Scientific
Research

Matrix Functions of Exponential Order

Mithat Idemen

Engineering Faculty, Okan University, Istanbul, Turkey
Email: midemen@gmail.com

ABSTRACT

Both the theoretical and practical investigations of various dynamical systems need to extend the definitions of various functions *defined on the real axis* to the set of matrices. To this end one uses mainly three methods which are based on 1) the Jordan canonical forms, 2) the polynomial interpolation, and 3) the Cauchy integral formula. All these methods give the same result, say $g(A)$, when they are applicable to given function $g(t)$ and matrix A. But, unfortunately, each of them puts certain restrictions on $g(t)$ and/or A, and needs tedious computations to find explicit exact expressions when the eigen-values of A are not simple. The aim of the present paper is to give an alternate method which is more logical, simple and applicable to all functions (continuous or discontinuous) of exponential order. It is based on the two-sided Laplace transform and analytical continuation concepts, and gives the result as a linear combination of certain n matrices determined *only* through A. Here n stands for the order of A. The coefficients taking place in the combination in question are given through the analytical continuation of $g(t)$ (and its derivatives if A has multiple eigen-values) to the set of eigen-values of A (numerical computation of inverse transforms is not needed). Some illustrative examples show the effectiveness of the method.

1. Introduction

For many theoretical and practical applications one needs to extend the definitions of functions defined on the real axis to the set of matrices. The history of the subject goes back to the second half of the nineteenth century when Cayley, who is the main instigator of the modern notation and terminology, introduced the concept of the square-root of a matrix A [1]. Since then a huge work has been devoted to the definitions, numerical computations and practical applications of the matrix functions. A rather detailed history (including a large reference list) and important results (especially those concerning the numerical computation techniques) are extensively discussed in the book by Higham [2]. So, we eschew here of making a review of the historical development and giving a large reference list.

To extend the definition of a scalar function $g(t)$, defined for $t \in \Re$, to the set of matrices, one starts from an explicit expression of $g(t)$, which can be continued analytically into the complex plane $\mathbb{C}$, and replaces there t by A. If the result is meaningful as a matrix, then it is defined to be $g(A)$. Before going into further detail, it is worthwhile to clarify the meaning of the word "defined" appearing in the expression of "defined for

$t \in \Re$". It is especially important when $g(t)$ consists of a multi-valued inverse function. To this end consider, for example, the square-root function $g(t) = t^{1/2}$. Its definition requires, first of all, a cut connecting the branch point $t = 0$ to the other branch point $t = \infty$ in the complex plane $\mathbb{C}$. Then, by choosing one of its possible values at a given point, for example $g(1)$, one defines it completely. The result consists of a well-defined branch of the square-root function. If one replaces t in this expression by A, then one gets a (unique) matrix to be denoted by $A^{1/2}$. This matrix satisfies the equation $X^2 = A$ which may have many solutions denoted also by $A^{1/2}$. The above-mentioned function $g(A)$, which consists merely of the extension of the above-mentioned well-defined branch of the square-root function, can not permit us to find all these solutions. For example the equation $X^2 = I$, where I denotes the unit 2×2 matrix, has infinitely many solutions given by

$$X = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix},$$

where θ stands for any complex angle. All these matrices are defined to be the square-root $\sqrt{I}$. But the above-mentioned matrix $g(A)$ gives only one of them, namely I

or $(-\mathbf{I})$.

The known classical methods used in this context are grouped as follows (see [2], Sections 1, 2):

1) Methods based on the Jordan canonical formula;

2) Methods based on the Hermite interpolation formula;

3) Methods based on the Cauchy integral formula.

All these methods are applicable when the function $g(t)$, defined on the real axis, can be analytically continued into a domain of the complex-plane, which involves the spectrum of the matrix $\mathbf{A}$ (see def. 1.2 and def. 1.4 in [2]). Consider, for example, the Heaviside unit step function $H(t)$ defined on the real axis by

$$H(t) = \begin{cases} 1, & t > 0 \\ 1/2, & t = 0 \\ 0, & t < 0. \end{cases} \tag{1a}$$

It is obvious that the analytical continuation of $H(t)$ into the complex z-plane, if it is exists, has the point $z = 0$ as a singular point. To reveal $H(z)$, let us try to find its Taylor expansion about any point $a > 0$. This expansion is valid in the circle with center at the point $z = a$ and radius equal to $r = a$. Since all the coefficients except the first one are equal to naught, one gets $H(z) \equiv 1$ at all points inside the circle in question. By letting $a \to \infty$ one concludes that $H(z)$ is regular in the right half-plane $\Re z > 0$. If the above-mentioned Taylor expansion were made about a point $a < 0$, then one would get $H(z) \equiv 0$ for all z with $\Re z < 0$. This shows that $H(z)$ is a sectionally regular (holomorphic) function (see [3], Section 2.15). On the basis of the Plemelj-Sokhotiskii formulas (see [3], Section 2.17), for the points on the imaginary axis one writes $H = 1/2$, which yields

$$H(z) = \begin{cases} 1, & \Re z > 0 \\ 1/2, & \Re z = 0 \\ 0, & \Re z < 0. \end{cases} \tag{1b}$$

Notice that (1a) and (1b) can also be obtained by computing the improper integral

$$H(z) = \frac{z}{2\pi} \int\limits_{-\infty}^{\infty} \frac{\mathrm{d}\zeta}{\zeta^2 + z^2} + \frac{1}{2}, \tag{1c}$$

where the bar on the integral sign stands for the Cauchy principal value.

From (1b) one concludes that the seemingly general and elegant method 3), which is based on the Cauchy integral

$$g(\mathbf{A}) = \frac{1}{2\pi i} \int\limits_{C} (\zeta \mathbf{I} - \mathbf{A})^{-1} \mathrm{d}\zeta, \tag{2}$$

where C stands for a closed contour such that the domain bounded by C involves all the eigen-values of $\mathbf{A}$ and the function $g(z)$ is regular there, can not be applicable to

find $H(\mathbf{A})$ (and other functions expressible through $H(t)$) when $\mathbf{A}$ has eigen-values having both positive and negative real parts.

As to the methods 1) and 2), they need, in general, some tedious and cumbersome computations if $\mathbf{A}$ has multiple eigen-values.

In the present note we will consider the case when the function $g(t)$, defined on the real axis, is of the exponential order at both $t = +\infty$ and $t = -\infty$, and give a new method which seems to be more logical and effective especially when the matrix $\mathbf{A}$ has multiple eigen-values. It gives the result as a linear combination of n matrices determined only by the matrix $\mathbf{A}$. To this end we consider the Laplace transforms of $g(t)$ on the right and left halves of the real axis, namely:

$$\hat{g}^+(s) \equiv \Lambda^+\{g\} = \int\limits_{0}^{\infty} g(t)\mathrm{e}^{-st}\mathrm{d}t, \tag{3a}$$

$$\hat{g}^-(s) \equiv \Lambda^-\{g\} = \int\limits_{-\infty}^{0} g(t)\mathrm{e}^{-st}\mathrm{d}t \tag{3b}$$

and write [4]

$$g(t) = \frac{1}{2\pi i} \int\limits_{L^+} \hat{g}^+(s)\mathrm{e}^{ts}\mathrm{d}s + \frac{1}{2\pi i} \int\limits_{L^-} \hat{g}^-(s)\mathrm{e}^{ts}\mathrm{d}s,$$
$$t \in (-\infty, \infty). \tag{3c}$$

If the orders of the function $g(t)$ for $t \to \infty$ and $t \to (-\infty)$ are c^+ and c^-, respectively, then the function $\hat{g}^+(s)$ is a regular function of s in the right-half plane $\Re s > c^+$ and the integration path L^+ appearing in (3c) consists of any vertical straight-line located in this half-plane (see **Figure 1**). Similarly, $\hat{g}^-(s)$ is regular in the half-plane $\Re s < c^-$ and the integration path L^- is any vertical straight-line in this half-plane (if $c^+ < c^-$, then one can assume $L^+ \equiv L^-$). Furthermore, if $g(t)$ as well as its derivatives up to the order $(m-1)$ are all naught at $t = 0$, *i.e.* when

$$g(0) = g'(0) = \cdots = g^{(m-1)}(0) = 0, \tag{4a}$$

then one has
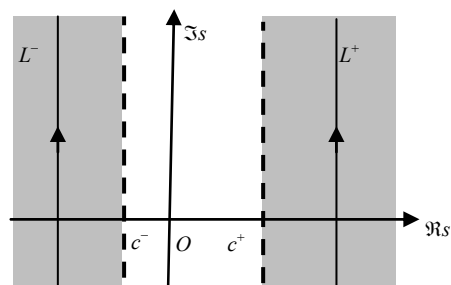
Figure 1. Regularity domains of $\hat{g}^\pm(s)$ and the integration lines $L^\pm$ when $c^- < c^+$.

AM

$$\Lambda^+ \left\{ g^{(m)} \right\} = s^m \hat{g}^+ (s) \qquad (4b)$$

and

$$\Lambda^- \left\{ g^{(m)} \right\} = s^m \hat{g}^- (s), \qquad (4c)$$

which yield inversely

$$\frac{1}{2\pi i} \int_{L^+} \hat{g}^+ (s) s^m \mathrm{e}^{ts} \mathrm{d}s + \frac{1}{2\pi i} \int_{L^-} \hat{g} (s) s^m \mathrm{e}^{ts} \mathrm{d}s = g^{(m)} (t),$$

$$t \in (-\infty, \infty). \qquad (4d)$$

It is worthwhile to remark here that the formula (4d) permits us to compute $g^{(m)}(t)$ only at points on the real axis although $g(t)$ and its derivatives are defined in (or can be continued analytically into) the complex t-plane. Therefore, when the point t is replaced by a complex $\lambda \in \mathbb{C}$, in what follows we will replace the left hand side of (4d) by the analytical continuation of $g^{(m)}(t)$ to the point λ and write

$$\frac{1}{2\pi i} \int_{L^+} \hat{g}^+ (s) s^m \mathrm{e}^{\lambda s} \mathrm{d}s + \frac{1}{2\pi i} \int_{L^-} \hat{g}^- (s) s^m \mathrm{e}^{\lambda s} \mathrm{d}s \rightarrow g^{(m)} (\lambda),$$

$$\lambda \in \mathbb{C}. \qquad (4e)$$

The formulas (3c), (4d) and (4e) will be the basis of our approach.

Let A be a square matrix of dimensions $n \times n$. We will define $g(A)$ by replacing t in (3c) by A, namely:

$$g(A) = \frac{1}{2\pi i} \int_{L^+} \hat{g}^+ (s) \mathrm{e}^{As} \mathrm{d}s + \frac{1}{2\pi i} \int_{L^-} \hat{g}^- (s) \mathrm{e}^{As} \mathrm{d}s. \quad (5)$$

Thus the computation of $g(A)$ becomes reduced to the computation of $\exp\{At\}$. As we will see later on, the latter consists of a linear combination of certain constant matrices Γ^j ($j = 1, \cdots, n$ = order of A). Hence $g(A)$ will also be a linear combination of these Γ^j's for every $g(t)$. It is important to notice that to compute the coefficients in the combinations in question we will never need to compute the transform functions $\hat{g}^\pm (s)$ as well as the integrals of the form (5) if the analytical continuation of $g^{(m)}(t)$ is known at the eigen-values of A (see the examples to be given in Section 4). These points constitute the essential properties of the definition (5):

1) It unifies the definition of $g(A)$ for all functions $g(t)$ of exponential order;

2) It gives an expression of $g(A)$ in terms of certain matrices which take place in the expression of $\exp(At)$ and are determined only by A;

3) It reduces the computation of $g(A)$ to the computation of $\exp(At)$ together with some scalar constants to be determined in terms of $g(t)$ (and its derivatives when A has multiple eigen-values) at the eigen-values of A.

The details are given in the theorems that follow.

2. Basic Results

In what follows we will denote a square matrix A of entries a_{jk} by $A = [a_{jk}]$. Here the first and second indices show, respectively, the row and column where a_{jk} is placed. The transpose of A will be indicated, as usual, by a super index T such as $A^{\mathrm{T}} = \left[a_{jk} \right]^{\mathrm{T}}$. The characteristic polynomial of A will be denoted by $f(\lambda)$, i.e. $f(\lambda) = \det(A - \lambda I)$.

Theorem-1. Let $A = [a_{jk}]$ be an $n \times n$ matrix with characteristic polynomial $f(\lambda)$. Then,

1) when all the zeros of $f(\lambda)$, say $\lambda_1, \cdots, \lambda_n$, are distinct, one has

$$\exp(At) = \sum_{\alpha=1}^{n} \Gamma_0^{(\alpha)} \mathrm{e}^{\lambda_\alpha t} \qquad (6a)$$

with $\Gamma_0^{(\alpha)}$ given by

$$\Gamma_0^{(\alpha)} = -\frac{1}{f'(\lambda_\alpha)} \left[\frac{\partial}{\partial a_{jk}} f(\lambda_\alpha) \right]^{\mathrm{T}}, \qquad (6b)$$

2) when $f(\lambda)$ has p distinct zeros, say $\lambda_1, \cdots, \lambda_p$, with multiplicities $m_1, \cdots, m_p$, respectively, one has

$$\exp(At)$$
$$= \sum_{\alpha=1}^{p} \left\{ \Gamma_{m_\alpha-1}^{(\alpha)} t^{m_\alpha-1} + \Gamma_{m_\alpha-2}^{(\alpha)} t^{m_\alpha-2} + \cdots + \Gamma_0^{(\alpha)} \right\} \mathrm{e}^{\lambda_\alpha t} \qquad (6c)$$

with the matrices $\Gamma_k^{(\alpha)} (k = 0, \cdots, m_a - 1)$ given by

$$\Gamma_k^{(\alpha)} = -\frac{1}{k! \, (m_\alpha - 1 - k)!}$$
$$\cdot \left[\frac{\mathrm{d}^{m_\alpha-1-k}}{\mathrm{d}s^{m_\alpha-1-k}} \left\{ \frac{(s-\lambda_\alpha)^{m_\alpha}}{f(s)} \left[\frac{\partial}{\partial a_{jk}} f(s) \right]^{\mathrm{T}} \right\} \right]_{s=\lambda_\alpha}. \qquad (6d)$$

Theorem-2. Let $A = [a_{jk}]$ be an n×n matrix while $g(z)$, defined in the complex z-plane, is regular at all the eigen-values of A and its restriction to the real axis is of exponential order at both $t = +\infty$ and $t = -\infty$. If all the eigen-values of A, say $\lambda_1, \cdots, \lambda_n$, are distinct, then one has

$$g(A) = \sum_{\alpha=1}^{n} g(\lambda_\alpha) \Gamma_0^{(\alpha)}. \qquad (7)$$

Here $\Gamma_0^{(\alpha)} (a = 1, \cdots, n)$ stands for the matrix taking place in the expression of $\exp\{At\}$.

Theorem-3. Let $A = [a_{jk}]$ be an $n \times n$ regular matrix which has p distinct eigen-values $\lambda_1, \cdots, \lambda_p$ with multiplicities $m_1, \cdots, m_p$, respectively, while $g(z)$, defined in the complex z-plane, is regular at all the eigen-values of A and its restriction to the real axis is of exponential order at both $t = +\infty$ and $t = -\infty$. Let the non-zero matrices taking place in the expression of $\exp\{At\}$ be $\Gamma_k^{(\alpha)} (a = 1, \cdots, p; \, k = 0, \cdots, K)$. Then one has

$$g(A) = A^{-N} \sum_{\alpha=1}^{p} \left\{ \Gamma_{m_\alpha-1}^{(\alpha)} G^{(m_\alpha-1)}(\lambda_\alpha) + \Gamma_{m_\alpha-2}^{(\alpha)} G^{(m_\alpha-2)}(\lambda_\alpha) + \cdots + \Gamma_0^{(\alpha)} G(\lambda_\alpha) \right\}, \quad (8)$$

where N stands for an integer such that $G(t) \equiv t^N g(t)$ and its derivatives up to the order $(K-1)$ are all naught at $t = 0$.

Before going into detail of proofs of the above-mentioned theorems, it is worthwhile to draw the attention to the fact that theorems 2 and 3 give the matrix function $g(A)$ as a combination of the n matrices $\Gamma_k^{(\alpha)}$ which appear in the expression of exp(At). They are the same (invariant) for all $g(t)$.

Proof of theorem-1. Our basic matrix function exp$\{At\}$ is defined, as usual, through the infinite series

$$\exp(at) = 1 + at + \frac{1}{2!} a^2 t^2 + \cdots, \quad |t| < \infty$$

by replacing there the scalar constant a by the square matrix A, namely:

$$\exp(At) = I + At + \frac{1}{2!} A^2 t^2 + \cdots \equiv X(t), \quad |t| < \infty.$$

It is obvious that $X(t)$ defined as above is the unique solution to the differential equation $X'(t) = AX(t)$ under the initial condition $X(0) = I$. Hence, by applying the Laplace transform to this equation one gets $\hat{X}(s) = -(A - sI)^{-1}$ which permits us to write

$$\exp(At) = -\frac{1}{2\pi i} \int_L (A - sI)^{-1} e^{ts} ds. \quad (9)$$

Let the characteristic polynomial of A be $f(s)$:

$$f(s) = \begin{vmatrix} a_{11} - s & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - s & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - s \end{vmatrix}. \quad (10)$$

Then the entry of the inverse matrix $(A - sI)^{-1}$, which is placed at the k-th row and j-th column, can be computed through the polynomial $f(s)$ as follows:

$$\frac{1}{f(s)} \frac{\partial}{\partial a_{jk}} f(s), \quad j, k = 1, 2, \cdots, n. \quad (11)$$

Thus (9) yields

$$\exp(At) = -\frac{1}{2\pi i} \int_L \frac{1}{f(s)} \left[\frac{\partial f(s)}{\partial a_{jk}} \right]^{T} e^{ts} ds. \quad (12)$$

Here the integration line L is any vertical straight-line such that all the eigen-values of A are located in the left side of L.

If the eigen-values are all simple, then the integral in (12) is computed by residues and gives (6a). When some of the eigen-values are multiple, as stated in theorem-1b, the residue method gives

$$\exp(At) = -\sum_{\alpha=1}^{p} \frac{1}{(m_\alpha-1)!} \left[\frac{d^{m_\alpha-1}}{ds^{m_\alpha-1}} \left\{ \frac{(s-\lambda_\alpha)^{m_\alpha}}{f(s)} \left[\frac{\partial}{\partial a_{jk}} f(s) \right]^{T} e^{st} \right\} \right]_{s=\lambda_\alpha}. \quad (13)$$

It is obvious that the derivatives in (13) yields a polynomial in t of degree $(m_\alpha - 1)$. Hence the final expression of exp(At) can be arranged as what is given in (6c).

Proof of theorem-3. Now consider the case when A

Proof of theorem-2. When the eigen-values of A are all simple, in (5) one replaces exp(As) by its expression given in (6a) and obtains

$$g(A) = -\sum_{\alpha=1}^{n} \frac{1}{f'(\lambda_\alpha)} \left[\frac{\partial}{\partial a_{jk}} f(\lambda_\alpha) \right]^{T} \left\{ \frac{1}{2\pi i} \int_{L^+} \hat{g}^+(s) e^{s\lambda_\alpha} ds + \frac{1}{2\pi i} \int_{L^-} \hat{g}^-(s) e^{s\lambda_\alpha} ds \right\}. \quad (14)$$

If all the eigen-values are real, then (3c) reduces (14) to (7). When some or all of the eigen-values are complex, we replace (3c) by (4e) with $m = 0$ and arrive again (7).

has multiple non-zero eigen-values and define $G(t) \equiv t^N g(t)$ where the integer N will be determined appropriately later on. If in (5) one replaces $g(t)$ by $G(t)$ and exp$\{As\}$ by (6c), then we get

$$A^N g(A) = \sum_{\alpha=1}^{p} \left\{ c_{m_\alpha-1}^{(\alpha)} \Gamma_{m_\alpha-1}^{(\alpha)} + c_{m_\alpha-2}^{(\alpha)} \Gamma_{m_\alpha-2}^{(\alpha)} + \cdots + c_0^{(\alpha)} \Gamma_0^{(\alpha)} \right\}, \quad (15a)$$

where

$$c_k^{(\alpha)} = \frac{1}{2\pi i} \int_{L^+} \hat{G}^+(s) s^k e^{s\lambda_\alpha} ds + \frac{1}{2\pi i} \int_{L^-} \hat{G}^-(s) s^k e^{s\lambda_\alpha} ds \quad (15b)$$

with $\alpha = 1, \cdots, p$ and $k = 0, \cdots, m_a - 1$. Remark that some of the coefficients $\Gamma_k^{(\alpha)}$ may be equal to naught (see ex.-2). Let the largest sub-index k be K for which one has $\Gamma_k^{(\alpha)} \neq 0$. Then, by considering the requirements in (4a), we will choose the integer N such that $G(t)$ and its derivatives up to the order $(K\text{-}1)$ are all naught for $t = 0$. In this case all terms existing in (15b) are computed through (4d) or (4e) and give (8).

3. A Corollary (Cayley-Hamilton Theorem)

Let $g(t)$ be the characteristic polynomial of A (i.e. $g(t) \equiv f(t)$). In this case all the terms taking place in (7) or (8) are equal to zero. When the eigen-values are all simple, from (7) one gets directly $f(A) = 0$, which is valid for both regular and singular matrices. In the case of multiple eigen-values, (8) gives $f(A) = 0$ if A is not singular. We remark that the Cayley-Hamilton theorem is correct for all matrices. We will use this theorem to compute the factor A^{-N} taking place in the formula (8) (see ex.-2).

4. Some Illustrative Examples

In order to show the application and effectiveness of the method, in what follows we will consider some simple examples.

Ex.-1 As a first example consider the case where A is as follows:

$$A = \begin{bmatrix} 8 & 12 & -2 \\ -3 & -4 & 1 \\ -1 & -2 & 2 \end{bmatrix}.$$

One can easily check that A has a triple eigen-value $\lambda = 2$. Therefore the theorems 1b and 3 are applicable directly for all functions of exponential order. The quanti-

ties p and m_α mentioned in those theorems are: $p = 1$, $m_1 = 3$. On the other hand from

$$f(\lambda) = \begin{vmatrix} 8-\lambda & 12 & -2 \\ -3 & -4-\lambda & 1 \\ -1 & -2 & 2-\lambda \end{vmatrix} = -(\lambda-2)^3$$

one computes

$$\left[\frac{\partial}{\partial a_{jk}} f(\lambda)\right]^{\mathrm{T}} = \begin{bmatrix} \lambda^2+2\lambda-6 & 12\lambda-20 & 4-2\lambda \\ 5-3\lambda & \lambda^2-10\lambda+14 & \lambda-2 \\ 2-\lambda & 4-2\lambda & \lambda^2-4\lambda+4 \end{bmatrix}$$

which gives (see (6c) and (6d))

$$\exp(At) = \left\{ \Gamma_2 t^2 + \Gamma_1 t + \Gamma_0 \right\} e^{2t}$$

where

$$\Gamma_0 = I, \ \Gamma_1 = \begin{bmatrix} 6 & 12 & -2 \\ -3 & -6 & 1 \\ -1 & -2 & 0 \end{bmatrix}$$

and

$$\Gamma_2 = \frac{1}{2}\begin{bmatrix} 2 & 4 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since $\Gamma_2 \neq 0$, one has $K = 2$ which shows that the formula (8) is applicable with $0 \leq N \leq 2$. For example, in order to find the expressions of A^ν and $\sin A^\nu$ (with $\Re \nu \geq 2$), one can choose $N = 0$ while $\sqrt{A}$, $\sin A$, $\sin \sqrt{A}$, $\sin \sqrt[3]{A}$, $\arcsin A$ etc. needs $N = 1$. To compute $\cos A$, $\log A$, $\cos \sqrt{A}$, $\operatorname{sign} A$ and $\arccos A$ one has to choose $N = 2$.

To check the formulas, we would like to compute first A^n ($n =$ integer ≥ 2) through the formula (7) which gives

$$A^n = n(n-1)2^{n-2}\Gamma_2 + n \cdot 2^{n-1}\Gamma_1 + 2^n \Gamma_0 = \begin{bmatrix} 2^{n-2}[n(n-1)+12n+4] & 2^{n-1}[n(n-1)+12n] & -2^n n \\ -2^{n-3}[n(n-1)+12n] & -2^{n-2}[n(n-1)+12n-4] & 2^{n-1}n \\ -2^{n-1}n & -2^n n & 2^n \end{bmatrix}.$$

Thus for $n = 2$ one gets

$$A^2 = \begin{bmatrix} 30 & 52 & -8 \\ -13 & -22 & 4 \\ -4 & -8 & 4 \end{bmatrix}.$$

Notice that by a direct multiplication of A with itself one gets the same result.
Similarly, one gets also

$$g(t) = \sqrt[3]{t} \Rightarrow A\sqrt[3]{A} = \left[\{t^{4/3}\}'' \Gamma_2 + \{t^{4/3}\}' \Gamma_1 + t^{4/3}\Gamma_0\right]_{t=2} = \frac{4}{9}\alpha^{-2/3}\Gamma_2 + \frac{4}{3}\alpha^{1/3}\Gamma_1 + \alpha^{4/3}I, \ \alpha^3 = 2$$

$$g(t) = \log(xt) \Rightarrow A^2 \log(Ax) = \left[\{t^2 \log(xt)\}'' \, \Gamma_2 + \{t^2 \log(xt)\}' \, \Gamma_1 + t^2 \log(xt) \Gamma_0 \right]_{t=2}$$

$$= \{2\log(2x) + 3\} \Gamma_2 + \{4\log(2x) + 2\} \Gamma_1 + 4\log(2x) I,$$

$$g(t) = \cos(x\sqrt{t}) \Rightarrow A^2 \cos(\sqrt{A}\, x) = \left[\{t^2 \cos(x\sqrt{t})\}'' \, \Gamma_2 + \{t^2 \cos(x\sqrt{t})\}' \, \Gamma_1 + t^2 \cos(x\sqrt{t}) \Gamma_0 \right]_{t=2}$$

$$= \left\{ \left(2 - \frac{x^2}{2}\right) \cos(\sqrt{2}x) - \frac{7x}{\sqrt{2}} \sin(\sqrt{2}x) \right\} \Gamma_2 + \left\{ 4\cos(\sqrt{2}x) - \sqrt{2}x \sin(\sqrt{2}x) \right\} \Gamma_1 + 4\cos(\sqrt{2}x) I,$$

$$g(t) = \sin(x\sqrt{t}) \Rightarrow A \sin(\sqrt{A}\, x) = \left[\{t \sin(x\sqrt{t})\}'' \, \Gamma_2 + \{t \sin(x\sqrt{t})\}' \, \Gamma_1 + t \sin(x\sqrt{t}) \Gamma_0 \right]_{t=2}$$

$$= \left\{ \frac{3x}{4\sqrt{2}} \cos(\sqrt{2}x) - \frac{x^2}{4} \sin(\sqrt{2}x) \right\} \Gamma_2 + \left\{ \frac{x}{\sqrt{2}} \cos(\sqrt{2}x) + \sin(\sqrt{2}x) \right\} \Gamma_1 + 2\sin(\sqrt{2}x) I,$$

$$g(t) = \operatorname{sign} t \Rightarrow A^2 \operatorname{sign} A = 2\Gamma_2 + 4\Gamma_1 + 4\Gamma_0 = \begin{bmatrix} 30 & 52 & -8 \\ -13 & -22 & 4 \\ -4 & -8 & 4 \end{bmatrix} \equiv A^2 \Rightarrow \operatorname{sign} A = I.$$

Remark that for different branches of $\sqrt{t}$, $\sqrt[3]{t}$ and $\log(xt)$ one gets different expressions for $\sqrt[3]{A}$, $\log(Ax)$, $\cos(\sqrt{A}x)$ and $\sin(\sqrt{A}x)$ (see Section 5 and theorem-4).

Finally, let us consider the branches of the inverse trigonometric functions $g(t) = \arcsin t$ and $h(t) = \arccos t$ which map the t-plane cut as shown in **Figure 2** into the regions in the g- and h-planes shown in **Figures 3** and **4** (the so-called principal branches of these functions!).

For the first function we have to choose $N = 1$ while the second one needs $N = 2$. The matrices A^{-1} and A^{-2}

$$\arcsin A = \frac{\gamma}{24} \begin{bmatrix} 8 & 16 & 0 \\ -4 & -8 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{\alpha + 2\gamma}{8} \begin{bmatrix} 20 & 40 & -8 \\ -10 & -20 & 4 \\ -4 & -8 & 0 \end{bmatrix} + \frac{\alpha}{4} \begin{bmatrix} -6 & -20 & 4 \\ 5 & 14 & -2 \\ 2 & 4 & 4 \end{bmatrix}$$

and

$$\arccos A = \frac{\beta - 8\gamma/3}{64} \begin{bmatrix} 32 & 64 & 0 \\ -16 & -32 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{\beta - \gamma}{16} \begin{bmatrix} 64 & 128 & -32 \\ -32 & -64 & 16 \\ 16 & -32 & 0 \end{bmatrix} + \frac{\beta}{16} \begin{bmatrix} -56 & -144 & 32 \\ 36 & 88 & -16 \\ 16 & 32 & 16 \end{bmatrix}$$

with

$$\alpha = \arcsin 2, \ \beta = \arccos 2, \ \gamma = 1/\sqrt{-3}.$$

It is an interesting exercise to check that

$$\arcsin A + \arccos A = \gamma 0 + \alpha I + \beta I = \frac{\pi}{2} I.$$

Ex.-2 Now consider the case where A is as follows:

taking place in the formula (8) can be computed rather easily by using the Cayley-Hamilton theorem as follows:

$$A^{-1} = \frac{1}{8} \begin{bmatrix} -6 & -20 & 4 \\ 5 & 14 & -2 \\ 2 & 4 & 4 \end{bmatrix},$$

$$A^{-2} = \frac{1}{64} \begin{bmatrix} -56 & -144 & 32 \\ 36 & 88 & -16 \\ 16 & 32 & 16 \end{bmatrix}.$$

Thus, by starting from $G(t) = t \arcsin t$ and $H(t) = t^2 \arccos t$ one gets from (8)

$$A = \begin{bmatrix} 7 & 0 & 4 \\ 8 & 3 & 8 \\ -8 & 0 & -5 \end{bmatrix}.$$

In this case one has

$$f(\lambda) = -(\lambda + 1)(\lambda - 3)^2,$$

which shows again that the theorems 1b and 3 are appli-
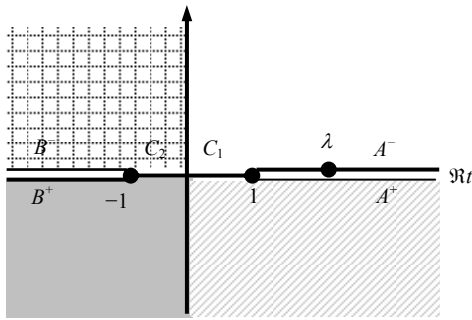
Figure 2. Complex plane-*t* cut along the lines (−∞) < *t* < −1 and 1 < *t* < ∞.
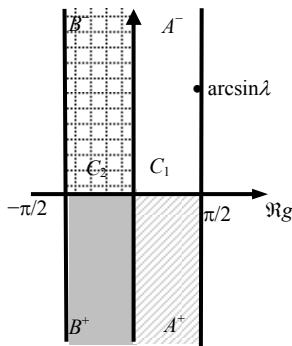

Figure 3. Mapping of the *t*-plane into the *g*-plane through the principal branch of the function *g* = arcsin*t*.
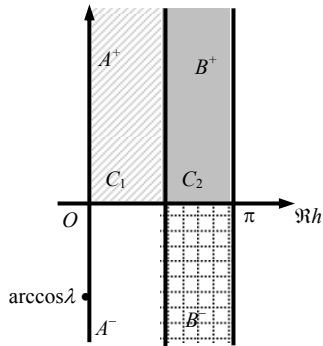

Figure 4. Mapping of the *t*-plane into the *h*-plane through the principal branch of the function *h* = arccos*t*.

cable with $p = 2$, $\lambda_1 = -1$, $\lambda_2 = 3$, $m_1 = 1$ and $m_2 = 2$. From the expression of $f(\lambda)$ one gets easily

$$\mathbf{\Gamma}_0^{(1)} = \begin{bmatrix} -1 & 0 & -1 \\ -2 & 0 & -2 \\ 2 & 0 & 2 \end{bmatrix}, \mathbf{\Gamma}_0^{(2)} = \begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \\ -2 & 0 & -1 \end{bmatrix}$$

and

$$\mathbf{\Gamma}_1^{(2)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0}.$$

Therefore from (6c) and (8) we write

$$\mathrm{e}^{At} = \mathrm{e}^{-t} \mathbf{\Gamma}_0^{(1)} + \mathrm{e}^{3t} \mathbf{\Gamma}_0^{(2)}.$$

Since $K = 0$, (8) is applicable with $N = 0$ and gives, for example,

$$\begin{aligned} \mathrm{sign} A &= \mathrm{sign}(\lambda_1) \mathbf{\Gamma}_0^{(1)} + \mathrm{sign}'(\lambda_2) \mathbf{\Gamma}_1^{(2)} + sign \ (\lambda_2) \mathbf{\Gamma}_0^{(2)} \\ &= -\mathbf{\Gamma}_0^{(1)} + \mathbf{\Gamma}_0^{(2)} \\ &= \begin{bmatrix} 3 & 0 & 2 \\ 4 & 1 & 4 \\ -4 & 0 & -3 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} H(A) &= H(\lambda_1) \mathbf{\Gamma}_0^{(1)} + H'(\lambda_2) \mathbf{\Gamma}_1^{(2)} + H \ (\lambda_2) \mathbf{\Gamma}_0^{(2)} = \mathbf{\Gamma}_0^{(2)} \\ &= \begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \\ -2 & 0 & -1 \end{bmatrix}. \end{aligned}$$

By direct multiplication of these matrices by themselves one gets

$$\{\mathrm{sign} A\}^2 = I, \{H(A)\}^2 = H(A),$$

which are the basic properties for the original functions sign*t* and *H*(*t*).

Similarly, one gets also

$$\mathrm{arcsin} \, A = \begin{bmatrix} 2\beta - \alpha & 0 & \beta - \alpha \\ 2\beta - 2\alpha & \beta & 2\beta - 2\alpha \\ 2\alpha - 2\beta & 0 & 2\alpha - \beta \end{bmatrix}$$

and

$$\mathrm{arccos} \, A = \begin{bmatrix} 2\delta - \gamma & 0 & \delta - \gamma \\ 2\delta - 2\gamma & \delta & 2\delta - 2\gamma \\ 2\gamma - 2\delta & 0 & 2\gamma - \delta \end{bmatrix}$$

with

$$\alpha = \mathrm{arcsin}(-1), \ \beta = \mathrm{arcsin} 3,$$

$$\gamma = \mathrm{arccos}(-1), \ \delta = \mathrm{arccos} 3.$$

Here the functions $\mathrm{arcsin} \, t \equiv g(t)$ and $\mathrm{arccos} \, t \equiv h(t)$ consist of the functions considered in the example-1 above. One can easily check that

$$\sin(\mathrm{arcsin} A) = A \ \text{ and } \ \mathrm{arcsin} \, A + \mathrm{arccos} \, A = \frac{\pi}{2} I,$$

which are the basic properties for the original functions arcsin*t* and arccos*t*.

Ex.-3 Finally, we want to give an example with complex-valued eigen-values. To this end consider the case where

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

with any real a and b. In this case one has

$$\lambda_1 = a + ib, \quad \lambda_2 = a - ib$$

and $\quad \mathbf{\Gamma}_0^{(1)} = \frac{1}{2}\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}, \quad \mathbf{\Gamma}_0^{(2)} = \frac{1}{2}\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}.$

Thus, for the following functions one writes

$$\arcsin \mathbf{A} = \begin{bmatrix} \frac{1}{2}\{\arcsin(a+ib)+\arcsin(a-ib)\} & \frac{1}{2}i\{\arcsin(a+ib)-\arcsin(a-ib)\} \\ -\frac{1}{2}i\{\arcsin(a+ib)-\arcsin(a-ib)\} & \frac{1}{2}\{\arcsin(a+b)+\arcsin(a-ib)\} \end{bmatrix}.$$

Here arcsin$t$ stands for the function defined in ex.-1 above. It is interesting to check that

$$[\text{sign}\mathbf{A}]^2 = \mathbf{I} \text{ if } a \neq 0, \quad \sin(\arcsin\mathbf{A}) = \mathbf{A}$$

5. Inverse Power of a Matrix

From the example 1 considered above one observes that there are many inverse powers $A^{1/m} (= \mathbf{B})$ which correspond to a given matrix A through the relation $\mathbf{B}^m = \mathbf{A}$. The next lemma and theorem concern this case.

Lemma. Let the values of the function $g(t) = t^{1/m}$, defined on a Riemann surface, at given p points $\lambda_1, \lambda_2, \cdots, \lambda_p$ be $g(\lambda_j) = w_j^{(k)} \neq 0$, $k = 1, \cdots, m$. Then there is a Riemann surface such that on one of its sheets $g(t)$ takes previously chosen values at the points $\lambda_1, \lambda_2, \cdots, \lambda_p$, namely:

$$g(\lambda_j) = w_j^{(k)} \text{ with arbitrarily chosen}$$

$$k_j \in \{1, 2, \cdots, m\}, \quad j = 1, 2, \cdots, p.$$

Proof. Let at the first r points $(1 \leq r < p)$ one has $g(\lambda_j) = w_j^{(1)}$ while at λ_{r+1} one wants to have $g(\lambda_{r+1}) = w_{r+1}^{(q)}$, where $q \in \{2, \cdots, m\}$ is any number. Then, we define the new cut line as a spiral curve which starts from the branch point $t = 0$, encircles the points $\lambda_1, \lambda_2, \cdots, \lambda_r$ (see the **Figure 4**) and $(q-1)$ times passes between the points λ_r and λ_{r+1}. Thus the analytic continuation of $g(\lambda_1) = w_1^{(1)}$ to λ_j $(j = 2, \cdots, r)$ becomes $w_j^{(1)}$ while at λ_{r+1} it is equal to $w_{r+1}^{(q)}$. We continue this process to adjust also the values at the points $\lambda_{r+2}, \cdots, \lambda_n$. Notice that when the super index k_{j+1} in $w_{j+1}^{(k_{j+1})}$ is smaller than k_j, we can replace k_{j+1} by $(k_{j+1} + m)$ because one has $w_{j+1}^{(k_{j+1})} = w_{j+1}^{(k_{j+1}+m)}$. At the end we arrive at a plane cut along an appropriate (spiral) curve joining the point $t = 0$ to $t = \infty$ such that at the given points $(\lambda_1, \lambda_2, \cdots, \lambda_p)$ the function $g(t)$ has desired values $w_j^{(k_j)}$. For example on the Riemann sheet shown in **Figure 5** one has

$$g(\lambda_1) = w_1^{(1)}, \quad g(\lambda_2) = w_2^{(1)}, \cdots, g(\lambda_5) = w_5^{(1)},$$

$$\text{sign}\mathbf{A} = \text{sign}(a+ib)\mathbf{\Gamma}_0^{(1)} + \text{sign}(a-ib)\mathbf{\Gamma}_0^{(2)}$$

$$= \begin{cases} \mathbf{\Gamma}_0^{(1)} + \mathbf{\Gamma}_0^{(2)} = \mathbf{I}, & \text{if } a > 0 \\ -\mathbf{\Gamma}_0^{(1)} - \mathbf{\Gamma}_0^{(2)} = -\mathbf{I}, & \text{if } a < 0 \\ \mathbf{0}, & \text{if } a = 0 \end{cases}$$

and

$$g(\lambda_6) = w_6^{(4)}, \quad g(\lambda_7) = w_7^{(4)}, \quad g(\lambda_8) = w_8^{(5)}.$$

Theorem-4. Let A be an $n \times n$ regular matrix with p different eigen-values $(1 \leq p \leq n)$ while m is an integer. Then there exist at least m^p different matrices $\mathbf{B}$ such that $\mathbf{B}^m = \mathbf{A}$.

Proof. Let the eigen-values of A be $\lambda_j (j = 1, \cdots, p)$. Consider a Riemann surface which corresponds to the function $g(t) = t^{1/m}$ and denote, on the k-th sheet of this Riemann surface, the values of the function $g(t)$ at points λ_j by $\left[(\lambda_j)^{1/m}\right]^{(k)} = w_j^{(k)}, (k = 1, \cdots, m)$. On the basis of the lemma, we can change the cut line appropriately such that the values of $t^{1/m}$ at the points $(\lambda_1, \lambda_2, \cdots, \lambda_p)$ become equal to previously chosen values. For example, $(\lambda_1)^{1/m} = w_1^{(1)}, \quad (\lambda_2)^{1/m} = w_2^{(3)}, \quad (\lambda_3)^{1/m} = w_3^{(2)}, \quad (\lambda_4)^{1/m} = w_4^{(1)}$ etc. Since these values can be arranged in m^p different forms, the right-hand side of (7) (or (8)) can have m^p different values. This proves thorem-4.

Ex. Let $a \in \mathbb{C}$ and $b \in \mathbb{C}$ any numbers which differ from zero. Then the matrix

$$A = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}$$
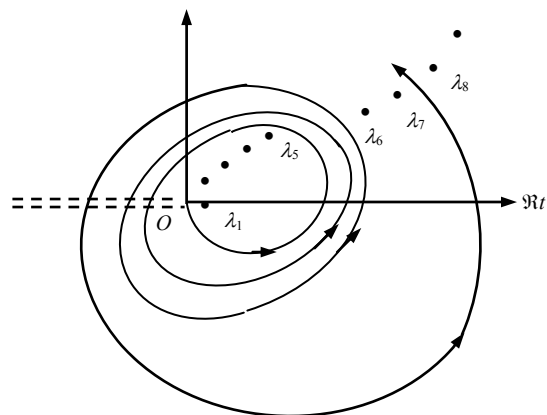
Figure 5. A cut appropriate for the particular case when $r = 5$ and $q = 4$.

has four square-roots given as follows:

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \begin{bmatrix} -a & 0 \\ 0 & b \end{bmatrix}, \begin{bmatrix} a & 0 \\ 0 & -b \end{bmatrix}, \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}.$$

REFERENCES

[1] A. Cayley, "A Memoir on the Theory of Matrices," *Philosophical Transactions of the Royal Society of London*, Vol. 148, 1858, pp. 17-37. doi:10.1098/rstl.1858.0002

[2] N. J. Higham, "Functions of Matrices: Theory and Computation," Society for Industrial and Applied Mathematica (SIAM), Philadelphia, 2008.

[3] N. I. Mushkhelishvili, "Singular Integral Equations," P. Noordhoff Ltd., Holland, 1958.

[4] E. C. Titchmarsh, "Introduction to the Theory of Fourier Integrals," Chelsea Publishing Company, 1986, Chapter 1.3.