CONSIDERATIONS ON COMPUTING REAL LOGARITHMS OF MATRICES, HAMILTONIAN LOGARITHMS, AND SKEW-SYMMETRIC LOGARITHMS *.

Luca Dieci
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332 U.S.A.
(e-mail: dieci@math.gatech.edu)

AMS Subject Classification: 15, 65

Key Words: real logarithm of a matrix, Padé approximants, Newton's method, series expansions, Hamiltonian logarithm, skew-symmetric logarithm.

ABSTRACT.

In this work, the issue of computing a real logarithm of a real matrix is addressed. After a brief review of some known methods, more attention is paid to three methods: (i) Padé approximation techniques, (ii) Newton's method, and (iii) a series expansion method. Newton's method has not been previously treated in the literature; we address commutativity issues, and simplify the algorithmic formulation. We also address general structure preserving issues for two applications in which we are interested: finding the real Hamiltonian logarithm of a symplectic matrix, and the skew-symmetric logarithm of an orthogonal matrix. The diagonal Padé approximants and the proposed series expansion technique are proven to be structure preserving. Some algorithmic issues are discussed.

Notation and a few known Facts. A matrix $M \in \mathbb{R}^{2n\times 2n}$ is called *Hamiltonian* if $M^TJ + JM = 0$, where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$; equivalently, M has the block structure $M = \begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$, where all blocks are $(n \times n)$ and B and C are symmetric. A matrix T is called *symplectic* if $T^TJT = J$; equivalently, $T^{-1} = -JT^TJ$, so that if $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, then $T^{-1} = \begin{pmatrix} D^T & -B^T \\ -C^T & A^T \end{pmatrix}$. A matrix $S \in \mathbb{R}^{n\times n}$ is *skew-symmetric* if $S^T = -S$, and $Q \in \mathbb{R}^{n\times n}$ is orthogonal if $Q^TQ = I$. A symplectic similarity transformation of a symplectic (Hamiltonian) matrix is symplectic (Hamiltonian). Hamiltonian and skew-symmetric matrices are closed under sum, multiplication by a scalar, transposition, and commutator operator. Symplectic and orthogonal matrices are closed under inversion, transposition, and multiplication.

1. INTRODUCTION

In this work, we consider some of the issues associated with computing the logarithm of a matrix. In general, the problem consists of the following: "Given a $(n \times n)$ matrix T, to find a $(n \times n)$ matrix X such that

* Work supported in part under NSF Grant DMS-9306412

$e^X = T$, where e^X is the matrix exponential of X. Any matrix X satisfying this relation is called a *logarithm* of T, and we write $X = \log(T)$."

The issue of computing logarithms of matrices has been treated by system engineers for quite some time, in connection with the continuization process, how to convert a discrete process into a continuous one (see [LS1], [LS2], [V]). It also has applications in stability of differential equations (see [Si], [YS]). Mathematically, it is well known (e.g., see [HJ]) that any invertible matrix has at least one logarithm. We will henceforth assume that T is invertible.

The logarithm of a matrix is just one instance of a map between the $(n \times n)$ matrices into themselves. Just as with other functions, in principle there are two possible types of solutions X to $e^X = T$. Those which are functions of T ([GvL], [H1]), called *primary matrix functions* in [HJ], and which are in fact polynomials in T, and those which are not. For computational purposes, it is more convenient to restrict to the case of logarithms which are primary matrix functions of T, and -unless otherwise stated- we restrict attention to these logarithms. The usual definition of the logarithm (or any other matrix function) goes through the Jordan canonical form of T, or, equivalently, the Cauchy integral formula (see [G], [GvL], [HJ]).

DEFINITION 1.1. Let T be an $(n \times n)$ matrix with Jordan decomposition $T = VJV^{-1}$, where

$$J = \begin{pmatrix} J_1(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_s(\lambda_s) \end{pmatrix},$$

and each $J_k(\lambda)$ is a Jordan block with eigenvalue λ. Then, one has

$$\log(T) = V \log(J) V^{-1} = V \begin{pmatrix} \log(J_1(\lambda_1)) & & 0 \\ & \ddots & \\ 0 & & \log(J_s(\lambda_s)) \end{pmatrix} V^{-1}, \tag{1.1}$$

where each block $\log(J_k(\lambda))$ is given by ($f(\lambda) := \log(\lambda)$ here)

$$\log(J_k(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \dots & \frac{1}{(k-1)!}f^{(k-1)}(\lambda) \\ 0 & f(\lambda) & f'(\lambda) & \dots & \frac{1}{(k-2)!}f^{(k-2)}(\lambda) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & f(\lambda) \end{pmatrix}. \tag{1.2}$$

Alternatively, the logarithms of T can be characterized by the following contour integral:

$$\log(T) = \frac{1}{2\pi i} \oint_\Gamma \log(z)(zI - T)^{-1} dz, \tag{1.3}$$

where the contour Γ is any simple curve enclosing all the eigenvalues of T. □

Our own interest is for the case in which $T \in \mathbb{R}^{n\times n}$, and $X = \log(T)$ is a real matrix as well. A complete existence result in this case is the following.

THEOREM 1.2. ([HJ], but see also [V]). *Let $T \in \mathbb{R}^{n\times n}$, nonsingular. Then, there exist $X \in \mathbb{R}^{n\times n}$, $X = \log(T)$, if and only if T has an even number of Jordan blocks of each size for every negative eigenvalue. If T has any eigenvalue on the negative real axis, then no real logarithm of T can be a primary matrix function of T.* □

From the point of view of applications, there are a number of more specific cases which are of interest. For instance, it is known ([HJ]) that if T is positive definite, then there exists a unique symmetric logarithm of T. Our own motivation in logarithms of matrices stems from the desire to "invert" the discretization process which occurs when solving systems of differential equations. Ideally, that way one would be able to precisely determine what is being solved. More precisely, the following linear problems motivated our interest

$$
\begin{aligned}
&(a) && \dot{Y}(t) = A(t)Y(t),\ t \geq 0, \\
& && Y(0) = Y_0,\ Y(t), A(t) \in \mathbb{R}^{n\times n}, \\
&(b) && \dot{Y}(t) = M(t)Y(t),\ t \geq 0, \\
& && Y(0) = Y_0,\ Y(t), M(t) \in \mathbb{R}^{2n\times 2n},\ Y_0 \quad \text{symplectic}, \quad M(t) \quad \text{Hamiltonian}, \\
&(c) && \dot{Q}(t) = S(t)Q(t),\ t \geq 0, \\
& && Q(0) = Q_0,\ Q(t), S(t) \in \mathbb{R}^{n\times n},\ Q_0 \quad \text{orthogonal}, \quad S(t) \quad \text{skew}-\text{symmetric}.
\end{aligned}
\tag{1.4}
$$

Case (1.4a) has no particular structure we are interested to maintain, but of course we will assume that the computed approximations to $Y(t)$ are all invertible matrices. For (1.4b) and (1.4c), instead, it is well known that their solutions are symplectic and orthogonal, respectively, for all times t. It is also known that if Runge-Kutta schemes at Gaussian points (GRK, for short) are used for the integration of (1.4b) ((1.4c)), then the solutions at the grid-points will also be symplectic (orthogonal) matrices (see [K], [SS], [DRV]). GRK schemes correspond to the diagonal Padé approximants to the exponential for constant coefficients problems. In the present context, we will think that approximations to (1.4b-c) have been computed which are symplectic and orthogonal, respectively (e.g., we have used GRK schemes).

It is easy to see that the exponential of any Hamiltonian (skew-symmetric) matrix is symplectic (orthogonal). But, in general, it is not true that the logarithm of a symplectic (orthogonal) matrix is Hamiltonian (skew-symmetric). The result below, given in [Si] and [YS], tells when this is true. The method of proof in [YS] (Lemma I, p.211, vol. 1) uses formula (1.3), and the result is proven only for the skew-symmetric case, but in fact the proof for the Hamiltonian case is the same. Also, the results in [Si] and [YS] do not explicitly contain part (b) below, but this follows easily from their proofs.

THEOREM 1.3. ([Si], [YS]). *Suppose that the matrix T is real and symplectic (orthogonal) and does not have any eigenvalue on the negative real axis (-1 is not an eigenvalue). Then:*

(a) there exists a real Hamiltonian (skew-symmetric) matrix X, such that $X = \log(T)$.

(b) $X = \log(T)$ can be uniquely specified if, corresponding to the eigenvalues of T, we specify which branch of the log *we take. For example, there is a unique X such that all of its eigenvalues z satisfy $-\pi < \mathrm{Im}(z) < \pi$.*

□

Under the assumptions of Theorem 1.3, we would like approximation techniques for the log which surely deliver the Hamiltonian, skew-symmetric, logarithms in question. This fact is of key importance if we want

the assurance that our results be qualitatively correct. As we will see, this is possible with the diagonal Padé approximants, and by truncating an appropriate series expansion.

In Section 2 we first review some of the existing methods. We then give a new result, concerning structure preserving properties of the diagonal Padé approximants. We also discuss how (and when) to incorporate an initial guess in the approximation of the log. In Section 3 we discuss the use of Newton's method for computing $\log(T)$. In Section 4 we consider a simple series expansion technique, which was used in [LS2], and which is the matrix analog of the recommended strategy in Calculus books (e.g., see [St]) for computing logarithms of real numbers. This technique enjoys some nice features. Some conclusions are in Section 5.

2. SOME METHODS

At present, there has been less interest in computation of the logarithm of a matrix than there has been in the "inverse" problem, the one of computing the exponential of a matrix. What makes computation of the logarithm more difficult than that of the exponential is the lack of uniqueness. This basic difficulty reflects in the methods. In any case, the three basic approaches which have been used to find the log of a matrix are close relatives of those used for the matrix exponential: (i) series expansion techniques ([GvL], [LS1-2]), (ii) eigendecomposition approaches ([GvL], [Matlab]), and (iii) Padé approximation methods ([KL1-2]). These three approaches are not mutually exclusive, and it is conceivable, for example, to use (ii) and (iii) together when the eigenvalues of T are close to each other.

Series Expansion. The simplest series expressing $\log(T)$ is the Taylor series

$$A := I - T\,, \quad \log(I - A) = -\sum_{k=1}^{\infty} \frac{A^k}{k}\,. \tag{2.1}$$

Of course, for (2.1) to make sense, the restriction $\rho(A) < 1$ is needed. Typically, this series converges rather slowly, and an algorithm based on the partial sums of (2.1) is not practical; a better algorithm can be based on the expansion in Section 4. It is easily seen that any partial sum of (2.1) is symmetric when T is positive definite, but generally not Hamiltonian (skew-symmetric) when T is symplectic (orthogonal).

Eigendecomposition Approaches. These are based on the fact that if the matrix T has the decomposition $T = URU^{-1}$, then also $\log(T) = U\log(R)U^{-1}$. A diagonalization approach is too prone to being unstable, except when T is positive definite. The Schur approach, as implemented in Matlab, is a much safer choice: T is reduced to Schur form (so, U is unitary and R is triangular), and then a fast recursion on the triangular portion is applied, in the way explained in [GvL] (Algorithm 11.1.1 of [GvL]). This approach suffers when there are repeated (or close together) eigenvalues of T. Moreover, complex arithmetic needs to be performed, and there is no guarantee to eventually obtain a real logarithm. Of course, the method is very reliable in recovering the symmetric logarithm of a positive definite matrix. But a drawback of the method is that it

treats any matrix the same way, and generally one has no guarantee to obtain a desired matrix structure for the logarithm (say, Hamiltonian). Finally, some decisions must be made as to which branch of the log one should ultimately select. On the other hand, the strength of the method is its generality. Aside from the coalescing eigenvalues' case, the approach can in principle be used on any invertible matrix. Appropriate measures to make this approach more reliable, and structure preserving for skew-symmetric logarithms, have been recently considered in [DMP], to which we refer for details.

Padé Approximants. An interesting approach is the Padé approximation technique used by Kenney and Laub in [KL1-2]. The starting point is the identity $\log(T) = \log((T^{1/2^k})^{2^k}) = 2^k \log(T^{1/2^k})$. So, given T, first one progressively takes square roots of T, say up to $T_{2k} := T^{1/2^k}$, so that T_{2k} is as close to the identity as desired. At this point, a Padé approximant is used for computing $\log(T_{2k}) = \log(I - A)$, $A = I - T_{2k}$. Finally, one recovers $\log(T) = 2^k \log(T_{2k})$. The "inverse squaring and scaling" procedure is needed, since Padé approximants are more accurate close to the origin (above, A is close to the origin), and for practical reasons one does not want to consider high order Padé approximants. In principle, any Padé approximant could be used for $\log(I-A)$. In their work, Kenney and Laub propose to choose k so that $\|A\| = \|I-T_{2k}\| \leq 1/4$, and then take the eighth order diagonal Padé approximant $R(A)$ to $\log(I - A)$; they show that $R(A)$ is within 10^{-18} to $\log(I - A)$. For future reference, this Padé approximant is

$$\begin{aligned} R(A) &= P(A)(Q(A))^{-1}\,, \quad A := I - T_{2k}\,, \\ P(A) &= -A + \frac{7}{2}A^2 - \frac{73}{15}A^3 + \frac{41}{12}A^4 - \frac{743}{585}A^5 + \frac{31}{130}A^6 - \frac{37}{1925}A^7 + \frac{761}{1801800}A^8\,, \\ Q(A) &= I - 4A + \frac{98}{15}A^2 - \frac{28}{5}A^3 + \frac{35}{13}A^4 - \frac{28}{39}A^5 + \frac{14}{143}A^6 - \frac{4}{715}A^7 + \frac{1}{12870}A^8\,. \end{aligned} \tag{2.2}$$

This algorithm is the analog of the well known "scaling and squaring" with Padé approximants algorithm used for computing exponentials of a matrix: given $T \in \mathbb{R}^{n\times n}$, to compute e^T, one uses the identity $e^T = (e^{\frac{T}{2k}})^{2k}$, and a Padé approximant for $e^{\frac{T}{2k}}$. This is one of the most successful ways to compute e^T, and it is the recommended approach implemented in Matlab. What makes it so successful is that scaling and squaring are (relatively speaking) straightforward and inexpensive; moreover, there are no uniqueness issues to resolve. In the case of the log, there is a key difference: the computation of square roots of a matrix. To this end, one might once more use a Schur decomposition approach, with the usual strengths and drawbacks of these methods ([H1]), or Newton's method ([H2]). The Schur approach has the nontrivial advantage of requiring only one decomposition (see [KL2]). In any case, the overall procedure for computing $\log(T)$ is more expensive than the analogous one for computing e^T, and choices must be made which affect which branch of the log one eventually computes.

REMARK 2.1. As we know, when approximating $\log(T)$ one has to select a proper branch for the log. The choice adopted in [KL1-2], and apparently also in [Matlab], is to approximate the principal log, that is to require that the eigenvalues z of $X = \log(T)$ satisfy $-\pi < \mathrm{Im}(z) < \pi$. This is the most sensible strategy. Naturally, also (2.1) approximates this principal log. If one wants different branches of the log, this can in principle be done with the eigendecomposition approaches, taking the appropriate log values for the

eigenvalues of T. It is not clear how to do it for the Padé approximation methods, and for series expansion methods.

It is easy to see, in case $A = I - T$ and T is a positive definite matrix, that any Padé approximant used for $\log(I - A)$ would give us a symmetric matrix. It is also clear that, in general, non-diagonal Padé approximants do not recover Hamiltonian (or skew-symmetric) structure (e.g., truncating (2.1) gives the first column in the Padé table for $\log(I - A)$). The question becomes: "Do the diagonal Padé approximants recover such structure?" The answer is positive, as we show next.

Let us begin with two simple examples. Let $A = I - T$ here below. Borrowing from Table 1 p.715 of [KL1], we have that the $(1, 1)$ Padé approximant is

$$R_{1,1}(A) = -2A(2 - A)^{-1} = -2(I - T)(I + T)^{-1}\,,$$

and this has the desired structure as a consequence of Lemma 2.5 below. The $(2, 2)$ Padé approximant is

$$R_{2,2}(A) = (-6A + 3A^2)(6I - 6A + A^2)^{-1}\,,$$

and after a bit of manipulation, expressed in terms of T this becomes

$$R_{2,2} = -3(I - T)(I + T)(T + aI)^{-1}(T + bI)^{-1}\,,\ a := 2 + \sqrt{3},\, b := 2 - \sqrt{3}\,.$$

This also recovers the desired structure. For example, when T is orthogonal, $R_{2,2}$ is a skew-symmetric matrix, since

$$R_{2,2}^T = 3(I - T)(I + T)(I + aT)^{-1}(I + bT)^{-1}\,,$$

and $(bI + T)(aI + T) = (bT + I)(aT + I)$, because $ab = 1$. Similarly, we can show by direct verification that if T is symplectic, then $R_{2,2}$ is Hamiltonian. Of course, this is not a sensible general strategy. In general, we have:

THEOREM 2.2. *Let $X = \log(T)$, $A = I - T$, and let $\rho(A) < 1$. Let $R_{m,m}(A)$ be the diagonal Padé approximants to $\log(I - A)$, $m = 0, 1, \ldots$. Then, we have*

(i) If T is orthogonal, $R_{m,m}(A)$ is skew-symmetric.

(ii) If T is symplectic, $R_{m,m}(A)$ is Hamiltonian.

Proof. The key result which makes things work is a theorem in Padé approximants, known as the "homographic invariance under argument transformations" (this is Theorem 1.5.2 in [BG-M], vol I). Let x be a scalar, and consider $f(x) = \log(1-x)$; from the equality $\log(1-x) = -\log(1-\frac{x}{x-1})$, this theorem in [BG-M] implies that also $R_{m,m}(x) = -R_{m,m}(\frac{x}{x-1})$, a fact already noted in Lemma 1 of [KL1]. In our matrix case, this relation reads

$$R_{m,m}(A) = -R_{m,m}(A(A - I)^{-1})\,, \tag{2.3}$$

and with this we can now show (i) and (ii) of our Theorem. Let

$$R_{m,m}(A) := P(A)(Q(A))^{-1}\,,\ P(A) = \sum_{k=0}^{m} a_k A^k\,,\ Q(A) = \sum_{k=0}^{m} b_k A^k\,.$$

(i) Let $T\,:\,T^T T = I$. From (2.3), we have

$$R_{m,m}(I-T) = -R_{m,m}(I-T^T)\,, \qquad \text{or} \qquad R_{m,m}(A) = -R_{m,m}(A^T)\,.$$

On the other hand, $R_{m,m}(A^T) = (R_{m,m}(A))^T$, since

$$\sum_{k=0}^{m} a_k B^k\,(\sum_{k=0}^{m} b_k B^k)^{-1} \;=\; (\sum_{k=0}^{m} b_k B^k)^{-1}\,\sum_{k=0}^{m} a_k B^k\,,$$

for any matrix B (in our case it would be $B = A^T$). Therefore, $R_{m,m}$ is skew-symmetric and (i) follows.

(ii) Let $T\,:\,T^{-1} = -JT^T J$. From (2.3), we have

$$R_{m,m}(I-T) = -R_{m,m}(I-T^{-1}) = -R_{m,m}(J(T^T-I)J) = -R_{m,m}(J^T(I-T^T)J)\,,$$

$$\text{or} \qquad R_{m,m}(A) = -R_{m,m}(J^T A^T J)\,.$$

On the other hand, $R_{m,m}(J^T A^T J) = J^T R_{m,m}(A^T) J$, because

$$J^T[\sum_{k=0}^{m} a_k B^k\,(\sum_{k=0}^{m} b_k B^k)^{-1}]J \;=\; \sum_{k=0}^{m} a_k (J^T B J)^k\,(\sum_{k=0}^{m} b_k (J^T B J)^k)^{-1}\,,$$

for any matrix B (in our case it would be $B = A^T$). Moreover, from the proof of part (i), we have $R_{m,m}(A^T) = (R_{m,m}(A))^T$. Therefore, we have $R_{m,m}(A) = -J^T(R_{m,m}(A))^T J$ and $R_{m,m}(A)$ is Hamiltonian. □

REMARKS 2.3.

(i) Theorem 2.2 extends to $\log(I-X)$ known results about e^X. It has been known for a while (e.g., see [K] and [DRV]) that if X is Hamiltonian or skew-symmetric, then the diagonal Padé approximants to e^X are symplectic and orthogonal, respectively. Theorem 2.2 extends this result to the inverse function, the log.

(ii) Theorem 2.2 relies on the result on homographic invariance of Padé approximants under change of coordinates. This fact only holds for diagonal Padé approximants. Still, we think that our Theorem 2.2 is optimal, that is we think that non-diagonal Padé approximants to $\log(A)$, $A = I - T$, will not recover, say, Hamiltonian structure for an arbitrary symplectic matrix T.

(iii) For the issue of approximation error with Padé approximants, we refer to [KL1].

Finally, we would like to discuss the issue of how and when to incorporate information on an approximate logarithm into any algorithm to find $\log(T)$. This basic issue has not been adequately treated in the literature before, but in fact information on an approximate log is often available, e.g. when one integrates (1.4). Suppose we need to find $X = \log(T)$, and have X_0, which is expected to be a good approximation to X, in the sense that $T_0 := e^{X_0}$ is close to T. One might want to use this information by considering the new matrix TT_0^{-1}, computing its log, and recovering $\log(T)$. But, in general, this is not possible unless T and X_0 (equivalently, T and T_0) commute. We summarize this simple fact.

LEMMA 2.4. *Let $X : e^X = T$. Let X_0 be such that $X_0 T = T X_0$, and let $T_0 = e^{X_0}$. Then, also $T T_0^{-1} = T_0^{-1} T$ and $\log(T) = \log(T T_0^{-1}) + X_0$.* □

So, one needs to have X_0 commuting with T in order to make use of it. We next discuss some possibilities. Our motivation is to consider simple and inexpensive strategies. A more expensive alternative to those below is given by taking high-order diagonal Padé approximants to give an initial guess (this is so, because it can be easily shown that if a matrix commutes with T, then it also commutes with any Padé approximant to $\log(T)$).

(A) $X_0 = dI\,, d \in \mathbb{R}$, a constant diagonal matrix. This simple choice can be useful for a matrix T having real positive eigenvalues (for example a positive definite matrix) in order to rescale it into a new matrix $\widehat{T} = e^{-d} T$ for which $I - \widehat{T}$ has all eigenvalues less than 1.

(B) Here, the idea is to "invert the implicit midpoint rule"; equivalently, taking the $(1,1)$ diagonal Padé approximant, the Cayley transform of T. The following simple result holds.

LEMMA 2.5. *Let -1 not be an eigenvalue of T. Consider the matrix*

$$H = c(I - T)(I + T)^{-1}, \quad c \in \mathbb{R},\, c \neq 0\,. \tag{2.4}$$

Then:

(i) $HT = TH$ and also $e^{-H} T = T e^{-H}$;

(ii) If $X = \log(T)$, then $XH = HX$ and so also $X = \log(e^{-H} T) + H$;

(iii) If $T^T = T$ then also $H^T = H$.

(iv) If T is symplectic then H is Hamiltonian, and also the converse is true when the inverse transformation $T = (cI + H)(cI - H)^{-1}$ is well defined. Finally, T is orthogonal if and only if H is skew-symmetric.

Proof. For completeness, we supply these simple proofs by direct verification. Since they are identical for all c, we just set $c = 1$.

(i) We have $H(I+T) = I - T$; since $(I-T)(I+T)^{-1} = (I+T)^{-1}(I-T)$, then we also have $(I+T)H = I - T$, so that $HT = TH$. The second implication follows from $e^{-H} = \displaystyle\sum_{k=0}^{\infty} \frac{(-H)^k}{k!}$.

(ii) Let $T(t) = e^{Xt}$, $t \geq 0$, so that $H = (I - T(1))(I + T(1))^{-1}$. We have $T(t)T(1) = T(1)T(t)$, from which it follows that $(I + T(1))^{-1} T(t) = T(t)(I + T(1))^{-1}$ and so $T(t)H = HT(t)$ and $XH = HX$.

(iii) This is obvious.

(iv) The first part is in [LM]. We show the second part. Let T be orthogonal. Then,

$$H^T = (T^T(T + I))^{-1}(I - T^T) = (T + I)^{-1}(T - I) = -(I - T)(I + T)^{-1} = -H\,.$$

Conversely, if $H^T = -H$, we immediately get $T^T T = I$. □

The motivation for considering (2.4) comes from the discretization process for (1.4). Suppose that some linear one step method is used to integrate (1.4); that is, with Y_k being the available approximation

at the point t_k, one computes $Y_{k+1} = S(h, A)Y_k$, $h = t_{k+1} - t_k$. What we want to find is the matrix $\widehat{A}$ such that $S(h, A) = e^{h\widehat{A}}$. If we had used the implicit midpoint rule to discretize (1.4), we would have $hA(t_k + h/2) = -2(I - S)(I + S)^{-1}$, and this is expected to be a very good approximation to $h\widehat{A}$, and it is of the type (2.4).

(C) The inverse scaling and squaring procedure of [KL2] can also be seen as a strategy to get a good initial guess for a modified problem (the guess being the zero matrix), and can be generally used for any of the algorithms considered in this work. Of course, for the symplectic and orthogonal cases, one would then need to be sure that also all square roots are such. In practice, this might be not easy. For example, if we use the eigendecomposition approach for square roots, and use Schur reductions, then we might loose symplectic structure. Relying on symplectic similarity transformation would be essential in this case. Anyhow, this is also an interesting issue to consider in the future.

REMARK 2.6. In principle, all methods discussed above, and those to be discussed next, can be modified to incorporate information on an initial guess X_0. In practice, however, the eigendecomposition approach does not benefit from this. The series (2.1) does, as do all other Padé approximants and the methods discussed in the next two Sections.

3. NEWTON'S METHOD

Here, we consider using Newton's method for the nonlinear matrix equation

$$F(X) = 0\,, \qquad F(X) := e^X - T\,. \tag{3.1}$$

Just for a moment, neglect the issue of the initial guess, and consider the formal Newton iteration

$$[F'(X)]_{X=X_k}\, Y = -F(X_k), \quad X_{k+1} = Y + X_k\,, \tag{3.2}$$

where the Fréchet derivative $F'(X)$ can be given as

$$F'(X):\ Y \to Y + \frac{XY + YX}{2} + \frac{X^2Y + XYX + YX^2}{3!} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=0}^{n-1} X^k Y X^{n-1-k}\,. \tag{3.3}$$

If X and Y commute, a major simplification occurs in (3.3), and one obtains

$$F'(X):\ Y \to Ye^X = e^X Y\,. \tag{3.4}$$

Since X_k commutes with itself, we propose to consider the following iteration (Newton's method)

$$Y = -I + e^{-X_k}T\,, \quad X_{k+1} = Y + X_k\,. \tag{3.5}$$

To justify (3.5), we need to restrict to initial guesses X_0 such that $XX_0 = X_0X$; therefore, such that $X_0T = TX_0$. We have

LEMMA 3.1. *Let $X : e^X = T$. Suppose that $X_0X = XX_0$. Then, also $X_{k+1}X = XX_{k+1}$, where X_{k+1} are the iterates from (3.5). Moreover, also $X_kX_{k+1} = X_{k+1}X_k$ in this case.*

Proof. By induction. For $k = 0$, since $XX_0 = X_0X$, then we have $X_1 = (X_0 - I) + e^{-X_0}T$. Therefore, $XX_1 = X_1X$ if $Xe^{-X_0}T = e^{-X_0}TX$ which is true since $XT = TX$. Next, if $XX_k = X_kX$, then $X_{k+1} = X_k - I + e^{-X_k}T$, and so $X_{k+1}X = XX_{k+1}$ again is true as above. Finally, $X_{k+1}X_k = X_k^2 - X_k + e^{-X_k}TX_k = X_k^2 - X_k + X_ke^{-X_k}T = X_kX_{k+1}$. □

We now have a concise formulation for *Newton's method*

$$\begin{aligned} X_{k+1} &= X_k - I + e^{-X_k}T\,, \quad k = 0, 1, \ldots \\ X_0 &\quad \text{s.t.} \quad X_0X = XX_0\,. \end{aligned} \tag{3.6}$$

Next, we show quadratic, and norm-monotóne, convergence of (3.6), provided the initial guess is close enough to the solution. From (3.6) we have

$$X_{k+1} - X = X_k - I + e^{-X_k}T - X\,,$$

and since $X_0X = XX_0$, then $e^{-X_k}T = e^{X-X_k}$. Using the expansion of the exponential function with error term as in [HJ], we have

$$e^{X-X_k} = I + (X - X_k) + (X - X_k)^2 \int_0^1 ue^{(X-X_k)(1-u)}du\,.$$

Now,

$$\int_0^1 ue^{A(1-u)}du = \sum_{n=0}^{\infty} \frac{A^n}{n!} \int_0^1 u(1-u)^n du = \sum_{n=0}^{\infty} \frac{A^n}{n!} \frac{1}{(n+1)(n+2)}\,,$$

so that we obtain

$$\|X_{k+1} - X\| < \frac{1}{2}\|X_k - X\|^2\, \|e^{(X-X_k)}\|\,. \tag{3.7}$$

So, quadratic convergence is apparent if X_0 is close enough to X. Moreover, if $\|X_0 - X\| < z_1$, where z_1 is the root of the equation $ze^z = 2$ ($z_1 \approx .85$), then we also have $\|X - X_1\| < \|X - X_0\|$ and monotóne convergence (in norm) follows.

REMARKS 3.2.

(i) Commutativity is essential in deriving convergence.

(ii) In the above, we have assumed to have the matrix exponentials exactly, but in practice they must be computed as well. This is the key expense with (3.6).

(iii) Unlike the general case with Newton's method, using a frozen Jacobian (quasi-Newton) approach does not lead to computational savings in this context. In fact, this quasi-Newton method would read

$$[F'(X)]_{X=X_0}\,(X_{k+1} - X_k) = -e^{X_k} + T\,,$$

and one still needs to compute e^{X_k}, which is the bulk of the expense.

(iv) It is easy to see that (3.6) does not usually lead to a sequence of Hamiltonian (skew-symmetric) matrices when finding the log of symplectic (orthogonal) matrices, even if X_0 has the desired structure. Of course, if we can block-diagonalize the symplectic matrix T by a symplectic transformation, that is bring it to the form $STS^{-1} := \widehat{T} = \begin{pmatrix} T_1 & 0 \\ 0 & T_1^{-T} \end{pmatrix}$, and then we use Newton's method to find $X_1 = \log(T_1)$, at the end we can recover the Hamiltonian $\log(T) = S \begin{pmatrix} X_1 & 0 \\ 0 & -X_1^T \end{pmatrix} S^{-1}$. This is sometimes possible.

(v) It is immediate, instead, to realize that (when convergent) (3.6) can be used to find the symmetric log of a positive definite matrix, if $X_0 = X_0^T$. In fact, when $X_k = X_k^T$, one immediately has

$$X_{k+1}^T = X_k - I + Te^{-X_k} = X_k - I + e^{-X_k}T = X_{k+1}\,.$$

One interesting aspect of Newton's method (3.6) is that one can also approximate logs which are not primary matrix functions of T. Consider the following example.

EXAMPLE 3.3. Consider the matrix $X = \begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix}$. It is easy to see that X solves $e^X = I$, the identity matrix. Since the two eigenvalues of X are not identical, X cannot be a primary matrix function of I. Consider the initial guess $X_0 = \begin{pmatrix} c & 2\pi - b \\ -2\pi + b & c \end{pmatrix}$, $b, c \in \mathbb{R}$, so that $XX_0 = X_0X$. Some results, obtained with a simple Matlab program for the Newton iteration, are: (i) for $b = c = 0.5$, convergence to X to full machine precision occurred in 6 iteration, (ii) for $b = c = 1$, 7 iterations were needed, and (iii) for $c = \pi$, $b = 0$ we needed 9 iterations. □

4. ANOTHER SERIES EXPANSION

Series expansions for the log of a matrix are based upon series expansions for the log of real numbers. The series (2.1) is based on the well known Taylor's expansion

$$\log(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}\,, \quad : x :< 1\,,$$

(where we note once more that partial sums of this correspond to the first column in a Padé table of $\log(1-x)$). A simple algebraic manipulation of this formula gives

$$\log(\frac{1+x}{1-x}) = 2(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots)\,,$$

and the change of variables $x = \frac{s-1}{s+1}$ then gives

$$\log(s) = 2\sum_{k=0}^{\infty} \frac{1}{2k+1}(\frac{s-1}{s+1})^{2k+1}\,,$$

which is convergent $\forall s > 0$. These basic manipulations are given in elementary Calculus books (e.g, see [St]). Based upon this last formula, we can take this series expansion for $\log(T)$ (already in [LS2])

$$\log(T) = 2\sum_{k=0}^{\infty} \frac{1}{2k+1}[(T-I)(T+I)^{-1}]^{2k+1}\,. \tag{4.1}$$

Clearly, (4.1) converges for all matrices T whose eigenvalues have positive real parts. Moreover, in our (so far, limited) experience, convergence is generally much faster than that of (2.1), and the expense to obtain the final result is often less than with the eigendecomposition or direct Padé approximants approaches. This is actually not surprising, since the above manipulations are an instance of the Eüler transformation, whose effect is at once to enlarge region of convergence and to enhance convergence rate of the series (e.g., see [BR], and references there). In fairness, this same viewpoint can be applied to the other approaches, say to Padé approximants, and it is possible that it would lead to enhanced convergence properties for the Padé approximants as well. For example, rather than considering Padé approximants based on the series expansion of $\log(1-x)$ we might want to consider approximants based on the expansion of $\log(\frac{1+x}{1-x})$ above. This remains to be done.

In any case, our interest in (4.1) originates from a property of its partial sums. We have

THEOREM 4.1. *The following facts hold:*

(i) If T is positive definite, then all partial sums of (4.1) are symmetric, and so is the limit (4.1).

(ii) If T is symplectic, then all partial sums of (4.1) are Hamiltonian.

(iii) If T is orthogonal, then all partial sums of (4.1) are skew-symmetric.

Proof. Fact (i) is clear, and convergence is assured since T is positive definite. Let us show (ii) and (iii). Since Hamiltonian and skew-symmetric matrices are closed under sum, it suffices to show that $T_k := [(T-I)(T+I)^{-1}]^{2k+1}$, $k = 0, 1, \ldots$, has the desired structure, and then the result will follow by induction. From Lemma 2.5, we know that T_0 has the desired structure. Consider case (ii). We need to show that $T_k^T J + JT_k = 0$, and since $T_k = T_0^{2k+1}$, this is equivalent to showing $(T_0^{2k+1})^T J + JT_0^{2k+1} = 0$. That is, we have to show that

$$(T_0^2)^T \cdots (T_0^2)^T T_0^T J + JT_0(T_0^2)\cdots T_0^2 = 0\,. \tag{4.2}$$

Now, T_0 has the block structure $T_0 = \begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$, for some A, B, C, with B, C symmetric. Therefore, we get

$$JT_0^2 = \begin{pmatrix} CA - A^T C & CB + (A^T)^2 \\ -A^2 - BC & BA^T - AB \end{pmatrix} = (T_0^2)^T J\,;$$

using this fact over and over, and since $T_0^T J = -JT_0$, we eventually get that (4.2) is satisfied. To show (iii) we proceed similarly. We have

$$(T_0^{2k+1})^T = ((T_0^2 \cdots T_0^2)T_0)^T = T_0^T((T_0^2)^T \cdots (T_0^2)^T) = -T_0(T_0^2 \cdots T_0^2) = -T_0^{2k+1}\,,$$

where we have used the elementary fact that T_0^2 is symmetric. □

REMARKS 4.2.

(i) Of course, the issue of enhancing convergence by a suitable approximation applies here as well.

(ii) It is worthwhile to point out that an algorithm based on taking partial sums of (4.1) can be made quite efficient from the computational point of view. The major expense is of course given by taking powers of the

matrix $A := (T-I)(T+I)^{-1}$. Binary powering (see [GvL]) can be nicely exploited in this context, possibly rewriting the k-th partial sum as $2A(I + \frac{A^2}{3} + \ldots + \frac{A^{2k}}{2k+1})$.

5. CONCLUSIONS AND FUTURE WORK

In this paper we have considered some old and some new methods for approximating real logarithms of matrices. One of our goals has been to lay ground work for a forthcoming comparison of the methods. Such a comparison has been recently carried out in [DMP].

A main motivation for this paper was to devise techniques which preserved known structural properties of the analytic logarithm. In particular, to devise approximation methods which recovered the skew-symmetric and Hamiltonian logarithms associated to orthogonal and symplectic matrices, respectively. We have proven that the diagonal Padé approximants enjoy this property, as do the partial sums of a certain series.

We have also considered Newton's method. With a suitable choice of initial guess, the formulation of Newton's method becomes manageable. As usual, quadratic convergence is recovered. A considerable expense with this method is given by the need to compute a matrix exponential at each iteration. Nonetheless, the method can be useful for approximating logarithms of a matrix T, which are not primary matrix functions of T.

Amongst several things which still have to be done, it seems to us of both theoretical and practical interest to investigate the usefulness of such "log" approach in the context of backward error analysis of ODEs, a fact which provided our initial motivation.

6. REFERENCES

[BG-M] G.A. Baker, P. Graves-Morris, *Padé Approximants, Part I & Part II*, Encyclopedia of Mathematics vol.s 13-14, Addison-Wesley (1981).

[BR] O. Bruno, F. Reitich, "Approximation of Analytic Functions: A method of Enhanced Convergence", *Math. Comp.* **63** (1994), pp. 195-213.

[DMP] L. Dieci, B. Morini, A. Papini, "Computational Techniques for Real Logarithms of Matrices", submitted to *SIAM J. Matrix Anal. & Applic.s* (1994).

[DRV] L. Dieci, R. D. Russell, E. S. Van Vleck, "Unitary Integrators and Applications to Continuous Orthonormalization Techniques", *SIAM J. Numer. Anal.* **31** (1994), pp. 261-281.

[G] F.R. Gantmacher, *Théorie des Matrices* **1**&**2**, Dunod, Paris (1966).

[GvL] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 2nd edition, The Johns Hopkins University Press (1989).

[H1] N.J. Higham, "Computing real Square Roots of a Real Matrix", *Lin. Alg. and Its Applic.s*, **88/89** (1987), pp. 405-430.

[H2] N.J. Higham, "Newton's Method for the Matrix Square Root", *Math. Comp.*, **46** (1986), pp. 537-550.

[HJ] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press (1991).

[K] Feng Kang, "Difference Schemes for Hamiltonian Formalism and Symplectic Geometry", *J. of Computational Mathematics*, **4** (1986), pp. 279-289.

[KL1] C. Kenney and A.J. Laub, "Padé Error Estimates for the Logarithm of a Matrix", *International Journal of Control*, **50-3** (1989), pp. 707-730.

[KL2] C. Kenney and A.J. Laub, "Condition Estimates for Matrix Functions", *SIAM J. Matrix Anal. & Applic.s*, **10** (1989), pp. 191-209.

[LS1] G.J. Lastman and N.K. Sinha, "Transformation Algorithm for Identification of Continuous-Time Multivariable Systems from Discrete Data", *Electronics Letters*, **17**, (1981), pp 779-780.

[LS2] G.J. Lastman and N.K. Sinha, "Infinite Series for Logarithm of Matrix, Applied to Identification of Linear Continuous-Time Multivariable Systems from Discrete-Time Models", *Electronics Letters*, **27-16**, (1991), pp 1468-1470.

[LM] A.J. Laub and K.R. Meyer, "Canonical Forms for Symplectic and Hamiltonian Matrices", *J. Celestial Mechanics* **9** (1974), pp. 213-238.

[Matlab] *Matlab Reference Guide*, The MathWorks, Inc. (1992).

[SS] J.M. Sanz-Serna, "Runge-Kutta Schemes for Hamiltonian Systems", *BIT*, **28** (1988), pp. 877-883.

[Si] Y. Sibuya, "Note on Real Matrices and Linear Dynamical Systems with Periodic Coefficients", sl J. Mathematical Analysis and Applications, **1** (1960), pp. 363-372.

[St] G. Strang, *Calculus*, Wellesley-Cambridge Press (1991).

[V] E.I. Verriest, "The matrix logarithm and the continuization of a discrete process", *Proc.s 1991 American Control Conference* (1991), pp.184-189.

[YS] V.A. Yakubovich and V.M. Starzhinskii, *Linear Differential Equations with Periodic Coefficients* **1**&**2**, John-Wiley, New York (1975).