

## THE EXPONENTIAL OF A CONSTANT MATRIX ON TIME SCALES

A. ZAFER<sup>1</sup>

(Received 8 March, 2006; revised 1 June, 2006)

### Abstract

In this paper we describe an elementary method for calculating the matrix exponential on an arbitrary time scale. An example is also given to illustrate the result.

2000 *Mathematics subject classification*: primary 34A30.

*Keywords and phrases*: time scale, matrix exponential, linear dynamic equation.

### 1. Introduction

There are several methods in the literature to calculate the matrix exponential  $e^{tA}$  and the matrix  $A^k$  for any given  $n \times n$  matrix  $A$ , due to their appearance in the study of systems of linear differential and difference equations, respectively. In 1996, I. E. Leonard [7] presented an elementary but powerful method to calculate the matrix exponential  $e^{At}$  which uses *only* knowledge of homogeneous linear differential equations with constant coefficients and the Cayley-Hamilton theorem. Two years later, by using a similar approach and employing homogeneous linear difference equations with constant coefficients, M. Kwapisz [6] derived an alternative method to determine  $A^k$ , the  $k$ -th power of  $A$ . The methods are especially useful when the matrix  $A$  is not diagonalisable.

A time scale  $\mathbb{T}$  is a nonempty closed subset of the set of real numbers  $\mathbb{R}$ . Time-scale calculus, introduced by Hilger [3], has recently gained considerable interest as it unifies continuous and discrete analysis. The most well-known examples of a time scale are  $\mathbb{T} = \mathbb{R}$ ,  $\mathbb{T} = \mathbb{Z}$  and  $\mathbb{T} = \overline{q^{\mathbb{Z}}}$ , where  $q^{\mathbb{Z}} = \{t : t = q^k, k \in \mathbb{Z}\}$  with  $q > 1$ .

In the light of time-scale calculus, our aim is to develop an elementary method to calculate the matrix exponential on an arbitrary time scale, and thereby unify the

---

<sup>1</sup>Department of Mathematics, Middle East Technical University, Ankara, 06531, Turkey; e-mail: zafer@metu.edu.tr

results in [6, 7]. To the best of our knowledge the Putzer algorithm is the only method available to calculate the exponential of a constant matrix on an arbitrary time scale, see [1, Theorem 5.35]. Here in this article we propose an alternative method which is more elementary and easier to apply than the Putzer algorithm.

For some basic aspects of time-scale calculus we refer to the monographs [1, 2, 5]. In the following lines we provide only some essential ingredients to be used in this paper.

The forward and backward jump operators  $\sigma, \rho : \mathbb{T} \rightarrow \mathbb{T}$  defined by

$$\sigma(t) = \inf\{s \in \mathbb{T} : s > t\} \quad \text{and} \quad \rho(t) = \sup\{s \in \mathbb{T} : s < t\}$$

together with the convention  $\inf \emptyset = \sup \mathbb{T}$  and  $\sup \emptyset = \inf \mathbb{T}$  are the main tools for time-scale calculus. A point  $t \in \mathbb{T}$  is called *right-scattered*, *right-dense*, *left-scattered*, *left-dense*, if  $\sigma(t) > t$ ,  $\sigma(t) = t$ ,  $\rho(t) < t$ ,  $\rho(t) = t$  is satisfied, respectively. The graininess at  $t$  is then defined by  $\mu(t) = \sigma(t) - t$ . The set  $\mathbb{T}^\kappa$  is defined as  $\mathbb{T} \setminus \{m\}$  if  $\mathbb{T}$  has a left-scattered maximum  $m$ , and as  $\mathbb{T}$  otherwise. A function  $f : \mathbb{T} \rightarrow \mathbb{R}$  is called *regressive* if  $1 + \mu(t)f(t) \neq 0$  for all  $t \in \mathbb{T}^\kappa$ . A matrix-valued function  $A(t)$  of size  $n \times n$  defined on a time scale  $\mathbb{T}$  is called *regressive* if for all  $t \in \mathbb{T}^\kappa$ ,  $\det[I + \mu(t)A(t)] \neq 0$ , where  $I$  is the  $n \times n$  unit matrix. It is not difficult to prove that an  $n \times n$  matrix  $A(t)$  is regressive if and only if the eigenvalues  $\lambda_i(t)$ ,  $i = 1, \dots, n$ , of  $A(t)$  are all regressive. A function  $f : \mathbb{T} \rightarrow \mathbb{R}$  is called *rd-continuous*, if it is continuous at every right-dense point and if the left-sided limit exists (finite) at every left-dense point. The set of all rd-continuous and regressive functions defined on  $\mathbb{T}$  is denoted by  $\mathcal{R} = \mathcal{R}(\mathbb{T})$ . The function  $f$  is called (delta) *differentiable* at  $t$  with the (delta) derivative denoted by  $f^\Delta(t)$  if for any given  $\varepsilon > 0$  there is a neighbourhood  $U$  of  $t$  such that

$$|f(\sigma(t)) - f(s) - f^\Delta(t)(\sigma(t) - s)| \leq \varepsilon|\sigma(t) - s| \quad \text{for all } s \in U.$$

The following theorem is due to Hilger [3], see also [1, Theorem 1.16].

**THEOREM 1.1.** *Assume that  $f : \mathbb{T} \rightarrow \mathbb{R}$  and let  $t \in \mathbb{T}$ .*

- (a) *If  $f$  is differentiable at  $t$ , then  $f$  is continuous at  $t$ .*
- (b) *If  $f$  is continuous at  $t$  and  $t$  is right-scattered, then  $f$  is differentiable at  $t$  with*

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

- (c) *If  $f$  is differentiable and  $t$  is right-dense, then*

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

(d) *If  $f$  is differentiable at  $t$ , then  $f(\sigma(t)) = f(t) + \mu(t)f^\Delta(t)$ .*

Let  $t_0 \in \mathbb{T}$  and  $A \in \mathcal{R}$  be an  $n \times n$  matrix-valued function. The matrix exponential  $e_A(t, t_0)$  on  $\mathbb{T}$ , [1, Definition 5.18], is defined to be the unique solution of the matrix initial value problem

$$X^\Delta = A(t)X, \quad X(t_0) = I, \tag{1.1}$$

where  $I$  denotes as usual the  $n \times n$  identity matrix.

It is not difficult to verify that

$$e_A(t, t_0)^{-1} = e_A(t_0, t), \quad e_A(\sigma(t), t_0) = [I + \mu(t)A(t)]e_A(t, t_0).$$

Further properties of the matrix exponential  $e_A(t, t_0)$  including the ones above are contained in [1, Theorem 5.21].

When  $A$  is an  $n \times n$  constant matrix, one can easily show that  $e_A(t, t_0) = e^{A(t-t_0)}$  for  $\mathbb{T} = \mathbb{R}$ , while if  $\mathbb{T} = \mathbb{Z}$  and  $A$  is regressive, then  $e_A(t, t_0) = (I + A)^{t-t_0}$ . In the next section we describe an elementary method to calculate  $e_A(t, t_0)$  for constant matrices by using homogeneous linear dynamic equations with constant coefficients and the Cayley-Hamilton theorem.

## 2. The main result

Let  $A$  be an  $n \times n$  constant matrix and denote by  $p(\lambda)$  its characteristic polynomial. For certain real numbers  $c_0, \dots, c_{n-1}$  we may write that

$$p(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_0.$$

The Cayley-Hamilton theorem states that

$$p(A) = 0. \tag{2.1}$$

LEMMA 2.1. *If  $A \in \mathcal{R}$ , then  $\Phi = e_A(\cdot, t_0)$  is the unique solution of the matrix initial value problem*

$$X^{\Delta^n} + c_{n-1}X^{\Delta^{n-1}} + \dots + c_0X = 0, \tag{2.2}$$

$$X^{\Delta^j}(t_0) = A^j, \quad j = 0, 1, \dots, n - 1. \tag{2.3}$$

PROOF. Let us first show that the initial value problem (2.2), (2.3) has a unique solution. Define  $Z_{j+1}(t) = X^{\Delta^j}(t)$  for  $j = 0, 1, \dots, n - 1$ , and  $Z = (Z_1, \dots, Z_n)^T$ . Then the matrix initial value problem (2.2), (2.3) becomes

$$Z^\Delta = MZ, \quad Z(t_0) = Z_0, \tag{2.4}$$

where

$$M = \begin{bmatrix} 0 & I & 0 & \dots & 0 \\ 0 & 0 & I & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I \\ -c_0 I & -c_1 I & -c_2 I & \dots & -c_{n-1} I \end{bmatrix}, \quad Z_0 = \begin{bmatrix} I \\ A \\ A^2 \\ \vdots \\ A^{n-1} \end{bmatrix}.$$

Since the regressivity condition

$$\det(I + \mu(t)M) = 1 + \sum_{i=1}^n (-\mu(t))^i c_{n-i} \neq 0$$

is satisfied due to  $A \in \mathcal{R}$ , it follows from [1, Theorem 5.8] that the matrix initial value problem (2.4) has a unique solution.

Now from (1.1) we have

$$\Phi^{\Delta^j}(t) = A^j \Phi(t), \quad j = 1, 2, \dots, n. \tag{2.5}$$

In view of (2.1) and (2.5) we obtain

$$\Phi^{\Delta^n} + c_{n-1} \Phi^{\Delta^{n-1}} + \dots + c_0 \Phi = p(A)\Phi = 0$$

and

$$\Phi^{\Delta^j}(t_0) = A^j \Phi(t_0) = A^j, \quad j = 0, 1, \dots, n - 1,$$

which means that  $\Phi = e_A(\cdot, t_0)$  is the unique solution of the matrix initial value problem (2.2)–(2.3). □

We may now state and prove our main result.

**THEOREM 2.2.** *Let  $A \in \mathcal{R}$  with the characteristic polynomial*

$$p(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_0.$$

*If  $x_i(t)$ ,  $i = 0, 1, \dots, n - 1$ , is the unique solution of the initial value problem*

$$Lx \stackrel{\text{def}}{=} x^{\Delta^n} + c_{n-1}x^{\Delta^{n-1}} + \dots + c_0x = 0, \tag{2.6}$$

$$x^{\Delta^j}(t_0) = \delta_{ij}, \quad j = 0, 1, \dots, n - 1, \tag{2.7}$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases},$$

then

$$e_A(t, t_0) = x_0(t)I + x_1(t)A + \dots + x_{n-1}(t)A^{n-1}. \tag{2.8}$$

PROOF. Define  $\Psi(t) = x_0(t)I + x_1(t)A + \dots + x_{n-1}(t)A^{n-1}$ . It is easy to see that

$$\begin{aligned} \Psi^{\Delta^n} + c_{n-1}\Psi^{\Delta^{n-1}} + \dots + c_0\Psi &= (Lx_0)I + (Lx_1)A + \dots + (Lx_{n-1})A^{n-1} \\ &= 0 \cdot I + 0 \cdot A + \dots + 0 \cdot A^{n-1} = 0. \end{aligned}$$

Also,  $\Psi^{\Delta^j}(t_0) = \sum_{i=0}^{n-1} x_i^{\Delta^j}(t_0)A^i = A^j, j = 0, \dots, n-1$ . Thus  $\Psi(t)$  is also a solution of the initial value problem (2.2)–(2.3). In view of Lemma 2.1 and the uniqueness of the solutions we conclude that  $\Psi(t) = \Phi(t)$  for all  $t \in \mathbb{T}$ . □

REMARK 1. The characteristic polynomial  $p(\lambda)$  may be replaced by the minimal polynomial of  $A$  or by any polynomial satisfied by  $A$ .

REMARK 2. If we take  $\mathbb{T} = \mathbb{R}$  and  $t_0 = 0$  then Theorem 2.2 coincides with [7, Theorem 2]. Note that the regressivity condition holds trivially in this case since  $\mu(t) \equiv 0$ .

REMARK 3. If we set  $\mathbb{T} = \mathbb{Z}, t_0 = 0$ , and  $A = B - I$ , then we can compute the  $k$ -th power of  $B$  and hence obtain Theorem 2 of [6]; see the example below.

### 3. Application

In this section we show how to calculate the matrix exponential  $e_A(t, t_0)$  by employing Theorem 2.2 for a given constant matrix  $A \in \mathcal{R}$ . The special cases  $\mathbb{T} = \mathbb{R}, \mathbb{T} = \mathbb{Z}$  and  $\mathbb{T} = \overline{q^{\mathbb{Z}}}$  will be deduced from the general result.

Let  $a, b \in \mathbb{T}$  satisfy  $a \neq b, 1 + \mu(t)a \neq 0, 1 + \mu(t)b \neq 0$ , and consider as in [7] the matrix

$$A = \begin{bmatrix} a & 0 & 1 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}, \tag{3.1}$$

for which  $p(\lambda) = (\lambda - a)^2(\lambda - b), \lambda_1 = a, \lambda_2 = b$ . Since  $\lambda_1, \lambda_2 \in \mathcal{R}, A \in \mathcal{R}$ . The corresponding scalar equation (2.6) then becomes

$$x^{\Delta^3} - (2a + b)x^{\Delta^2} + a(2b + a)x^{\Delta} - a^2bx = 0.$$

‘ It follows that, see [1, 2] for details,

$$x(t) = c_1e_a(t, t_0) + c_2k(t, t_0)e_a(t, t_0) + c_3e_b(t, t_0)$$

is a general solution, where

$$k(t, t_0) = \int_{t_0}^t \frac{\Delta s}{1 + a\mu(s)}.$$

From the general solution one can easily get the special solutions  $x_i(t)$ ,  $i = 0, 1, 2$ , satisfying (2.7). Indeed, it turns out that

$$\begin{aligned} x_0(t) &= \left[ 1 - \frac{a^2}{(a-b)^2} \right] e_a(t, t_0) + \frac{ab}{a-b} k(t, t_0) e_a(t, t_0) + \frac{a^2}{(a-b)^2} e_b(t, t_0), \\ x_1(t) &= \frac{2a}{(a-b)^2} e_a(t, t_0) - \frac{a+b}{a-b} k(t, t_0) e_a(t, t_0) - \frac{2a}{(a-b)^2} e_b(t, t_0), \\ x_2(t) &= -\frac{1}{(a-b)^2} e_a(t, t_0) + \frac{1}{a-b} k(t, t_0) e_a(t, t_0) + \frac{1}{(a-b)^2} e_b(t, t_0). \end{aligned}$$

Now since

$$A^2 = \begin{bmatrix} a^2 & 0 & a+b \\ 0 & a^2 & 0 \\ 0 & 0 & b^2 \end{bmatrix},$$

we obtain from (2.8) that

$$\begin{aligned} e_A(t, t_0) &= \left\{ \left[ 1 - \frac{a^2}{(a-b)^2} \right] e_a(t, t_0) + \frac{ab}{a-b} k(t, t_0) e_a(t, t_0) + \frac{a^2}{(a-b)^2} e_b(t, t_0) \right\} I \\ &+ \left\{ \frac{2a}{(a-b)^2} e_a(t, t_0) - \frac{a+b}{a-b} k(t, t_0) e_a(t, t_0) - \frac{2a}{(a-b)^2} e_b(t, t_0) \right\} A \\ &+ \left\{ -\frac{1}{(a-b)^2} e_a(t, t_0) + \frac{1}{a-b} k(t, t_0) e_a(t, t_0) + \frac{1}{(a-b)^2} e_b(t, t_0) \right\} A^2 \\ &= \begin{bmatrix} e_a(t, t_0) & 0 & (e_a(t, t_0) - e_b(t, t_0))/(a-b) \\ 0 & e_a(t, t_0) & 0 \\ 0 & 0 & e_b(t, t_0) \end{bmatrix}. \end{aligned} \tag{3.2}$$

If we take  $\mathbb{T} = \mathbb{R}$ ,  $t_0 = 0$ ,  $a = 2$  and  $b = 3$ , then our example coincides with the one given by Leonard [7].

The  $k$ -th power of  $A$  for  $a = 2$  and  $b = 3$  was calculated by Kwapisz [6] as

$$A^k = \begin{bmatrix} 2^k & 0 & 3^k - 2^k \\ 0 & 2^k & 0 \\ 0 & 0 & 3^k \end{bmatrix}, \quad k = 0, 1, \dots,$$

by considering an analogous method applied to the system

$$x(k+1) = Ax(k).$$

If we set  $\mathbb{T} = \mathbb{Z}$ ,  $t_0 = 0$ ,  $a = 2$  and  $b = 3$  we may write the above system of difference equations as follows:

$$x^\Delta = Bx, \quad B = A - I.$$

Replacing  $a$  by  $a - 1$  and  $b$  by  $b - 1$  in (3.2) we obtain

$$\begin{aligned}
 A^t &= (I + B)^t = e_B(t) \\
 &= \begin{bmatrix} e_{a-1}(t, 0) & 0 & (e_{a-1}(t, 0) - e_{b-1}(t, 0))/(a - b) \\ 0 & e_{a-1}(t, 0) & 0 \\ 0 & 0 & e_{b-1}(t, 0) \end{bmatrix} \\
 &= \begin{bmatrix} a^t & 0 & (a^t - b^t)/(a - b) \\ 0 & a^t & 0 \\ 0 & 0 & b^t \end{bmatrix}
 \end{aligned}$$

as desired. Furthermore, since the regressivity condition is satisfied ( $\mu(t) = 1, a = 2, b = 3$ ) the above computation is valid for  $t = -1, -2, \dots$  as well. In particular, we can obtain the inverse of  $A$  by setting  $t = -1$ .

In quantum calculus ( $q$ -difference equations theory) [4], one is often interested in the problem of finding the solution of the initial value problem

$$D_q x(t) = Ax(t), \quad x(t_0) = x_0, \tag{3.3}$$

where

$$D_q x(t) \stackrel{\text{def}}{=} \frac{x(qt) - x(t)}{(q - 1)t}, \quad q > 1.$$

If we take  $\mathbb{T} = \overline{q^{\mathbb{Z}}}$  then it turns out that  $\sigma(t) = qt, \mu(t) = (q - 1)t$ , and hence

$$D_q x(t) = x^\Delta(t).$$

Let  $A$  be defined by (3.1). Assume that

$$1 + a(q - 1)t \neq 0 \quad \text{and} \quad 1 + b(q - 1)t \neq 0 \quad \text{for all } t \in \mathbb{T},$$

which means that  $A$  is regressive. Clearly,

$$\begin{aligned}
 x(t) &= e_A(t, t_0)x_0 \\
 &= \begin{bmatrix} e_a(t, t_0) & 0 & (e_a(t, t_0) - e_b(t, t_0))/(a - b) \\ 0 & e_a(t, t_0) & 0 \\ 0 & 0 & e_b(t, t_0) \end{bmatrix} x_0
 \end{aligned}$$

is the unique solution of (3.3), where (see [1, page 74])

$$e_a(t, t_0) = \prod_{s \in [t_0, t)} [1 + (q - 1)\alpha s], \quad t > t_0.$$

From (d) of Theorem 1.1 one can also easily see that

$$e_a(t, t_0) = \prod_{s \in [t, t_0)} \frac{1}{1 + (q - 1)\alpha s}, \quad t < t_0.$$

REMARK 4. It is easy to verify that  $m(\lambda) = (\lambda - a)(\lambda - b)$  is the minimal polynomial of  $A$ . This fact could have been used to considerably reduce the amount of work done above. We leave this to the reader. Of course, one should always prefer using the minimal polynomial provided it can be obtained without too much effort.

### Acknowledgements

I am grateful to the referee(s) for the suggestions and comments on the original manuscript.

### References

- [1] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales* (Birkhäuser, Boston, 2001).
- [2] M. Bohner and A. Peterson, *Advances in Dynamic Equations on Time Scales* (Birkhäuser, Boston, 2003).
- [3] S. Hilger, "Analysis on measure chains—a unified approach to continuous and discrete calculus", *Results Math.* **18** (1990) 18–56.
- [4] V. Kac and P. Cheung, *Quantum Calculus* (Springer, New York, 2002).
- [5] B. Kaymakçalan, V. Lakshmikantham and S. Sivasundaram, *Dynamic Systems on Measure Chains* (Kluwer, Dordrecht, 1996).
- [6] M. Kwapisz, "The power of a matrix", *SIAM Rev.* **40** (1998) 703–705.
- [7] I. E. Leonard, "The matrix exponential", *SIAM Rev.* **38** (1996) 507–512.