DEPARTMENT OF MATHEMATICS
RADBOUD UNIVERSITY NIJMEGEN The Netherlands

Linear maps A^z expressed in terms of $A^{-1}, A^{-2}, \ldots, A^{-n}$ and analytic functions of z.

Stefan Maubach[1]

September 2006

DEPARTMENT OF MATHEMATICS
RADBOUD UNIVERSITY NIJMEGEN
Toernooiveld
6525 ED Nijmegen
The Netherlands

[1]Funded by Veni-grant of council for the physical sciences, Netherlands Organisation for scientific research (NWO)

Linear maps A^z expressed in terms of $A^{-1}, A^{-2}, \ldots, A^{-n}$ and analytic functions of z.

Stefan Maubach[†]

Radboud University Nijmegen
Toernooiveld 1, The Netherlands
s.maubach@math.ru.nl

October 26, 2006

Abstract

Suppose $A \in GL_n(\mathbb{C})$ has a relation $A^p = c_{p-1}A^{p-1} + \ldots + c_1A + c_0I$ where the $c_i \in \mathbb{C}$. This article describes how to construct analytic functions $c_i(z)$ such that $A^z = c_{p-1}(z)A^{p-1} + \ldots + c_1(z)A + c_0(z)I$. One of the theorems gives a possible description of the $c_i(z)$: $c_i(z) = C^z\alpha$ where $C \in Mat_p(\mathbb{C})$ is (similar to) the companion matrix of $X^p - c_{p-1}X^{p-1} - \ldots - c_1X - c_0I$, and $\alpha := (c_{p-1}, \ldots, c_1, c_0)^t$.

1 Introduction

1.1 Motivation

The original motivation, resulting in this article, was a problem not concerning linear maps, but the more general polynomial maps $F : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ that satisfy relations $F^n = c_{n-1}F^{n-1} + \ldots c_1F + c_0I$ where $c_i \in \mathbb{C}$. It was conjectured in [1] that such maps are an exponent of a so-called locally finite derivation (for more details, see [1], or [2] paragraph 4.3). Equivalently, it was enough to show that for such a polynomial map there existed an exponent map, i.e. a set of polynomial maps $\{F_t \mid t \in \mathbb{C}\}$ such that $F_0 = I, F_1 = F$, and $F_tF_u = F_{t+u}$. The author found a surprising formula that gave this exponent map in all concrete cases:

[†]Funded by Veni-grant of council for the physical sciences, Netherlands Organisation for scientific research (NWO)

$$F_t := (F^{-1}, F^{-2}, \ldots, F^{-n}) \begin{pmatrix} c_{n-1} & 1 & 0 & \cdots & 0 \\ c_{n-2} & 0 & 1 & \cdots & 0 \\ \ldots & \vdots & \vdots & \ddots & \ldots \\ c_1 & 0 & 0 & \cdots & 1 \\ c_0 & 0 & 0 & \cdots & 0 \end{pmatrix}^t \begin{pmatrix} c_{n-1} \\ \ldots \\ c_1 \\ c_0 \end{pmatrix}.$$

However, the author was not able to prove that this formula always worked. In fact, the question was not quite trivial for linear maps F, though it was clear that it should be correct for this case, and that it should be provable. This paper is exactly that proof. Of course, it is imperative to understand the linear case before any attempt at the more general case can be done, which gives a clear motivation.

On a different note, it can be quite a joy to run into such a problem - for if a problem *is* linear algebra, then one can have good confidence that it a solution is within grasp, and the problems appearing are most of the time beautiful puzzles. At least, this is how the author experienced it, and he hopes that the reader will have some of the same joy when reading (parts of) this article.

1.2 Basic facts and definitions

Consider the linear map induced by the matrix

$$A := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then one can notice that $A^2 = 2A - I$. Using this relation one sees $A^3 = A(2A-I) = 2A^2 - A = 2(2A-I) - A = 3A - 2I$, $A^4 = 4A - 3I$,etc In fact, $A^n = nA - (n-1)I$ for each $n \in \mathbb{N}$. Now, a natural way to define A^z for $z \in \mathbb{C}$ seems to be to extend the above formula, replacing n by z: $zA - (z-1)I$. And indeed, $A^z = zA - (z-1)I$. Well- that depends. It depends on how A^z is defined! But in this case we can say that it works. But *why*? And more general: can one always find such a formula for A^z by studying a relation as $A^2 = 2A - I$?

A short answer is: no, as no well-defined map $z \longrightarrow A^z$ can be defined for singular matrices A. But if one assumes that A is invertible, we need a long answer.

In order to answer the question, it is necessary to recall and notice some basic linear algebra facts, which we do in section 2. Then, in section 3 we will discuss basic facts on constructing the map $z \longrightarrow A^z$ in a well-defined way (i.e. as a true exponent map). In section 4 we will briefly discuss the set $V_A := \oplus_{i\in\mathbb{N}} \mathbb{C} A^i$. This will be used in section 5 to arrive at the main result, how to express A^z in $A^{-1}, \ldots A^{-p}$ and analytic functions. In section 6 we will discuss how one could compute the analytic functions in practice.

2 Preliminaries

2.1 Notations

I_n will denote the identity map in $GL_n(\mathbb{C})$. N_n will denote the map in $Mat_n(\mathbb{C})$ which is zero everywhere except at the $n-1$ positions directly above the diagonal, where it is 1. If n is known, it may be omitted, writing I and N.

If $Q(T) := T^p - c_{p-1}T^{p-1} - c_{p-2}T^{p-2} - \ldots - c_1 T - c_0$, then we define the *companion matrix*[1] of Q as

$$C_Q := \begin{pmatrix} c_1 & 1 & 0 & \ldots & 0 \\ c_2 & 0 & 1 & \ldots & 0 \\ & \vdots & & \vdots & \\ c_{p-1} & 0 & 0 & \ldots & 1 \\ c_p & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

If $A \in Mat_n(\mathbb{C})$, and $Q(T)$ is the characteristic polynomial of A, then we define the companion matrix of A as $C_A := C_Q$.

2.2 Jordan normal form and minimum polynomial

Remark 2.1. We have equivalence between

(i). $A \in GL_n(\mathbb{C})$ has minimum polynomial of degree n;

(ii). The m blocks of the Jordan-normal-form of $A \in GL_n(\mathbb{C})$ have eigenvalues $\lambda_1, \ldots, \lambda_m$ which all differ;

(iii). $A \in GL_n(\mathbb{C})$ is a conjugate of its companion matrix $C_A \in GL_n(\mathbb{C})$.

Proof. First, note that conjugating does not change the minimum/characteristic polynomial. Using this fact, *(i)*$\Longleftrightarrow$ *(ii)* is an easy consequence. Second, the minimum polynomial of C_A is the same as the characteristic polynomial of C_A, which in turn is the same as the characteristic polynomial of A (see [4]). From this, the equivalence *(iii)* $\Longleftrightarrow$ *(i)* follows easily. $\square$

Remark 2.2. Let $m(X)$ be the minimum polynomial of $A \in Mat_n(\mathbb{C})$ of degree p. Then there exist $J \in Mat_n(\mathbb{C})$, $\tilde{J} \in Mat_p(\mathbb{C})$, such that

1. J is a Jordan normal form of A, and $\tilde{J}$ is a Jordan normal form of C_m,
2. $\tilde{J}$ is the upper left $p \times p$ minor of J, consisting of a subset of the blocks appearing in J,

[1] Though usually defined slightly different.

3. per eigenvalue λ occurring in A (or C_m), there is only one block having this eigenvalue in $\tilde{J}$, and it has the maximum of the size of all the blocks having eigenvalue λ in A.

Proof. Let J be a Jordan normal form of A, i.e. J consists of blocks $B_1, \ldots, B_k$. Let $n_i \in \mathbb{N}$ be the size of B_i (i.e. $n_1 + n_2 + \ldots + n_k = n$). Now, notice that if $P(X) \in \mathbb{C}[X]$, then $P(J) = 0$ if and only if $P(B_i) = 0$ for each $1 \leq i \leq m$. It is easy to see that the minimum polynomial of B_i is $m_i(X) := (X - \lambda_i)^{n_i}$, where λ_i is the eigenvalue of B_i. Now: $P(J) = 0$ if and only if $m(X)|P(X)$ if and only if $m_i(X)|P(X)$ for all $1 \leq i \leq k$ if and only if $gcd(m_1(X), \ldots, m_k(X))|P(X)$. Therefore, $gcd(m_1(X), \ldots, m_k(X)) = m(X)$.

The results now follow easily: C_m has $m(X)$ as minimum polynomial (see [4])), so has $\tilde{J}$. Therefore, C_m and $\tilde{J}$ are equivalent by remark 2.1, yielding (1). (2) Can be achieved by choosing an appropriate J with the right order of blocks, and (3) is then automatic. $\square$

3 Exponents of linear maps

It is well-known how to make well-defined exponent maps $\mathbb{C} \longrightarrow \mathbb{C} : z \longrightarrow A^z$ for a given $A \in GL_n(\mathbb{C})$. This section provides some of these details, and points out a pitfall which one should keep in mind: the freedom in choosing an exponent map.

Let us start with a seemingly trivial definition:

Definition 3.1. Given an invertible map $f : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, we say that the map $\varphi : \mathbb{C} \times \mathbb{C}^n \longrightarrow \mathbb{C}^n$ is an exponent map of f if it satisfies
(1) $\varphi(z, -)\varphi(w, -) = \varphi(z + w, -)$ for all $z, w \in \mathbb{C}$,
(2) $\varphi(1, -) = f$,
(3) $\varphi(0, -) = I$.

If one has such a map φ, then one can rightfully define the notation $f^z := \varphi(z, -)$, as this coincides with the normal notation $f^n = f \circ f \circ \cdots \circ f$ when $z = n \in \mathbb{N}$, and $\varphi(-1, -)$ indeed is the same map as f^{-1} since it is the inverse of $\varphi(1, -) = f$. To take a trivial example, let $f = I$, and take $f^z = I$. To take a little less trivial example for the same f, take $f^z := e^{Im(z)} I$ (and $\varphi(z, x) = f^z(x) = e^{Im(z)} x$). This shows one of the problems in finding exponent maps: they are not unique. For some specific functions, like the map $x \longrightarrow ex$, a standard choice has been made by literature ($\varphi(z, x) \longrightarrow e^z x$). But if one wants to define an exponent of the map $x \longrightarrow \lambda x$ where $\lambda \in \mathbb{C}^*$, then one has to fix a value $\log(\lambda)$ such that $e^{\log(\lambda)} = \lambda$, and then define $\varphi(z, x) := e^{z \log(\lambda)} x$.

In this section, we assume that we have **fixed** for every $\lambda \in \mathbb{C}$ one such value $\log(\lambda)$ and a corresponding map $z \longrightarrow \lambda^z$. We will now shortly tell how this lets us fix maps $z \longrightarrow A^z$ for each $A \in GL_n(\mathbb{C})$.

Definition 3.2. Define $g_0(z) = 1$, $g_i(z) := \frac{1}{i!}z(z-1)\cdots(z-i+1)$ for all $i \in \mathbb{N}^*$. Note that if $k \in \mathbb{N}$, then $g_i(k) = \binom{k}{i}$ (even if $k < i$).
Given $\lambda \in \mathbb{C}^*$, define $B_n(\lambda) := \lambda I + N \in GL_n(\mathbb{C})$.

Lemma 3.3. *Let $\varphi(z, -) := \sum_{i=0}^{n} g_i(z)\lambda^{z-i}N^i$. Then φ is an exponent map of $B_n(\lambda)$.*

Proof. *(i)* Define $P_k(z, w) := \sum_{l=0}^{k} g_l(z)g_{k-l}(w) - g_k(z + w) = 0$ for all $z, w \in \mathbb{C}, k \in N$. Note that $P_k(z, w)$ is a polynomial. If $z, w \in \mathbb{N}$ then $P_k(z, w)$ equals $\sum_{l=0}^{k} \binom{z}{l}\binom{w}{k-l} - \binom{z+w}{k}$, which in turn is equal to "The coefficient of X^k in $(1+X)^z(1+X)^w$ minus the coefficient of X^k in $(1+X)^{z+w}$, which is obviously equal to zero. Thus, $P_k(z, w) = 0$ for all $z, w \in \mathbb{N}$, but for a polynomial this is only possible if the polynomial is zero. Thus $P_k(z, w) = 0$ for all $z, w \in \mathbb{C}$, and therefore $\sum_{l=0}^{k} g_l(z)g_{k-l}(w) = g_k(z+w)$.
(ii) $\varphi(0, -) = \sum_{i=0}^{n} g_i(0)\lambda^{-i}N^i = I$ and $\varphi(1, -) = \sum_{i=0}^{n} g_i(1)\lambda^{1-i}N^i = \lambda^{1-0}I + g_1(1)\lambda^{1-1}N = \lambda I + N = B$, therefore only left to prove is: $\varphi(z, -)\varphi(w, -) = \varphi(z + w, -)$. Let us do some computation:

$$\begin{aligned}
\varphi(z,-)\varphi(w,-) &= \left(\sum_{i=0}^{n} g_i(z)\lambda^{z-i}N^i\right)\left(\sum_{i=0}^{n} g_i(w)\lambda^{w-i}N^i\right) \\
&= \sum_{k=0}^{n}\sum_{l=0}^{k} g_l(z)\lambda^{z-l}N^l g_{k-l}(w)\lambda^{w-k+l}N^{k-l} \\
&= \sum_{k=0}^{n}\left(\sum_{l=0}^{k} g_l(z)g_{k-l}(w)\right)\lambda^{z+w-k}N^k \\
&=^{(i)} \sum_{k=0}^{n} g_k(z+w)\lambda^{z+w-k}N^k \\
&= \varphi(z+w,-).
\end{aligned}$$

□

Remark 3.4. The notation B^z is now justified. Note that one may see B^z as an element in $GL_n(H(z))$ where $H(z)$ is the set of holomorphic functions on $\mathbb{C}$. The only nonzero coefficients appearing in B^z are the functions $g_i(z)\lambda^{z-i}$ for $i = 0, \dots, n-1$.

Now, let $J \in GL_n(\mathbb{C})$ be on Jordan normal form, having blocks $B_1, \dots, B_k$, where each $B_i = \lambda_i I_{n_i} + N_{n_i}$ for some $n_i \in \mathbb{N}$ (and $\sum_{i=1}^{m} n_i = n$). Now one can easily define J^z by exponentiating each component B_i.

Definition 3.5. Let $A \in GL_n(\mathbb{C})$, and J a Jordan normal form of A (i.e. $M^{-1}AM = J$ for some $M \in GL_n(\mathbb{C})$). Then define $A^z := MJ^zM^{-1}$.

The definition is well-defined as the choice of Jordan-normal form is irrelevant (it will have the same blocks, only permuted).

Remark 3.6. If one chooses different exponential maps λ^z, then one obtains different exponential maps A^z, but when the maps λ^z are fixed, the maps A^z are also fixed (if one uses the above construction).
Also, the only nonzero coefficients appearing in J^z are the only nonzero coefficients appearing in the B_j^z : $g_i(z)\lambda_j^{z-i}$ for $j = 1, \dots, m, i = 0, \dots, n_i-1$.

4 The ring V_A

Definition 4.1. If $A \in GL_n(\mathbb{C})$, then define $V_A := \oplus_{i\in\mathbb{Z}}\mathbb{C}A^i$. Define $V_A^* := \{M \in Mat_n(\mathbb{C}) \mid M(V_A) \subseteq V_A\}$.

Remark 4.2. If the minimum polynomial $m_A(X)$ has degree p, then $dim(V_A) = p$. Also, $V_A^* = V_A \cong \mathbb{C}[X]/m_A(X)$.

Proof. $dim(V_A) = p$ is trivial. Now notice that $V_A \subseteq V_A^*$. Suppose $V_A \subsetneqq V_A^*$. A p-dimensional vector space can only have a p-dimensional set of linear endomorphisms. So, since $V_A \subsetneqq V_A^*$, there must be $M, M' \in Mat_n(\mathbb{C})$ such that $M \neq M'$, but $M = M'$ if restricted to V_A, i.e. $M - M'$ is the zero map on V_A. Thus, if one substitutes $A \in V_A$ in the map $M - M'$, it results in the zero matrix. Thus $(M - M')A = 0$ but since A is invertible, $M - M' = 0$. Contradiction, as $M \neq M'$. Thus $V_A = V_A^*$.

Since $A^{p-1}, \ldots, A, I$ are linearly independent (otherwise the minimum polynomial is not of degree p) this must be a basis of V_A^*. Notice that V_A is a commutative ring. Also, the morphism $\sigma : \mathbb{C}[X] \longrightarrow V_A$ given by $\sigma(X) = A$ has $m_A(X)$ in its kernel. Since $dim(\mathbb{C}[X]/m_A(X)) = dim(V_A)$ as $\mathbb{C}$-vector spaces we have an isomorphism between $\mathbb{C}[X]/m_A(X)$ and V_A. $\square$

Now we come close towards the example in the introduction: let the minimum polynomial of $A \in GL_n(\mathbb{C})$ have degree p, and $A^p = \sum_{i=1}^p c_i A^{p-i}$, or equivalently, $I = \sum_{i=1}^p c_i A^{-i}$. Now one can describe the map $A : V_A \longrightarrow V_A$ as a linear map with respect to the basis $A^{-1}, \ldots, A^{-p}$. First, define $\vec{\lambda} = (\lambda_1, \ldots, \lambda_m)^t$ and $\vec{A} := (A^{-1}, \ldots, A^{-m})$.

Lemma 4.3. *If $A \in GL_n(\mathbb{C})$ and $m(X)$ has degree p, then the map $A : V_A \longrightarrow V_A$, seen as a linear map with respect to the basis $A^{-1}, \ldots, A^{-p}$, is the matrix $C_m \in Mat_p(\mathbb{C})$. In other words, $A < \vec{\lambda}, \vec{A} > = < C_m\vec{\lambda}, \vec{A} >$*

Proof.

$$\begin{aligned} A(\textstyle\sum_{i=1}^p \lambda_i A^{-i}) &= \textstyle\sum_{i=2}^p \lambda_i A^{-i+1} + \lambda_1 I \\ &= \textstyle\sum_{i=2}^p \lambda_i A^{-i+1} + \lambda_1 \sum_{i=1}^n c_i A^{-i} \end{aligned}$$

thus the matrix of $A : V_A \longrightarrow V_A$ w.r.t. the basis $A^{-1}, \ldots, A^{-p}$ is the matrix belonging to the map $\vec{\lambda} \longrightarrow C_m\vec{\lambda}$. $\square$

5 Expressing A^z in $A^{-1}, \ldots, A^{-n}$.

In this section, we assume all maps $z \longrightarrow \lambda^z$ to be fixed, and all exponents A^z to be fixed by fixing the exponent of its Jordan normal form. We will keep on using the notations $\vec{\lambda} := (\lambda_1, \ldots, \lambda_q)^t$ and $\vec{A} := (A^{-1}, \ldots, A^{-q})$.

Lemma 5.1. $A^z \in V_A$ *for all* $z \in \mathbb{C}$.

Proof. (i) First, let $B = I + \lambda N \in Mat_n(\mathbb{C})$ for some $\lambda \in \mathbb{C}$. Let $W := \mathbb{C}I + \mathbb{C}N + \ldots + \mathbb{C}N^{n-1}$. Then it is not difficult to see that W is spanned by $I, B, B^2, \ldots, B^{n-1}$. By lemma 3.3 $B^z \in W$, so it follows that $B^z = \sum_{i=0}^{n-1} b_i B^i$ for some $b_i \in \mathbb{C}$.
(ii) Now, let $J \in Mat_n(\mathbb{C})$ be a Jordan normal form matrix with m blocks B_i of size n_i. Then J^z is a matrix with blocks B_i^z. We hope to find $a_j \in \mathbb{C}$ such that $J^z = \sum_{j=0}^{n-1} a_j J^j$. This last equation is equivalent to the m equations $B_i^z = \sum_{j=0}^{n-1} a_j B_i^j$. Using (i), we see that these are m equations in n unknowns. The i-th equation has degree of freedom $n - n_i$, so the total system has degree of freedom $\geq n - \sum_{i=1}^{m} n_i = 0$, so there exists at least one solution $a_0, \ldots, a_n$ such that $J^z = \sum_{j=0}^{n} a_j J^j$.
(iii) Finally, the theorem follows from the fact that $A^z = \sum_{i=0}^{n} a_i A^i$ is equivalent to $J^z = \sum_{i=0}^{n} a_i J^i$ where J is a Jordan normal form of A. □

Definition 5.2. Let $A \in GL_n(\mathbb{C})$, and assume that the minimum polynomial of A, $m(X)$, has degree p. Define $\tau : V_A \longrightarrow V_{C_m} \subseteq Mat_p(\mathbb{C})$ as the $\mathbb{C}$-linear ring isomorphism map sending $A \in V_A^*$ to the map $C_m \in Mat_p(\mathbb{C})$.

A short remark: V_{C_m} indeed is isomorphic to $\mathbb{C}[X]/m(X)$ since $m(X)$ is the minimum polynomial of C_m by lemma 2.1, and notice that τ is indeed a well-defined isomorphism.

Corollary 5.3. *The map τ projects maps $M : V_A \longrightarrow V_A$ onto their matrix representation w.r.t. the basis $A^{-1}, \ldots, A^{-p}$. In other words, $M < \vec{\lambda}, \vec{A} > = < \tau(M)\vec{\lambda}, \vec{A} >$.*

Proof. Lemma 4.3 states that $A < \vec{\lambda}, \vec{A} > = < \tau(A)\vec{\lambda}, \vec{A} >$. Since $M \in V_A^*$ one has $M = \sum_{i=0}^{p-1} \alpha_i A^i$ for some $\alpha_i \in \mathbb{C}$. Thus,

$$\begin{aligned} M < \vec{\lambda}, \vec{A} > = & \sum_{i=0}^{p-1} \alpha_i A^i < \vec{\lambda}, \vec{A} > \\ = & \sum_{i=0}^{p-1} \alpha_i < (\tau(A))^i \vec{\lambda}, \vec{A} > \\ = < \tau(M)\vec{\lambda}, \vec{A} > . & \end{aligned}$$

□

Write $(I0) := (I_n, 0_{(p \times n)})$ as the matrix of size $(n + p) \times n$ which has an identity matrix in the first n colums and zero entries anywhere else. Use M^t to denote the transpose of a matrix M.

Proposition 5.4. *Let the characteristic polynomial $m(X)$ of $A \in GL_n(\mathbb{C})$ be of degree p.*
(i) Then τ has the following form: There exists $S \in Mat_{p \times n}(\mathbb{C}), W \in Mat_{n \times p}$ satisfying $WS = I_d$, such that for all $M \in V_A^*$, $\tau(M) = WMS$.
(ii) Let $\tilde{J} := (I0)J(I0)^t$ be as in lemma 2.2, $V \in GL_n(\mathbb{C}), U \in GL_p(\mathbb{C})$ such that $V^{-1}JV = A, U^{-1}\tilde{J}U = C_m$. Then one can take $W = U^{-1}(I0)V, S = V^{-1}(I0)^t U$.

Proof. $A = V^{-1}JV$ for some $V, J \in GL_n(\mathbb{C})$ where J is a Jordan normal form of A. Furthermore, $C_m = U^{-1}\tilde{J}U$ for some $U, \tilde{J} \in GL_p(\mathbb{C})$ where $\tilde{J}$ is a Jordan normal form of C_m. Using remark 2.2, we can conclude that $\tilde{J} = (I0)J(I0)^t$. Now for each $i \in \mathbb{Z}$:

$$\begin{aligned} C_m^i = &\quad U^{-1}\tilde{J}^iU = U^{-1}((I0)J^i(I0)^t)U \\ = &\quad U^{-1}(I0)VA^iV^{-1}(I0)^tU. \end{aligned}$$

Define $W := U^{-1}(I0)V, S := V^{-1}(I0)^tU$. Then the above equations say $C_q^i = WA^iS$. Also, $WS = U^{-1}(I0)VV^{-1}(I0)^tU = U^{-1}(I0)(I0)^tU = U^{-1}I_pU = U^{-1}U = I_p$. Now if $M \in V_A$, then

$$\begin{aligned} \tau(M) = &\quad \tau(\textstyle\sum_{i=1}^p \lambda_i A^{-i}) \text{ for some } \lambda_i \in \mathbb{C} \\ \textstyle\sum_{i=1}^p \lambda_i \tau(A^{-i}) & \\ \textstyle\sum_{i=1}^p \lambda_i W A^{-i} S & \\ W(\textstyle\sum_{i=1}^p \lambda_i A^{-i})S & \\ = WMS. & \end{aligned}$$

□

Proposition 5.5. *Let $A \in GL_n(\mathbb{C})$ having minimum polynomial $m(X) := X^p - c_{p-1}X^{p-1} - \ldots - c_1X - c_0$. Let $c = (c_{p-1}, \ldots, c_0)^t$, and $\mu_i(z) := (C_m^z c)_i$. Then $A^z = \sum_{i=1}^p \mu_i(z)A^{-i}$.*

Proof. *(i)* $\tau(A^z) = \tau(A)^z$ for each $z \in \mathbb{C}$: By lemma 5.1 we know that for each $z \in \mathbb{C}$, $A^z(V_A) \subseteq V_A$. Therefore, each map A^z can be seen as a linear map on V_A. By definition 5.2 we see that $\tau(A^i) = (\tau A)^i = C_m^i$ for all $i \in \mathbb{Z}$. Now beware: we cannot immediately state that $\tau(A^z) = \tau(A)^z = C_m^z$ for all $z \in \mathbb{C}$. For this, consider $M(z) := \tau(A^z) - C_m^z \in Mat_p(H(\mathbb{C}))$, i.e. see $M(z)$ as a matrix of size p with holomorphic functions as entries. We are done if we show $M(z) = 0$. Let $J, \tilde{J}$ be as in lemma 2.2. Now using proposition 5.4:

$$\begin{aligned} M(z) = &\quad \tau(A^z) - C_m^z \\ = &\quad WA^zS - C_m^z \\ = &\quad U^{-1}(I0)V^{-1}VJ^zVV^{-1}(I0)^tU - U^{-1}\tilde{J}^zU \\ = &\quad U^{-1}(I0)J^z(I0)^tU - U^{-1}\tilde{J}^zU \end{aligned}$$

By assumption (lemma 2.2) we have $(I0)J(I0)^t = \tilde{J}$, thus $M(z) = 0$.
(ii) Notice that $I =< c, \vec{A} >$. Now using corollary 5.3 and *(i)*, we see that $A^z = A^z < c, \vec{A} >=< \tau(A^z)c, \vec{A} >=< C_m^z c, \vec{A} >$. This gives $A^z = \sum_{i=1}^p (C_m c)_i A^{-i}$. □

We are now almost able to prove the main theorem, we need just one more lemma:

Lemma 5.6. *Let $A \in Mat_n(\mathbb{C})$, let $m(X)$ be the minimum polynomial of A of degree p, and let $f(X) \in \mathbb{C}[X]$ be a nonzero polynomial of degree d. Then there exists $\tilde{A} \in Mat_{n+d}(\mathbb{C})$ such that it (1) has minimum polynomial $m(X)f(X)$, (2) the upper left $n \times n$ part of $\tilde{A}^z$ is equal to A^z for every $z \in \mathbb{C}$.*

Proof. We will replace (2) by (2') and (2") where (2') is: "the upper left $n \times n$ part of $\tilde{A}$ is equal to A" and (2") is "the rows $n+1, \ldots n+d$ have zeroes below the diagonal". If we can guarantee (2') and (2"), then (2) will hold. The proof will go in some steps.
(i) It is enough to prove (1),(2') and (2") for A on Jordan normal form: Let $A = T^{-1}JT$ where J is a Jordan normal form of A. Let $\tilde{T} \in Mat_{n+p}(\mathbb{C})$ be the canonical extension of T: the upper left $n \times n$ part equals T, and the rest of the coefficients equal the coefficients of an identity matrix. Let $\tilde{J}$ be satisfying (1), (2') and (2"). Then one can take $\tilde{A} := \tilde{T}\tilde{J}\tilde{T}^{-1}$.
(ii) It is enough to prove the theorem for $f(X) = X - a$ of degree one: the full theorem follows by induction. We will split into the case that a is an eigenvalue of A, and the case that it is not.
(iii) Suppose a is not an eigenvalue of A: then define

$$\tilde{A} := \begin{pmatrix} A & 0 \\ 0 & a \end{pmatrix}.$$

It is easy to check that this matrix satisfies the criteria.
(iv) Suppose a is an eigenvalue of A. We assume A on Jordan normal form, and let $B_1, \ldots, B_m$ be the blocks of A. We assume that $B_m := B(a, n_m)$ is the largest block having a as eigenvalue (or one of the largest, if there are more). Now define the $(n+1) \times (n+1)$ matrix

$$\tilde{A} := \begin{pmatrix} B_1 & \ldots & 0 & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \ldots & B_{m-1} & 0 \\ 0 & \ldots & 0 & \tilde{B}_m \end{pmatrix}$$

where $\tilde{B}_m := B(a, n_m + 1)$, i.e. a block with the same eigenvalue as B_m but one size larger. We leave it to the reader to verify the following three steps: The minimum polynomial of $\tilde{A}$ will now be (1) of degree one more than the minimum polynomial of A; (2) divisible by the minimum polynomial $m(X)$ of A; (3) divisible by the minimum polynomial of $\tilde{B}_m$: $(X-a)^{n_m+1}$. Since $(X-a)^{n_m}$ divides $m(X)$ but $(X-a)^{n_m+1}$ does not, the minimum polynomial of A is $m(X)(X-a)$.

□

Theorem 5.7. *Let $A \in GL_n(\mathbb{C})$, and $Q(X) := X^p - c_{p-1}X^{p-1} - \ldots - c_1X - c_0 \in \mathbb{C}[X]$ such that $Q(A) = 0$. Let $c = (c_{p-1}, \ldots, c_0)^t$, and $\mu_i(z) := (C_Q^z c)$. Then $A^z = \sum_{i=1}^{p} \mu_i(z) A^{-i}$.*

Proof. Suppose the minimum polynomial of A has degree $q \leq p$. By lemma 5.6 we can find a matrix $\tilde{A} \in Mat_{n+p-q}(\mathbb{C})$ such that it (1) has minimum polynomial $Q(X)$, (2) the upper left $n \times n$ part of $\tilde{A}^z$ is equal to A^z for every $i \in \mathbb{Z}$.

By proposition 5.4 we have $\tilde{A}^z = \sum_{i=1}^{p} \mu_i(z)\tilde{A}^{-i}$. Now restricting to the upper left $n \times n$ coefficients we have the equality $A^z = \sum_{i=1}^{p} \mu_i(z) A^{-i}$. □

6 Practical computation of the analytic functions

When one would like to compute the analytic functions $\mu_i(z)$ in a practical situation, given a relation $Q(A) := A^p - c_{p-1}A^{p-1} - \ldots - c_1 A - c_0 = 0$, one may now use theorem 5.7 as a basis for a more efficient computation. In stead of computing C_Q^z, it may be more easy to know which form the analytic functions have, and compute it directly. Since C_Q is a conjugation of a Jordan-normal form J, we know that C_Q^z is a conjugation of J^z. The Jordan normal form J of C_Q can be computed by finding the roots with multiplicity of $Q(X)$: $\lambda_1, \ldots, \lambda_m$ with multiplicity $n_1, \ldots, n_m$. Now one knows which analytic functions occur in J^z: let $B_1, \ldots, B_m$ be the blocks of J, then block $B_i := \sum_{j=0}^{p-1} g_j(z)\lambda^{z-j}N^j$ where $g_j(z) := z(z-1)\ldots(z-j+1)$ if $j > 0$ and $g_0(z) = 1$ (See definition 3.2 and lemma 3.3). Define the functions $f_1(z), \ldots, f_p(z)$ as the p functions appearing as coefficients in J^z: $g_j(z)\lambda_i^{z-j}$ where $1 \leq i \leq m$, $0 \leq j \leq n_i - 1$, one knows that C_Q^z has entries which are linear combinations of the f_i, and thus so does $C_Q^z c$. So, in a practical situation it may be easier to compute e_{ij} such that

$$A^z = \sum_{i=1}^{p} \Big(\sum_{j=1}^{p} e_{ij} f_j(z) \Big) A^{-i}.$$

In fact, we claim that the $p \times p$ matrix (e_{ij}) can be computed as the inverse of the $p \times p$ matrix $(f_j(-i))$:

Lemma 6.1. *The matrix $B := (f_i(-j))$ (where the rows i and the colums j run from 1 to p) is invertible, and*

$$A^z = \sum_{i=1}^{p} \Big(\sum_{j=1}^{p} (B^{-1})_{ij} f_j(z) \Big) A^{-i}.$$

Proof. One wants to find e_{ij} such that

$$A^z = \sum_{i=1}^{p} \Big(\sum_{j=1}^{p} e_{ij} f_j(z) \Big) A^{-i}.$$

Now notice that this is equivalent to

$$A^z = \begin{pmatrix} f_1(z), \ldots, f_p(z) \end{pmatrix} \begin{pmatrix} e_{11} & \ldots & e_{p1} \\ \vdots & & \vdots \\ e_{1p} & & e_{pp} \end{pmatrix} \begin{pmatrix} A^{-1} \\ \vdots \\ A^{-p} \end{pmatrix}.$$

Substituting $z = -1, \ldots, -p$ one gets

$$\begin{pmatrix} A^{-1} \\ \vdots \\ A^{-p} \end{pmatrix} = \begin{pmatrix} f_1(-1) & \ldots & f_p(-1) \\ \vdots & & \vdots \\ f_1(-p) & \ldots & f_p(-p) \end{pmatrix} \begin{pmatrix} e_{11} & \ldots & e_{p1} \\ \vdots & & \vdots \\ e_{1p} & & e_{pp} \end{pmatrix} \begin{pmatrix} A^{-1} \\ \vdots \\ A^{-p} \end{pmatrix}.$$

In case $(f_i(-j))_{ij}$ is an invertible matrix, then the e_{ij} are unique and thus must be the solution that exists according to theorem 5.7. Such a matrix $(f_i(-j))_{ij}$ is called a *generlized vandermonde matrix* and it is proven in [3] theorem 1 that these matrices have a determinant which is nonzero. Thus , $f_j(-i)$ (rows i, colums j) is the inverse of e_{ij} (rows j, colums i), which gives the result. $\square$

7 Summary

For the reader's convenience, we will summarize the results of the last section in a concise and clear way.

Suppose a matrix A satisfies a relation

$$A^p - c_{p-1}A^{p-1} - \ldots - c_1 A - c_0 I = 0.$$

Then one can define a formula for A^z in the following way:

- Compute the roots of $X^p - c_{p-1}X^{p-1} - \ldots - c_1 X - c_0 \in \mathbb{C}[X]$ and count their multiplicity. Name the roots $\lambda_1, \ldots, \lambda_m$ with multiplicities $n_1, \ldots, n_m$.
- Choose for each $1 \leq i \leq m$ a function λ_i^z.
- Define for each $1 \leq i \leq m$, $0 \leq j \leq n_i - 1$ the functions $h_{ij} := \lambda^{z-j} g_j(z)$ where $g_0(z) = 1, g_j(z) = \frac{1}{j!} z(z-1)\cdots(z-j+1)$ if $j \geq 1$. Rename the p functions h_{ij} obtained in this way as $f_1(z), \ldots, f_p(z)$.
- Compute the inverse of $(f_i(-j))_{ij}$ where i, j run from 1 to p, and name the inverse e_{ij}.
- Now
$$A^z = \sum_{i=1}^{p} \sum_{j=1}^{p} e_{ij} f_j(z) A^{-i}.$$

Acknowledgements
The author would like to express his gratitude to prof. Jean-Philippe Furter, prof. Arno van den Essen and dr.ing. A.J.E.M.Janssen for some priceless (email) discussion and corrective work.

References

[1] J-Ph. Furter, S.Maubach, *Locally Finite Polynomial Endomorphisms and an extension of the Cayley-Hamilton Theorem*, accepted (2006) in Journal of Pure and Applied Algebra.

[2] S.Maubach *Polynomial maps and Kernels of Derivations*, PhD thesis, University of Nijmegen, (2003)

[3] G. Sobczyk, *Generalized vandermonde determinants and applications*, Serie Comunicaciones **30** (2002) 3-17

[4] http://mathworld.wolfram.com/CompanionMatrix.html