The space of all $p$-th roots of a nilpotent complex matrix is path-connected

Clément de Seguins Pazzis

March 17, 2022

Abstract

Let $p$ be a positive integer and $A$ be a nilpotent complex matrix. We prove that the set of all $p$-th roots of $A$ is path-connected.

AMS Classification: 15A24; 54D05.

Keywords: Matrices, Nilpotency, Path-connectedness, Jordan normal form.

1 Introduction

Let $U$ be an open subset of the field $\mathbb{C}$ of complex numbers, $f : U \to \mathbb{C}$ be an analytic function and $n$ be a positive integer. Given a matrix $A \in M_n(\mathbb{C})$, it is natural to ask whether the matrix equation $f(X) = A$, with unknown $X \in M_n(\mathbb{C})$, has at least one solution. By using the fact that $X$ commutes with $f(X)$, and by using the characteristic subspaces of $A$, this problem can be reduced to the one of deciding whether the equation $g(X) = N$ has a solution, where $N$ is a given nilpotent matrix, and $g$ is a given analytic function.

There is a (not very satisfying) answer to that question, and we shall recall it in short notice. Given a nilpotent matrix $A \in M_n(\mathbb{C})$ and a positive integer $k$, we denote by $m_k(A)$ the number of Jordan cells of size $k$ in the Jordan normal form of $A$. The sequence $(m_k(A))_{k \geq 1}$ is called the (Jordan) profile of $A$. It belongs to the additive semigroup $\mathbb{N}^{(\mathbb{N}^*)}$ of all sequences of non-negative integers.
with finite support and indexed over the positive integers (here, \( \mathbb{N} \) denotes the set of all non-negative integers, and \( \mathbb{N}^* \) the one of all positive integers). More generally, any element of \( \mathbb{N}^{(\mathbb{N}^*)} \) is called a profile. Two nilpotent matrices are similar if and only if they have the same Jordan profile. Throughout the article, profiles will be seen as elements of the abelian group \( \mathbb{Z}^{(\mathbb{N}^*)} \) of all sequences of integers with finite support.

Given \( k \in \mathbb{N}^* \), we denote by \( J_k \in M_k(\mathbb{C}) \) the Jordan cell of size \( k \) (i.e. the matrix of \( M_k(\mathbb{C}) \) in which the entry at the \((i, i+1)\)-spot equals 1 for all \( i \in [1, k-1] \), and all the other entries equal 0), and we denote its profile by \( e_k \) (so that \((e_k)_i = 1 \) if \( i = k \), and \((e_k)_i = 0 \) otherwise). We convene that \( J_0 \) is the 0-by-0 matrix and that \( e_0 \) is the zero sequence in \( \mathbb{N}^{(\mathbb{N}^*)} \).

The following result is folklore:

**Lemma 1.** Let \( k \) and \( p \) be positive integers. Then \( J_k^p \) is similar to the direct sum of \( k - pa \) copies of \( J_{a+1} \) and of \( p(a+1) - k \) copies of \( J_a \), for every non-negative integer \( a \) such that \( pa \leq k \leq p(a+1) \) (in particular, this holds when \( a \) is the quotient of \( k \) modulo \( p \)).

From there, one proves (see Appendix A for details) that, given a nilpotent matrix \( A \in M_n(\mathbb{C}) \), the equation \( f(X) = A \) has a solution if and only if the profile of \( A \) belongs to the sub-semigroup of \( \mathbb{N}^{(\mathbb{N}^*)} \) generated by the profiles of the form \( r \cdot e_{a+1} + (p-r) \cdot e_a \) – where \( a \) is a non-negative integer, \( p \) is the finite multiplicity of some zero of \( f \), and \( r \in [0, p] \) – and the profile \( e_1 \) if some zero of \( f \) has infinite multiplicity (i.e. \( f \) is constant on the connected component of that zero). In particular, if \( f \) has at least one simple zero then the equation \( f(X) = A \) has a solution for every nilpotent matrix \( A \).

The above characterization is not very convenient though. In very special cases, one can formulate an equivalent one that can easily be tested: a nilpotent matrix \( A \) has a \( p \)-th root if and only if, for all \( k \in \mathbb{N}^* \), the integer \( p - m_k(A) \) is less than or equal to the remainder of \( \sum_{j=k+1}^{+\infty} m_j(A) \) modulo \( p \) provided that this remainder is non-zero (for example, if \( p = 2 \) this means that \( m_k(A) > 0 \) whenever \( \sum_{j=k+1}^{+\infty} m_j(A) \) is odd). Moreover this result holds not only over the field of complex numbers, but over any skew field. If \( f \) has exactly two zeroes, one with multiplicity 2 and one with multiplicity 3 (e.g. if \( f : z \mapsto z^3(z-1)^2 \)), then, given a nilpotent matrix \( A \in M_n(\mathbb{C}) \), the equation \( f(X) = A \) has a solution if and only if there is no pair \((k, l)\) of positive integers for which \( m_k(A) = m_{k+2l}(A) = 0 \).
and $m_{k+i}(A) = 1$ for all $i \in [1, 2l - 1]$. We leave these results as exercises for the reader.

Here, we will stick to the equation $X^p = A$ for a fixed nilpotent complex matrix $A$ and a fixed positive integer $p$. When this equation has a solution, we are interested in the topological structure of its solution set $A^{1/p}$, i.e. the set of all $p$-th roots of $A$. Note that all the matrices in $A^{1/p}$ are nilpotent.

A very ambitious goal is to understand the homotopy type of $A^{1/p}$. As a first step towards that goal, we will consider here its path-connectedness. Here is our main theorem:

**Theorem 2.** Let $p$ be a positive integer and $A$ be a nilpotent complex matrix. Then the set $A^{1/p}$ is path-connected.

The case $p = 1$ is straightforward. In the remainder of this section, we fix an integer $p > 1$ and a nilpotent matrix $A \in M_n(\mathbb{C})$. Given $m \in \mathbb{N}^{(N^*)}$, we denote by $A^{1/p}_m$ the subset of all $N \in A^{1/p}$ with profile $m$ (of course this subset may be empty). We denote by $P_p(A)$ the set of all profiles $m$ such that $A^{1/p}_m$ is non-empty. Hence, the family $(A^{1/p}_m)_{m \in P_p(A)}$ yields a partition of $A^{1/p}$.

Two profiles $m$ and $m'$ are called $p$-adjacent, and we write $m \sim_p m'$, when there exist non-negative integers $a, k, l$ such that $p^a \leq k < l \leq p^{a+1}$ and

$$m - m' = \pm(e_k + e_l - e_{k+1} - e_{l-1}).$$

Finally, we denote by $A_p(A)$ the set of all pairs $\{m, m'\}$ of distinct $p$-adjacent elements of $P_p(A)$. Thus, we have defined a non-oriented graph $(P_p(A), A_p(A))$.

The definition of $p$-adjacency is motivated by the following basic result:

**Lemma 3.** Let $a, k, l$ be integers such that $0 \leq pa \leq k < l \leq p(a + 1)$. Then the matrices $(J_k \oplus J_l)^p$ and $(J_{k+1} \oplus J_{l-1})^p$ are similar.

**Proof.** Denote respectively by $r$ and $s$ the remainders of $k$ and $l - 1$ modulo $p$. By Lemma 1, we find that $(J_k \oplus J_l)^p$ is similar to the direct sum of $r + (s + 1)$ copies of $J_a$ and of $p - r + (p - s - 1)$ copies of $J_a$. Likewise, $(J_{k+1} \oplus J_{l-1})^p$ is similar to the direct sum of $(r + 1) + s$ copies of $J_a$, and of $(p - r - 1) + (p - s)$ copies of $J_a$. The claimed result ensues.

We are now able to state the three steps of our proof of Theorem 2:

**Lemma 4.** Let $m \in P_p(A)$. Then the space $A^{1/p}_m$ is path-connected.
Lemma 5. Let \( m, m' \) be adjacent profiles in \( P_p(A) \). Then there exist \( N \in A^{1/p}_m \) and \( N' \in A^{1/p}_{m'} \) together with a path from \( N \) to \( N' \) in \( A^{1/p} \).

Lemma 6. The graph \( (P_p(A), A_p(A)) \) is connected.

Combining those three results readily yields Theorem 2.

2 Proof of Theorem 2

Throughout this part, we let \( A \in M_n(\mathbb{C}) \) be a nilpotent matrix and \( p \) be a positive integer.

2.1 Proof of Lemma 4

Let \( m \) belong to \( P_p(A) \). Let \( X \) and \( Y \) belong to \( A^{1/p}_{m} \). The matrices \( X \) and \( Y \) are nilpotent with the same profile, and hence they are similar. Thus we have some \( P \in \text{GL}_n(\mathbb{C}) \) such that \( Y = PXP^{-1} \). Since \( X^p = Y^p = A \), we obtain that \( P \) belongs to the centralizer \( C(A) \) of \( A \) in the algebra \( M_n(\mathbb{C}) \). As \( C(A) \cap \text{GL}_n(\mathbb{C}) \) is a Zariski-open subset of the complex finite-dimensional vector space \( C(A) \), it is path-connected (see Lemma 7.2 in [5]). Choose a path \( Q : t \in [0, 1] \mapsto Q(t) \in C(A) \cap \text{GL}_n(\mathbb{C}) \) from \( I_n \) to \( P \). Then, one checks that \( q : t \in [0, 1] \mapsto Q(t)XQ(t)^{-1} \) is a path from \( X \) to \( Y \), and \( q(t)^p = Q(t)AQ(t)^{-1} = A \) for all \( t \in [0, 1] \). Finally, \( q(t) \) is similar to \( X \) for all \( t \in [0, 1] \), and hence its profile is \( m \). Hence, there is a path from \( X \) to \( Y \) in \( A^{1/p}_m \). This completes the proof of Lemma 4.

2.2 Proof of Lemma 5

As we will see, the proof of Lemma 5 boils down to the following basic result:

Lemma 7. Let \( a, k, l \) be integers such that \( 0 \leq pa \leq k < l \leq p(a+1) \). Set \( N := k + l \). Then there exists a path \( \gamma : [0, 1] \to M_N(\mathbb{C}) \) such that:

(i) \( \gamma(0) = J_k \oplus J_l \);

(ii) \( \gamma(1) \) is similar to \( J_{k+1} \oplus J_{l-1} \);

(iii) the mapping \( t \in [0, 1] \mapsto \gamma(t)^p \) is constant.
Proof. We shall think in terms of endomorphisms of $\mathbb{C}^N$: denote by $u$ the endomorphism of $\mathbb{C}^N$ represented by $J_k \oplus J_l$ in the standard basis $(x_k, \ldots, x_1, y_1, \ldots, y_l)$ of $\mathbb{C}^N$. We convene that $y_j = 0$ for all $j > l$, and that $x_i = 0$ for all $i > k$. Hence, $u$ maps $x_i$ to $x_{i+1}$ for all $i > 0$, and it maps $y_j$ to $y_{j+1}$ for all $j > 0$. Given $t \in [0,1]$, define $u_t$ as the endomorphism of $\mathbb{C}^N$ on the standard basis by $u_t(y_1) = (1-t)y_2 + tx_1$, and by mapping any other vector $z$ of that basis to $u(z)$. Clearly, $t \in [0,1] \mapsto u_t$ is a path in the space of all endomorphisms of $\mathbb{C}^N$, and $u_0 = u$.

Next, one sees that $u_1$ is represented by the matrix $J_{k+1} \oplus J_{l-1}$ in the basis $(x_k, \ldots, x_1, y_1, \ldots, y_l)$.

Next, let $t \in (0,1)$. One checks that $(x_k, \ldots, x_1, (1-t)y_l + tx_{l-1}, \ldots, (1-t)y_2 + tx_1, y_1)$ is a basis of $\mathbb{C}^N$, and the matrix of $u_t$ in that basis is $J_k \oplus J_l$. Hence, $u_t$ is similar to $u_0$, and it follows that $u_t^p$ is similar to $u_0^p$. Besides, Lemma 3 shows that $u_t^p$ is also similar to $u_0^p$.

Now, for $t \in [0,1]$, denote by $U_t$ the matrix of $u_t$ in the standard basis of $\mathbb{C}^N$. It follows from the above that $t \in [0,1] \mapsto U_t$ is a path, in the space $M_N(\mathbb{C})$, from $J_k \oplus J_l$ to a matrix that is similar to $J_{k+1} \oplus J_{l-1}$, and that the path $t \in [0,1] \mapsto (U_t)^p$ takes its values in the similarity class $S(U_0^p)$ of the matrix $U_0^p$.

It is folklore that the mapping $P \in GL_N(\mathbb{C}) \mapsto PU_0^pP^{-1} \in S(U_0^p)$ is a fibration (it is a principal fibre bundle whose structural group is the group of all invertible elements of the centralizer of $U_0^p$): see Appendix B for a short elementary proof, and the combination of Theorem 1.4.3 and Proposition 1.4.6 of [1] and Proposition 8.3 of [2] for a more sophisticated one. Hence, there is a path $q : [0,1] \rightarrow GL_N(\mathbb{C})$ such that

$$\forall t \in [0,1], \quad U_t^p = q(t)U_0^p q(t)^{-1} \quad \text{and} \quad q(0) = I_N.$$

Finally, we consider the path $\gamma : t \in [0,1] \mapsto q(t)^{-1}U_t q(t) \in M_N(\mathbb{C})$. The above properties of $q$ show that $t \mapsto \gamma(t)^p$ is constant. Next, $\gamma(0) = U_0 = J_k \oplus J_l$. Finally, $\gamma(1)$ is similar to $U_1$ and hence to $J_{k+1} \oplus J_{l-1}$. \qed

Now, we can prove Lemma 5. Let $m, m'$ be distinct adjacent profiles in $P_p(A)$. We wish to prove that some element of $A_1^1/m$ is path-connected in $A_1^1/p$ to some element of $A_1^1/m'$. Without loss of generality, we can assume that there is a non-negative integer $a$ together with elements $k < l$ of $[pa, p(a+1)]$ such that $m - m' = e_k + e_l - e_{k+1} - e_{l-1}$. As $m \neq m'$, we must have $l > k + 1$, and it follows that $m_k > 0$ and $m_l > 0$. Let us choose $N \in A_1^1/m$. Then $N$ has at
least one Jordan cell of each size $k$ and $l$. Hence, $N = P(B \oplus J_k \oplus J_l)P^{-1}$ for some nilpotent matrix $B$ and some $P \in \text{GL}_n(\mathbb{C})$. The profile of $B$ is obviously $m - e_k - e_l$.

Let us take a path $\gamma$ that satisfies the conclusion of Lemma 7 for the pair $(k, l)$: then, $q : t \in [0, 1] \mapsto P(B \oplus \gamma(t))P^{-1}$ is a path in $M_n(\mathbb{C})$, and we see from condition (iii) in Lemma 7 that $t \mapsto q(t)p$ is constant with value $q(0)p = Np = A$. In other words, $q$ is a path in $A^{1/p}$. Finally, $q(1)$ is similar to $B \oplus \gamma(1)$, and hence to $B \oplus J_{k+1} \oplus J_{l-1}$, whose profile equals $(m - e_k - e_l) + e_{k+1} + e_{l-1} = m'$. Hence, $q(1) \in A^{1/p}_{m'}$. This completes the proof of Lemma 5.

2.3 Proof of Lemma 6

We start with some preliminary notation. Given an element $m \in \mathbb{Z}(N^*)$, we set $S(m) := \sum_{k=1}^{+\infty} km_k$ (called the size of $m$), and $m^{[p]} := \left( \sum_{p<k<p} (p - |k|) m_{pa+k} \right)_{a \geq 1}$, which is an element of $\mathbb{Z}([N^*)$. Note that both maps $S : \mathbb{Z}(N^*) \rightarrow \mathbb{Z}$ and $m \in \mathbb{Z}([N^*) \mapsto m^{[p]} \in \mathbb{Z}([N^*)$ are group homomorphisms.

Using the results recalled in the introduction, one sees that if $m$ is the profile of some nilpotent matrix $N$, then $m^{[p]}$ is the profile of $N^p$, while $S(m)$ is obviously the number of rows of $N$, and hence $S(m^{[p]}) = S(m)$. Besides, using Lemma 3, we find that $m^{[p]} = (m')^{[p]}$ for any two $p$-adjacent profiles $m$ and $m'$.

Given profiles $m$ and $m'$, a p-chain of profiles from $m$ to $m'$ is a list $(a^{(0)}, \ldots, a^{(N)})$ of profiles such that $a^{(i)}_p \sim a^{(i+1)}_p$ for all $i \in [0, N-1]$, and $m = a^{(0)}$ and $m' = a^{(N)}$.

From there, Lemma 6 can be seen as a reformulation of the following result:

**Lemma 8.** Let $m, m'$ be two profiles such that $m^{[p]} = (m')^{[p]}$. Then there is a p-chain of profiles from $m$ to $m'$.

**Proof.** Note that the assumptions yield $S(m) = S(m^{[p]}) = S((m')^{[p]}) = S(m')$. We will prove the result by induction on the size of $m$. 

6
The result is obvious if \( S(m) = 0 \): in that case both \( m \) and \( m' \) equal the zero sequence, and we simply take the trivial chain \((m)\). Assume now that \( S(m) > 0 \).

Assume first that there exists an integer \( k \geq 1 \) such that \( m_k > 0 \) and \( m'_k > 0 \). Then \( m - e_k \) and \( m' - e_k \) obviously satisfy the assumptions, and their size equals \( S(m) - k \). By induction, there is a \( p \)-chain \((a^{(0)}, \ldots, a^{(N)})\) of profiles from \( m - e_k \) to \( m' - e_k \). Clearly, \((a^{(0)} + e_k, \ldots, a^{(N)} + e_k)\) is a \( p \)-chain of profiles from \( m \) to \( m' \).

Hence, in the remainder of the proof we assume that \( m_k m'_k = 0 \) for all \( k \geq 1 \). Denote by \( q \) the greatest positive integer such that \( m_q + m'_q > 0 \). Without loss of generality, we can assume that \( m'_q > 0 \) (and hence \( m_q = 0 \)). Denote by \( a \) the least (non-negative) integer such that \( q \in \left[p a, p(a + 1)\right] \), so that \( q > p a \). Hence, \( m_{a+1}^q = (m'_a)^q \geq q - p a \). In particular, \( m_k > 0 \) for some \( k \in \left[p a + 1, p(a + 1)\right] \), and we consider the greatest such integer \( k \). Note that \( p a < k < q \). If \( m_k > 1 \), we note that \( m - 2 e_k + e_{k+1} + e_{k-1} \) is still a profile that is \( p \)-adjacent to \( m \). If \( m_k = 1 \), then having \( m_{a+1}^q \geq q - p a \) we must also have \( m_l > 0 \) for some \( l \in \left[p a + 1, k - 1\right] \), and then we note that \( m - e_k - e_l + e_{k+1} + e_{l-1} \) is a profile. In any case, we have found a profile \( a^{(k+1)} \) that is \( p \)-adjacent to \( m \) and for which \( k + 1 \) is the greatest integer \( i \) such that \( a^{(k+1)} \) is \( \geq 0 \). Continuing by finite induction, we create a \( p \)-chain \((a^{(k)}, a^{(k+1)}, \ldots, a^{(q)})\) of profiles from \( m \) to some profile \( a^{(q)} \) such that \( (a^{(q)})^q > 0 \). Hence \( (a^{(q)})^p = \cdots = a^{(k)})^p = m^p = (m')^p \). As \((a^{(q)})^q > 0 \), the first case tackled in the above yields a \( p \)-chain of profiles from \( a^{(q)} \) to \( m' \). Linking those \( p \)-chains yields a \( p \)-chain of profiles from \( m \) to \( m' \).

Lemmas 4 to 6 are now proved, and hence Theorem 2 is established.

3 Further questions

Now that Theorem 2 has been proved, we wish to suggest several related open problems. First, given an analytic function \( f : U \to \mathbb{C} \), what are the nilpotent complex matrices \( A \) for which the set of all solutions of the equation \( f(X) = A \) is path-connected? More precisely, is there a simply characterization of such matrices in terms of the profile of \( A \) and the zeroes of \( f \) (and their multiplicities)?

Next, given a positive integer \( p \), we wonder about the homotopy type of \( A^{1/p} \). For example, if \( A = 0 \) then \( A^{1/p} \) is contractible (since it is star-shaped around 0). However, for \( E := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \), one checks that \( E^{1/2} \) is the set of all matrices
of the form \[
\begin{bmatrix}
0 & x & y \\
0 & 0 & x^{-1} \\
0 & 0 & 0
\end{bmatrix},
\] a space that is homeomorphic to \((\mathbb{C} \setminus \{0\}) \times \mathbb{C}\) and hence homotopy equivalent to the circle \(S^1\) (and not contractible!). Is there a simple way to compute the homotopy type of \(A^{1/p}\) as a function of \(p\) and the profile of \(A\)? Computing the fundamental group of \(A^{1/p}\) would be interesting, for a start.

There are other interesting open questions related to the real and quaternionic cases. The set of all square roots of \(E\) with real entries is homeomorphic to \((\mathbb{R} \setminus \{0\}) \times \mathbb{R}\), and hence it has exactly two path-connected components. Is there a sensible way to compute the number of path-connected components of the set of all \(p\)-th roots of \(A\) (with real entries) as a function of \(p\) and of the profile of \(A\)? In that prospect, it is worthwhile to note that the real equivalent of Lemmas 5 and 6 holds (with the same proof): the only step that fails is the real equivalent of Lemma 4. Nevertheless, the set of all real \(p\)-th roots of \(A\) is a real affine variety, and hence it has finitely many path-connected components (alternatively, one can adapt the proof of Lemma 4 to yield that \(A^{1/p}\) has finitely many path-connected components, using the fact that \(C(A) \cap GL_n(\mathbb{R})\) is a Zariski open subset of a finite-dimensional real vector space, see [4], Section 2.4).

Finally, there are similar issues in the quaternionic case: in that one however we have not succeeded in finding a single example of a nilpotent quaternionic matrix \(A\) and of a positive integer \(p\) such that the set of all \(p\)-th roots of \(A\) is not path-connected.

**Appendix**

**A When does the equation \(f(X) = N\) have a solution?**

Let \(U\) be an open subset of \(\mathbb{C}\) and \(f : U \to \mathbb{C}\) be an analytic function. Let \(N \in M_n(\mathbb{C})\) be nilpotent. We wish to characterize the existence of a solution to the equation \(f(X) = N\) with unknown \(X \in M_n(\mathbb{C})\).

**Lemma 9.** Let \(N \in M_n(\mathbb{C})\) be a Jordan cell and \(x\) be a zero of \(f\) with finite multiplicity \(p\). Write \(n = mp + r\) the Euclidean division of \(n\) by \(p\). Then \(f(xI_n + N)\) is similar to the direct sum of \(r\) Jordan cells of size \(m + 1\) and of \(p - r\) Jordan cells of size \(m\).

**Proof.** This result is known by Lemma 4 if \(f : z \mapsto (z - x)^p\), in which case \(f(xI_n + N) = X_1 \oplus \ldots \oplus X_r\), where \(X_i = (xI_{m+1} + N_i)\) for some matrices \(N_i\) of size \(m + 1\). If \(f : z \mapsto (z - x)^p\), then \(f(xI_n + N) = (xI_n + N)^p\), which is similar to \(X_1 \oplus \ldots \oplus X_r\), where \(X_i = (xI_{m+1} + N_i)\) for some matrices \(N_i\) of size \(m + 1\).
In the general case we factorize \( f : z \mapsto (z - x)^pg(z) \) for some analytic function \( g \) on \( U \). Using the commutation of \( P := g(xI_n + N) \) with \( N \), we see that \( N^pP \) is nilpotent and \( \text{rk}( (N^pP)^k ) = \text{rk}( (N^p)^k P^k ) = \text{rk}( (N^p)^k ) \) for every non-negative integer \( k \). Classically, the similarity class of a nilpotent matrix \( M \) is characterized by the sequence of ranks \( (\text{rk}(M^k))_{k \geq 0} \), and hence \( N^pP \simeq N^p \), which completes the proof.

If, on the other hand, \( x \) is a zero of \( f \) with infinite multiplicity (i.e. \( f \) vanishes on a whole neighborhood of \( x \)) then \( f(xI_n + N) = \sum_{k=0}^{+\infty} \frac{f^{(k)}(x)}{k!} N^k = 0 \) for every nilpotent matrix \( N \) of \( M_n(\mathbb{C}) \), so \( f(xI_n + N) \) is the direct sum of \( n \) Jordan cells of size 1.

Now, let \( X \in M_n(\mathbb{C}) \) be such that \( f(X) = N \). The eigenvalues of \( f(X) \) are the images under \( f \) of those of \( X \), and hence the eigenvalues of \( X \) are zeroes of \( f \). Using the Jordan reduction theorem, we obtain

\[
X \simeq (x_1 I_{d_1} + N_1) \oplus \cdots \oplus (x_p I_{d_p} + N_p)
\]

where \( x_1, \ldots, x_p \) are zeroes of \( f \) and \( N_1, \ldots, N_p \) are Jordan cells with respective positive sizes \( d_1, \ldots, d_p \). Therefore

\[
N = f(X) \simeq f(x_1 I_{d_1} + N_1) \oplus \cdots \oplus f(x_p I_{d_p} + N_p)
\]

and it follows that the Jordan profile of \( N \) is the sum of the Jordan profiles of the matrices \( f(x_k I_{d_k} + N_k) \). Using Lemma 9 and the remark thereafter, we deduce the “only if” part in the following statement:

**Theorem 10.** Let \( N \in M_n(\mathbb{C}) \) be nilpotent. The following conditions are equivalent:

(i) There exists a matrix \( X \in M_n(\mathbb{C}) \) such that \( f(X) = N \).

(ii) The Jordan profile of \( N \) belongs to the sub-semigroup of \( \mathbb{N}^{(N^p)} \) generated by the elements of the form \( (p - r) \cdot e_a + r \cdot e_{a+1} \) where \( p \) is the (finite) multiplicity of some zero of \( f \), \( a \) is an arbitrary non-negative integer and \( r \) belongs to \([0, p]\), together with the additional element \( e_1 \) if \( f \) has a zero with infinite multiplicity.

The “if” part of the above statement is proved in a similar fashion as the “only if” part.
B The fibration $P \mapsto PAP^{-1}$

Here, $\mathbb{F}$ denotes one of the fields $\mathbb{R}$ or $\mathbb{C}$. Let $A \in M_n(\mathbb{F})$. Denote by $C(A)$ the centralizer of $A$ in the algebra $M_n(\mathbb{F})$, by $C(A)^\times$ its group of invertible elements, and by $S(A)$ the similarity class of $A$. We wish to prove that the mapping $\pi : P \in GL_n(\mathbb{F}) \mapsto PAP^{-1} \in S(A)$ defines a $C(A)^\times$-principal bundle. For the continuous left-action $(P, M) \mapsto PMP^{-1}$ of $GL_n(\mathbb{F})$ on $M_n(\mathbb{F})$, the stabilizer of $A$ is $C(A)^\times$, and hence classically it suffices to prove that the mapping $\pi$ admits a local cross-section around $A$.

The proof is based upon the following elementary lemma:

**Lemma 11.** Let $V$ be a finite-dimensional vector space over $\mathbb{F}$. Let $u \in \text{End}(V)$, and let $x_0 \in V$ be a non-zero vector such that $u(x_0) = 0$. Then there exists a neighborhood $U$ of $u$ in $\text{End}(V)$, together with a continuous mapping $f : U \to V$ such that $v[f(v)] = 0$ for all $v \in U$ with the same rank as $u$, and $f(u) = x_0$.

**Proof.** Denote by $n$ the dimension of $V$, and by $p$ the rank of $u$. Let us extend $x_0$ first into a basis $(e_{n-p}, \ldots, e_n)$ of the kernel of $u$, with $e_n = x_0$, and then into a basis $B := (e_1, \ldots, e_n)$. We extend the linearly independent $p$-tuple $(u(e_1), \ldots, u(e_p))$ into a basis $C := (u(e_1), \ldots, u(e_p), f_{p+1}, \ldots, f_n)$ of $V$. In the bases $B$ and $C$, the matrix of $u$ reads

$$\begin{bmatrix} I_p & 0_{p \times (n-p)} \\ 0_{(n-p) \times p} & 0_{(n-p) \times (n-p)} \end{bmatrix}.$$ 

For any $v \in \text{End}(V)$, let us write its matrix in the bases $B$ and $C$ as

$$M(v) = \begin{bmatrix} A(v) & C(v) \\ B(v) & D(v) \end{bmatrix}$$

along the same pattern. The mapping $v \in \text{End}(V) \mapsto A(v) \in M_p(\mathbb{F})$ is linear, and hence continuous. It follows that

$$U := \{ v \in \text{End}(V) : A(v) \in GL_p(\mathbb{F}) \}$$

is an open subset of $\text{End}(V)$ that contains $u$.

Next, let $v \in U$. Consider the invertible matrix

$$N(v) := \begin{bmatrix} I_p & -A(v)^{-1}C(v) \\ 0_{(n-p) \times p} & I_{n-p} \end{bmatrix} \in \text{GL}_n(\mathbb{F}),$$

10
so that \( M(v)N(v) = \begin{bmatrix} A(v) & 0_{p \times (n-p)} \\ B(v) & 0 \end{bmatrix} \) has the same rank as \( M(v) \). Assume that \( v \) has rank \( p \). Since \( A(v) \) has rank \( p \), it follows that the last \( n - p \) columns of \( M(v)N(v) \) equal zero, and in particular \( M(v) \) annihilates the last column of \( N(v) \).

For \( v \in U \), denote by \( f(v) \) the vector of \( V \) whose matrix in \( \mathcal{B} \) is the last column of \( N(v) \); obviously \( f : U \to V \) is continuous, and the previous study shows that \( v[f(v)] = 0 \) for all \( v \in U \) with rank \( p \). Finally, \( f(u) = e_n = x_0 \). □

Remark 1. Set \( p := \text{rk } u \) and define \( \text{End}_p(V) \) as the set of all endomorphisms of \( V \) with rank \( p \), and \( \xi : (u, x) \in \text{End}_p(V) \times V \mapsto u \in \text{End}_p(V) \) as the trivial vector bundle with fiber \( V \) and base space \( \text{End}_p(V) \). The mapping \( f : (u, x) \mapsto (u, u(x)) \) is obviously a \( \text{End}_p(V) \)-bundle morphism from \( \xi \) to itself with constant rank \( p \), therefore its kernel, which equals

\[
\left\{ \begin{array}{l}
\{(u, x) \in \text{End}_p(V) \times V : u(x) = 0\} \\
(u, x) \end{array} \right\} \to \text{End}_p(V) \to u,
\]

is also a vector bundle: see [3], Chapter 3 Theorem 8.2. The above result can then be obtained by using a local trivialization of this bundle.

We are now ready to construct the claimed local cross-section. Consider the endomorphism \( \text{ad}_A : M \mapsto AM - MA \) of the vector space \( \mathcal{M}_n(\mathbb{F}) \). Denote by \( p \) its rank. applying the above lemma, we find a neighborhood \( U \) of \( \text{ad}_A \) in \( \text{End}(\mathcal{M}_n(\mathbb{F})) \) together with a continuous mapping \( f : U \to \mathcal{M}_n(\mathbb{F}) \) such that \( f(\text{ad}_A) = I_n \) and \( v(f(v)) = 0 \) for all \( v \in U \) with rank \( p \). The mapping

\[ \Phi : B \in \mathcal{M}_n(\mathbb{F}) \mapsto [M \mapsto BM - MA] \in \text{End}(\mathcal{M}_n(\mathbb{F})) \]

is affine, and hence continuous: thus \( U_0 := \Phi^{-1}(U) \) is a neighborhood of \( A \) in \( \mathcal{M}_n(\mathbb{F}) \). We set

\[ g : B \in U_0 \cap S(A) \mapsto f(\Phi(B)) \in \mathcal{M}_n(\mathbb{F}), \]

so that \( g(A) = I_n \). Since \( g \) is continuous, \( U'_0 := g^{-1}(\text{GL}_n(\mathbb{F})) \) is a neighborhood of \( A \) in \( S(A) \).

We will conclude the proof by showing that the restriction \( g|_{U'_0} \) is a local cross-section for the mapping \( P \in \text{GL}_n(\mathbb{F}) \mapsto PAP^{-1} \in S(A) \).

Let \( B \in U'_0 \). Since \( B \in S(A) \), there is a matrix \( Q \in \text{GL}_n(\mathbb{F}) \) such that \( B = QAQ^{-1} \). It follows that \( \Phi(B) = L_Q \circ \text{ad}_A \circ L_Q^{-1} \) where \( L_N : M \mapsto NM \) for all \( N \in \mathcal{M}_n(\mathbb{F}) \). Hence, \( \text{rk } \Phi(B) = \text{rk}(\text{ad}_A) = p \). It follows that \( \Phi(B)[g(B)] = 0 \), that is
$Bg(B) = g(B)A$. Moreover, $g(B)$ is invertible, and hence $B = g(B)Ag(B)^{-1}$, as claimed.

References

[1] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*. Graduate Texts in Mathematics 98, Springer-Verlag, Berlin Heidelberg New York, 1985.

[2] J. Humphreys, *Linear algebraic groups*. Graduate Texts in Mathematics 21, Springer-Verlag, Berlin Heidelberg New York, 1975.

[3] D. Husemoller, *Fibre bundles, 3rd ed.*, Graduate Texts in Mathematics 20, Springer-Verlag, Berlin Heidelberg New York, 1994.

[4] J. Bochnak, M. Coste, M.-F. Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete 36, Springer-Verlag, Berlin Heidelberg New York, 1998.

[5] I.R. Shafarevich, *Basic algebraic geometry 2, 3rd ed.*, Springer-Verlag, Berlin Heidelberg New York, 2013.