Linear Algebra via Complex Analysis

Alexander P. Campbell   Daniel Daners

Corrected Version January 24, 2014*

Abstract

The resolvent $(\lambda I - A)^{-1}$ of a matrix A is naturally an analytic function of $\lambda \in \mathbb{C}$, and the eigenvalues are isolated singularities. We compute the Laurent expansion of the resolvent about the eigenvalues of A. Using the Laurent expansion, we prove the Jordan decomposition theorem, prove the Cayley-Hamilton theorem, and determine the minimal polynomial of A. The proofs do not make use of determinants, and many results naturally generalise to operators on Banach spaces.

1 Introduction

The Jordan decomposition theorem for square matrices with coefficients in $\mathbb{C}$ is most commonly proved by means of algebraic methods. Every good theorem has several proofs, which give different insights and generalise into different directions. The aim of this exposition is to present an approach using complex analysis. We derive the Jordan decomposition theorem, the Cayley-Hamilton theorem, and the minimal polynomial from the Laurent expansions about the eigenvalues of the matrix.

The approach is known to experts in operator theory and functional calculus and is outlined in [5, Section I.5]. It shows unexpected connections between topics usually treated separately in undergraduate mathematics. We rely on elementary properties of vector spaces and basic theorems of complex analysis such as the Cauchy integral formula and Laurent expansions. These theorems are valid for vector valued functions ; see [4, Sections 3.10 & 3.11]. They can also be applied entry by entry in a matrix or vector.

Let V be a finite dimensional normed vector space over $\mathbb{C}$ and let $A\colon V \to V$ be a linear operator. In the simplest case, we have $V = \mathbb{C}^n$ with the Euclidean norm, and A is a $n \times n$ matrix with entries in $\mathbb{C}$. We first demonstrate why it is natural to analyse the structure of the *resolvent*

$$\lambda \mapsto (\lambda I - A)^{-1}$$

*The original publication appears in *American Mathematical Monthly* **120** (2013) No. 10, 877–892. doi:10.4169/amer.math.monthly.120.10.877
This version contains some corrections thanks to Robert Burckel.

using complex analysis. In one dimension, $A = a$ is a complex number, $I = 1$, and the resolvent corresponds to $(\lambda - a)^{-1}$. Expanding by a geometric series about $\lambda_0 \neq a$ we get

$$(\lambda - a)^{-1} = \frac{1}{\lambda - a} = \frac{1}{(\lambda - \lambda_0) + (\lambda_0 - a)} = \frac{(\lambda_0 - a)^{-1}}{1 + (\lambda - \lambda_0)(\lambda_0 - a)^{-1}}$$
$$= \sum_{k=0}^{\infty} (-1)^k (\lambda_0 - a)^{-(k+1)} (\lambda - \lambda_0)^k \quad (1.1)$$

if $|(\lambda_0 - a)^{-1}(\lambda - \lambda_0)| < 1$. For a linear operator A, there might be several points for which $(\lambda I - A)^{-1}$ does not exist. We define the *resolvent set* of A by

$$\varrho(A) := \{\lambda \in \mathbb{C} \colon \lambda I - A \text{ is invertible}\} \quad (1.2)$$

and the *spectrum* of A by

$$\sigma(A) := \mathbb{C} \setminus \varrho(A). \quad (1.3)$$

For a matrix A, $\sigma(A)$ is the set of eigenvalues because the rank-nullity theorem implies that $\ker(\lambda I - A) = \{0\}$ if and only if $\lambda I - A$ is invertible. Replacing $\lambda - a$ by $\lambda I - A$ in (1.1) and absolute value by operator norm (see (2.1)) we might expect for $\lambda_0 \in \varrho(A)$ that

$$(\lambda I - A)^{-1} = \sum_{k=0}^{\infty} (-1)^k (\lambda_0 I - A)^{-(k+1)} (\lambda - \lambda_0)^k \quad (1.4)$$

if

$$|\lambda - \lambda_0| < \frac{1}{\|(\lambda_0 I - A)^{-1}\|}. \quad (1.5)$$

That is, if $\lambda_0 \in \varrho(A)$, then $(\lambda I - A)^{-1}$ can be expanded in a power series about λ_0 with radius of convergence at least $1/\|(\lambda_0 I - A)^{-1}\|$. Hence $(\lambda I - A)^{-1}$ is an analytic (holomorphic) function of $\lambda \in \varrho(A)$, with Taylor series (1.4) at $\lambda_0 \in \varrho(A)$, and $\varrho(A)$ is open. In Section 2 we make these arguments rigorous.

A matrix has only finitely many eigenvalues, so they are isolated singularities of the resolvent. Hence, it is natural to use Laurent expansions about the eigenvalues to analyse the structure of the resolvent. If $\lambda_1, \dots, \lambda_q$ are the distinct eigenvalues of A, then the expansion turns out to be

$$(\lambda I - A)^{-1} = \sum_{k=1}^{m_j - 1} \frac{N_j^k}{(\lambda - \lambda_j)^{k+1}} + \frac{P_j}{\lambda - \lambda_j} + \sum_{k=0}^{\infty} (-1)^k B_j^{k+1} (\lambda - \lambda_j)^k, \quad (1.6)$$

where P_j is the projection parallel to $W_j := \ker(P_j)$ onto the generalised eigenspace associated with λ_j, $m_j = \dim(\operatorname{im}(P_j))$ and N_j is nilpotent with $\operatorname{im}(N_j) \subseteq \operatorname{im}(P_j)$. Moreover, $\lambda_j \in \varrho(A|_{W_j})$ and $B_j = (\lambda_j I - A|_{W_j})^{-1}$, which is consistent with (1.4). Note that P_j is the residue of $(\lambda I - A)^{-1}$ at λ_j, so

$$P_j = \frac{1}{2\pi i} \int_{C_j} (\lambda I - A)^{-1} \, d\lambda, \quad (1.7)$$

where C_j is a positively oriented circle about λ_j, not containing any other eigenvalue of A. Moreover,

$$N_j = \frac{1}{2\pi i}\int_{C_j} \frac{(\lambda I - A)^{-1}}{\lambda - \lambda_j}\, d\lambda. \tag{1.8}$$

The Laurent expansion (1.6) is the core of our exposition and is discussed in Section 3.

In Section 4 we prove the Jordan decomposition $A = D + N$, where $D = \lambda_1 P_1 + \cdots + \lambda_q P_q$ is diagonalisable, $N := N_1 + \cdots + N_q$ is nilpotent, and $DN = ND$. By (1.6) the eigenvalues λ_j are poles of $(\lambda I - A)^{-1}$ of order

$$n_j := \min\{k \geq 1\colon N_j^k = 0\} \leq m_j. \tag{1.9}$$

Hence A is diagonalisable if and only if all of its eigenvalues are simple poles. The order of λ_j as a pole of the resolvent is therefore a measure for how far an operator is from being diagonalisable. We show in Section 5 that $p(\lambda) = \prod_{j=1}^{k}(\lambda - \lambda_j)^{n_j}$ is the minimal polynomial of A, and we prove the Cayley-Hamilton theorem.

2 The resolvent as an analytic map

Let V be a finite dimensional normed vector space over $\mathbb{C}$ and let $A\colon V \to V$ be a linear operator. To deal with convergent series such as (1.4), we need a metric or norm on the space of linear operators. We define the *operator norm* by

$$\|A\| := \sup_{\|x\|\leq 1} \|Ax\|. \tag{2.1}$$

This number is finite for every linear operator on finite dimensional spaces. Note that every finite dimensional space has a norm induced by the Euclidean norm on $\mathbb{C}^n$ and some isomorphism from V to $\mathbb{C}^n$. As all norms on finite dimensional vector spaces are equivalent, it does not matter which one we use; see [8, Sections II.1–3].

The expansion (1.4) was motivated by a geometric series. The counterpart of the geometric series in operator theory is the *Neumann series*.

Proposition 2.1 (Neumann Series)**.** *Let $B\colon V \to V$ be a linear operator and let*

$$r := \limsup_{n\to\infty} \|B^n\|^{1/n}.$$

Then $\sum_{k=0}^{\infty} B^k$ converges if $r < 1$ and diverges if $r > 1$. Moreover, $r \leq \|B\|$. If the series converges, then $(I - B)^{-1}$ exists and

$$(I - B)^{-1} = \sum_{k=0}^{\infty} B^k. \tag{2.2}$$

Proof. The root test for the absolute convergence of series implies that $\sum_{k=0}^{n} B^k$ converges if $r < 1$ and diverges if $r > 1$; see [1, Theorem 8.5]. The partial sum $\sum_{k=0}^{n} B^k$ satisfies the identity

$$(I - B)\sum_{k=0}^{n} B^k = \Big(\sum_{k=0}^{n} B^k\Big)(I - B) = \sum_{k=0}^{n} B^k - \sum_{k=0}^{n} B^{k+1} = I - B^{n+1}. \quad (2.3)$$

If $\sum_{k=0}^{\infty} B^k$ converges, then $B^{n+1} \to 0$, and letting $n \to \infty$ in (2.3)

$$(I - B)\sum_{k=0}^{\infty} B^k = \Big(\sum_{k=0}^{\infty} B^k\Big)(I - B) = I.$$

To pass to the limit in (2.3), we use the continuity of multiplication (composition) of linear operators on V. Hence $I - B$ is invertible and (2.2) holds. Since $\|B^n\| \leq \|B\|^n$, we have $r \leq \|B\|$. Hence (2.2) holds if $\|B\| < 1$. □

We can now justify the power series expansion (1.4).

Theorem 2.2 (analyticity of resolvent)**.** *The resolvent set $\varrho(A)$ is open and the map $\lambda \mapsto (\lambda I - A)^{-1}$ is analytic on $\varrho(A)$. If $\lambda_0 \in \varrho(A)$, then the power series expansion* (1.4) *is valid whenever λ satisfies* (1.5)*.*

Proof. We use a calculation similar to (1.1) with a replaced by A. The difficulty is that we need to show that $(\lambda I - A)$ is invertible for λ close to λ_0, so we cannot start with $(\lambda I - A)^{-1}$. In the spirit of (1.1) we write

$$\lambda I - A = (\lambda_0 I - A) + (\lambda - \lambda_0)I = \big(I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}\big)(\lambda_0 I - A) \quad (2.4)$$

and then show that we can invert. The first term in (2.4) is of the form $I - B$ with $B := -(\lambda - \lambda_0)(\lambda_0 I - A)^{-1}$. By Proposition 2.1, $I - B$ is invertible if

$$\|(\lambda - \lambda_0)(\lambda_0 I - A)^{-1}\| = |\lambda - \lambda_0| \|(\lambda_0 I - A)^{-1}\| < 1,$$

which is equivalent to (1.5). Hence if λ satisfies (1.5), then by Proposition 2.1

$$\big(I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}\big)^{-1} = \sum_{k=0}^{\infty} (-1)^k (\lambda_0 I - A)^{-k} (\lambda - \lambda_0)^k.$$

We can therefore invert (2.4) to get (1.4). □

We next prove that $\sigma(A) \neq \emptyset$ by giving an operator theory version of a simple proof of the fundamental theorem of algebra from [7]. The proof relies only on the Cauchy integral formula and a decay estimate for $(\lambda I - A)^{-1}$. Having proved that $\sigma(A) \neq \emptyset$, it makes sense to define the *spectral radius*

$$\mathrm{spr}(A) := \sup\{|\lambda| \colon \lambda \in \sigma(A)\}$$

of A.

Theorem 2.3. *If A is a linear operator on a finite dimensional vector space over $\mathbb{C}$, then $\sigma(A) \neq \emptyset$. Moreover, $\operatorname{spr}(A) = \limsup_{n\to\infty} \|A^n\|^{1/n}$ and for $|\lambda| > \operatorname{spr}(A)$ we have the Laurent series expansion*

$$(\lambda I - A)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}. \tag{2.5}$$

Proof. Let $r := \limsup_{n\to\infty} \|A^n\|^{1/n}$ and note that

$$\limsup_{n\to\infty} \left\| \frac{A^n}{\lambda^n} \right\|^{1/n} = \frac{1}{|\lambda|} \limsup_{n\to\infty} \|A^n\|^{1/n} = \frac{r}{|\lambda|}.$$

By Proposition 2.1, the series $\sum_{k=0}^{\infty} A^k/\lambda^k$ converges if $|\lambda| > r$ and diverges if $|\lambda| < r$. Moreover, (2.5) holds for $|\lambda| > r$ because

$$(\lambda I - A)^{-1} = \frac{1}{\lambda}\Big(I - \frac{1}{\lambda}A\Big)^{-1} = \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{A^k}{\lambda^k} = \sum_{k=0}^{\infty}\frac{A^k}{\lambda^{k+1}}.$$

Hence $\lambda \in \varrho(A)$ if $|\lambda| > r$ and (2.5) is the Laurent expansion of $(\lambda I - A)^{-1}$ about zero in that region. Because the Laurent expansion is valid in the largest annulus about zero in $\varrho(A)$, either $\sigma(A) = \emptyset$ or there exists $\lambda_0 \in \sigma(A)$ with $|\lambda_0| = r$. Hence $r = \operatorname{spr}(A)$ if $\sigma(A) \neq \emptyset$.

It remains to prove that $\sigma(A) \neq \emptyset$. As $r \leq \|A\|$ we get from (2.5) that

$$\|(\lambda I - A)^{-1}\| \leq \frac{1}{|\lambda|}\sum_{k=0}^{\infty}\frac{\|A\|^k}{|\lambda|^k} = \frac{1}{|\lambda|}\frac{1}{1 - \|A\||\lambda|^{-1}} = \frac{1}{|\lambda| - \|A\|} \tag{2.6}$$

for $\lambda \in \mathbb{C}$ with $|\lambda| > \|A\|$. Suppose that $\varrho(A) = \mathbb{C}$. As $\lambda \mapsto (\lambda I - A)^{-1}$ is analytic on $\mathbb{C}$, the Cauchy integral formula yields

$$A^{-1} = -\frac{1}{2\pi i}\int_{|\lambda|=R} \frac{(\lambda I - A)^{-1}}{\lambda}\, d\lambda$$

for all $R > 0$. Using the decay estimate (2.6), we obtain

$$\|A^{-1}\| \leq \frac{2\pi R}{2\pi} \sup_{|\lambda|\geq R} \frac{\|(\lambda I - A)^{-1}\|}{R} \leq \frac{1}{R - \|A\|}$$

for all $R > \|A\|$. Letting $R \to \infty$, we see that $\|A^{-1}\| = 0$ which is impossible. Hence $\sigma(A) \neq \emptyset$ as claimed. □

3 The Laurent expansion about an eigenvalue

We have established that the resolvent is an analytic function on $\varrho(A)$ and know that the eigenvalues are isolated singularities of the resolvent. The centerpiece of our exposition is the Laurent expansion of $(\lambda I - A)^{-1}$ about an eigenvalue $\lambda_0 \in \sigma(A)$.

Theorem 3.1. *Let $\lambda_0 \in \sigma(A)$. Then there exist operators P_0, N_0 and B_0 so that for λ in a neighbourhood of λ_0*

$$(\lambda I - A)^{-1} = \sum_{k=1}^{\infty} \frac{N_0^k}{(\lambda-\lambda_0)^{k+1}} + \frac{P_0}{\lambda - \lambda_0} + \sum_{k=0}^{\infty} (-1)^k B_0^{k+1} (\lambda - \lambda_0)^k. \tag{3.1}$$

Moreover, the operators P_0, N_0 and B_0 have the following properties:

(i) $P_0^2 = P_0$, *that is, P_0 is a projection;*

(ii) $N_0 P_0 = P_0 N_0 = N_0$;

(iii) $B_0 P_0 = P_0 B_0 = 0$;

(iv) $\operatorname{spr}(N_0) = 0$;

(v) $AP_0 = P_0 A = \lambda_0 P_0 + N_0$;

(vi) $(\lambda_0 I - A) B_0 = B_0(\lambda_0 I - A) = I - P_0$.

We defer the proof of the theorem to Section 6 and now discuss some consequences. We show that N_0 is nilpotent and deduce that every eigenvalue of A is a pole of the resolvent.

Remark 3.2. By (ii) $\operatorname{im}(N_0) \subseteq \operatorname{im}(P_0)$, so there exists $n_0 \leq m_0 := \dim(\operatorname{im}(P_0)) \leq \dim V < \infty$ so that

$$\ker(P_0) \subset \ker(N_0) \subset \ker(N_0^2) \subset \cdots \subset \ker(N_0^{n_0}) = \ker(N_0^{n_0+1}) = \dots \tag{3.2}$$

with proper inclusions up to the n_0-th power, and then equality; see [2, Prop. 8.5 & 8.6]. If we show that $N_0^{n_0} = 0$, then N_0 is nilpotent, $N_0^{n_0-1} \neq 0$, and n_0 is the order of λ_0 as a pole of $(\lambda I - A)^{-1}$. Once we know this it follows that $N_0 \colon \operatorname{im}(N_0^{n_0}) \to \operatorname{im}(N_0^{n_0})$ is invertible. To show that $N_0^{n_0} = 0$ note that from (iv) zero is the only eigenvalue of N_0. Hence $\operatorname{im}(N_0^{n_0}) = \{0\}$ as claimed, as otherwise N_0 restricted to $\operatorname{im}(N_0^{n_0})$ has a non-zero eigenvalue by Theorem 2.3. Further note that (3.2) also implies that $\{P_0, N_0, N_0^2, \dots, N_0^{n_0-1}\}$ is linearly independent.

Next we discuss the structure of the regular and singular parts of the Laurent expansion.

Remark 3.3. Since P_0 is a projection we have the direct sum decomposition

$$V = \operatorname{im}(P_0) \oplus \ker(P_0).$$

Choose bases of $\operatorname{im}(P_0)$ and $\ker(P_0)$ to form a basis of V. With respect to that basis, P_0 can be written as a block matrix

$$P_0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Similarly, with respect to the basis introduced, (ii) and (v) of the theorem imply that N_0 and A are block matrices of the form

$$N_0 = \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} \lambda_0 I + N & 0 \\ 0 & A_0 \end{bmatrix}.$$

In particular, $A|_{\ker(P_0)} = A_0$ and (vi) shows that $\lambda_0 \in \varrho(A_0)$ with $B_0|_{\ker(P_0)} = (\lambda_0 I - A_0)^{-1}$. Moreover, by (iii), $B_0|_{\operatorname{im}(P_0)} = 0$, so B_0 is of the form

$$B_0 = \begin{bmatrix} 0 & 0 \\ 0 & (\lambda_0 I - A_0)^{-1} \end{bmatrix}.$$

In particular, the regular part of the Laurent expansion (3.4) is consistent with (1.4), and coincides with the power series expansion of the resolvent $(\lambda I - A_0)^{-1}$ about λ_0. Further note that the singular part of the Laurent expansion is trivial on $\ker(P_0)$ and the regular part is trivial on $\operatorname{im}(P_0)$, so the singular and regular parts live on complementary subspaces.

The above remark proves the following corollary.

Corollary 3.4. *Let λ_0 be an eigenvalue of A and let P_0, N_0 and let B_0 be as in Theorem 3.1. If $m_0 := \dim(\operatorname{im}(P_0)$, then*

$$P_0(\lambda I - A)^{-1} = (\lambda I - A)^{-1} P_0 = \frac{P_0}{\lambda - \lambda_0} + \sum_{k=1}^{m_0} \frac{N_0^k}{(\lambda - \lambda_0)^{k+1}} \tag{3.3}$$

is the singular part of the Laurent expansion (3.1), and is valid for all $\lambda \in \mathbb{C} \setminus \{\lambda_0\}$. Moreover,

$$(I - P_0)(\lambda I - A)^{-1} = (\lambda I - A)^{-1}(I - P_0) = \sum_{k=0}^{\infty} (-1)^k B_0^{k+1} (\lambda - \lambda_0)^k \tag{3.4}$$

is the regular part of the Laurent expansion (3.1) and valid for λ in a neighbourhood of λ_0. Moreover,

(i) *$V = \operatorname{im}(P_0) \oplus \ker(P_0)$ is a direct sum;*

(ii) *λ_0 is the only eigenvalue of $A \colon \operatorname{im}(P_0) \to \operatorname{im}(P_0)$;*

(iii) *$\lambda_0 I - A \colon \ker(P_0) \to \ker(P_0)$ is invertible.*

In Theorem 3.1(v) we already see how the Jordan decomposition arises from the Laurent expansion about λ_0 since $\lambda_0 P_0$ is diagonalisable on $\operatorname{im}(P_0)$ and N_0 is nilpotent. The following proposition is useful to prove uniqueness of the Jordan decomposition.

Proposition 3.5. *Suppose that $A = D + N$, where D and N are such that $DN = ND$ and* $\operatorname{spr}(N) = 0$*. Then $\varrho(A) = \varrho(D)$ and*

$$(\lambda I - A)^{-1} = \sum_{k=0}^{\infty} N^k (\lambda I - D)^{-(k+1)} \tag{3.5}$$

uniformly with respect to λ in compact subsets of $\varrho(A)$. If $\lambda_0 \in \sigma(A)$, then

$$P_0 = \frac{1}{2\pi i} \int_{C_r} (\lambda I - A)^{-1} \, d\lambda = \frac{1}{2\pi i} \int_{C_r} (\lambda I - D)^{-1} \, d\lambda, \tag{3.6}$$

where C_r is a circle centred at λ_0 not containing any other eigenvalues of A.

Proof. If $\lambda \in \varrho(D)$, then

$$\lambda I - A = \lambda I - D - N = (\lambda I - D)\big(I - (\lambda I - D)^{-1} N\big). \tag{3.7}$$

By assumption $(\lambda I - D)N = N(\lambda I - D)$. Applying $(\lambda I - D)^{-1}$ from the left and from the right we get $N(\lambda I - D)^{-1} = (\lambda I - D)^{-1} N$, and so

$$\big(N(\lambda I - D)^{-1}\big)^n = N^n (\lambda I - D)^{-n}$$

for all $n \in \mathbb{N}$. Therefore, if $K \subseteq \varrho(D)$ is compact, then there exists $M > 0$ such that for all $n \in \mathbb{N}$ and $\lambda \in K$

$$\big\|\big(N(\lambda I - D)^{-1}\big)^n\big\|^{1/n} \leq \|N^n\|^{1/n} \|(\lambda I - D)^{-1}\| \leq M \|N^n\|^{1/n}.$$

As $\operatorname{spr}(N) = 0$, for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that, if $n > n_0$ and $\lambda \in K$, then

$$\big\|\big(N(\lambda I - D)^{-1}\big)^n\big\|^{1/n} \leq M \|N^n\|^{1/n} < \varepsilon. \tag{3.8}$$

In particular $\operatorname{spr}\big(N(\lambda I - D)^{-1}\big) = 0$. By Proposition 2.1, we can invert (3.7) to get (3.5) and $\lambda \in \varrho(A)$. The convergence is uniform on K because of (3.8). If $\lambda \in \varrho(A)$, then $D = A - N$ has the same structure with $AN = NA$, so we can interchange the roles of D and A and conclude that $\lambda \in \varrho(D)$. This proves that $\varrho(A) = \varrho(D)$.

If $\lambda_0 \in \sigma(A)$, then by the uniform convergence of (3.5) on the compact set C_r we have that

$$P_0 = \frac{1}{2\pi i} \int_{C_r} (\lambda I - A)^{-1} \, d\lambda = \frac{1}{2\pi i} \sum_{k=0}^{\infty} N^k \int_{C_r} (\lambda I - D)^{-(k+1)} \, d\lambda. \tag{3.9}$$

From the Taylor series expansion (1.4)

$$\frac{d^k}{d\lambda^k} (\lambda I - D)^{-1} = k! (\lambda I - D)^{-(k+1)},$$

so $(\lambda I - D)^{-(k+1)}$ has primitive $(\lambda I - D)^{-k}$ on $\varrho(A)$ for all $k \geq 1$. Hence, all integrals in (3.9) vanish except for the first one, and (3.9) reduces to (3.6). $\square$

4 The Jordan decomposition theorem

In the previous section we looked at the Laurent expansion about a single eigenvalue of A. Here we look at the expansions about all distinct eigenvalues $\lambda_1, \dots, \lambda_q$ of A and use them to derive the Jordan decomposition theorem. For $j = 1, \dots, q$ we look at the projections P_j given by (1.7). We choose C_j to be mutually disjoint positively oriented circles centred at λ_j, not containing any other eigenvalues.

Proposition 4.1. *For $j = 1, \dots, q$ let P_j be the projection defined by (1.7). Then*

$$I = P_1 + \dots + P_q = \frac{1}{2\pi i} \int_{C_R} (\lambda I - A)^{-1} \, d\lambda, \tag{4.1}$$

where C_R is a circle of radius $R > \operatorname{spr}(A)$ centred at zero. Moreover,

$$V = \operatorname{im}(P_1) \oplus \operatorname{im}(P_2) \oplus \dots \oplus \operatorname{im}(P_q), \tag{4.2}$$

and this direct sum completely reduces A. Finally, for $j = 1, \dots, q$,

$$\operatorname{im}(P_j) = \ker(\lambda_j I - A)^{m_j}, \tag{4.3}$$

where $m_j = \dim(\operatorname{im}(P_j))$.

Proof. As $R > \operatorname{spr}(A)$ the circle C_R encloses all eigenvalues. By the residue theorem and the Laurent expansion (2.5), we get

$$\begin{aligned}\sum_{j=1}^{q} P_j = \sum_{j=1}^{q} \frac{1}{2\pi i} \int_{C_j} (\lambda I - A)^{-1} \, d\lambda &= \frac{1}{2\pi i} \int_{C_R} (\lambda I - A)^{-1} \, d\lambda \\ &= \frac{1}{2\pi i} \sum_{k=0}^{\infty} \int_{C_R} \frac{A^k}{\lambda^{k+1}} \, d\lambda = \frac{I}{2\pi i} \int_{C_R} \frac{1}{\lambda} \, d\lambda = I\end{aligned}$$

as all terms in the series are zero except the one with $k = 0$. We next show that P_j is a projection parallel to P_k if $k \neq j$. We have

$$\begin{aligned}(2\pi i)^2 P_j P_k = \int_{C_j} (\lambda I - A)^{-1} \, d\lambda \int_{C_k} (\mu I - A)^{-1} \, d\mu \\ = \int_{C_j} \int_{C_k} (\lambda I - A)^{-1} (\mu I - A)^{-1} \, d\mu \, d\lambda.\end{aligned}$$

Using the resolvent identity from Proposition 6.1(ii) below we get

$$\begin{aligned}(2\pi i)^2 P_j P_k &= \int_{C_j} \int_{C_k} \frac{(\lambda I - A)^{-1} - (\mu I - A)^{-1}}{\mu - \lambda} \, d\mu \, d\lambda \\ &= \int_{C_j} (\lambda I - A)^{-1} \int_{C_k} \frac{1}{\mu - \lambda} \, d\mu \, d\lambda - \int_{C_k} (\mu I - A)^{-1} \int_{C_j} \frac{1}{\mu - \lambda} \, d\lambda \, d\mu = 0,\end{aligned}$$

since the circle C_j is outside C_k and vice versa. This completes the proof of (4.2). The fact that the direct sum reduces A follows from Corollary 3.4.

To prove (4.3) note that Theorem 3.1(v) implies that $(A - \lambda_j I)^{m_j} P_j = N_j^{m_j} = 0$ and so $\operatorname{im}(P_j) \subseteq \ker(\lambda_j I - A)^{m_j}$. By Corollary 3.4 $(A - \lambda_j I)^{m_j}(I - P_j)$ is injective on $\ker(P_j)$, so $\ker(\lambda_j I - A)^{m_j} \subseteq \operatorname{im}(P_j)$, proving (4.3). □

From Corollary 3.4 we know that $A\colon \operatorname{im}(P_j) \to \operatorname{im}(P_j)$ has λ_j as its only eigenvalue. This motivates the following definition.

Definition 4.2. We call $\operatorname{im}(P_j)$ the *generalised eigenspace* associated with the eigenvalue λ_j and $m_j = \dim(\operatorname{im}(P_j))$ the *algebraic multiplicity* of λ_j.

The identity (4.3) ensures that Definition 4.2 agrees with the usual definition of the generalised eigenspace. We now derive a formula for the resolvent in terms of N_j and P_j similar to a partial fraction decomposition of a rational function.

Theorem 4.3. *For every $\lambda \in \varrho(A)$ we have the representation*

$$(\lambda I - A)^{-1} = \sum_{j=1}^{q} \left(\frac{P_j}{\lambda - \lambda_j} + \sum_{k=1}^{m_j - 1} \frac{N_j^k}{(\lambda - \lambda_j)^{k+1}} \right).$$

Proof. Using Corollary 3.4 and Proposition 4.1, for every $\lambda \in \varrho(A)$

$$\begin{aligned}(\lambda I - A)^{-1} = (\lambda I - A)^{-1} \sum_{j=1}^{q} P_j &= \sum_{j=1}^{q} (\lambda I - A)^{-1} P_j \\ &= \sum_{j=1}^{q} \left(\frac{P_j}{\lambda - \lambda_j} + \sum_{k=1}^{m_j - 1} \frac{N_j^k}{(\lambda - \lambda_j)^{k+1}} \right)\end{aligned}$$

as claimed. □

An operator D is called *diagonalisable* if for some direct sum decomposition D acts by scalar multiplication on each subspace. These scalars are the eigenvalues of D. We are now ready to prove the Jordan decomposition theorem.

Theorem 4.4 (Jordan decomposition)**.** *Let V be a finite dimensional vector space over $\mathbb{C}$ and $A\colon V \to V$ a linear operator. Then there exists a diagonalisable operator D and a nilpotent operator N such that $A = D + N$ and $DN = ND$. If $\lambda_1, \dots, \lambda_q$ are the distinct eigenvalues of A, then*

$$D = \sum_{j=1}^{q} \lambda_j P_j \qquad \text{and} \qquad N = \sum_{j=1}^{q} N_j, \tag{4.4}$$

where P_j and N_j are as defined before. In particular D and N are uniquely determined by A. Finally, A is diagonalisable if and only if all eigenvalues of A are simple poles of $(\lambda I - A)^{-1}$.

Proof. By Theorem 3.1(v) $AP_j = \lambda_j P_j + N_j$ for $j = 1, \dots, q$ and therefore by Proposition 4.1

$$A = A\sum_{j=1}^{q} P_j = \sum_{j=1}^{q} AP_j = \sum_{j=1}^{q}(\lambda_j P_j + N_j) = \sum_{j=1}^{q} \lambda_j P_j + \sum_{j=1}^{q} N_j.$$

Hence if we define D and N as in (4.4), then $A = D + N$. It is clear that D is diagonalisable. By Theorem 3.1(ii) and Proposition 4.1 it follows that

$$P_k N_j = P_k P_j N_j = \delta_{kj} N_j = N_j P_k,$$

so the direct sum (4.2) reduces N. In particular, N is nilpotent since each N_j is nilpotent and also $DN = ND$ since this is the case on $\operatorname{im}(P_j)$.

To show the uniqueness of the decomposition, assume that $A = \tilde{D} + \tilde{N}$ with $\tilde{D}$ diagonalisable, $\tilde{N}$ nilpotent and $\tilde{D}\tilde{N} = \tilde{N}\tilde{D}$. Proposition 3.5 implies that $\varrho(\tilde{D}) = \varrho(A)$ and that the spectral projections are equal. Hence $\lambda_1, \dots, \lambda_q$ are the eigenvalues of $\tilde{D}$. As $\tilde{D}$ is diagonalisable, $\tilde{D}P_j = \lambda_j P_j$ and $\tilde{N}P_j = AP_j - \lambda_j P_j = N_j$ for $j = 1, \dots, q$. Hence $\tilde{D} = D$ and $\tilde{N} = N$ as claimed.

The last assertion of the theorem follows since $N = 0$ if and only if $N_j = 0$ in the Laurent expansion (1.6) for all $j = 1, \dots, q$, which means that all eigenvalues are simple poles. □

To obtain the Jordan canonical form for matrices, it is sufficient to construct a basis of $\operatorname{im}(P_j)$ such that the matrix representation of AP_j consists of Jordan blocks; see e.g. [2, Theorem 8.47]. For many purposes the full Jordan canonical form is not needed as examples in [6] show.

5 The Cayley-Hamilton theorem and the minimal polynomial

If $p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$ is a polynomial, we define

$$p(A) = a_n A^n + a_{n-1}A^{n-1} + \dots + a_1 A + a_0 I.$$

The Cayley-Hamilton theorem asserts that $p_A(A) = 0$ if $p_A(\lambda) := \det(\lambda I - A)$ is the characteristic polynomial of A. We start by finding a representation of $p(A)$ reminiscent of the Cauchy integral formula.

Lemma 5.1. *If $p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$ is a polynomial, then*

$$p(A) = \frac{1}{2\pi i}\int_{C_R} p(\lambda)(\lambda I - A)^{-1}\, d\lambda,$$

where C_R is a positively oriented circle centred at zero with radius $R > \operatorname{spr}(A)$.

Proof. By the linearity of integrals, it is sufficient to consider $p(\lambda) = \lambda^k$. Using (4.1), we get from the Cauchy integral theorem that

$$\begin{aligned} 2\pi i A^k &= \int_{C_R} A^k(\lambda I - A)^{-1}\, d\lambda = \int_{C_R} \big(\lambda I - (\lambda I - A)\big)^k (\lambda I - A)^{-1}\, d\lambda \\ &= \sum_{\ell=0}^{k} (-1)^\ell \binom{k}{\ell} \int_{C_R} \lambda^{k-\ell}(\lambda I - A)^{\ell-1}\, d\lambda = \int_{C_R} \lambda^k (\lambda I - A)^{-1}\, d\lambda \end{aligned}$$

as required. $\square$

Theorem 4.3 allows us to derive a formula for $p(A)$.

Theorem 5.2. *If p is a polynomial and $p^{(k)}$ its k-th derivative, then*

$$p(A) = \sum_{j=1}^{q} \left(p(\lambda_j) P_j + \sum_{k=1}^{m_j - 1} \frac{p^{(k)}(\lambda_j)}{k!} N_j^k \right). \tag{5.1}$$

Proof. From Lemma 5.1 and Theorem 4.3

$$\begin{aligned} p(A) &= \frac{1}{2\pi i} \int_{C_R} p(\lambda)(\lambda I - A)^{-1}\, d\lambda = \frac{1}{2\pi i} \sum_{j=1}^{q} \int_{C_j} p(\lambda)(\lambda I - A)^{-1}\, d\lambda \\ &= \frac{1}{2\pi i} \sum_{j=1}^{q} \left(\int_{C_j} \frac{p(\lambda)}{\lambda - \lambda_j}\, d\lambda P_j + \sum_{k=1}^{m_j - 1} \int_{C_j} \frac{p(\lambda)}{(\lambda - \lambda_j)^{k+1}}\, d\lambda N_j^k \right). \end{aligned} \tag{5.2}$$

By the Cauchy integral formula

$$p(\lambda_j) = \frac{1}{2\pi i} \int_{C_j} \frac{p(\lambda)}{\lambda - \lambda_j}\, d\lambda$$

and therefore

$$p^{(k)}(\lambda_j) = \frac{k!}{2\pi i} \int_{C_j} \frac{p(\lambda)}{(\lambda - \lambda_j)^{k+1}}\, d\lambda.$$

Substitution into (5.2) yields (5.1). $\square$

We are now ready to prove the Cayley-Hamilton theorem.

Corollary 5.3 (Cayley-Hamilton)**.** *If p_A is the characteristic polynomial of A, then $p_A(A) = 0$.*

Proof. The characteristic polynomial is given by $p_A(\lambda) = \prod_{j=1}^{q} (\lambda - \lambda_j)^{m_j}$, where m_j is the algebraic multiplicity of the eigenvalue λ_j. Hence $p_A^{(k)}(\lambda_j) = 0$ for $0 \le k \le m_j - 1$ and the representation (5.1) ensures that $p_A(A) = 0$. $\square$

The monic polynomial p of smallest degree such that $p(A) = 0$ is called the *minimal polynomial* of A. According to Theorem 5.2 it is the polynomial p of smallest degree with

$$p(A) = \sum_{j=1}^{q} \left(p(\lambda_j) P_j + \sum_{k=1}^{n_j-1} \frac{p^{(k)}(\lambda_j)}{k!} N_j^k \right) = 0, \tag{5.3}$$

where n_j is the order of λ_j as a pole of the resolvent given by (1.9). By Remark 3.2 and Proposition 4.1, the set of operators

$$\{P_j \colon 1 \leq j \leq q\} \cup \{N_j^k \colon 1 \leq k \leq n_j, 1 \leq j \leq q\}$$

is linearly independent, so (5.3) holds if and only if $p^{(k)}(\lambda_j) = 0$ for all $j = 1, \dots, q$ and $0 \leq k \leq n_j - 1$. This proves the following theorem.

Theorem 5.4 (minimal polynomial)**.** *Let A be a matrix over $\mathbb{C}$ with distinct eigenvalues $\lambda_1, \dots \lambda_q$. Then the minimal polynomial of A is given by*

$$p(\lambda) = \prod_{j=1}^{q} (\lambda - \lambda_j)^{n_j},$$

where n_j is the order of λ_j as a pole of $(\lambda I - A)^{-1}$.

The preceding theorem shows that the minimal polynomial determines the order of the poles of the resolvent and vice versa.

6 Computation of the Laurent expansion

In this section we prove Theorem 3.1 on the Laurent expansion of the resolvent about $\lambda_0 \in \sigma(A)$. The arguments in this section do not use that λ_0 is an eigenvalue, nor that $\dim(V) < \infty$. We first need some elementary properties of the resolvent.

Proposition 6.1. *If $(\lambda I - A)^{-1}$ and $(\mu I - A)^{-1}$ exist, then*

(i) $A(\lambda I - A)^{-1} = (\lambda I - A)^{-1} A = \lambda(\lambda I - A)^{-1} - I$;

(ii) $(\mu I - A)^{-1} - (\lambda I - A)^{-1} = (\lambda - \mu)(\mu I - A)^{-1}(\lambda I - A)^{-1}$;

(iii) $(\lambda I - A)^{-1}(\mu I - A)^{-1} = (\mu I - A)^{-1}(\lambda I - A)^{-1}$.

Proof. For (i) we write

$$A(\lambda I - A)^{-1} = \big(\lambda I - (\lambda I - A)\big)(\lambda I - A)^{-1} = \lambda(\lambda I - A)^{-1} - I$$

and similarly

$$(\lambda I - A)^{-1} A = (\lambda I - A)^{-1}\big(\lambda I - (\lambda I - A)\big) = \lambda(\lambda I - A)^{-1} - I.$$

For (ii) we note that

$$(\mu I - A)\big((\mu I - A)^{-1} - (\lambda I - A)^{-1}\big)(\lambda I - A) \\ = (\lambda I - A) - (\mu I - A) = (\lambda - \mu)I.$$

Applying $(\mu I - A)^{-1}$ from the left and $(\lambda I - A)^{-1}$ from the right, we get (ii). Finally, (iii) follows from (ii) by interchanging the roles of μ and λ. □

Property (ii) is often referred to as the *resolvent identity*. It corresponds to the partial fraction decomposition

$$\frac{\lambda - \mu}{(\mu - a)(\lambda - a)} = \frac{1}{\mu - a} - \frac{1}{\lambda - a}.$$

The Laurent series about λ_0 representing $(\lambda I - A)^{-1}$ is

$$(\lambda I - A)^{-1} = \sum_{n=-\infty}^{\infty} B_n(\lambda - \lambda_0)^n, \tag{6.1}$$

where

$$B_n = \frac{1}{2\pi i}\int_{C_r} \frac{(\lambda I - A)^{-1}}{(\lambda - \lambda_0)^{n+1}}\, d\lambda \tag{6.2}$$

and C_r is a positively oriented circle of radius r centred at λ_0, not enclosing any other eigenvalue of A; see [3, Theorem 8.3.1] or [4, Section 3.11]. We next prove some recursion relations between the B_n. The aim is to be able to express all B_n in terms of B_{-2}, B_{-1} and B_0.

Lemma 6.2. *The coefficients B_n in (6.1) satisfy the relation*

$$B_m B_n = B_n B_m = \begin{cases} -B_{n+m+1} & \text{if } n, m \geq 0, \\ B_{n+m+1} & \text{if } n, m < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6.3}$$

Moreover,

$$AB_n = B_n A = \begin{cases} B_{n-1} + \lambda_0 B_n & \text{if } n \neq 0, \\ B_{n-1} + \lambda_0 B_n - I & \text{if } n = 0. \end{cases} \tag{6.4}$$

Proof. By replacing A by $\lambda_0 I - A$, we may assume that $\lambda_0 = 0$. Let C_r and C_s be circles of radius $0 < r < s$, both centred at zero such that C_s does not enclose any other eigenvalue of A as shown in Figure 6.1. Then

$$(2\pi i)^2 B_n B_m = \left(\int_{C_r} \frac{(\lambda I - A)^{-1}}{\lambda^{n+1}}\, d\lambda\right)\left(\int_{C_s} \frac{(\mu I - A)^{-1}}{\mu^{m+1}}\, d\mu\right) \\ = \int_{C_r}\int_{C_s} \frac{(\lambda I - A)^{-1}(\mu I - A)^{-1}}{\lambda^{n+1}\mu^{m+1}}\, d\mu\, d\lambda.$$

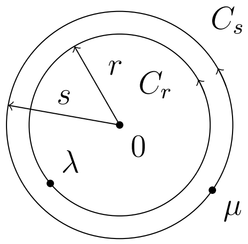

Figure 6.1: The circles C_r and C_s about 0

In the spirit of a partial fraction decomposition, we use the resolvent identity from Proposition 6.1(ii) to get

$$\begin{aligned}(2\pi i)^2 B_n B_m &= \int_{C_r}\int_{C_s} \frac{(\mu I - A)^{-1} - (\lambda I - A)^{-1}}{(\lambda-\mu)\lambda^{n+1}\mu^{m+1}}\, d\mu\, d\lambda \\ &= \int_{C_s} \frac{(\mu I - A)^{-1}}{\mu^{m+1}} \int_{C_r} \frac{1}{\lambda^{n+1}(\lambda-\mu)}\, d\lambda\, d\mu \\ &\qquad - \int_{C_r} \frac{(\lambda I - A)^{-1}}{\lambda^{n+1}} \int_{C_s} \frac{1}{\mu^{m+1}(\lambda-\mu)}\, d\mu\, d\lambda.\end{aligned} \tag{6.5}$$

If $n, m \geq 0$, then we use the partial fraction decompositions

$$\frac{1}{\lambda^{n+1}(\lambda-\mu)} = \frac{\lambda^{n+1} - (\lambda^{n+1} - \mu^{n+1})}{\lambda^{n+1}\mu^{n+1}(\lambda-\mu)} = \frac{1}{\mu^{n+1}(\mu-\lambda)} - \sum_{k=0}^{n} \frac{1}{\lambda^{n-k+1}} \frac{1}{\mu^{k+1}} \tag{6.6}$$

and

$$\frac{1}{\mu^{m+1}(\lambda-\mu)} = \frac{1}{\lambda^{m+1}(\lambda-\mu)} + \sum_{k=0}^{m} \frac{1}{\mu^{m-k+1}} \frac{1}{\lambda^{k+1}} \tag{6.7}$$

to evaluate the inner integrals. Note that $\mu \in C_s$ is outside the circle C_r. Using (6.6) if $n \geq 0$ and the Cauchy integral theorem if $n < 0$ we get

$$\frac{1}{2\pi i}\int_{C_r} \frac{1}{\lambda^{n+1}(\lambda-\mu)}\, d\lambda = \begin{cases} -\dfrac{1}{\mu^{n+1}} & \text{if } n \geq 0, \\ 0 & \text{if } n < 0. \end{cases}$$

For the other integral note that if $m \geq 0$, then both $\mu = 0$ and $\mu = \lambda$ are singularities enclosed by C_s. Hence, using (6.7) and the residue theorem, we obtain

$$\frac{1}{2\pi i}\int_{C_s} \frac{1}{\mu^{m+1}(\mu-\lambda)}\, d\mu = \frac{1}{\lambda^{m+1}} - \frac{1}{\lambda^{m+1}} = 0.$$

If $m < 0$, then only $\mu = \lambda$ is a singularity, and by the Cauchy integral formula

$$\frac{1}{2\pi i}\int_{C_s} \frac{1}{\mu^{m+1}(\lambda-\mu)}\, d\mu = \frac{1}{\lambda^{m+1}}.$$

Hence if $m, n \geq 0$, then the second of the inner integrals on the right hand side of (6.5) is zero, and the other is $2\pi i \mu^{-(n+1)}$. Therefore

$$B_n B_m = \frac{1}{2\pi i} \int_{C_s} \frac{(\mu I - A)^{-1}}{\mu^{m+n+2}} \, d\mu = B_{n+m+1}.$$

If $m, n < 0$, then the first of the inner integrals on the right hand side of (6.5) is zero, and the other is $2\pi i \lambda^{-(m+1)}$ and therefore

$$B_n B_m = -\frac{1}{2\pi i} \int_{C_r} \frac{(\lambda I - A)^{-1}}{\lambda^{m+n+2}} \, d\lambda = -B_{n+m+1}.$$

If $n \geq 0$ and $m < 0$, then a similar argument shows that $B_n B_m = B_{n+m+1} - B_{m+n+1} = 0$. In the remaining case both inner integrals in (6.5) are zero, so $B_n B_m = 0$. From (6.5) it is also clear that $B_n B_m = B_m B_n$. To prove (6.4) we use Proposition 6.1(i) to conclude that

$$\begin{aligned} AB_n &= \frac{1}{2\pi i} \int_{C_r} \frac{A(\lambda I - A)^{-1}}{\lambda^{n+1}} \, d\lambda = \frac{1}{2\pi i} \int_{C_r} \frac{\lambda(\lambda I - A)^{-1} - I}{\lambda^{n+1}} \, d\lambda \\ &= \frac{1}{2\pi i} \int_{C_r} \frac{(\lambda I - A)^{-1}}{\lambda^{n}} \, d\lambda - \frac{I}{2\pi i} \int_{C_r} \frac{1}{\lambda^{n+1}} \, d\lambda = B_{n-1} - \frac{I}{2\pi i} \int_{C_r} \frac{1}{\lambda^{n+1}} \, d\lambda. \end{aligned}$$

This completes the proof of the lemma since the last integral is zero if $n \neq 0$, and is $2\pi i$ if $n = 0$. $\square$

Proof of Theorem 3.1. To prove Theorem 3.1 we set

$$P_0 := B_{-1} \quad \text{and} \quad N_0 := B_{-2}.$$

First we deduce from (6.3) that

$$P_0^2 = B_{-1} B_{-1} = B_{-1-1+1} = B_{-1} = P_0$$

which proves (i). Similarly, applying (6.3) again we get (ii) since

$$P_0 N_0 = B_{-1} B_{-2} = B_{-1-2+1} = B_{-2} = N_0$$

and B_{-1} and B_{-2} commute. Similarly we get (iii) since by (6.3)

$$P_0 B_0 = B_{-1} B_0 = 0 = B_0 B_{-1} = B_0 P_0.$$

We next use induction to show that $B_n = (-1)^n B_0^{n+1}$. This is obvious for $n = 0$, so assume that $n \geq 1$. Then by (6.3) and the induction assumption

$$B_{n+1} = B_{n+0+1} = -B_n B_0 = -(-1)^n B_0^{n+1} B_0 = (-1)^{n+1} B_0^{(n+1)+1}$$

as claimed. In a similar manner, we prove that $B_{-n} = B_{-2}^{n-1} = N_0^{n-1}$ for $n \geq 2$. Again this is obvious for $n = 2$. By (6.3) and the induction assumption

$$B_{-(n+1)} = B_{-n-2+1} = B_{-n} B_{-2} = B_{-2}^{n-1} B_{-2} = B_{-2}^{(n+1)-1} = N_0^{(n+1)-1}$$

as claimed. Hence (3.1) follows. To prove (iv) note that from (6.2)

$$\|N_0^n\| = \|B_{-(n+1)}\| \leq r^{n+1} \sup_{|\lambda|=r} \|(\lambda I - A)^{-1}\| \leq K r^{n+1}$$

for some constant $K > 0$ depending on r. The constant K is finite since the circle $|\lambda| = r$ is compact and the resolvent is continuous. Hence

$$\operatorname{spr}(N_0) \leq r \lim_{n\to\infty} (Kr)^{1/n} = r. \tag{6.8}$$

As we can choose r as small as we like, we conclude that $\operatorname{spr}(N_0) = 0$. Finally, note that (v) and (vi) are special cases of (6.4) for $n = -1$ and $n = 0$, respectively. □

References

[1] H. Amann and J. Escher, *Analysis I*, Birkhäuser Verlag, Basel, 2005. doi:10.1007/b137107

[2] S. Axler, *Linear Algebra Done Right*, second edition, Springer-Verlag, New York, 1997. doi:10.1007/b97662

[3] E. Hille, *Analytic Function Theory*, Vol. 1, Ginn and Company, Boston, 1959.

[4] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, revised edition. American Mathematical Society, Providence, R. I., 1957,

[5] T. Kato, *Perturbation Theory for Linear Operators*, second edition. Springer-Verlag, Berlin, 1976. doi:10.1007/978-3-642-66282-9

[6] R. R. London and H. P. Rogosinski, Decomposition theory in the teaching of elementary linear algebra, *Amer. Math. Monthly* **97** (1990) 478–485. doi:10.2307/2323830

[7] A. R. Schep, A simple complex analysis and an advanced calculus proof of the fundamental theorem of algebra, *Amer. Math. Monthly* **116** (2009) 67–68. http://www.jstor.org/stable/27642665

[8] A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis*, second edition. John Wiley & Sons, New York-Chichester-Brisbane, 1980.

Alexander Campbell
School of Mathematics & Statistics, The University of Sydney, NSW 2006, Australia
Current address: *Department of Mathematics, Macquarie University, NSW 2109, Australia*

Daniel Daners
School of Mathematics & Statistics, The University of Sydney, NSW 2006, Australia
daniel.daners@sydney.edu.au