a posteriori* error bounds (Section 5) that are essential for practical purposes.

The basic idea of the Krylov subspace techniques considered in this paper is to approximately project the exponential of the large matrix onto a small Krylov subspace. The only matrix exponential operation performed is therefore with a much smaller matrix. This technique has been successfully used in several applications, in spite of the lack of theory to justify it. For example, it is advocated in the paper by Park and Light [10] following the work by Nauts and Wyatt [7, 8]. Also, the idea of exploiting the Lanczos algorithm to evaluate terms of the exponential of Hamiltonian operators seems to have been first used in Chemical Physics, by Nauts and Wyatt [7]. The general Krylov subspace approach for nonsymmetric matrices was used in [4] for solving parabolic equations, and more recently, Friesner et al. [2] demonstrated that this technique can

* University of Minnesota, Computer Science department, EE/CS building, Minneapolis, MN 55455.
Work supported by the NAS Systems Division and/or DARPA via Cooperative Agreement NCC 2-387 between NASA and the University Space Research Association (USRA).

be extended to solving systems of stiff nonlinear differential equations. In the work by Nour-Omid [9] systems of ODEs are solved by first projecting them onto Krylov subspaces and then solving reduced tridiagonal systems of ODEs. This is equivalent to the method used in [10] which consists of projecting the exponential operator on the Krylov subspace. The evaluation of arbitrary functions of a matrix with Krylov subspaces has also been briefly mentioned by van der Vorst [13].

The purpose of this paper is to explore these techniques a little further both from a practical and a theoretivcal viewpoint. On the practical side we introduce new schemes that can be viewed as simple extensions or slight improvements of existing ones. We also provide a-posteriori error estimates which can be of help when developing integration schemes for ODE's. On the theoretical side, we prove some characterization results and a few additional a priori error bounds.

2. Krylov subspace approximations for $e^A v$. We are interested in approximations to the matrix exponential operation $exp(A)v$ of the form

$$e^A v \approx p_{m-1}(A)v, \tag{1}$$

where A is a matrix of dimension N, v an arbitrary nonzero vector, and p_{m-1} is a polynomial of degree $m-1$. Since this approximation is an element of the Krylov subspace

$$K_m \equiv span\{v, Av, \ldots, A^{m-1}v\},$$

the problem can be reformulated as that of finding an element of K_m that approximates $u = exp(A)v$. In the next subsections we consider three different possibilities for finding such approximations. The first two are based on the usual Arnoldi and nonsymmetric Lanczos algorithms respectively. Both reduce to the same technique when the matrix is symmetric. The third technique presented can be viewed as a corrected version of either of these two basic approaches.

2.1. Exponential propagation using the Arnoldi algorithm. In this section we give a brief description of the method presented in [3] which is based on Arnoldi's algorithm. The procedure starts by generating an orthogonal basis of the Krylov subspace with the well-known Arnoldi algorithm using $v_1 = v/\|v\|_2$ as an initial vector.

Algorithm: Arnoldi

1. *Initialize:* Compute $v_1 = v/\|v\|_2$.
2. *Iterate:* Do $j = 1, 2, ..., m$
   (a) Compute $w := Av_j$
   (b) Do $i = 1, 2, \ldots, j$
      i. Compute $h_{i,j} := (w, v_i)$
      ii. Compute $w := w - h_{i,j}v_i$
   (c) Compute $h_{j+1,j} := \|w\|_2$ and $v_{j+1} := w/h_{j+1,j}$.

Note that step 2-(b) is nothing but a modified Gram-Schmidt process. The above algorithm produces an orthonormal basis $V_m = [v_1, v_2, \ldots, v_m]$ of the Krylov subspace K_m. If we denote the $m \times m$ upper Hessenberg matrix consisting of the coefficients h_{ij} computed from the algorithm by H_m, we have the relation

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T, \tag{2}$$

from which we get $H_m = V_m^T A V_m$. Therefore H_m represents the projection of the linear transformation A onto the subspace K_m, with respect to the basis V_m. Based on this, the following approximation was introduced in [3]:

$$e^A v \approx \beta V_m e^{H_m} e_1. \tag{3}$$

Since $V_m^T(\tau A)V_m = \tau H_m$ and the Krylov subspaces associated with A and τA are identical we can also write

$$e^{\tau A} v \approx \beta V_m e^{\tau H_m} e_1, \tag{4}$$

for an arbitrary scalar τ. The practical evaluation of the vector $exp(\tau H_m)e_1$ is discussed in Section 2.4. Formula (4) approximates the action of the operator $exp(\tau A)$ which is sometimes referred to as the *exponential propagation operator*. It has been observed in several previous articles that even with relatively small values of m a remarkably accurate approximation can be obtained from (3); see for example [3, 2, 9].

2.2. Exponential propagation using the Lanczos algorithm. Another well-known algorithm for building a convenient basis of K_m is the well-known Lanczos algorithm. The algorithms starts with two vectors v_1 and w_1 and generates a bi-orthogonal basis of the subspaces $K_m(A, v_1)$ and $K_m(A^T, w_1)$.

Algorithm: Lanczos

1. *Start:* Compute $v_1 = v/\|v\|_2$ and select w_1 so that $(v_1, w_1) = 1$.
2. *Iterate:* For $j = 1, 2, \ldots, m$ do:
   - $\alpha_j := (Av_j, w_j)$
   - $\hat{v}_{j+1} := Av_j - \alpha_j v_j - \beta_j v_{j-1}$
   - $\hat{w}_{j+1} := A^T w_j - \alpha_j w_j - \delta_j w_{j-1}$
   - $\beta_{j+1} := \sqrt{|(\hat{v}_{j+1}, \hat{w}_{j+1})|}$, $\delta_{j+1} := \beta_{j+1}.sign[(\hat{v}_{j+1}, \hat{w}_{j+1})]$
   - $v_{j+1} := \hat{v}_{j+1}/\delta_{j+1}$
   - $w_{j+1} := \hat{w}_{j+1}/\beta_{j+1}$

If we set, as before, $V_m = [v_1, v_2, ..., v_m]$ and, similarly, $W_m = [w_1, w_2, ..., w_m]$ then

$$W_m^T V_m = V_m^T W_m = I \ , \tag{5}$$

where I is the identity matrix. Let us denote by T_m the tridiagonal matrix

$$T_m \equiv Tridiag[\delta_i, \alpha_i, \beta_{i+1}].$$

An analogue of the relation (2) is

$$AV_m = V_m T_m + \delta_{m+1} v_{m+1} e_m^T, \tag{6}$$

and an approximation similar to the one based on Arnoldi's method is given by

$$e^A v \approx \beta V_m e^{T_m} e_1, \tag{7}$$

in which as before $\beta = \|v\|_2$. The fact that V_m is no longer orthogonal may cause some nonnegligible numerical difficulties since the norm of V_m may, in some instances, be very large.

What can be said of one of these algorithms can often also be said of the other. In order to unify the presentation we will sometimes refer to (3) in which we use the same symbol H_m to represent either the Hessenberg matrix H_m in Arnoldi's method or the tridiagonal matrix T_m in the Lanczos algorithm. In summary, the features that distinguish the two approaches are the following:

- In the Arnoldi approximation V_m is orthogonal and H_m is upper Hessenberg;
- In the Lanczos approximation H_m is tridiagonal and V_m is not necessarily orthogonal;
- In the symmetric case V_m is orthogonal and H_m is tridiagonal and symmetric.

2.3. Corrected schemes . In this section we will use the same unified notation (3) to refer to either the Arnoldi based approximation (3) or the Lanczos based approximation (7). In the approximations described above, we generate $m+1$ vectors $v_1, ..., v_{m+1}$ but only the first m of them are needed to define the approximation. A natural question is whether or not one can use the extra vector v_{m+1} to obtain a slightly improved approximation at minimal extra cost. The answer is yes and the alternative scheme which we propose is obtained by making use of the function

$$\phi(z) = \frac{e^z - 1}{z}. \tag{8}$$

To approximate $e^A v$ we write

$$e^A v = v + A\phi(A)v$$

and we approximate $\phi(A)v$ by

$$\phi(A)v \approx V_m \phi(H_m)\beta e_1. \tag{9}$$

Denoting by s_m the error $s_m \equiv \phi(A)v_1 - V_m\phi(H_m)e_1$ and making use of the relation (2) we then have

$$\begin{aligned}
e^A v_1 &= v_1 + A\phi(A)v_1 \\
&= v_1 + A(V_m\phi(H_m)e_1 + s_m) \\
&= v_1 + (V_m H_m + h_{m+1,m} v_{m+1} e_m^T)\phi(H_m)e_1 + As_m \\
&= V_m[e_1 + H_m\phi(H_m)e_1] + h_{m+1,m} e_m^T \phi(H_m) e_1 v_{m+1} + As_m \\
&= V_m e^{H_m} e_1 + h_{m+1,m} e_m^T \phi(H_m) e_1 v_{m+1} + As_m. \qquad (10)
\end{aligned}$$

The relation (10) suggests using the approximation

$$e^A v \approx \beta \left[V_m e^{H_m} e_1 + h_{m+1,m} e_m^T \phi(H_m) e_1 \; v_{m+1} \right] \tag{11}$$

which can be viewed as a corrected version of (3). As is indicated by the expression (10), the error in the new approximation is $\beta A s_m$.

The next question we address concerns the practical implementation of these corrected schemes. The following proposition provides a straightforward solution.

PROPOSITION 2.1. *Define the* $(m+1) \times (m+1)$ *matrix*

$$\overline{H}_m \equiv \begin{pmatrix} H_m & 0 \\ c & 0 \end{pmatrix} \tag{12}$$

where c *is any row vector of length* m*. Then,*

$$e^{\overline{H}_m} = \begin{pmatrix} e^{H_m} & 0 \\ c\phi(H_m) & 1 \end{pmatrix}. \tag{13}$$

The proof of the proposition is an easy exercise which relies on the Taylor series expansion of the exponential function. Taking $c = h_{m+1,m} e_m^T$, it becomes clear from the proposition and (11) that the corrected scheme is mathematically equivalent to the approximation:

$$e^A v \approx \beta V_{m+1} e^{\overline{H}_m} e_1 \tag{14}$$

whose computational cost is very close to that of the original methods (3) or (7). In the remainder of the paper the matrix $\overline{H}_m$ is defined with $c = h_{m+1,m} e_m^T$.

Note that the approximation of $\phi(A)v$ is an important problem in its own right. It can be useful when solving a system of ordinary differential equations $\dot{w} = -Aw + r$, whose solution si given by $w(t) = w_0 + t\phi(-tA)(r - Aw_0)$. A transposed version of the above proposition provides a means for computing vectors of the form $\phi(H_m)c$ which can be exploited for evaluating the right hand side of (9). In addition, the ϕ function also plays an important role in deriving a-posterior error estimates, see Section 5.

2.4. Use of rational functions. In this section we briefly discuss some aspects related to the practical computation of $exp(H_m)e_1$ or $\phi(H_m)e_1$. Additional details may be found in [3]. Since m will, in general, be much smaller than N, the dimension of A, it is clear that any of the standard methods for computing $exp(H_m)e_1$ may be acceptable. However, as was discussed in [3], for large m the cost may become nonnegligible if an algorithm of order m^3 is used especially when the computations are carried out on a supercomputer. In addition, the discussion in this section is essential for the development of the a posteriori error estimates described in Section 5.

Consider the problem of evaluating $exp(H)b$ where H is a small, possibly dense, matrix. One way in which the exponential of a matrix is evaluated is by using rational approximation to the exponential function [3, 4]. Using a rational approximation $R(z)$,

of the type (p, p), i.e., such that the denominator has the same degree p as the numerator, it is more viable and economical to use the partial fraction expansion of R, see [3],

$$R(z) = \alpha_0 + \sum_{i=1}^{p} \frac{\alpha_i}{z - \theta_i}, \tag{15}$$

where the θ_i's are the poles of R. Expressions for the α_i's and the θ_i's for the Chebyshev approximation to e^{-z} on the positive real line are tabulated once and for all. The approximation to $y = exp(H)b$ for some vector b can be evaluated via

$$y = \alpha_0 b + \sum_{i=1}^{p} \alpha_i (H - \theta_i I)^{-1} b \ , \tag{16}$$

which normally requires p factorizations. In fact since the roots θ_i seem to come generally in complex conjugate pairs we only need about half as many factorizations to be carried out in complex arithmetic. The rational functions that are used in [3] as well as in the numerical experiments section in this paper are based on the Chebyshev approximation on $[0, \infty)$ of the function e^{-x}, see [1] and the references therein. This approximation is remarkably accurate and, in general, a small degree approximation, as small as $p = 14$, suffices to obtain a good working accuracy. We also observe that the Hessenberg or tridiagonal form are very desirable in this context: assuming the matrix H is of size m it costs $O(pm^2)$ to compute $exp(H)e_1$ when H is Hessenberg and only $O(pm)$ when it is tridiagonal. This is to be compared with a cost of $O(m^3)$ for a method that would compute $exp(H_m)$ based on the Schur decomposition.

Concerning the computation of vectors of the form $\phi(A)b$ it is easily seen that if the rational approximation $R(z)$ to e^z is exact at $z = 0$, i.e., if $R(0) = 1$, then we can derive from the definition (8) the following rational approximation to $\phi(z)$:

$$R_\phi(z) = \sum_{i=1}^{p} \frac{\alpha_i}{\theta_i(z - \theta_i)} \ . \tag{17}$$

As a result, the same tables for the coefficients α_i and the roots θ_i can be used to compute either $exp(H)b$ or $\phi(H)b$. This is exploited in Section 5.

3. Characterization and Exactness. Throughout this section we will refer to the eigenvalues of the matrix H_m (resp. T_m) as the Ritz values [1]. We will denote by $\sigma(X)$ the spectrum of a given square matrix X.

There are two theoretical questions we would like to address. First, we would like to examine the relation between the methods described earlier and the problem of approximating the exponential function. As will be seen, these methods are mathematically equivalent to interpolating the exponential function over the Ritz values. Second, we would like to determine when the Krylov approximation (3) becomes exact.

[1] This constitutes a slight abuse of the terminology since this term is generally used in the Hermitian case only.

3.1. Characterization. The following lemma is fundamental in proving a number of results in this paper.

LEMMA 3.1. *Let A be any matrix and V_m, H_m the results of m steps of the Arnoldi or Lanczos method applied to A. Then for any polynomial p_j of degree $j \leq m-1$ the following equality holds*

$$p_j(A)v_1 = V_m p_j(H_m)e_1 \tag{18}$$

Proof. Consider the Arnoldi case first. Let $\pi_m = V_m V_m^T$ be the orthogonal projector onto K_m as represented in the original basis. We will prove by induction that $A^j v_1 = V_m H_m^j e_1$, for $j = 0, 1, 2, ..., m-1$. The result is clearly true for $j = 0$. Assume that it is true for some j with $j \leq m-2$. Since the vectors $A^{j+1}v_1$ and $A^j v_1$ belong to K_m we have

$$A^{j+1}v_1 = \pi_m A^{j+1}v_1 = \pi_m A A^j v_1 = \pi_m A \pi_m A^j v_1 \ .$$

The relation (2) yields $\pi_m A \pi_m = V_m H_m V_m^T$. Using the induction hypothesis we get,

$$A^{j+1}v_1 = V_m H_m V_m^T V_m H_m^j e_1 = V_m H_m^{j+1} e_1.$$

This proves the result for the Arnoldi case. The proof for the Lanczos approximation (7) is identical except that π_m must be replaced by the oblique projector $V_m W_m^T$. □

Let now q_ν be the minimal polynomial of an arbitrary matrix A, where ν is its degree. We know that any power of the matrix A can be expressed in terms of a polynomial in A, of degree not exceeding $\nu - 1$. An immediate consequence is that if $f(z)$ is an entire function then $f(A) = p_{\nu-1}(A)$ for a certain polynomial of degree at most $\nu - 1$. The next lemma determines this polynomial. Recall that a polynomial p interpolates a function f in the Hermite sense at a given point x repeated k times if f and p as well as their $k-1$ first derivatives agree at x.

LEMMA 3.2. *Let A be any matrix whose minimal polynomial is of degree ν and $f(z)$ a function in the complex plane which is analytic in an open set containing the spectrum of A. Moreover, let $p_{\nu-1}$ be the interpolating polynomial of the function $f(z)$, in the Hermite sense, at the roots of the minimal polynomial of A, repeated according to their multiplicities. Then:*

$$f(A) = p_{\nu-1}(A) \ . \tag{19}$$

See [5] for a proof.

Consider now a Hessenberg matrix H. It is known that whenever $h_{j+1,j} \neq 0$, $j = 1, 2, ..., m-1$, the geometric multiplicity of each eigenvalue is one, i.e., the minimal polynomial of H is simply its characteristic polynomial. Therefore, going back to the Hessenberg matrix provided by the Arnoldi process, we can state that

$$e^{H_m} = p_{m-1}(H_m). \tag{20}$$

where the polynomial p_{m-1} interpolates the exponential function on the spectrum of H_m in the Hermite sense. We are now ready to state our main characterization theorem.

THEOREM 3.3. *The approximations (3) and (7) are mathematically equivalent to approximating $exp(A)v$ by $p_{m-1}(A)v$, where p_{m-1} is the (unique) polynomial of degree $m-1$ which interpolates the exponential function in the Hermite sense on the set of Ritz values repeated according to their multiplicities.*

Proof. In both the Arnoldi and Lanczos cases, the approximation to $exp(A)v$ is defined by $\beta V_m e^{H_m} e_1$, where H_m must be replaced by the tridiagonal matrix T_m in the Lanczos case. Using the previous lemma, we have

$$\beta V_m e^{H_m} e_1 = \beta V_m p_{m-1}(H_m) e_1$$

where p_{m-1} is the polynomial defined in the theorem. Using lemma 3.1 this becomes,

$$\beta V_m e^{H_m} e_1 = \beta p_{m-1}(A) v_1 = p_{m-1}(A) v \ .$$

□

As is well-known Krylov subspace methods tend to provide better approximations for the eigenvalues located in the outermost part of the spectrum than for those located in the interior. Therefore, from the above result, one can expect that the components of the approximations (3) in the eigenvectors associated with the outermost part of the spectrum will be more accurate.

We now state an analogous result concerning the corrected schemes of Section (2.3).

THEOREM 3.4. *The corrected approximation (14) used with either the Arnoldi or Lanczos method is mathematically equivalent to approximating $exp(A)v$ by $p_m(A)v$, in which p_m is the (unique) polynomial of degree m which interpolates the exponential function, in the Hermite sense, on the set*

$$\sigma(H_m) \bigcup \{0\}$$

in which $\sigma(H_m)$ is the set of Ritz values with each eigenvalue repeated according to its multiplicity.

Proof. By a proof analogous to that of the previous theorem we can easily see that the corrected scheme consists of approximating $exp(A)$ by $I + Ap_{m-1}(A)$ where $p_{m-1}(z)$ interpolates the function $\phi(z)$ on $\sigma(H_m)$ in the Hermite sense. We notice that $p_m(z) \equiv 1 + zp_{m-1}(z)$ is such that

$$\begin{aligned} p_m(0) &= 1 \\ p_m(\tilde{\lambda}_i) &= 1 + \tilde{\lambda}_i p_{m-1}(\tilde{\lambda}_i) = 1 + \tilde{\lambda}_i \phi(\tilde{\lambda}_i) = e^{\tilde{\lambda}_i} \ , \ \ \forall \ \tilde{\lambda}_i \ \in \sigma(H_m). \end{aligned}$$

Let ν_i be the multiplicity of the Ritz value $\tilde{\lambda}_i$. Then, in addition, we have

$$p_m^{(k)}(\tilde{\lambda}_i) = e^{\tilde{\lambda}_i} \ , \ \ \forall \ \tilde{\lambda}_i \ \in \sigma(H_m) \ , \ k = 1, 2, ..., \nu_i - 1$$

as can be seen by comparing the successive derivatives of $p_m(z) \equiv 1 + zp_{m-1}(z)$ and $e^z = 1 + z\phi(z)$ and the fact that $\phi^{(k)}(\tilde{\lambda}_i) = p_{m-1}^{(k)}(\tilde{\lambda}_i)$ for $k \leq \nu_i - 1$. This implies that $p_m(z)$ interpolates e^z on $\sigma(H_m) \bigcup \{0\}$, in the Hermite sense. □

The above result justifies the term "corrected schemes" used for the techniques derived in Section 2.3, since these correspond to adding one additional constraint to the approximations. They can therefore be expected to be slightly more accurate than the parent schemes while their costs are essentially identical.

3.2. Exactness. By analogy with conjugate gradient algorithms, we might wonder if the degree of the minimal polynomial of v with respect to A is also an upper bound on the dimension m needed to get the exact solution by the formula (3). As will be shown the answer is yes. In this section we consider only the Arnoldi based method of Section 2.1.

Consider the case where at some step m we have $h_{m+1,m} = 0$ in the Arnoldi process. In this situation, the algorithm stops. This is referred to as a 'lucky breakdown' because (2) simplifies into $AV_m = V_m H_m$, which implies that K_m is an invariant subspace. Moreover, it is clear that K_m will be the first of the sequence of Krylov subspaces K_i for which this happens; otherwise the algorithm would have stopped at an earlier step. As a result of the above relation we will have $A^k V_m = V_m H_m^k$ for all k and the Taylor series expansion of the exponential function shows that the approximation (3) is exact. Clearly, (4) is also exact for any τ in this situation. As is well-known, this 'lucky breakdown' occurs if and only if the minimum degree of v is equal to m and we have therefore proved the following result.

PROPOSITION 3.5. *When the minimal polynomial of v is of degree equal to m, the Krylov approximation (4) is exact for all τ.* In particular, since the minimum degree of v cannot exceed N, the algorithm will deliver the exact answer for $m \leq N$. The above exactness condition is also sufficient, as is shown next.

THEOREM 3.6. *The following three conditions are equivalent:*

(i) $e^{\tau A} v = \beta V_m e^{\tau H_m} e_1 \ , \quad \forall \tau$;

(ii) The Krylov subspace K_m is invariant under A, and no subspace K_i with $i < m$ is invariant;

(iii) The minimal polynomial of v is of degree equal to m.

Proof. That (iii) is equivalent to (ii) is a simple and well-known result on Krylov subspaces and was essentially proved above, see also [11]. In addition, the previous proposition shows that (ii) $\rightarrow$ (i).

It remains to show that (i) $\rightarrow$ (ii). Taking the derivative of order k (with respect to τ) of the condition in (i) we get,

$$A^k e^{\tau A} v = \beta V_m H_m^k e^{\tau H_m} e_1 \ , \quad \forall \ \tau.$$

In particular, taking $\tau = 0$ leads to the relation

$$A^k v = \beta V_m H_m^k e_1 \ ,$$

which establishes that $A^k v$ is in K_m for all k. This proves in particular that any vector in AK_m which is a linear combination of the vectors $A^k v, k = 1, \ldots, m$ is in K_m. As a result K_m is invariant under A. Moreover, we cannot have a K_i invariant under A with $i<m$; otherwise the dimension of K_m would be less than m, which would be a contradiction since V_m is an orthogonal basis of K_m. □

We note that the characteristic properties shown in the previous section could also have been used to prove the above theorem.

4. General a priori error bounds. In this section we establish general a priori error bounds for the approximations defined in Section 2. Related results have been shown in [3] for the Arnoldi-based algorithm. Here we aim at showing general error bounds for all the algorithms presented earlier. Some analysis on their optimality is also provided.

4.1. Preliminary Lemmas. The following lemma shows how to systematically exploit polynomial approximations to e^x, in order to establish a priori error bounds. They generalize a result shown in [3].

LEMMA 4.1. *Let A be an arbitrary matrix and V_m, H_m the results of m steps of the Arnoldi or the Lanczos process. Let $f(z)$ be any function such that $f(A)$ and $f(H_m)$ are defined. Let p_{m-1} be any polynomial of degree $\leq m-1$ approximating $f(z)$, and define the remainder $r_m(z) \equiv e^z - p_{m-1}(z)$. Then,*

$$f(A)v - \beta V_m f(H_m)e_1 = \beta\left[r_m(A)v_1 - V_m r_m(H_m)e_1\right] \tag{21}$$

Proof. As a result of the relation $f(z) = p_{m-1}(z) + r_m(z)$ we have

$$f(A)v_1 = p_{m-1}(A)v_1 + r_m(A)v_1. \tag{22}$$

Lemma 3.1 implies that

$$p_{m-1}(A)v_1 = V_m p_{m-1}(H_m)e_1. \tag{23}$$

Similarly to (22) we can write for H_m

$$p_{m-1}(H_m)e_1 = f(H_m)e_1 - r_m(H_m)e_1. \tag{24}$$

To complete the proof, we substitute (24) in (23) and the resulting equation in (22) to get, after multiplying through by β

$$f(A)v = \beta V_m f(H_m)e_1 + \beta[r_m(A)v_1 - V_m r_m(H_m)e_1].$$

□

Note that for Arnoldi's method the orthogonality of V_m immediately yields

$$\|f(A)v - \beta V_m f(H_m)e_1\|_2 \leq \beta(\|r_m(A)\|_2 + \|r_m(H_m)\|_2) \tag{25}$$

whereas for the Lanczos process we can only say that

$$\|f(A)v - \beta V_m f(H_m)e_1\|_2 \leq \beta(\|r_m(A)\|_2 + \|V_m\|_2\|r_m(T_m)\|_2). \tag{26}$$

The following lemma will be useful in proving the next results.

LEMMA 4.2. *Let s_{m-1} be the polynomial of degree $m-1$ obtained from the partial Taylor series expansion of e^x, i.e.,*

$$s_{m-1}(x) = \sum_{j=0}^{m-1} \frac{x^j}{j!},$$

and let $t_m(x) = e^x - s_{m-1}(x)$ be the remainder. Then, for any nonnegative real number x and any nonnegative integer m:

$$t_m(x) = \frac{x^m}{m!}(1 + o(1)) \leq \frac{x^m e^x}{m!} \tag{27}$$

Proof. The integral form of the remainder in the Taylor series for e^x is

$$t_m(x) = \frac{x^m}{(m-1)!} \int_0^1 e^{(1-\tau)x} \tau^{m-1} d\tau \tag{28}$$

Following the argument in the proof of Theorem 1 in [12], we notice that as m increases the term inside the integrand becomes an increasingly narrow spike centered around $\tau = 1$. As a result the function $e^{(1-\tau)x}$ in that narrow interval can be assimilated to the constant function whose value is one and we can write

$$\int_0^1 e^{(1-\tau)x} \tau^{m-1} d\tau = (1 + o(1)) \int_0^1 \tau^{m-1} d\tau = \frac{1}{m}(1 + o(1))$$

which proves the first part of (27). The inequality in the right-hand-side of (27) is obtained by using the inequality $e^{(1-\tau)x} \leq e^x$ in the integral (28). □

4.2. A priori error bounds for the basic schemes. We will now prove the following theorem.

THEOREM 4.3. *Let A be any matrix and let $\rho = \|A\|_2$. Then the error of the approximation (3) obtained using Arnoldi's method is such that*

$$\|e^A v - \beta V_m e^{H_m} e_1\|_2 \leq 2\beta t_m(\rho) \leq 2\beta \frac{\rho^m e^\rho}{m!}. \tag{29}$$

Similarly, the error of the approximation (7) obtained using the Lanczos method is such that

$$\|e^A v - \beta V_m e^{T_m} e_1\|_2 \leq \beta[t_m(\rho) + \|V_m\|_2 t_m(\tilde{\rho})] \leq \left(\rho^m e^\rho + \|V_m\|_2 \tilde{\rho}^m e^{\tilde{\rho}}\right) \tag{30}$$

in which $\tilde{\rho} = \|T_m\|_2$.

Proof. The proof of the theorem makes use of the particular polynomial $s_m(z)$ where s_m is defined in Lemma 4.2. For this polynomial, we have,

$$\|r_m(A)v_1\|_2 = \|\sum_{j=m}^{\infty} \frac{1}{j!} A^j v_1\|_2 \leq \sum_{j=m}^{\infty} \frac{1}{j!} \|A^j v_1\|_2 \leq \sum_{j=m}^{\infty} \frac{\rho^j}{j!} = t_m(\rho). \tag{31}$$

We consider first the Arnoldi case. We can prove that similarly to (31) we have

$$\|r_m(H_m)e_1\|_2 \leq t_m(\tilde{\rho}) \ , \tag{32}$$

where $\tilde{\rho} = \|H_m\|_2 = \|V_m^T A V_m\|_2$. Observing (e.g. from (28)) that t_m is an increasing function of x for x positive, and that $\tilde{\rho} \leq \rho$, we obtain

$$t_m(\tilde{\rho}) \leq t_m(\rho). \tag{33}$$

Substituting (33), (32) and (31) in inequality (25) yields the first part of the result (29). The second part of (29) is a consequence of the inequality in Lemma 4.2.

For the Lanczos case, (32) becomes

$$\|r_m(T_m)e_1\|_2 \leq t_m(\tilde{\rho})$$

where $\tilde{\rho} = \|T_m\|_2 = \|W_m^T A V_m\|_2$. Unfortunately, in this case we cannot relate $\tilde{\rho}$ and ρ. Moreover, the norm of V_m is no longer equal to one. The result (30) follows immediately from (26), (31) and the above inequality. □

We would like now to briefly discuss an asymptotic analysis in an attempt to determine how sharp the result may be in general. First, for m large enough in the above proof, the norm $\|A\|_2$ can be replaced by the spectral radius of A. This is because $\|A^j\|_2$ is asymptotically equivalent to $|\lambda_{max}|^j$, where $|\lambda_{max}|$ is the spectral radius. Secondly, as is indicated by Lemma 4.2, $t_m(\rho)$ is in fact of the order of $\rho^m/m!$. In reality, we therefore have

$$\|Error\|_2 \leq 2\beta \frac{|\lambda_{max}|^m}{m!}(1 + o(1))$$

Moreover, as is indicated in [12], the polynomial s_{m-1} will be (asymptotically) near the polynomial that best approximates the exponential function on a circle centered at the origin. Therefore, if the eigenvalues of A are distributed in a circle of radius $|\lambda_{max}|$ one can expect the above inequality to be asymptotically optimal.

In fact we now address the case where the matrix A has a spectrum that is enclosed in a ball of center α located far away from the origin. In such a case, the polynomial used in the above proof is inappropriate, in that it is likely to be far from optimal, and the following lemma provides a better polynomial. In the following $D(\alpha, \rho)$ denotes the closed disk of the complex plane centered at α and with radius ρ.

LEMMA 4.4. *Let s_{m-1} be the polynomial defined in Lemma 4.2 and let for any real α,*

$$c_{m-1,\alpha}(z) = e^{\alpha} s_{m-1}(z - \alpha).$$

Then

$$\max_{z \in D(\alpha,\rho)} |e^z - c_{m-1,\alpha}(z)| \leq e^{\alpha} t_m(\rho) \leq \frac{\rho^m e^{\rho+\alpha}}{m!}. \tag{34}$$

Proof. From Lemma 4.2 for any z with $|z-\alpha| \leq \rho$ we have

$$|e^{z-\alpha} - s_{m-1}(z-\alpha)| \leq \sum_{j=m}^{\infty} \frac{\rho^j}{j!} = t_m(\rho) \leq \frac{\rho^m e^{\rho}}{m!} \tag{35}$$

Multiplying both sides by e^{α} yields (34). ∎

As an application, we can prove the following theorem.

THEOREM 4.5. *Let A be any matrix and $\rho_\alpha = \|A - \alpha I\|_2$ where α is any real scalar. Then the error of the approximation (3) satisfies*

$$\|e^A v - \beta V_m e^{H_m} e_1\|_2 \leq 2\beta e^{\alpha} t_m(\rho_\alpha) \leq 2\beta \frac{\rho_\alpha^m e^{\rho_\alpha + \alpha}}{m!}. \tag{36}$$

Proof. The proof is identical with that of Theorem (4.3) except that it uses the polynomial $c_{m-1,\alpha}$ defined in the previous lemma. ∎

Ideally, we would like to utilize the best α possible in order to minimize the right-hand side of the inequality (36). However, this seems to be a rather difficult minimization problem to solve. The important factor in the inequality of the theorem is the term ρ_α, which may be made much smaller than ρ, the norm of A, by a proper choice of α. As a result we can nearly minimize the right-hand side of (36) by minimizing ρ_α with respect to α, since the term e^{ρ_α} is asymptotically unimportant. According to (36) and to the discussion prior to the theorem, we can say that in the case where the eigenvalues of A are distributed in a disk of center α, and radius r, the error will behave asymptotically like $e^{\alpha} r^m/m!$, and this is likely to be an asymptotically sharp result. In practice the interesting cases are those where the spectrum is contained in the left half of the complex plane. An important particular case in some applications is when the matrix A is symmetric negative definite.

COROLLARY 4.6. *Let A be a symmetric definite negative matrix and let $\rho = \|A\|_2$. Then the error of the approximation (3) satisfies*

$$\|e^A v - \beta V_m e^{H_m} e_1\|_2 \leq \beta \frac{\rho^m}{m! 2^{m-1}}. \tag{37}$$

Proof. We apply the previous theorem with $\alpha = -\|A\|_2/2$ and use the fact that the 2-norm of a Hermitian matrices is equal to its spectral radius. Since A is symmetric definite negative, we have that $\rho_\alpha = |\rho/2| = -\alpha$. The proof follows immediately from (36). ∎

In [3] it was shown that the term e^{ρ} in the right-hand side of (29) can be replaced by the term $\max(1, e^{\eta})$ where $\eta = \max\{\lambda_i(A + A^T)/2)\}$ is the 'logarithmic norm' of A.

4.3. A priori error bounds for the corrected schemes. The corrected schemes are based on approximations using polynomials of the form $p(z) = 1 + zp_{m-1}(z)$, i.e., polynomials of degree m satisfying the constraint $p(0) = 1$. In this section we prove results similar to those of the Section 4.2, for the Arnoldi-based corrected scheme. We

do not consider the Lanczos-based approximation for which similar results can also be proved.

THEOREM 4.7. *Let A be any matrix and let $\rho = \|A\|_2$. Then the error of the corrected approximation (14) obtained using Arnoldi's method satisfies*

$$\|e^A v - \beta V_{m+1} e^{\overline{H}_m} e_1\|_2 \leq 2\beta t_{m+1}(\rho) \leq 2\beta \frac{\rho^{m+1} e^\rho}{(m+1)!}. \tag{38}$$

Proof. From (10) we see that the error in the corrected approximation satisfies

$$e^A v - \beta V_{m+1} e^{\overline{H}_m} e_1 = A s_m \equiv A(\phi(A)v - V_m \phi(H_m) e_1).$$

By Lemma 4.1 for any polynomial p_{m-1} of degree $m-1$ we have

$$s_m \equiv \phi(A)v - \beta V_m \phi(A) e_1 = \beta[r_m(A)v_1 - V_m r_m(H_m) e_1] \tag{39}$$

in which $r_m(z) \equiv \phi(z) - p_{m-1}(z)$. We now take the particular polynomial

$$p_{m-1}(z) \equiv \frac{s_m(z) - 1}{z} = \sum_{j=1}^{m-1} \frac{z^j}{(j+1)!}$$

which is such that

$$r_m(z) = \phi(z) - p_{m-1}(z) = \frac{e^z - 1}{z} - \frac{s_m(z) - 1}{z} = \frac{t_{m+1}(z)}{z} \tag{40}$$

where $t_{m+1}(z)$ is defined as in Lemma 4.2. Then, continuing as in the proof of Theorem 4.3 and using the result of Lemma 4.2 we get

$$\|r_m(A)v_1\|_2 \leq \sum_{j=m}^{\infty} \frac{\rho^j}{(j+1)!} = \frac{t_{m+1}(\rho)}{\rho}. \tag{41}$$

Similarly, we have for H_m

$$\|r_m(H_m)e_1\|_2 \leq \frac{t_{m+1}(\tilde{\rho})}{\tilde{\rho}},$$

where $\tilde{\rho}$ is the norm of H_m. Since once again the function $t_{m+1}(x)/x$ is an increasing function for x real positive we also have, as in the proof of Theorem 4.3,

$$\frac{t_{m+1}(\tilde{\rho})}{\tilde{\rho}} \leq \frac{t_{m+1}(\rho)}{\rho}.$$

Multiplying (39) by A and taking the 2-norms we get

$$\|A s_m\|_2 \leq \|A(\phi(A)v - \beta V_m \phi(A) e_1)\|_2 \leq \rho\beta \left[\|r_m(A)v_1\| + \|r_m(H_m)e_1\|\right],$$

which, upon combination with the above inequalities, yields the desired result. □

An interesting observation is that the upper bound for the m-th approximation in the corrected scheme (i.e., the one derived from m steps of Arnoldi's method) is identical with that obtained for the the $(m+1)$-st approximation of the basic scheme. There are also analogues of Theorem 4.2 and Corollary 4.1 in which, similarly, m is replaced by $m+1$.

5. A posteriori error estimates. The error bounds established in the previous section may be computed, provided we can estimate the norm of A or $A-\alpha I$. However, the fact that they may not be sharp leads us to seek alternatives. A posteriori error estimates are crucial in practical computations. It is important to be able to assess the quality of the result obtained by one of the techniques described in this paper in order to determine whether or not this result is acceptable. If the result is not acceptable then one can still utilize the Krylov subspace computed to evaluate, for example, $exp(A\delta)v$ where δ is chosen small enough to yield a desired error level. We would like to obtain error estimates that are at the same time inexpensive to compute and sharp. Without loss of generality the results established in this section are often stated for the scaled vector v_1 rather than for the original vector v.

5.1. Expansion of the error. We seek an expansion formula for the error in terms of $A^k v_{m+1}, k = 0, 1, ..., \infty$. To this end we need to define functions similar to the ϕ function used in earlier sections. We define by induction,

$$\begin{aligned} \phi_0(z) &= e^z \\ \phi_{k+1}(z) &= \frac{\phi_k(z) - \phi_k(0)}{z} \ , \quad k \geq 0 \end{aligned}$$

We note that $\phi_{k+1}(0)$ is defined by continuity. Thus, the function ϕ_1 is identical with the function ϕ used in Section 2. It is clear that these functions are well-defined and analytic for all k. We can now prove the following theorem. Recall that we are using a unified notation: H_m may in fact mean T_m in the case where Lanczos is used instead of the Arnoldi algorithm.

THEOREM 5.1. *The error produced by the Arnoldi or Lanczos approximation (3) satisfies the following expansion:*

$$e^A v_1 - V_m e^{H_m} e_1 = h_{m+1,m} \sum_{k=1}^{\infty} e_m^T \phi_k(H_m) e_1 A^{k-1} v_{m+1} \ , \tag{42}$$

while the corrected version (14) satisfies,

$$e^A v_1 - V_{m+1} e^{\overline{H}_m} e_1 = h_{m+1,m} \sum_{k=2}^{\infty} e_m^T \phi_k(H_m) e_1 A^{k-1} v_{m+1} \ . \tag{43}$$

Proof. The starting point is the relation (10)

$$e^A v_1 = V_m e^{H_m} e_1 + h_{m+1,m} e_m^T \phi(H_m) e_1 v_{m+1} + A s_m^1 \tag{44}$$

in which we now define, more generally,

$$s_m^j = \phi_j(A) v_1 - V_m \phi_j(H_m) e_1 \ .$$

Proceeding as for the development of the formula (10), we write

$$\phi_1(A) v_1 = \phi_1(0) v_1 + A \phi_2(A) v_1$$

$$
\begin{aligned}
&= \phi_1(0)v_1 + A(V_m\phi_2(H_m)e_1 + s_m^2) \\
&= \phi_1(0)v_1 + (V_mH_m + h_{m+1,m}v_{m+1}e_m^T)\phi_2(H_m)e_1 + As_m^2 \\
&= V_m[\phi_1(0)e_1 + H_m\phi_2(H_m)e_1] + h_{m+1,m}e_m^T\phi_2(H_m)e_1v_{m+1} + As_m^2 \\
&= V_m\phi_1(H_m)e_1 + h_{m+1,m}e_m^T\phi_2(H_m)e_1v_{m+1} + As_m^2 \qquad (45)
\end{aligned}
$$

Which gives the relation

$$s_m^1 = h_{m+1,m}e_m^T\phi_2(H_m)e_1v_{m+1} + As_m^2.$$

We can continue expanding s_m^2 and then repeat the above development for $s_m^3, s_m^4, ..$ and so forth, in the same manner. Combining (44) and (10) we get

$$
\begin{aligned}
e^Av_1 &= V_me^{H_m}e_1 + h_{m+1,m}e_m^T\phi_1(H_m)e_1v_{m+1} + As_m^1 \\
&= V_me^{H_m}e_1 + h_{m+1,m}e_m^T\phi_1(H_m)e_1v_{m+1} + h_{m+1,m}e_m^T\phi_2(H_m)e_1Av_{m+1} + A^2s_m^2 \\
&= .... \\
&= V_me^{H_m}e_1 + h_{m+1,m}\sum_{k=1}^{j} e_m^T\phi_k(H_m)e_1A^{k-1}v_{m+1} + A^js_m^j \qquad (46)
\end{aligned}
$$

It remains to show that $A^js_m^j$ converges to zero. This is a consequence of the following inequality satisfied for any nonnegative x:

$$0 \leq \phi_j(x) \leq \frac{e^x}{j!}$$

which immediately follows from the Taylor series expansion of ϕ_j. As a consequence the norm of the vector s_m^j as defined above can be bounded above by a term in the form $C/j!$ where C is a constant, and the two results in the theorem follow from (46). □

Notice by comparing (42) and (43) that the first term in the first expansion is nothing but the correcting term used in the corrected schemes. In the next subsection we will show how to exploit these expansions to derive computable estimates of the error.

5.2. Practical error estimates. The first one or two terms in the expansion formulas established in the previous theorem will usually suffice to provide a good estimate for the error. Thus, for the Arnoldi and Lanczos processes the first term of this series is $h_{m+1,m}e_m^T\phi_1(H_m)e_1v_{m+1}$ whose norm yields the error estimate

$$Er_1 = h_{m+1,m}|e_m^T\phi_1(H_m)\beta e_1|$$

As a result of Proposition 2.1, this is rather straightforward to implement in practice. It suffices to use any means, e.g. rational approximations, to compute $\bar{y} = exp(\overline{H}_m)\beta e_1$. Then the vector consisting of the first m components of $\bar{y}$ is equal to $y = exp(H_m)\beta e_1$ which is needed to compute the actual approximation via $u = V_my$, while the last component multiplied by $h_{m+1,m}$ is the desired error estimate.

If a rough estimate of the error is desired, one may avoid invoking the ϕ_1 function by replacing $\phi_1(H_m)$ by $exp(H_m)$ leading to the estimate

$$Er_2 = h_{m+1,m}|e_m^T y_m|.$$

The above expression is reminiscent of a formula used to compute the residual norm of linear systems solved by the Arnoldi or the Lanczos/ Bi-Conjugate Gradient algorithms.

For the corrected schemes a straightforward analogue of Er_2 is

$$Er_3 = h_{m+1,m}|e_m^T \phi_1(H_m)\beta e_1|.$$

We can also use the first term in the expansion (43) to get the more elaborate error estimate for the corrected schemes:

$$Er_4 = h_{m+1,m}|e_m^T \phi_2(H_m)\beta e_1|\|Av_{m+1}\|_2.$$

The above formula requires the evaluation of the norm of Av_{m+1}. Although this is certainly not a high price to pay there is one compelling reason why one would like to avoid it: if we were to perform one additional matrix-vector product then why not compute the next approximation which belongs the Krylov subspace K_{m+2}? We would then come back to the same problem of seeking an error estimate for this new approximation. Since we are only interested in an estimate of the error, it may be sufficient to approximate Av_{m+1} using some norm estimate of A provided by the Arnoldi process. One possibility is to resort to the scaled Frobenius norm of H_m:

$$\|Av_{m+1}\|_2 \approx \|H_m\|_F \equiv \left[\frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{j+1} h_{i,j}^2\right]^{1/2}$$

which yields

$$Er_5 = h_{m+1,m}|e_m^T \phi_2(H_m)\beta e_1|\|H_m\|_F.$$

There remains the issue of computing $e_m^T \phi_2(H_m)(\beta e_1)$ for Er_4 and Er_5. One might use a separate calculation. However, when rational approximations are used to evaluate $exp(\overline{H}_m)(\beta e_1)$, a more attractive alternative exists. Assume that the vector $\bar{y} = exp(\overline{H}_m)(\beta e_1)$ is computed from the partial fraction expansion

$$\bar{y} = \sum_{i=1}^{p} \alpha_i (\overline{H}_m - \theta_i I)^{-1}(\beta e_1). \tag{47}$$

A consequence of Proposition 2.1 is that the last component of the result is $e_m^T \phi_1(H_m)(\beta e_1)$ but we need $e_m^T \phi_2(H_m)(\beta e_1)$. Now observe that if instead of the coefficients α_i we were to use the coefficients α_i/θ_i in (47) then, according to the formula (17) in Section 2.4, we would have calculated the vector $\phi_1(H_m)(\beta e_1)$ and *the last component* would be the desired $e_m^T \phi_2(H_m)(\beta e_1)$. This suggests that all we have to do is to accumulate the scalars

$$\sum_{i=1}^{p} \frac{\alpha_i}{\theta_i} e_{m+1}^T (\overline{H}_m - \theta_i I)^{-1} e_1, \tag{48}$$

a combination of the last components of the consecutive solutions of (47), as these systems are solved one after the other. The rest of the computation is in no way disturbed since we only accumulate these components once the solution of each linear system is completed. Notice that once more the cost of this error estimator is negligible. We also point out that with this strategy there is no difficulty in getting more accurate estimates of the error which correspond to using more terms in the expansion (43). For example, it suffices to replace α_i/θ_i in (48) by α_i/θ_i^2 to get $e_m^T\phi_3(H_m)e_1$, and therefore a similar technique of accumulation can be used to compute the higher order terms in the expansion.

These error estimates have been compared on a few examples described in Section 6.

5.3. Error with respect to the rational approximation. The vector $e^{H_m}e_1$ in (3) may be evaluated with the help of a rational approximation to the exponential of the form (15), i.e., the actual approximation used is of the form

$$e^A v_1 \approx V_m R(H_m) e_1 \tag{49}$$

In the symmetric case, the Chebyshev rational approximation to the exponential yields a known error with respect to the exact exponential, i.e., using the same rational function R,

$$\|e^A v_1 - R(A)v_1\|_2 = \eta_p$$

is known a priori from results in approximation theory. As a result one may be able to estimate the total error by simply estimating the norm of the difference $R(A)v_1 - V_m R(H_m)e_1$.

PROPOSITION 5.2. *The following is an exact expression of the error made when $R(A)v_1$ is approximated by $V_m R(H_m)e_1$.*

$$R(A)v_1 - V_m R(H_m)e_1 = \sum_{i=1}^{p} \alpha_i \epsilon_i (A-\theta_i I)^{-1} v_{m+1} \tag{50}$$

in which

$$\epsilon_i = -h_{m+1,m} e_m^T (H_m - \theta_i I)^{-1} e_1.$$

Proof. We have

$$\begin{aligned} R(A)v_1 - V_m R(H_m)e_1 &= \alpha_0(v_1 - V_m e_1) + \sum_{i=1}^{p} \alpha_i \left[(A-\theta_i I)^{-1}v_1 - V_m(H_m-\theta_i I)^{-1}e_1\right] \\ &= \sum_{i=1}^{p} \alpha_i (A-\theta_i I)^{-1}[v_1 - (A-\theta_i I)V_m(H_m-\theta_i I)^{-1}e_1] \qquad (51) \end{aligned}$$

Using once more the relation (2) we have

$$v_1 - (A-\theta_i I)V_m(H_m-\theta_i I)^{-1}e_1 = -h_{m+1,m}e_m^T(H_m-\theta_i I)^{-1}e_1 v_{m+1}$$

TABLE 1
Actual error and estimates for basic method on example 1.

m	Actual Err	Er_1	Er_2
3	0.301D-01	0.340D-01	0.889D-01
5	0.937D-04	0.102D-03	0.466D-03
6	0.388D-05	0.416D-05	0.232D-04
7	0.137D-06	0.146D-06	0.958D-06
8	0.424D-08	0.449D-08	0.339D-07
9	0.119D-09	0.123D-09	0.105D-08
10	0.220D-10	0.301D-11	0.287D-10

TABLE 2
Actual error and estimates for the corrected method on example 1.

m	Actual Err	Er_4	Er_5	Er_3
3	0.484D-02	0.571D-02	0.599D-02	0.340D-01
5	0.992D-05	0.112D-04	0.115D-04	0.102D-03
6	0.351D-06	0.389D-06	0.399D-06	0.416D-05
7	0.108D-07	0.119D-07	0.121D-07	0.146D-06
8	0.298D-09	0.323D-09	0.329D-09	0.449D-08
9	0.230D-10	0.788D-11	0.803D-11	0.123D-09
10	0.218D-10	0.174D-12	0.176D-12	0.301D-11

With the definition of ϵ_i the result follows immediately. □

Note that $|\epsilon_i|$ represents the residual norm for the linear system $(A - \theta_i I)x = v_1$ when the approximation $V_m(H_m - \theta_i I)^{-1}e_1$ is used for the solution, while $\epsilon_i(A - \theta_i I)^{-1}v_{m+1}$ is the corresponding error vector. Thus, an interpretation of the result of the proposition, is that the error $R(A)v_1 - V_m R(H_m)e_1$ is equal to a linear combination with the coefficients $\alpha_i, i = 1, ..., p$ of the errors made when the solution of each linear system $(A - \theta I)x = v_1$ is approximated by $V_m(H_m - \theta I)^{-1}e_1$. These errors may be estimated by standard techniques in several different ways. In particular the condition number of each matrix $(A - \theta_i I)$ can be estimated from the eigenvalues/ singular values of the Hessenberg matrix $H_m - \theta_i I$.

6. Numerical experiments. The purpose of this section is to illustrate with the help of two simple examples how the error estimates provided in the previous section behave in practice. In the first example we consider a diagonal matrix of size $N = 100$ whose diagonal entries $\lambda_i = (i + 1)/(N + 1)$, are uniformly distributed in the interval $[0, 1]$. In the second example the matrix is block diagonal with 2 x 2 blocks of the form

$$\begin{pmatrix} a_i & c \\ -c & a_i \end{pmatrix}$$

in which $c = \frac{1}{2}$ and $a_j = (2\ j - 1)/(N + 1)$ for $j = 1, 2, ..., N/2$. In the first example the vector v was selected so that the result $u = exp(A)w$ is known to be the vector

TABLE 3
Actual error and estimates for basic method on example 2.

m	Actual Err	Er_1	Er_2
3	0.114D+00	0.128D+00	0.337D+00
5	0.164D-02	0.178D-02	0.816D-02
6	0.124D-03	0.134D-03	0.744D-03
7	0.107D-04	0.114D-04	0.751D-04
8	0.593D-06	0.627D-06	0.473D-05
9	0.399D-07	0.419D-07	0.358D-06
10	0.181D-08	0.189D-08	0.180D-07
11	0.105D-09	0.102D-09	0.109D-08
12	0.346D-10	0.351D-11	0.425D-10

$(1, 1, 1..., 1)^T$. In the second example it was generated randomly.

To compute the coefficient vector y_m we used the rational approximation outlined in Section 2.4. Moreover, for the corrected schemes we used the same rational approximations and the a posteriori errors have been calculated following the ideas at the end of Section 5.2. The rational function used in all cases is the Chebyshev approximation to e^x of degree (14,14) on the real negative line [3].

Table 1 shows the results for example 1 using the Arnoldi method which in this case amounts to the usual symmetric Lanczos algorithm. The estimates are based on the formulas Er_2 and Er_1 of Section 5.2. Using the corrected version of the Arnoldi algorithm we get the results in Table 2. Tables 3 and 4 show similar results for example 2. Note that as expected the columns Er_1 and Er_3 are identical.

We can make the following observations. First, the estimate Er_1 is far more accurate than the estimate Er_2 for the basic Arnoldi-Lanczos methods. However, we start seeing some difficulties for large values of m. These are due to the fact that the rational approximation does not yield an accurate enough answer for $exp(H_m)e_1$. In fact for larger values of m the actual error stalls at the level 2.3 10^{-11} while the estimates continue to decrease sharply. Unfortunately, we did not compute the coefficients of the rational approximation, as described in Section 2, beyond the degree $(14, 14)$ used here. In fact one can speculate that the estimated errors in the tables should be close to the actual error that we would have obtained had we used more accurate rational approximations. Second, the estimates Er_1 and Er_5 based on just one term of the expansion of the error shown in Theorem 5.1, are surprisingly sharp in these examples. Finally, note that, as expected from the theory, the corrected approximations are slightly more accurate than the basic approximations. The tables indicate that, roughly speaking, for the larger values of m we gain one step for free. Note also that for these examples the scaled Frobenius norm seems to give a rather good estimate of the norm $\|Av_{m+1}\|_2$ as is revealed by comparing the values of Er_4 and Er_5.

7. Conclusion. The general analysis proposed in this paper shows that the technique based on Krylov subspaces can provide an effective tool for approximating the

TABLE 4
Actual error and estimates for the corrected method on example 2.

m	Actual Err	Er_4	Er_5	Er_3
3	0.245D-01	0.284D-01	0.270D-01	0.128D+00
5	0.225D-03	0.249D-03	0.249D-03	0.178D-02
6	0.146D-04	0.160D-04	0.159D-04	0.134D-03
7	0.110D-05	0.119D-05	0.119D-05	0.114D-04
8	0.536D-07	0.575D-07	0.577D-07	0.627D-06
9	0.323D-08	0.344D-08	0.345D-08	0.419D-07
10	0.137D-09	0.140D-09	0.140D-09	0.189D-08
11	0.349D-10	0.655D-11	0.655D-11	0.102D-09
12	0.344D-10	0.133D-12	0.136D-12	0.351D-11

exponential propagation operator. This has already been confirmed by numerous experiments elsewhere [3, 4, 9, 10].

The a posteriori error estimates of Section 5, are mandatory for implementing general purpose software for computing $exp(A)v$. The numerical experiments reported show that the simple estimates based on just one term of the expansion of the error developed in Section 5, are sufficient to provide enough accuracy for practical purposes. These estimates may be combined with step-reduction strategies to build efficient integrators for systems of ordinary differential equations with constant coefficients, with possible extensions to more general problems.

REFERENCES

[1] A. J. Carpenter, A. Ruttan, and R. S. Varga, Extended numerical computations on the 1/9 conjecture in rational approximation theory, in *Rational Approximation and Interpolation*, P. R. Graves-Morris, E. B. Saff, and R. S. Varga, eds., vol. 1105 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1984, pp. 383–411.

[2] R. A. Friesner, L. S. Tuckerman, B. C. Dornblaser, and T. V. Russo. A method for exponential propagation of large systems of stiff nonlinear differential equations. Journal of Scientific Computing, 4:327–354, 1989.

[3] E. Gallopoulos and Y. Saad. Parallel solution of parabolic equations using polynomial approximations to the exponential. Technical report, Research Institute for Advanced Computer Science, Technical report number 90-14. Submitted.

[4] E. Gallopoulos and Y. Saad. On the parallel solution of parabolic equations. In *Proc. 1989 ACM Int'l. Conference on Supercomputing*, pages 17–28, Herakleion, Greece, June 1989. Also CSRD Tech. Rep. 854.

[5] F. R. Gantmacher. *The Theory of Matrices*, Chelsea, New York, 1959.

[6] T. Kato. *Perturbation theory for Linear Operators* , Springer Verlag, New York, 1965.

[7] A. Nauts and R. E. Wyatt. New approach to many-state quantum dynamics: The recursive-residue-generation method, *Phys. Rev. Lett.*, (51):2238–2241, 1983.

[8] A. Nauts and R. E. Wyatt. *Theory of laser-module interaction: The recursive-residue-generation method*, Physical Rev., (30):872–883, 1984.

[9] B. Nour-Omid. Applications of the Lanczos algorithm. *Comp. Phy. Comm.*, 53, 1989.

[10] T. J. Park and J. C. Light. Unitary quantum time evolution by iterative Lanczos reduction. J. Chem. Phys. 85:5870–5876, 1986.

[11] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Statist. Comput, 7:856–869, 1986.

[12] L. N. Trefethen. The asymptotic accuracy of rational best approximation to e^z in a disk. *J. Approx. Theory*, 40:380-383, 1984.

[13] H. van der Vorst. An iterative solution method for solving $f(A) = b$ using Krylov subspace information obtained for the symmetric positive definite matrix A. *J. Comput. Appl. Math.*, 18:249–263, 1987.