Journal of Difference Equations and Applications, 2003 Vol. 9 (1), pp. 77–92

Taylor & Francis
Taylor & Francis Group

Floquet Theory for Time Scales and Putzer Representations of Matrix Logarithms

CALVIN D. AHLBRANDT[a,*] and JERRY RIDENHOUR[b,†]

[a]*Department of Mathematics, University of Missouri, Columbia, MO 65211, USA;* [b]*Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900, USA*

(Received 14 September 2001; Revised 9 October 2001; In final form 9 October 2001)

Dedicated to Allan Peterson on the occasion of his 60th birthday.

A Floquet theory is presented that generalizes known results for differential equations and for difference equations to the setting of dynamic equations on time scales. Since logarithms of matrices play a key role in Floquet theory, considerable effort is expended in producing case-free exact representations of the principal branch of the matrix logarithm. Such representations were first produced by Putzer as representations of matrix exponentials. Some representations depend on knowledge of the eigenvalues while others depend only on the coefficients of the characteristic polynomial. Logarithms of special forms of matrices are also considered. In particular, it is shown that the square of a nonsingular matrix with real entries has a logarithm with real entries.

Keywords: Floquet theory; Matrix logarithm; Principal branch; Putzer representation

2000 AMS Subject Classification: 39A10

1. INTRODUCTION

The theory of dynamic equations on time scales (or, more generally, measure chains) was introduced in Stefan Hilger's PhD thesis in 1988. The theory presents a structure where, once a result is established for general time scales, special cases include a result for differential equations (obtained by taking the time scale to be the real numbers) and a result for difference equations (obtained by taking the time scale to be the integers). A great deal of work has been done since 1988 unifying the theory of differential equations and the theory of difference equations by establishing the corresponding results in the time scale setting. The recent book of Bohner and Peterson [3] provides both an excellent introduction to the subject and up-to-date coverage of much of the linear theory.

In that spirit, the first part of this paper presents a time scale treatment of Floquet theory. Logarithms of matrices arise in describing the fundamental solution. Motivated by this fact, the second part of the paper presents some new representations of the principal branch of the matrix logarithm function. We call these Putzer representations since such results for matrix exponentials go back to Putzer [19] in 1966.

*Corresponding author.

†Visitor during 1998–1999 in the Department of Mathematics, University of Missouri, Columbia, MO 65211, USA.

ISSN 1023-6108 print/ISSN 1563-5120 online
DOI: 10.1080/10236190290015362

Finally, in the last section, we present some special logarithms where some property or structure in the choice of a logarithm of a given matrix is preserved.

2. FLOQUET THEORY ON TIME SCALES

A *time scale* $\mathbb{T}$ is a closed nonempty subset of the real numbers $\mathbb{R}$. We wish to study periodic behavior and we fix a positive number p to be the period of interest. We consider only time scales $\mathbb{T}$ that are p-periodic in the sense that $t \in \mathbb{T}$ implies $t + p \in \mathbb{T}$. Clearly, a p-periodic time scale is unbounded above. The *forward jump operator* $\sigma : \mathbb{T} \rightarrow \mathbb{T}$ is defined by

$$\sigma(t) := \inf\{s \in \mathbb{T} | s > t\}.$$

Similarly, the *backward jump operator* $\rho : \mathbb{T} \rightarrow \mathbb{T}$ is defined by

$$\rho(t) := \sup\{s \in \mathbb{T} | s < t\}.$$

Here we take the supremum of the empty set to be $\inf \mathbb{T}$ in the case that $\mathbb{T}$ is bounded below (so that $\rho(t) = t$ when $t = \min \mathbb{T}$). The *graininess function* $\mu : \mathbb{T} \rightarrow [0, \infty)$ defined by $\mu(t) := \sigma(t) - t$ measures the distance between a point t and its nearest neighbor to the right in the time scale $\mathbb{T}$. A point $t \in \mathbb{T}$ is *right-scattered* if $\sigma(t) > t$ and *left-scattered* when $\rho(t) < t$. Also, $t \in \mathbb{T}$ is *right-dense* if $\sigma(t) = t$ and *left-dense* if $\rho(t) = t$ and t is not a minimum of $\mathbb{T}$. A function $f : \mathbb{T} \rightarrow \mathbb{R}$ is said to be *rd-continuous* if it is continuous at all right-dense points of $\mathbb{T}$ and has finite left-hand limits at all left-dense points of $\mathbb{T}$. We refer the reader to p. 5 of Ref. [3] for the definition of the *delta derivative* (or *Hilger derivative*) f^{Δ} of a function $f : \mathbb{T} \rightarrow \mathbb{R}$. It follows from Theorem 1.16 of Ref. [3] that, when $f^{\Delta}(t)$ exists, it is given by

$$f^{\Delta}(t) = \lim_{s \rightarrow t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s}.$$

We note that this limit by itself is not a suitable definition of $f^{\Delta}(t)$ since the limit can exist without f being delta differentiable at t. (This happens when t is right-scattered and $\lim_{s \rightarrow t} f(s) = L \neq f(t)$). As usual, we differentiate vectors and matrices component-wise and say the vector or matrix is rd-continuous if each entry is rd-continuous.

Let $\mathbb{M}_n$ denote the set of $n \times n$ matrices with real entries or complex entries. If $A : \mathbb{T} \rightarrow \mathbb{M}_n$ is a matrix-valued function, we are interested in establishing a Floquet theory for the vector equation

$$x^{\Delta} = A(t)x \tag{1}$$

when both $A(t)$ and the time scale $\mathbb{T}$ are p-periodic. The coefficient function A is said to be *regressive* if the matrix

$$I + \mu(t)A(t)$$

is nonsingular for all $t \in \mathbb{T}$. When A is rd-continuous and regressive, it is a basic fact (e.g. see Ref. [3]) that solutions to initial value problems for *dynamic equations* such as Eq. (1) exist, are unique and extend to all of $\mathbb{T}$.

We begin with a preliminary lemma.

LEMMA 1 *Suppose that $A : \mathbb{T} \rightarrow \mathbb{M}_n$ is rd-continuous and regressive and that both $A(t)$ and the time scale $\mathbb{T}$ are p-periodic. If $\Phi(t)$ is a fundamental matrix for Eq. (1) and if Ψ is defined* *by $\Psi(t) := \Phi(t+p)$, then $\Psi^{\Delta}(t) = \Phi^{\Delta}(t+p)$ and $\Psi(t)$ is also a fundamental matrix for Eq. (1).*

Proof Since $\mathbb{T}$ is p-periodic, we first note that $\rho(t+p) = \rho(t) + p$ for all $t \in \mathbb{T}$. With Φ and Ψ as given, we have

$$\begin{aligned}\Psi^{\Delta}(t) &= \lim_{s \to t} \frac{\Psi(\sigma(t)) - \Psi(s)}{\sigma(t) - s} = \lim_{s \to t} \frac{\Phi(\sigma(t) + p) - \Phi(s+p)}{\sigma(t) - s} \\ &= \lim_{r \to t+p} \frac{\Phi(\sigma(t+p)) - \Phi(r)}{\sigma(t+p) - p - (r-p)} \quad (\text{since } \mathbb{T} \text{ is } p - \text{periodic}) \\ &= \Phi^{\Delta}(t+p).\end{aligned}$$

Then

$$\Psi^{\Delta}(t) = \Phi^{\Delta}(t+p) = A(t+p)\Phi(t+p) = A(t)\Psi(t),$$

completing the proof. □

We note that the derivative calculation in this proof establishes a simple chain rule in this setting. The validity of chain rule calculations in more general time scale settings is addressed in Ref. [2]. We are ready for the basic Floquet theorem.

THEOREM 1 (FLOQUET THEOREM ON TIME SCALES) *Suppose that $A : \mathbb{T} \to \mathbb{M}_n$ is rd-continuous and regressive and that both $A(t)$ and the time scale $\mathbb{T}$ are p-periodic. If $\Phi(t)$ is a fundamental matrix for Eq. (1), then there exists a p-periodic matrix $P(t)$ and a constant matrix R such that*

$$\Phi(t) = P(t)e^{Rt} \quad \textit{for all } t \in \mathbb{T}.$$

Proof For such a fundamental matrix Φ, let $\Psi(t) = \Phi(t+p)$. We have by Lemma 1 that

$$\Psi(t) = \Phi(t)C$$

for some nonsingular matrix C. Let S be a logarithm of C (that is, S is a matrix such that $e^S = C$). Define R by $R := (1/p)S$. Then $e^{pR} = C$ and

$$\Phi(t+p) = \Phi(t)e^{pR}.$$

Define $P(t)$ by

$$P(t) := \Phi(t)e^{-tR}.$$

Then

$$P(t+p) = \Phi(t+p)e^{-(t+p)R} = P(t).$$

Hence $P(t)$ is p-periodic with $\Phi(t) = P(t)e^{tR}$ for all $t \in \mathbb{T}$ proving the theorem. □

When $\mathbb{T} = \mathbb{R}$, Theorem 1 is the usual result for ordinary differential equations. The following example gives some of the details when $\mathbb{T}$ is the set of nonnegative integers.

Example 2 Suppose $\mathbb{T} = \mathbb{N} \cup \{0\}, p \in \mathbb{T}$ and $A(t)$ is both p-periodic and regressive. In this case, regressive means that $I + A(t)$ is nonsingular for $t \in \mathbb{T}$. Consider the matrix dynamic equation

$$X^{\Delta} = A(t)X. \tag{2}$$

Let Φ be the unique solution of Eq. (2) satisfying the initial condition $\Phi(0) = I$. Since $X^{\Delta}(t) = X(t+1) - X(t)$, it is easy to see iteratively that

$$\Phi(t) = \prod_{i=0}^{t-1}(I + A(i)) \text{ for } t \in \mathbb{N}$$

where the product from a lower index of 0 to an upper index of 0 is defined to be I. Define the matrix Q by

$$Q := \Phi(p) = \prod_{i=0}^{p-1}(I + A(i)).$$

By the Floquet Theorem, we get

$$\Phi(t) = P(t)e^{tR}$$

where $P(t+p) = P(t)$ and R is a constant matrix. It then follows that $P(0) = I$ and $e^{pR} = Q$ so $R = (1/p)S$ where S is a logarithm of Q. Writing $t \in \mathbb{T}$ uniquely in the form $t = kp + j$ where $0 \le j \le p - 1$, we obtain

$$\Phi(t) = \Phi(kp+j) = P(kp+j)e^{(kp+j)R} = P(j)e^{jR}(e^{pR})^k = \Phi(j)Q^k = \left[\prod_{i=0}^{j-1}(I + A(i))\right]Q^k.$$

An arbitrary solution of Eq. (2) is then of the form

$$X(t) = \left[\prod_{i=0}^{j-1}(I + A(i))\right]Q^k X(0).$$

This agrees with the representation of such a solution given in Eq. (3.130) of Ref. [1, p. 146].

The stability properties of Eq. (2) depend on the eigenvalues of Q. In particular, the zero solution of Eq. (2) is asymptotically stable if and only if $\rho(Q) < 1$ where $\rho(\cdot)$ is the spectral radius function.

3. PUTZER REPRESENTATIONS OF MATRIX LOGARITHMS

In his 1966 paper [19], Putzer proved two theorems which provide closed form representations of the matrix exponential function e^{At} where A is an $n \times n$ matrix. The first theorem expressed $F(t) = e^{At}$ in the form

$$F(t) = \sum_{i=0}^{n-1} c_{i+1}(t)A^i, \tag{3}$$

while the second gave $F(t) = e^{At}$ as

$$F(t) = \sum_{i=0}^{n-1} r_{i+1}(t)P_i, \tag{4}$$

where $\lambda_1, \ldots, \lambda_n$ is a listing of the eigenvalues of A with repeats indicating multiplicity and the matrices $P_0, P_1, \ldots, P_n$ are given by

$$\begin{cases} P_0 = I, \\ P_1 = (A - \lambda_1 I)P_0, \\ P_j = (A - \lambda_j I)P_{j-1} = \prod_{k=1}^{j}(A - \lambda_k I) \quad \text{for } 2 \leq j \leq n. \end{cases} \tag{5}$$

Here the coefficient functions $c_1(t), \ldots, c_n(t)$ in Eq. (3) and $r_1(t), \ldots, r_n(t)$ in Eq. (4) are determined by solving an initial-value problem for a simple first-order system of differential equations. For purposes of reference, we call the matrices in (5) *Putzer matrices*. Furthermore, since the right-hand side of Eq. (4) is a finite linear combination of the Putzer matrices with variable coefficients, we call Eq. (4) a *Putzer matrix representation* of $F(t)$. Similarly, since the right-hand side of Eq. (3) is a finite linear combination of powers of A with variable coefficients, we call Eq. (3) a *Putzer polynomial representation* of $F(t)$. In this section, we give both Putzer matrix representations and Putzer polynomial representations of an appropriately defined principal branch of the matrix logarithm function $F(t) = \text{Ln}(I + At)$.

Textbook presentations of Putzer's second theorem are given by Waltman [22, pp. 49–51] and Horn and Johnson [12, pp. 504–507]. Also, Kelley and Peterson [13] use a Putzer method for calculating certain matrix powers. Other recent works dealing with such representations are [7,17,18]. However, to the knowledge of the authors, the Putzer representations presented here for treating matrix logarithms are new.

Golub and Van Loan [9, p. 578] state that, "The computation of a logarithm of a matrix is an important area demanding much more work." Our methods produce case-free exact representations of the principal branch of the matrix logarithm function. When the order of the matrix is small, our exact formulas may be used as benchmarks for testing the accuracy of numerical algorithms for evaluating matrix logarithms. However, the primary uses of our representations are as theoretical tools for studying matrix logarithms, not for numerical estimation of them.

Logarithms of matrices occur in many engineering applications. Kenney and Laub [14] explain how logarithms may be used to recover an unknown coefficient matrix X in a system modeled by a linear differential equation $dy/dt = Xy$. The need to find logarithms of matrices also arises in control problems (see Refs. [6,15,16]), in parameter identification problems for Markov processes (see Ref. [21]), and in Floquet theory as shown in Section 2 of this paper. The discussion of matrix logarithms in most graduate level textbooks on ordinary differential equations, such as [10, p. 61], relies on the Jordan canonical form.

Given an $n \times n$ matrix A, we write the characteristic polynomial $p(\lambda)$ as

$$p(\lambda) = \det[\lambda I - A] = \lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_1\lambda + p_0, \tag{6}$$

or in factored form as

$$p(\lambda) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n) = \prod_{k=1}^{n}(\lambda - \lambda_k). \tag{7}$$

The Cayley–Hamilton Theorem [12, p. 86] states that an $n \times n$ matrix satisfies its characteristic equation $p(A) = 0$. A matrix U is unitary if $U^*U = I$ where U^* is the Hermitian or conjugate-transpose of U. The Schur Triangularization Theorem [12, p. 79] states that, given an $n \times n$ matrix B with eigenvalues $\lambda_1, \ldots, \lambda_n$, there exists a unitary matrix U such that U^*BU is upper triangular with $\lambda_1, \ldots, \lambda_n$ on the main diagonal. Both will be used below.

An $n \times n$ matrix Y is said to be a *logarithm* of the matrix X if and only if $e^Y = X$. Further, $\ln X$ will denote the set of all logarithms of X. Since e^Y is always nonsingular, only nonsingular matrices can have logarithms; hence, $\ln X$ is the empty set when X is singular. If X is nonsingular and $Y \in \ln X$, then $Y_n = Y + 2n\pi iI$ is also a logarithm of X for any n in the set of integers $\mathbb{Z}$. (This is since Y and $2n\pi iI$ commute so $e^{Y+2n\pi iI} = e^Y e^{2n\pi iI} = X$). Since any nonsingular matrix X has a logarithm, it follows that there are at least countably many such logarithms. In fact, it is known that the set $\ln X$ is uncountable when it is nonempty. In Gantmacher's advanced monograph on the theory of matrices, an algebraic characterization [8, p. 241] of all logarithms of a matrix X is given.

In spite of the severe nonuniqueness problem when it comes to choosing a matrix logarithm, the computation of logarithms is one-to-one in the following sense: if A and B are nonsingular $n \times n$ matrices with $A \neq B$, then the intersection of the sets $\ln A$ and $\ln B$ is empty.

The scalar geometric series $1 - x + x^2 - \cdots = 1/(1 + x)$ integrates term by term to give $\ln(1 + x) = x - x^2/2 + x^3/3 - \cdots$, valid for $|x| < 1$. By analogy, we expect that one choice of $\ln(I + tA)$ and its derivative, for small $|t|$, will be given by

$$X(t) = \ln(I + tA) = tA - t^2A^2/2 + t^3A^3/3 - \cdots \tag{8}$$

and

$$X'(t) = A - tA^2 + t^2A^3 - \cdots = (I - tA + t^2A^2 - \cdots)A = (I + tA)^{-1}A. \tag{9}$$

Thus $X(t)$ defined by the series in Eq. (8) satisfies the initial value problem

$$X'(t) = (I + tA)^{-1}A, \quad X(0) = 0. \tag{10}$$

Note that, since there are infinitely many possible logarithms of I, the constant in the series in Eq. (8) and the value of $X(0)$ in (10) could have been chosen differently. It is this choice of 0 as the value of $X(0)$ that will lead us to the principal branch of $\ln(I + tA)$.

Let $\mathbb{R}^-$ denote the set of real numbers t with $t \leq 0$. As above $\mathbb{M}_n$ denotes the set of $n \times n$ matrices with real (or complex) entries. Our principal branch of the logarithm will be defined only on the following subset of $\mathbb{M}_n$.

Definition 1 For each positive integer n, the set $\mathbb{A}_n$ of *admissible* matrices is defined to be the set of matrices in $\mathbb{M}_n$ which have no eigenvalues in $\mathbb{R}^-$.

Given any matrix $A \in \mathbb{M}_n$, we define the *dual polynomial* for A by $q(t) = \det(I + tA)$. The following lemma gives basic information about $q(t)$ that will be essential in all that follows.

LEMMA 2 *Suppose $B \in \mathbb{A}_n$ is given. Let $A = B - I$ and let $p(\lambda)$ and $q(t)$ be the characteristic and dual polynomials, respectively, for the matrix A. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. Then $q(t)$ may be written in factored form as*

$$q(t) = \prod_{i=1}^{n} (1 + \lambda_i t), \tag{11}$$

or in polynomial form as

$$q(t) = q_0 + q_1 t + \cdots + q_n t^n \quad \text{with } q_0 = q(0) = 1. \tag{12}$$

The degree of $q(t)$ is $n - m$ where m is the number of times 0 is an eigenvalue of A. The factors in Eq. (11) satisfy

$$1 + \lambda_i t \notin \mathbb{R}^- \quad \text{for } t \in [0, 1], \tag{13}$$

and consequently

$$q(t) \neq 0 \quad \text{for } t \in [0, 1]. \tag{14}$$

Also, $q(t)$ and $p(\lambda)$ are related by the identity

$$q(t) = (-t)^n p\left(-\frac{1}{t}\right) \quad \text{for } t \neq 0 \tag{15}$$

and the coefficients of $q(t)$ in Eq. (12) and of $p(\lambda)$ in Eq. (6) are related by

$$q_k = (-1)^k p_{n-k}, \quad 0 \leq k \leq n. \tag{16}$$

Proof From basic linear algebra, the eigenvalues of $I + tA$ are $1 + t\lambda_1, \ldots, 1 + t\lambda_n$ and the determinant of a matrix is the product of its eigenvalues, so Eq. (11) holds. Both Eq. (12) and the claim that the degree of $q(t)$ is $n - m$ follow directly from Eq. (11).

Now we prove (13). Fix an i with $1 \leq i \leq n$ and consider the function $f_i(t) = 1 + \lambda_i t$ for $t \in [0, 1]$. Let $\mu_i = \lambda_i + 1$. Then $\mu_i \notin \mathbb{R}^-$ since μ_i is an eigenvalue of $B = A + I$ and B is admissible. Both $f_i(0) = 1$ and $f_i(1) = \mu_i$ are not in $\mathbb{R}^-$. Suppose there is a number $t_0 \in (0, 1)$ with $f_i(t_0) = 1 + \lambda_i t_0 \in \mathbb{R}^-$; i.e., $1 + \lambda_i t_0 \leq 0$. But then λ_i is a real number with $\lambda_i \leq -1/t_0 < -1$, so $\mu_i = \lambda_i + 1$ must be in $\mathbb{R}^-$ contradicting that B is admissible. This proves (13). We obtain (14) directly from Eq. (11) and (13), since $1 + \lambda_i t \neq 0$ for $1 \leq i \leq n$, $t \in [0, 1]$.

Finally, Eq. (15) follows by putting $\lambda = -1/t$ in Eq. (7) and comparing that result to Eq. (11). From Eq. (15),

$$q(t) = (-t)^n \cdot p\left(-\frac{1}{t}\right) = 1 - p_{n-1} t + \cdots + (-1)^n p_0 t^n,$$

valid for $t \neq 0$. Then the polynomial difference $q(t) - [1 - p_n t + \cdots + (-1)^n p_0 t^n]$ is zero for all $t \neq 0$ and hence must be identically zero. This proves Eq. (16). □

Now suppose $B \in \mathbb{A}_n$ is given and $A = B - I$. From (14), the matrix $I + tA$ is invertible for all $t \in [0, 1]$ and, from the cofactor method of calculating inverses, all entries in $(I + tA)^{-1}$

are continuous on the interval $0 \le t \le 1$. Hence, the initial-value problem (10) has a unique solution $X(t)$ valid on the interval $0 \le t \le 1$. Furthermore, (10) is equivalent both to the initial-value problem

$$(I + tA)X'(t) = A, \quad X(0) = 0, \tag{17}$$

and to the integral expression

$$X(t) = \int_0^t (I + sA)^{-1} A \, ds. \tag{18}$$

Definition 2 If $B \in \mathbb{A}_n$ and $A = B - I$, then we define the *principal logarithm* of the matrix $I + tA$, denoted by $\mathrm{Ln}(I + tA)$ for $0 \le t \le 1$, to be the solution $X(t)$ of the initial-value problem (10). Then $\mathrm{Ln}\, B = X(1)$.

Before proceeding to the Putzer representations, we note that $\mathrm{Ln}\, z$ for $z \in \mathbb{C}$ will denote the principal branch of the logarithm of complex numbers (that is, $\mathrm{Ln}\, z = \log_e(|z|) + \theta$ where the argument θ of z is chosen in the interval $-\pi < \theta < \pi$). If $\lambda \in \mathbb{C}$ and $1 + \lambda \notin \mathbb{R}^-$, then $z(t) = 1 + \lambda t$, $0 \le t \le 1$ describes a smooth arc (line segment) from the point $z = 1$ to the point $z = 1 + \lambda$ which does not contain any point in $\mathbb{R}^-$. From basic complex variables (see Ref. [4]), for $0 \le t \le 1$ and k a positive integer,

$$\int_0^t \frac{\lambda}{(1 + \lambda s)^k} \, ds = \begin{cases} \mathrm{Ln}(1 + \lambda t) & \text{if } k = 1, \\ \frac{1}{k-1}\left(1 - \frac{1}{(1+\lambda t)^{k-1}}\right) & \text{if } k \ge 2. \end{cases} \tag{19}$$

We now give our first representation.

THEOREM 3 (PUTZER MATRIX REPRESENTATION) *Suppose $B \in \mathbb{A}_n$, $A = B - I$, $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, and $P_0, \ldots, P_n$ are the Putzer matrices in (5). Then $X(t) = \mathrm{Ln}(I + At)$ for $0 \le t \le 1$ has a Putzer sum representation $X(t) = \sum_{i=0}^{n-1} r_{i+1}(t) P_i$ provided that $r_1(t), \ldots, r_n(t)$ satisfy the first-order system of differential equations*

$$\begin{cases} (1 + t\lambda_1) r_1'(t) = \lambda_1, \\ (1 + t\lambda_2) r_2'(t) = -t r_1'(t) + 1, \\ (1 + t\lambda_{i+1}) r_{i+1}'(t) = -t r_i'(t) \quad \text{for } 2 \le i \le n - 1 \end{cases} \tag{20}$$

and the initial conditions

$$r_1(0) = \cdots = r_n(0) = 0. \tag{21}$$

Solutions for $r_1(t), \ldots, r_n(t)$ are given by

$$r_1(t) = \int_0^t \frac{\lambda_1}{1 + \lambda_1 s} \, ds = \mathrm{Ln}(1 + \lambda_1 t), \tag{22}$$

$$r_k(t) = \int_0^t \frac{(-1)^{k-2} s^{k-2}}{\prod_{j=1}^k (1 + \lambda_j s)} \, ds, \quad \text{for } 2 \le k \le n. \tag{23}$$

Proof The system (20) results from substituting $X(t) = \sum_{i=0}^{n-1} r_{i+1}(t) P_i$ into the differential equation (17), writing $A = \lambda_1 P_0 + P_1$ for the right-hand side, regrouping terms on

the left-hand side, and equating the coefficients of P_0 on both sides to get the first equation in (20), equating coefficients of P_1 to get the second equation, and so on. Clearly, the initial condition in (17) holds when (21) holds.
We solve for $r_1'(t)$ in the first equation and integrate as in Eq. (19) to get Eq. (22). We then substitute for $r_1'(t)$ on the right-hand of the second equation, and solve to get $r_2'(t) = 1/[(1 + t\lambda_1)(1 + t\lambda_2)]$. Continuing this recursive procedure leads to

$$r_k'(t) = \frac{(-1)^{k-2} t^{k-2}}{\prod_{j=1}^{k} (1 + \lambda_j t)}$$

and integration leads to (23). Note that all the denominators in the integrands in Eqs. (22) and (23) are nonzero by (13). □

The integrals in (23) can all be calculated explicitly by expanding the integrands in partial fractions and using Eq. (19). This results in closed-form exact representations of $\mathrm{Ln}(I + tA)$. Two special cases are covered in the following corollary.

COROLLARY 1 *Suppose* $X(t) = \sum_{i=0}^{n-1} r_{i+1}(t) P_i$ *is the Putzer sum in Theorem 3.*

1. *(Distinct Eigenvalues) If* $\lambda_1, \ldots, \lambda_n$ *are distinct, then (23) becomes*

$$r_k(t) = \sum_{j=1}^{k} \frac{1}{\prod\limits_{1 \le i \le k,\, i \ne j} (\lambda_j - \lambda_i)} \mathrm{Ln}(1 + \lambda_j t) \quad \text{for } 2 \le k \le n. \tag{24}$$

2. *(All Eigenvalues Equal) If* $\lambda_1 = \cdots = \lambda_n$, *then (23) becomes*

$$r_k(t) = \frac{(-1)^{k-2} t^{k-1}}{(k-1)(1 + \lambda_1 t)^{k-1}} \quad \text{for } 2 \le k \le n. \tag{25}$$

Proof In the case of distinct eigenvalues, use a partial fractions expansion to get (24) from (23). When all eigenvalues are equal, use mathematical induction together with an integration by parts to get (25). □

Example 4 If we use Part 2 of Corollary 1 to find an exact representation of $\mathrm{Ln}\, B$ when $B = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$, we find $A = B - I = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, $\lambda_1 = \lambda_2 = 2$, $P_0 = I$, $P_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $r_1(t) = \mathrm{Ln}(1 + 2t)$, $r_2(t) = t/(1 + 2t)$, and

$$\mathrm{Ln}(I + At) = r_1(t) P_0 + r_2(t) P_1 = \begin{bmatrix} \mathrm{Ln}(1 + 2t) & t/(1 + 2t) \\ 0 & \mathrm{Ln}(1 + 2t) \end{bmatrix}, \quad 0 \le t \le 1.$$

Letting $t = 1$ gives $\mathrm{Ln}\, B = \begin{bmatrix} \mathrm{Ln}\, 3 & 1/3 \\ 0 & \mathrm{Ln}\, 3 \end{bmatrix}$.

Numerical coding of these exact formulas provides a benchmark for the computation of matrix logarithms. Matlab's "Version 6" returns excellent numerical results when the routine **logm** is applied to the matrix B in Example 4, but "Version 4" and earlier Versions gave relatively poor estimates for the same matrix B. The difficulty in the numerical calculation of logarithms of Jordan blocks has been recently resolved in the work of Kenney and Laub [14]. We note here that conditions (20) and (21) are sufficient for the Putzer matrix representation $\sum_{i=0}^{n-1} r_{i+1}(t)P_i$ to equal $X(t) = \mathrm{Ln}(I + tA)$, but are not necessary because Putzer matrix representations do not always have unique coefficients. For example, if A is the zero matrix, then $r_2(t),\ldots,r_n(t)$ are arbitrary since $P_1,\ldots,P_{n-1}$ are zero matrices. The same remarks apply to the Putzer polynomial representations that we now develop.

THEOREM 5 (PUTZER POLYNOMIAL REPRESENTATION) *Suppose $B \in \mathbb{A}_n$ and $A = B - I$. Then there exist continuously differentiable functions $c_i(t)$, $1 \le i \le n$, such that*

$$\mathrm{Ln}(I + tA) = \sum_{i=0}^{n-1} c_{i+1}(t)A^i \tag{26}$$

for $0 \le t \le 1$. The initial conditions $c_i(0) = 0$ make $\mathrm{Ln}\, I = 0$. For $n = 1$ we have $c_1(t) = r_1(t)$ since this sum is the same as the Putzer matrix representation. For $n > 1$ matching coefficients in $(I + tA)X'(t) = A$ leads to the system

$$\begin{aligned} c_1'(t) &= tc_n'(t)p_0, \\ c_2'(t) &= 1 - tc_1'(t) + tc_n'(t)p_1, \\ c_{i+1}'(t) &= -tc_i'(t) + tc_n'(t)p_i, \quad 2 \le i \le n-1 \end{aligned} \tag{27}$$

which has solution

$$\begin{aligned} &c_1'(t) = -\frac{(-t)^{n-2}}{q(t)}[-tp_0], \quad c_k'(t) = (-t)^{k-2} - \frac{(-t)^{n-2}}{q(t)}\sum_{j=1}^{k}(-t)^j p_{k-j}, \\ &2 \le k \le n-1, \quad c_n'(t) = \frac{(-t)^{n-2}}{q(t)}, \end{aligned} \tag{28}$$

where $q(t)$ is the dual polynomial. Then $c_i(0) = 0$ implies $c_i(t) = \int_0^t c_i'(s)\,\mathrm{d}s$ for each i.

Proof The existence of the $c_i(t)$ is a corollary to the Putzer matrix representation of $\mathrm{Ln}(I + tA)$. Indeed, for $r_i(t)$ as in Theorem 3 and $X(t) = \sum_{i=0}^{n-1} r_{i+1}(t)P_i$, replace each P_i for $1 \le i \le n-1$ by $P_i = \prod_{i=1}^{n}(A - \lambda_i I)$. For $n \ge 2$, a direct choice of the coefficients is obtained by matching coefficients of A^i after substitution of $X(t) = \sum_{i=0}^{n-1} c_{i+1}(t)A^i$ in $(I + tA)X'(t) = A$ and use of the Cayley–Hamilton Theorem to write $A^n = -\sum_{i=0}^{n-1} p_i A^i$.

Now $c_1'(t)$ is in terms of $c_n'(t)$. The system is solved by solving for $c_2'(t)$, etc., in terms of $c_n'(t)$. This gives

$$\begin{aligned} c_1'(t) &= -c_n'(t)[-tp_0], \\ c_2'(t) &= 1 - c_n'(t)[(-t)^2p_0 + (-t)p_1], \\ c_k'(t) &= (-t)^{k-2} - c_n'(t)\sum_{j=1}^{k}(-t)^j p_{k-j} \quad \text{for } 2 \le k \le n. \end{aligned} \tag{29}$$

For the case of $k = n$, bring all the $c_n'(t)$ terms to the left side for the equation

$$c_n'(t)\left[1+\sum_{j=1}^{n} t^j(-1)^j p_{n-j}\right] = (-t)^{n-2} \tag{30}$$

and, using Eq. (16), the coefficient of $c_n'(t)$ is the dual polynomial $q(t)$. Replace $c_n'(t)$ in each equation for the final solution. □

It is important to note that this Putzer polynomial representation of the logarithm does not explicitly require the eigenvalues, but it needs only the similarity invariants p_i which are the coefficients of the characteristic polynomial of A.

Remark 6 We note that both the Putzer matrix representation and the Putzer polynomial representation are easily implemented as calculational algorithms. Here, one can use Computer algebra systems such as *Mathematica* and *Maple* to find closed-form solutions to the initial-value problem (20) and (21), to produce exact representations of the integrals (22) and (23), to first calculate $c_1'(t),\ldots,c_n'(t)$ via (28), and finally to calculate $c_i(t) = \int_0^t c_i'(s)\,\mathrm{d}s$.

If the eigenvalues of B are real, then the p_i are real and the $c_i(t)$ are real. Furthermore, this representation allows us to rewrite the integral definition (18) as [23, p. 134]

$$\mathrm{Ln}(I+tA) = \int_0^t A(I+sA)^{-1}\,\mathrm{d}s \tag{31}$$

for admissible $B = I + A$. Indeed, for $X(t) = \mathrm{Ln}(I + tA)$, we have from Eq. (26) that $X'(t) = \sum_{i=0}^{n-1} c_{i+1}'(t)A^i$ and consequently that A commutes with $X'(t)$. Hence, $(I + tA)X'(t) = X'(t)\times(I + tA) = A$ for $t \in [0, 1]$ which then gives Eq. (31).

So far, we have only shown the existence of a logarithm for admissible matrices. Given a matrix D with eigenvalues $\lambda_1 = r_1\mathrm{e}^{i\theta_1},\ldots,\lambda_n = r_n\mathrm{e}^{i\theta_n}$ and a real number δ, the matrix $\mathrm{e}^{i\delta}D$ has eigenvalues $r_1\mathrm{e}^{i(\delta+\theta_1)},\ldots,r_n\mathrm{e}^{i(\delta+\theta_n)}$ which means all eigenvalues of D are rotated through an angle δ to obtain eigenvalues of $\mathrm{e}^{i\delta}D$. Hence, if D is nonsingular, then zero is not an eigenvalue of D and it is easy to choose $\delta \in \mathbb{R}$ so that $\mathrm{e}^{i\delta}D$ is admissible.

Definition 3 Given $D \in \mathbb{M}_n$ and $\delta \in \mathbb{R}$, we say D is *δ-admissible* if the matrix $\mathrm{e}^{i\delta}D$ is admissible. If D is δ-admissible, we define $\mathrm{Ln}(D,\delta)$ by

$$\mathrm{Ln}(D,\delta) = \mathrm{Ln}(\mathrm{e}^{i\delta}D) - i\delta I.$$

When D is δ-admissible, D cannot have zero as an eigenvalue and must be nonsingular. The following theorem gives an assortment of logarithm properties.

THEOREM 7 *Suppose the matrices B, C, D, E, and P are in $\mathbb{M}_n$, B and C are admissible, D is δ-admissible, E is ε-admissible and P is nonsingular. Then the following hold:*

1. *Inverse Mapping Property.* $\mathrm{Ln}\,B \in \ln B$ *and* $\mathrm{Ln}(D,\delta) \in \ln D$; *i.e.,*

$$\mathrm{e}^{\mathrm{Ln}\,B} = B \text{ and } \mathrm{e}^{\mathrm{Ln}(D,\delta)} = D.$$

2. *Principal Log Property. If B has eigenvalues $\lambda_1,\ldots,\lambda_n$, then $\mathrm{Ln}\,B$ has eigenvalues $\mathrm{Ln}\,\lambda_1,\ldots,\mathrm{Ln}\,\lambda_n$ where $\mathrm{Ln}\,z$ is the principal branch of the complex logarithm. Hence, all eigenvalues λ of $\mathrm{Ln}\,B$ satisfy $-\pi < \mathrm{Im}\lambda < \pi$, and our principal branch of the matrix logarithm agrees with the principal branch as defined by others (e.g., [14, p. 644]).*

3. *General Properties.* $\bar{B}$, B^T, *and* B^* *are admissible with* $\operatorname{Ln}\bar{B} = \overline{\operatorname{Ln}B}$, $\operatorname{Ln}(B^T) = (\operatorname{Ln}B)^T$, *and* $\operatorname{Ln}(B^*) = (\operatorname{Ln}B)^*$. *Also* $P^{-1}BP$ *is admissible with* $\operatorname{Ln}(P^{-1}BP) = P^{-1}(\operatorname{Ln}B)P$. *Ln B is real if B is real, Ln B is real symmetric if B is real symmetric and Ln B is upper triangular if B is upper triangular. If* $B = diag(b_1, \ldots, b_n)$, *then* $\operatorname{Ln}B = diag(\operatorname{Ln}b_1, \ldots, \operatorname{Ln}b_n)$.
4. *Inverse Matrix Property.* B^{-1} *is admissible and* D^{-1} *is* $(-\delta)$*-admissible with* $\operatorname{Ln}(B^{-1}) = -\operatorname{Ln}B$ *and* $\operatorname{Ln}(D^{-1}, -\delta) = -\operatorname{Ln}(D, \delta)$.
5. *Commutativity. Any matrix which commutes with B must also commute with Ln B. Ln B and Ln C commute if and only if B and C commute. If B and C commute, then* $\operatorname{Ln}B + \operatorname{Ln}C \in \ln(BC)$. *If D and E commute, then* $\operatorname{Ln}(D, \delta) + \operatorname{Ln}(E, \varepsilon) \in \ln(DE)$.
6. *Monotoneity Properties. Assume additionally that B and C are real symmetric and* $B = I + A$ *with A nonsingular. Then*

$$\operatorname{Ln}(B) = \int_0^1 (A^{-1} + tI)^{-1}\mathrm{d}t.$$

If $B > I$, *then* $\operatorname{Ln}B > 0$. *If* $0 < B < I$, *then* $\operatorname{Ln}B < 0$. *If* $B > C > I$, *then* $\operatorname{Ln}B > \operatorname{Ln}C$. *If* $0 < B < C < I$, *then* $\operatorname{Ln}B < \operatorname{Ln}C < 0$.

Proof of the Inverse Mapping Property First we prove for any scalar differentiable function $g(t)$ of the real variable t that $(\mathrm{d}/\mathrm{d}t)[\mathrm{e}^{g(t)A}] = g'(t)Ae^{g(t)A}$. We use a Putzer polynomial representation (26) of e^{At} to write $\mathrm{e}^{At} = \sum_{i=0}^{n-1} c_{i+1}(t)A^i$. Then $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} = Ae^{At} = \sum_{i=0}^{n-1} c'_{i+1}(t)A^i$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{g(t)A}\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=0}^{n-1} c_{i+1}(g(t))A^i\right] = \sum_{i=0}^{n-1} c'_{i+1}(g(t))g'(t)A^i = g'(t)Ae^{g(t)A}.$$

Now let $X(t) = \sum_{i=0}^{n-1} r_{i+1}(t)P_i$ be a Putzer matrix representation of $\operatorname{Ln}(I + At)$. From (5), P_i commutes with P_j so we obtain

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)\mathrm{e}^{-X(t)} = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)\exp\left(-\sum_{i=0}^{n-1} r_{i+1}(t)P_i\right) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)\left(\prod_{i=0}^{n-1}\exp(-r_{i+1}(t)P_i)\right)$$

$$= \left(-\sum_{i=0}^{n-1} r'_{i+1}(t)P_i\right)\left(\prod_{i=0}^{n-1}\exp(-r_{i+1}(t)P_i)\right) = -X'(t)\mathrm{e}^{-X(t)}.$$

Thus,

$$\left[(I + tA)\mathrm{e}^{-X(t)}\right]' = [A - (I + tA)X'(t)]\mathrm{e}^{-X(t)} = 0$$

since $X'(t) = (I + tA)^{-1}A$. The function $(I + tA)\mathrm{e}^{-X(t)}$ then has the constant value of I, since that is its value at $t = 0$. Hence $\mathrm{e}^{X(t)} = (\mathrm{e}^{-X(t)})^{-1} = I + tA$. Then $\mathrm{e}^{\operatorname{Ln}(I+tA)} = I + tA$ which, for $t = 1$, gives $\mathrm{e}^{\operatorname{Ln}B} = B$. For D, $\mathrm{e}^{\operatorname{Ln}(D,\delta)} = \mathrm{e}^{\operatorname{Ln}(\mathrm{e}^{i\delta}D)-(i\delta)I} = \mathrm{e}^{\operatorname{Ln}(\mathrm{e}^{i\delta}D)}\mathrm{e}^{-i\delta I} = \mathrm{e}^{i\delta}D\cdot\mathrm{e}^{-i\delta}I = D$. □

Proof of Principal Log Property Let B have eigenvalues $\lambda_1, \ldots, \lambda_n$. Using Schur's Theorem, find a unitary matrix U so that $T = U^{-1}(B - I)U$ is an upper triangular matrix with the eigenvalues of $B-I$ as diagonal entries; i.e., $T_{ii}=\lambda_i-1$ for $1\le i\le n$. Let $S=U^{-1}(\operatorname{Ln}B)U$. Then S is similar to $\operatorname{Ln}B$ and

$$\begin{aligned}
S=U^{-1}(\operatorname{Ln}B)U&=\int_0^1 U^{-1}(I+s(B-I))^{-1}(B-I)U\,\mathrm{d}s\\
&=\int_0^1 U^{-1}(I+s(B-I))^{-1}UU^{-1}(B-I)U\,\mathrm{d}s=\int_0^1[U^{-1}(I+s(B-I))U]^{-1}T\,\mathrm{d}s\\
&=\int_0^1[I+sT]^{-1}T\,\mathrm{d}s.
\end{aligned}$$

Hence, S is upper triangular which, together with Eq. (19), leads to

$$S_{ii}=\int_0^1\frac{\lambda_i-1}{1+s(\lambda_i-1)}\,\mathrm{d}s=\operatorname{Ln}\lambda_i.$$

Then the eigenvalues of $\operatorname{Ln}B$ are $\operatorname{Ln}\lambda_i$, $1\le i\le n$, as claimed. □

Proof of Inverse Matrix Property Set $A\equiv B-I$ and $Y\equiv B^{-1}-I$. Then $Y=(I+A)^{-1}-I=(I+A)^{-1}(-A)=-(I+A)^{-1}A$ and

$$\begin{aligned}
\operatorname{Ln}(B^{-1})=\operatorname{Ln}(I+Y)&=\int_0^1(I+sY)^{-1}Y\,\mathrm{d}s=\int_0^1[I-s(I+A)^{-1}A]^{-1}(I+A)^{-1}(-A)\,\mathrm{d}s\\
&=\int_0^1[(I+A)(I-s(I+A)^{-1}A)]^{-1}(-A)\,\mathrm{d}s=\int_0^1[I+A-sA]^{-1}(-A)\,\mathrm{d}s\\
&=-\int_0^1[I+(1-s)A]^{-1}A\,\mathrm{d}s.
\end{aligned}$$

Then set $\tau=1-s$ for $\operatorname{Ln}(B^{-1})=\int_1^0[I+\tau A]^{-1}A\,\mathrm{d}\tau=-\operatorname{Ln}B$. Also, $\operatorname{Ln}(D^{-1},-\delta)=\operatorname{Ln}(\mathrm{e}^{-i\delta}D^{-1})+i\delta I=-(\operatorname{Ln}(\mathrm{e}^{i\delta}D)-i\delta I)=-\operatorname{Ln}(D,\delta)$. □

The proofs of the other parts of Theorem 7 are routine and are omitted.

4. SPECIAL LOGARITHMS

One often wants to preserve some property or structure in the choice of a logarithm of a given matrix. For example, in some applications, a real matrix arises and one wants to find a real logarithm. Producing these real logarithms by automated packages is difficult. In fact, as the following example shows, there may not be a real logarithm.

Example 8 A real (or complex) $n\times n$ matrix Y with $\det Y<0$ cannot have a real logarithm. In particular, the negative of an $n\times n$ identity matrix of odd dimension has no real logarithm. This follows from the trace formula [12, p. 439].

Culver [5] has given a complete characterization, using the Jordan normal form, of all real matrices that have real logarithms and we refer the interested reader to his paper for details. The following theorem gives a criterion under which a possibly complex matrix of the form

$C\bar{C}$ will have a real logarithm. It applies in particular to the square of a nonsingular real matrix.

THEOREM 9 (THE C^2 PROBLEM) *Suppose C is nonsingular $n \times n$ and commutes with its complex conjugate $\bar{C}$. Then there exist real $n \times n$ matrices U and V which commute such that*

$$C = e^{U+iV} = e^U e^{iV} \quad \text{and} \quad \bar{C} = e^{U-iV} = e^U e^{-iV}.$$

Then $e^{2U} = C\bar{C}$ so $2U$ is a real logarithm of $C\bar{C}$. In particular, if C is nonsingular and has real entries, then there exists a real matrix R such that $C^2 = e^R$.

Proof Suppose C is nonsingular with $C\bar{C} = \bar{C}C$. Choose $\alpha \in \mathbb{R}$ so that C is α-admissible. Let $B \equiv e^{i\alpha}C$ and $A = B - I$. Then the matrices $A = e^{i\alpha}C - I$ and $\bar{A} = e^{-i\alpha}\bar{C} - I$ commute and $B\bar{B} = \bar{B}B$ so $\mathrm{Ln}(\bar{B}) = \overline{\mathrm{Ln}\, B}$ commutes with $\mathrm{Ln}\, B$. Set $W = \mathrm{Ln}(C, \alpha) = \mathrm{Ln}\, B - i\alpha I$, $U = (1/2)[W + \overline{W}]$, and $V = (1/(2i))[W - \overline{W}]$. Then $W = U + iV = \mathrm{Ln}(C, \alpha)$ with U and V having real entries. Since $\overline{LnB}$ commutes with $\mathrm{Ln}\, B$, it follows that U commutes with V. Thus $C = e^{\mathrm{Ln}(C,\alpha)} = e^{\mathrm{Ln}\, B - i\alpha I} = e^{U+iV} = e^U e^{iV}$ and $\bar{C} = e^{U-iV} = e^U e^{-iV}$ as claimed. If C is real, then $\bar{C}$ commutes with C and $C^2 = e^{2U}$ so $R = 2U$ is a real logarithm of C^2. □

As in calculus, logarithms can be used to define roots and powers. We elaborate on that as follows.

Example 10 (Principal Roots) Suppose p is a positive real number and B is nonsingular. Define the set of p-th roots of B by

$$B^{1/p} = \left\{ Y : Y = e^{(1/p)X} \quad \text{for some } X \in \ln B \right\}.$$

Define for B admissible and $p > 0$, the principal branch of the p-th root by $\sqrt[p]{B} = e^{(1/p)\mathrm{Ln}\, B}$. If C is α-admissible and $p > 0$, define the function *Root* by

$$Root(C, p, \alpha) = e^{-i\alpha/p} \sqrt[p]{B} \quad \text{for } B = e^{i\alpha}C.$$

For $m \in \mathbb{Z}^+$, the set of positive integers, one can show that:

1. If B is admissible, then $\sqrt[m]{B} \in B^{1/m}$ and $(\sqrt[m]{B})^m = B$.
2. If C is α-admissible, then $(Root(C, m, \alpha))^m = C$.
3. If $B > 0$ then $\sqrt[m]{B} > 0$ and, if $X > 0$ satisfies $X^m = B$, then $X = \sqrt[m]{B}$. For the uniqueness of the m-th root, use the fact that a positive definite matrix is diagonalizable to find a nonsingular matrix Q such that $Q^{-1}XQ = D$, a diagonal matrix with positive real diagonal entries δ_k. Then $D^m = Q^{-1}X^mQ$, $Q^{-1}(\mathrm{Ln}\, B)Q = \mathrm{Ln}(D^m)$, $Q^{-1}\sqrt[m]{B}Q = e^{(1/m)\mathrm{Ln}(D^m)}$, and finally,

$$\sqrt[m]{B} = Q\mathrm{diag}\left\{ e^{(1/m)\mathrm{Ln}((\delta_k)^m)} \right\} Q^{-1} = QDQ^{-1} = X.$$

Householder matrices may be used to produce many interesting examples. Part 6 of the following shows that $\ln I$ is an uncountable set. Part 7 gives an example where A and B commute, $Y_1 \in \ln A$, $Y_2 \in \ln B$, and both A and B commute with Y_1 and with Y_2, yet $Y_1 + Y_2 \notin \ln(AB)$; a situation that strongly contrasts with properties of the principal branch of the matrix logarithm.

Example 11 (Householder Matrix Examples) Suppose that $u \in \mathbb{R}^n$ is a unit column vector in the 2-norm, i.e. $u^T u = 1$. Let H be the real Householder matrix (i.e. elementary reflector as in Ref. [9]) $H = I - 2uu^T$. Then the following hold:

1. uu^T has rank 1 and $uu^T u = u$. Thus uu^T has $n-1$ eigenvalues of 0 and one eigenvalue of 1.
2. H has one eigenvalue of -1 and $n-1$ eigenvalues of 1; hence $\det H = -1$.
3. $H^T = H$, $H^2 = I$, and $H^{-1} = (H^T)^{-1} = H$.
4. For $n = 2$, H is real Householder if and only if H has the form $\begin{bmatrix} c & s \\ s & -c \end{bmatrix}$ with $c^2 + s^2 = 1$.
5. For $n > 2$ and H as in Part (4), $\tilde{H} = \operatorname{diag}\{H, I_{n-2}\}$ is real Householder. Hence for $n \geq 2$, $I^{1/2}$ contains an uncountable set of real Householder matrices.
6. $X = \operatorname{Ln}(H, \pi/2)$ gives $e^X = H$ and $e^{2X} = I$. Thus $\ln I$ is an uncountable set. This result is due to Helton [11, p. 734].
7. Consider $H_1 = [0.6, 0.8; 0.8, -0.6]$ and $H_2 = [0.8, 0.6; 0.6, -0.8]$. Use of the following Matlab code (and Matlab's **expm**)

```
function Y=logput2d(X)
%gives Putzer log of 2 by 2 with distinct e.vals.
A=X-eye(2);lambda=eig(A);r1=log(1+lambda(1));
r2=(log(1+lambda(1))-log(1+lambda(2)))/(lambda(1)-lambda(2));
P0=eye(2);P1=A-lambda(1)*eye(2);
Y=r1*P0+r2*P1;
```

   leads to $B_1 = -i{*}H_1$, $B_2 = -i{*}H_2$, $Y_1 = 2(\operatorname{Ln}(B_1) + i(\pi/2)I)$, and $Y_2 = 2(\operatorname{Ln}(B_2) + i(\pi/2)I)$ give $e^{Y_1} = e^{Y_2} = I$ but $e^{Y_1+Y_2} \neq I$. Thus Y_1 and Y_2 are in $\ln I$, but $Y_1 + Y_2 \notin \ln I$. Hence, even when A and B commute with Y_1 in $\ln A$ and Y_2 in $\ln B$, it is not always true that $Y_1 + Y_2$ is in $\ln(AB)$.

The following example shows some of the structure of interest when finding logarithms.

Example 12 (Symplectic Versus Hamiltonian) We collect together some facts and leave the details to the reader. Suppose that J is the $2n \times 2n$ skew symmetric matrix $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. A matrix M is said to be **Hamiltonian** if it is $2n \times 2n$ and $M^* J + JM = 0$. A matrix T is said to be **symplectic** if it is $2n \times 2n$ and $T^* JT = J$.

1. *Radon* [20]. Suppose that for each real t, $M(t)$ is Hamiltonian and $Y(t)$ is the solution of the initial value problem $Y'(t) = M(t)Y(t)$, $Y(0) = I_{2n}$. Then $Y(t)$ is symplectic for each real t. This can be proved by showing that the expression $Y^*(t)JY(t) - J$ has derivative zero; hence is constant and therefore is the same for all t as for $t = 0$.
2. If $M_{2n\times 2n}$ is partitioned as $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, with $n \times n$ block entries A, B, C, D, then M is Hamiltonian if and only if B and C are Hermitian and $D = -A^*$.
3. If $T_{2n\times 2n}$ is partitioned as $T = \begin{bmatrix} E & F \\ G & H \end{bmatrix}$, with $n \times n$ block entries E, F, G, H, then T is symplectic if and only if the following three conditions are satisfied: $E^* G = G^* E$, $F^* H = H^* F$, $E^* H - G^* F = I$.

4. If M is a constant $2n \times 2n$ Hamiltonian matrix and t is real, then tM is Hamiltonian and $T(t)$ defined by $T(t) = e^{tM}$ is symplectic.
5. Finally J is Hamiltonian. It follows from consideration of the initial value problem $Y'(t) = JY(t)$, $Y(0) = I_{2n}$, that e^{tJ} is the symplectic matrix

$$e^{tJ} = \begin{bmatrix} (\cos t)I & (\sin t)I \\ -(\sin t)I & (\cos t)I \end{bmatrix}. \tag{32}$$

Then, selecting t properly, each of the matrices I_{2n}, $-I_{2n}$, J, and $-J$ has a countable infinity of real logs of the form tJ.

References

[1] C. D. Ahlbrandt and A. C. Peterson, *Discrete Hamiltonian Systems*, Kluwer Academic Publishers, Boston, 1996.
[2] C. D. Ahlbrandt, M. Bohner and J. Ridenhour, Hamiltonian systems on time scales, *J. Math. Anal. Appl.*, **250** (2000), 561–578.
[3] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser, Boston, 2001.
[4] J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, 6th Ed., McGraw-Hill, New York, 1996.
[5] W. J. Culver, On the existence and uniqueness of the real logarithms of a matrix, *Proc. Am. Math. Soc.*, **17** (1966), 1146–1151.
[6] W. J. Culver, An analytic theory of modeling for a class of minimal-energy control systems (disturbance-free case), *SIAM J. Control*, **2** (1964), 267–294.
[7] S. Elaydi and W. A. Harris Jr, On the computation of A^N, *SIAM Rev.*, **40** (1998), 965–971.
[8] F. R. Gantmacher, *The Theory of Matrices*, **Vol. 1** Chelsea, New York, 1960.
[9] G. Golub and C. Van Loan, *Matrix Computations*, 3rd Ed., Johns Hopkins Press, Baltimore, 1996.
[10] P. Hartman, *Ordinary Differential Equations*, 2nd Ed., Birkhäuser, Boston, 1982.
[11] B. W. Helton, Logarithms of matrices, *Proc. Am. Math. Soc.*, **19** (1968), 733–738.
[12] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[13] W. G. Kelley and A. C. Peterson, *Difference Equations, An Introduction with Applications*, Academic Press, San Diego, CA, 1991.
[14] C. S. Kenney and A. J. Laub, A Schur-Fréchet algorithm for computing the logarithm and exponential of a matrix, *SIAM J. Math. Anal. Appl.*, **19** (1998), 640–663.
[15] G. J. Lastman and N. K. Sinha, Transformation algorithm for identification of continuous-time multivariable systems from discrete data, *Electron. Lett.*, **17** (1981), 779–780.
[16] 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, *Electron. Lett.*, **27**(16) (1991), 1468–1470.
[17] I. E. Leonard, The matrix exponential, *SIAM Rev.*, **38** (1996), 507–512.
[18] E. Liz, A note on the matrix exponential, *SIAM Rev.*, **40** (1998), 700–702.
[19] E. J. Putzer, Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients, *Am. Math. Monthly*, **73** (1966), 2–7.
[20] J. Radon, *Zum Problem von Lagrange: vier Vortrage gehalten im Mathematischen Seminar der Hamburgischen Universitat, Hamburger Mathematische Einzelschriften*, B.G. Teubner, Leipzig, **6 Heft.**, 1928.
[21] B. Singer and S. Spilerman, The representation of social processes by Markov models, *Am. J. Sociol.*, **82** (1976), 1–54.
[22] P. Waltman, *A Second Course in Elementary Differential Equations*, Academic Press, New York, 1986.
[23] A. Wouk, Integral representation of the logarithm of matrices and operators, *J. Math. Anal. Appl.*, **11** (1965), 131–138.