Chapter 2
Square Matrices of Order 2

Abstract The main topic of this chapter is a detailed study of 2×2 matrices and their applications, for instance to linear recursive sequences and Pell's equations. The key ingredient is the Cayley–Hamilton theorem, which is systematically used in analyzing the properties of these matrices. Many of these properties will be proved in subsequent chapters by more advanced methods.

Keywords Cayley–Hamilton • Trace • Determinant • Pell's equation • Binomial equation

In this chapter we will study some specific problems involving matrices of order two and to make things even more concrete, we will work exclusively with matrices whose entries are real or complex numbers. The reason for doing this is that in this case one can actually perform explicit computations which might help the reader become more familiar with the material introduced in the previous chapter. Also, many of the results discussed in this chapter in a very special context will later on be generalized (very often with completely different methods and tools!). We should however warn the reader from the very beginning: studying square matrices of order 2 is very far from being trivial, even though it might be tempting to believe the contrary.

A matrix $A \in M_2(\mathbf{C})$ is **scalar** if it is of the form zI_2 for some complex number z. One can define the notion of scalar matrix in full generality: if F is a field and $n \geq 1$, the scalar matrices are precisely the matrices of the form cI_n, where $c \in F$ is a scalar.

2.1 The Trace and the Determinant Maps

We introduce now two fundamental invariants of a 2×2 matrix, which will be generalized and extensively studied in subsequent chapters for $n \times n$ matrices:

T. Andreescu, *Essential Linear Algebra with Applications: A Problem-Solving Approach*, DOI 10.1007/978-0-8176-4636-3_2

Definition 2.1. Consider a matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_2(\mathbf{C})$. We define

- the **trace** of A as

$$\mathrm{Tr}(A) = a_{11} + a_{22}.$$

- the **determinant** of A as

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

We also write

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

for the determinant of A.

We obtain in this way two maps

$$\mathrm{Tr}, \det : M_2(\mathbf{C}) \to \mathbf{C}$$

which essentially govern the theory of 2×2 matrices. The following proposition summarizes the main properties of the trace map. The second property is absolutely fundamental. Recall that ${}^tA$ is the transpose of the matrix A.

Proposition 2.2. *For all matrices $A, B \in M_2(\mathbf{C})$ and all complex numbers $z \in \mathbf{C}$ we have*

(a) $\mathrm{Tr}(A + zB) = \mathrm{Tr}(A) + z\mathrm{Tr}(B)$.
(b) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.
(c) $\mathrm{Tr}({}^tA) = \mathrm{Tr}(A)$.

Proof. Properties (a) and (c) are readily checked, so let us focus on property (b). Write

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

and

$$BA = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22}. \end{bmatrix}.$$

Thus

$$\mathrm{Tr}(AB) = a_{11}b_{11} + a_{22}b_{22} + a_{12}b_{21} + a_{21}b_{12} = \mathrm{Tr}(BA).$$

□

Remark 2.3. The map $\mathrm{Tr} : M_2(\mathbf{C}) \to \mathbf{C}$ is not multiplicative, i.e., generally $\mathrm{Tr}(AB) \neq \mathrm{Tr}(A)\mathrm{Tr}(B)$. For instance $\mathrm{Tr}(I_2 \cdot I_2) = \mathrm{Tr}(I_2) = 2$ and $\mathrm{Tr}(I_2) \cdot \mathrm{Tr}(I_2) = 2 \cdot 2 = 4 \neq 2$.

Let us turn now to properties of the determinant map:

Proposition 2.4. *For all matrices $A, B \in M_2(\mathbf{C})$ and all complex numbers α we have*

(1) $\det(AB) = \det A \cdot \det B$*;*
(2) $\det {}^tA = \det A$*;*
(3) $\det(\alpha A) = \alpha^2 \det A$.

Proof. Properties (2) and (3) follow readily from the definition of a determinant. Property (1) will be checked by a painful direct computation. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} x & y \\ z & t \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} ax + bz & ay + bt \\ cx + dz & cy + dt \end{bmatrix}$$

and so

$$\det(AB) = (ax + bz)(cy + dt) - (cx + dz)(ay + bt) =$$

$$acxy + adxt + bcyz + bdzt - acxy - bcxt - adyz - bdzt =$$

$$xt(ad - bc) - yz(ad - bc) = (ad - bc)(xt - yz) = \det A \cdot \det B,$$

as desired. □

Problem 2.5. Let $A \in M_2(\mathbf{R})$ such that

$$\det(A + 2I_2) = \det(A - I_2).$$

Prove that

$$\det(A + I_2) = \det(A).$$

Solution. Write $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The condition becomes

$$\begin{vmatrix} a+2 & b \\ c & d+2 \end{vmatrix} = \begin{vmatrix} a-1 & b \\ c & d-1 \end{vmatrix}$$

or equivalently

$$(a + 2)(d + 2) - bc = (a - 1)(d - 1) - bc.$$

Expanding and canceling similar terms, we obtain the equivalent relation $a+d=-1$. Using similar arguments, the equality $\det(A + I_2) = \det A$ is equivalent to $(a + 1)(d + 1) - bc = ad - bc$, or $a + d = -1$.The result follows. □

2.1.1 Problems for Practice

1. Compute the trace and the determinant of the following matrices

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} -3 & 6 \\ 2 & -4 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

2. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Compute the determinant of A^7.
3. The trace of $A \in M_2(\mathbf{C})$ equals 0. Prove that the trace of A^3 also equals 0.
4. Prove that for all matrices $A \in M_2(\mathbf{C})$ we have

$$\det A = \frac{(\mathrm{Tr}(A))^2 - \mathrm{Tr}(A^2)}{2}.$$

5. Prove that for all matrices $A, B \in M_2(\mathbf{C})$ we have

$$\det(A + B) = \det A + \det B + \mathrm{Tr}(A) \cdot \mathrm{Tr}(B) - \mathrm{Tr}(AB).$$

6. Let $f : M_2(\mathbf{C}) \to \mathbf{C}$ be a map with the property that for all matrices $A, B \in M_2(\mathbf{C})$ and all complex numbers z we have

$$f(A + zB) = f(A) + zf(B) \quad \text{and} \quad f(AB) = f(BA).$$

(a) Consider the matrices

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and define $x_{ij} = f(E_{ij})$. Check that $E_{12}E_{21} = E_{11}$ and $E_{21}E_{12} = E_{22}$, and deduce that $x_{11} = x_{22}$.

(b) Check that $E_{11}E_{12} = E_{12}$ and $E_{12}E_{11} = O_2$, and deduce that $x_{12} = 0$. Using a similar argument, prove that $x_{21} = 0$.

(c) Conclude that there is a complex number c such that

$$f(A) = c \cdot \mathrm{Tr}(A)$$

for all matrices A.

2.2 The Characteristic Polynomial and the Cayley–Hamilton Theorem

Let $A \in M_2(\mathbf{C})$. The **characteristic polynomial** of A is by definition the polynomial denoted $\det(XI_2 - A)$ and defined by

$$\det(XI_2 - A) = X^2 - \mathrm{Tr}(A)X + \det A.$$

We note straight away that AB and BA have the same characteristic polynomial for all matrices $A, B \in M_2(\mathbf{C})$, since AB and BA have the same trace and the same determinant, by results established in the previous section. In particular, if P is invertible, then A and PAP^{-1} have the same characteristic polynomial.

The notation $\det(XI_2 - A)$ is rather suggestive, and it is indeed coherent, in the sense that for any complex number z, if we evaluate the characteristic polynomial of A at z, we obtain precisely the determinant of the matrix $zI_2 - A$. More generally, we have the following very useful:

Problem 2.6. For any two matrices $A, B \in M_2(\mathbf{C})$ there is a complex number u such that

$$\det(A + zB) = \det A + uz + \det B \cdot z^2$$

for all complex numbers z. If A, B have integer/rational/real entries, then u is integer/rational/real.

Solution. Write $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$. Then

$$\det(A + zB) = \begin{vmatrix} a + z\alpha & b + z\beta \\ c + z\gamma & d + z\delta \end{vmatrix} =$$

$$(a+z\alpha)(d+z\delta)-(b+z\beta)(c+z\gamma) = z^2(\alpha\delta-\beta\gamma)+z(a\delta+d\alpha-\beta c-\gamma b)+ad-bc.$$

Since $\alpha\delta - \beta\gamma = \det B$ and $ad - bc = \det A$, the result follows. □

In other words, for any two matrices $A, B \in M_2(\mathbf{C})$ we can define a quadratic polynomial $\det(A + XB)$ which evaluated at any complex number z gives $\det(A + zB)$. Moreover, $\det(A + XB)$ has constant term $\det A$ and leading term B, and if A, B have rational/integer/real entries, then this polynomial has rational/integer/real coefficients. Before moving on, let us practice some problems to better digest these ideas.

Problem 2.7. Let $U, V \in M_2(\mathbf{R})$. Using the polynomial $\det(U + XV)$, prove that

$$\det(U + V) + \det(U - V) = 2 \det U + 2 \det V.$$

Solution. Write

$$f(X) = \det(XV + U) = \det V \cdot X^2 + mX + \det U,$$

for some $m \in \mathbf{R}$. Then

$$\det(U + V) + \det(U - V) = f(1) + f(-1) =$$

$$(\det V + m + \det U) + (\det V - m + \det U) = 2(\det U + \det V).$$

□

Problem 2.8. Let $A, B \in M_2(\mathbf{R})$. Using the previous problem, prove that

$$\det(A^2 + B^2) + \det(AB + BA) \geq 0.$$

Solution. As suggested, we use the identity

$$\det(U + V) + \det(U - V) = 2 \det U + 2 \det V.$$

from Problem 2.7, and take $U = A^2 + B^2$, $V = AB + BA$. Thus

$$\begin{aligned} &\det(A^2 + B^2 + AB + BA) + \det(A^2 + B^2 - AB - BA) \\ &= 2 \det(A^2 + B^2) + 2 \det(AB + BA). \end{aligned}$$

As $A^2 + B^2 + AB + BA = (A + B)^2$ and $A^2 + B^2 - AB - BA = (A - B)^2$, we obtain

$$2\det(A^2 + B^2) + 2\det(AB + BA) = \det(A + B)^2 + \det(A - B)^2 \geq 0.$$

□

Problem 2.9. Let $A, B \in M_2(\mathbf{R})$. Using the polynomial

$$f(X) = \det(I_2 + AB + x(BA - AB)),$$

prove that

$$\det\left(I_2 + \frac{2AB + 3BA}{5}\right) = \det\left(I_2 + \frac{3AB + 2BA}{5}\right).$$

Solution. As suggested, consider the polynomial of degree at most 2

$$f(X) = \det(I_2 + AB + x(BA - AB)).$$

We need to prove that $f\left(\frac{2}{5}\right) = f\left(\frac{3}{5}\right)$. We claim that $f(X) = f(1 - X)$, which clearly implies the desired result. The polynomial $g(X) = f(X) - f(1 - X)$ has degree at most 1 and satisfies $g(0) = g(1) = 0$. Indeed, we have

$$g(0) = f(0) - f(1) = \det(I_2 + AB) - \det(I_2 + BA) = 0,$$

since AB and BA have the same characteristic polynomial. Also, $g(1) = f(1) - f(0) = 0$. Thus g must be the zero polynomial and the result follows. □

We introduce now another crucial tool in the theory of matrices, which will be vastly generalized in subsequent chapters to $n \times n$ matrices (using completely different ideas and techniques).

Definition 2.10. The *eigenvalues* of a matrix $A \in M_2(\mathbf{C})$ are the roots of its characteristic polynomial, in other words they are the complex solutions λ_1, λ_2 of the equation

$$\det(tI_2 - A) = t^2 - \operatorname{Tr}(A)t + \det A = 0.$$

Note that

$$\lambda_1 + \lambda_2 = \operatorname{Tr}(A) \quad \text{and} \quad \lambda_1\lambda_2 = \det A,$$

i.e., **the trace is the sum of the eigenvalues and the determinant is the product of the eigenvalues**. Indeed, by definition of λ_1 and λ_2 the characteristic polynomial is $(X-\lambda_1)(X-\lambda_2)$, and identifying the coefficients of X and $X^0 = 1$ yields the desired relations.

The following result is **absolutely fundamental for the study of square matrices of order** 2.

Theorem 2.11 (Cayley–Hamilton). *For any $A \in M_2(\mathbf{C})$ we have*

$$A^2 - \mathrm{Tr}(A)\cdot A + (\det A)\cdot I_2 = O_2.$$

Proof. Write $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then a direct computation shows that

$$A^2 = \begin{bmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{bmatrix}.$$

Letting $x = \mathrm{Tr}(A)$, we obtain

$$A^2 - \mathrm{Tr}(A)\cdot A + (\det A)\cdot I_2 = \begin{bmatrix} a^2+bc & bx \\ cx & d^2+bc \end{bmatrix} - \begin{bmatrix} ax & bx \\ cx & dx \end{bmatrix}$$

$$+ \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} a^2+ad-ax & 0 \\ 0 & d^2+ad-dx \end{bmatrix} = 0,$$

since $a^2+ad-ax = a(a+d-x) = 0$ and similarly $d^2+ad-dx = d(a+d-x) = 0$. $\square$

Remark 2.12. (a) In other words, the matrix A is a solution of the characteristic equation

$$\det(tI_2 - A) = t^2 - \mathrm{Tr}(A)t + \det A = 0.$$

(b) Expressed in terms of the eigenvalues λ_1 and λ_2 of A, the Cayley–Hamilton theorem can be written either

$$A^2 - (\lambda_1+\lambda_2)A + \lambda_1\lambda_2 \cdot I_2 = O_2 \tag{2.1}$$

or equivalently

$$(A - \lambda_1 \cdot I_2)(A - \lambda_2 \cdot I_2) = O_2. \tag{2.2}$$

Both relations are extremely useful when dealing with square matrices of order 2, and we will see many applications in subsequent sections.

Problem 2.13. Let $A \in M_2(\mathbf{C})$ have eigenvalues λ_1 and λ_2. Prove that for all $n \geq 1$ we have

$$\mathrm{Tr}(A^n) = \lambda_1^n + \lambda_2^n.$$

Deduce that λ_1^n and λ_2^n are the eigenvalues of A^n.

Solution. Let $x_n = \mathrm{Tr}(A^n)$. Multiplying relation (2.1) by A^n and taking the trace yields

$$x_{n+2} - (\lambda_1 + \lambda_2)x_{n+1} + \lambda_1\lambda_2 x_n = 0.$$

Since $x_0 = 2$ and $x_1 = \mathrm{Tr}(A) = \lambda_1 + \lambda_2$, an immediate induction shows that $x_n = \lambda_1^n + \lambda_2^n$ for all n.

For the second part, let z_1, z_2 be the eigenvalues of A^n. By definition, they are the solutions of the equation $t^2 - \mathrm{Tr}(A^n)t + \det(A^n) = 0$. Since $\det(A^n) = (\det A)^n = \lambda_1^n\lambda_2^n$ and $\mathrm{Tr}(A^n) = \lambda_1^n + \lambda_2^n$, the previous equation is equivalent to

$$t^2 - (\lambda_1^n + \lambda_2^n)t + \lambda_1^n\lambda_2^n = 0 \quad \text{or} \quad (t - \lambda_1^n)(t - \lambda_2^n) = 0.$$

The result follows. □

Problem 2.14. Let $A \in M_2(\mathbf{C})$ be a matrix with $\mathrm{Tr}(A) \neq 0$. Prove that a matrix $B \in M_2(\mathbf{C})$ commutes with A if and only if B commutes with A^2.

Solution. Clearly, if $BA = AB$, then $BA^2 = A^2B$, so assume conversely that $BA^2 = A^2B$. Using the Cayley–Hamilton theorem, we can write this relation as

$$B(\mathrm{Tr}(A)A - \det A \cdot I_2) = (\mathrm{Tr}(A)A - \det A \cdot I_2)B$$

or

$$\mathrm{Tr}(A)(BA - AB) = O_2.$$

Since $\mathrm{Tr}(A) \neq 0$, we obtain $BA = AB$, as desired. □

Problem 2.15. Prove that for any matrices $A, B \in M_2(\mathbf{R})$ there is a real number α such that $(AB - BA)^2 = \alpha I_2$.

Solution. Let $X = AB - BA$. Since $\mathrm{Tr}(X) = \mathrm{Tr}(AB) - \mathrm{Tr}(BA) = 0$, the Cayley–Hamilton theorem yields $X^2 = -\det X I_2$ and so we can take $\alpha = -\det X$. □

Problem 2.16. Let $X \in M_2(\mathbf{R})$ be a matrix such that $\det(X^2 + I_2) = 0$. Prove that $X^2 + I_2 = O_2$.

Solution. We have $\det(X + iI_2) = 0$ or $\det(X - iI_2) = 0$, and since $\det(X - iI_2) = \overline{\det(X + iI_2)}$, we deduce that $\det(X + iI_2) = 0 = \det(X - iI_2)$. If $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the relation $\det(X + iI_2) = 0$ is equivalent to $(a + i)(d + i) - bc = 0$, i.e., $ad - bc = 1$ and $a + d = 0$. Thus $\det X = 1$ and $Tr(X) = 0$ and we conclude using the Cayley–Hamilton theorem. $\square$

An important consequence of the Cayley–Hamilton theorem is the following result (which can of course be proved directly by hand).

Theorem 2.17. *A matrix $A \in M_2(\mathbf{C})$ is invertible if and only if* $\det A \neq 0$. *If this is the case, then*

$$A^{-1} = \frac{1}{\det A}(\mathrm{Tr}(A) \cdot I_2 - A).$$

Proof. Suppose that A is invertible. Then taking the determinant in the equality $A \cdot A^{-1} = I_2$ we obtain

$$\det A \cdot \det A^{-1} = \det I_2 = 1,$$

thus $\det A \neq 0$.

Conversely, suppose that $\det A \neq 0$ and define

$$B = \frac{1}{\det A}(\mathrm{Tr}(A) \cdot I_2 - A).$$

Then using the Cayley–Hamilton theorem we obtain

$$AB = \frac{1}{\det A}(\mathrm{Tr}(A) \cdot A - A^2) = \frac{1}{\det A} \cdot \det A I_2 = I_2$$

and similarly $BA = I_2$. Thus A is invertible and $A^{-1} = B$. $\square$

Remark 2.18. One can also check directly that if $\det A \neq 0$, then A is invertible, its inverse being given by

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Problem 2.19. Let $A, B \in M_2(\mathbf{C})$ be two matrices such that $AB = I_2$. Then A is invertible and $B = A^{-1}$. In particular, we have $BA = I_2$.

Solution. Since $AB = I_2$, we have $\det A \cdot \det B = \det(AB) = 1$, thus $\det A \neq 0$. The previous theorem shows that A is invertible. Multiplying the equality $AB = I_2$ by A^{-1} on the left, we obtain $B = A^{-1}$. Finally, $BA = A^{-1}A = I_2$. $\square$

A very important consequence of the previous theorem is the following characterization of eigenvalues:

Theorem 2.20. *If $A \in M_2(\mathbf{C})$ and $z \in \mathbf{C}$, then the following assertions are equivalent:*

(a) *z is an eigenvalue of A;*
(b) $\det(zI_2 - A) = 0$.
(c) *There is a nonzero vector $v \in \mathbf{C}^2$ such that $Av = zv$.*

Proof. By definition of eigenvalues, (a) implies (b). Suppose that (b) holds and let $B = A - zI_2$. The assumption implies that $\det B = 0$ and so $b_{11}b_{22} = b_{12}b_{21}$. We need to prove that we can find $x_1, x_2 \in \mathbf{C}$ not both zero and such that

$$b_{11}x_1 + b_{12}x_2 = 0 \quad \text{and} \quad b_{21}x_1 + b_{22}x_2 = 0.$$

If $b_{11} \neq 0$ or $b_{12} \neq 0$, choose $x_2 = b_{11}$ and $x_1 = -b_{12}$, so suppose that $b_{11} = 0 = b_{12}$. If one of b_{21}, b_{22} is nonzero, choose $x_1 = -b_{22}$ and $x_2 = b_{21}$, otherwise choose $x_1 = x_2 = 1$. Thus (b) implies (c).

Suppose now that (c) holds. Then $A^2v = zAv = z^2v$ and using relation (2.1) we obtain

$$(z^2 - \mathrm{Tr}(A)z + \det A)v = 0.$$

Since $v \neq 0$, this forces $z^2 - \mathrm{Tr}(A)z + \det A = 0$ and so z is an eigenvalue of A. Thus (c) implies (a) and the theorem is proved. $\square$

Problem 2.21. Let $A \in M_2(\mathbf{C})$ have two distinct eigenvalues λ_1, λ_2. Prove that we can find an invertible matrix $P \in \mathrm{GL}_2(\mathbf{C})$ such that

$$A = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}.$$

Solution. By the previous theorem, we can find two nonzero vectors $X_1 = \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix}$ and $X_2 = \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}$ such that $AX_i = \lambda_i X_i$.

Consider the matrix $P = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ whose columns are X_1, X_2. A simple computation shows that the columns of AP are $\lambda_1 X_1$ and $\lambda_2 X_2$, which are the columns of $P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, thus $AP = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. It remains to see that if $\lambda_1 \neq \lambda_2$, then P is invertible (we haven't used so far the hypothesis $\lambda_1 \neq \lambda_2$).

Suppose that $\det P = 0$, thus $x_{11}x_{22} = x_{21}x_{12}$. This easily implies that the columns of P are proportional, say the second column X_2 is α times the first column, X_1. Thus $X_2 = \alpha X_1$. Then

$$\lambda_2 X_2 = AX_2 = \alpha AX_1 = \alpha\lambda_1 X_1 = \lambda_1 X_2,$$

forcing $(\lambda_1 - \lambda_2)X_2 = 0$. This is impossible as both $\lambda_1 - \lambda_2$ and X_2 are nonzero. The problem is solved. $\square$

Problem 2.22. Solve in $M_2(\mathbf{C})$ the following equations

(a) $A^2 = O_2$.
(b) $A^2 = I_2$.
(c) $A^2 = A$.

Solution. (a) Let A be a solution of the problem. Then $\det A = 0$ and the Cayley–Hamilton relation reduces to $\mathrm{Tr}(A)A = 0$. Taking the trace yields $\mathrm{Tr}(A)^2 = 0$, thus $\mathrm{Tr}(A) = 0$. Conversely, if $\det A = 0$ and $\mathrm{Tr}(A) = 0$, then the Cayley–Hamilton theorem shows that $A^2 = O_2$. Thus the solutions of the problem are the matrices

$$A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}, \quad \text{with} \quad a, b, c \in \mathbf{C} \quad \text{and} \quad a^2 + bc = 0.$$

(b) We must have $\det A = \pm 1$ and, by the Cayley–Hamilton theorem, $I_2 - \mathrm{Tr}(A)A + \det A I_2 = O_2$. If $\det A = 1$, then $\mathrm{Tr}(A)A = 2I_2$ and taking the trace yields $\mathrm{Tr}(A)^2 = 4$, thus $\mathrm{Tr}(A) = \pm 2$. This yields two solutions, $A = \pm I_2$. Suppose that $\det A = -1$. Then $\mathrm{Tr}(A)A = O_2$ and taking the trace gives $\mathrm{Tr}(A) = 0$. Conversely, any matrix A with $\mathrm{Tr}(A) = 0$ and $\det A = -1$ is a solution of the problem (again by Cayley–Hamilton). Thus the solutions of the equation are

$$\pm I_2 \quad \text{and} \quad A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}, \quad a, b, c \in \mathbf{C}, \quad a^2 + bc = 1.$$

(c) If $\det A \neq 0$, then multiplying by A^{-1} yields $A = I_2$. So suppose that $\det A = 0$. The Cayley–Hamilton theorem yields $A - \mathrm{Tr}(A)A = O_2$. If $\mathrm{Tr}(A) \neq 1$, this forces $A = O_2$, which is a solution of the problem. Thus if $A \neq O_2, I_2$, then $\det A = 0$ and $\mathrm{Tr}(A) = 1$. Conversely, all such matrices are solutions of the problem (again by Cayley–Hamilton). Thus the solutions of the problem are

$$O_2, \quad I_2 \quad \text{and} \quad A = \begin{bmatrix} a & b \\ c & 1-a \end{bmatrix}, \quad a, b, c \in \mathbf{C}, \quad a^2 + bc = a.$$

□

Problem 2.23. Let $A \in M_2(\mathbf{C})$ be a matrix. Prove that the following statements are equivalent:

(a) $\mathrm{Tr}(A) = \det A = 0$.
(b) $A^2 = O_2$.
(c) $\mathrm{Tr}(A) = \mathrm{Tr}(A^2) = 0$.
(d) There exists $n \geq 2$ such that $A^n = O_2$.

Solution. Taking the trace of the Cayley–Hamilton theorem, we see that $\mathrm{Tr}(A^2) = \mathrm{Tr}(A)^2 - 2\det A$. From this it is clear that (a) and (c) are equivalent.

The implication (a) implies (b) is just an application of the Cayley–Hamilton theorem. The implication (b) implies (d) is obvious. Thus we need only show (d) implies (a). If $A^n = O_2$ for some $n \geq 2$, then clearly $\det A = 0$. Thus the Cayley–Hamilton theorem reads $A^2 = \mathrm{Tr}(A)A$. Iterating this an immediate induction gives $A^n = \mathrm{Tr}(A)^{n-1}A$, hence $O_2 = \mathrm{Tr}(A)^{n-1}A$. Taking the trace of this identity gives $0 = \mathrm{Tr}(A)^n$ and hence $\mathrm{Tr}(A) = 0$. $\square$

Problem 2.24. Find all matrices $X \in M_2(\mathbf{R})$ such that $X^3 = I_2$.

Solution. We must have $(\det X)^3 = 1$ and so $\det X = 1$ (since $\det X \in \mathbf{R}$). Letting $t = \mathrm{Tr}(X)$, the Cayley–Hamilton theorem and the given equation yield

$$I_2 = X^3 = X(tX - I_2) = t(tX - I_2) - X = (t^2 - 1)X - tI_2.$$

If $t^2 \neq 1$, then the previous relation shows that X is scalar and since $X^3 = I_2$, we must have $X = I_2$. If $t^2 = 1$, then the previous relation gives $t = -1$. Conversely, any matrix $X \in M_2(\mathbf{R})$ with $\mathrm{Tr}(X) = -1$ and $\det X = 1$ satisfies $X^2 + X + I_2 = O_2$ and so also $X^3 = I_2$. We conclude that the solutions of the problem are

$$I_2 \quad \text{and} \quad \begin{bmatrix} a & b \\ c & -1-a \end{bmatrix}, \quad a, b, c \in \mathbf{R}, \quad a^2 + a + bc = -1.$$

$\square$

2.2.1 *Problems for Practice*

1. Let $A, B \in M_2(\mathbf{R})$ be commuting matrices. Prove that

$$\det(A^2 + B^2) \geq 0.$$

   Hint: check that $A^2 + B^2 = (A + iB)(A - iB)$.
2. Let $A, B \in M_2(\mathbf{R})$ be such that $AB = BA$ and $\det(A^2 + B^2) = 0$. Prove that $\det A = \det B$. Hint: use the hint of the previous problem and consider the polynomial $\det(A + XB)$.
3. Let $A, B, C \in M_2(\mathbf{R})$ be pairwise commuting matrices and let

$$f(X) = \det(A^2 + B^2 + C^2 + x(AB + BA + CA)).$$

   (a) Prove that $f(2) \geq 0$. Hint: check that

$$A^2 + B^2 + C^2 + 2(AB + BA + CA) = (A + B + C)^2.$$

(b) Prove that $f(-1) \geq 0$. Hint: denote $X = A - B$ and $Y = B - C$ and check that

$$A^2 + B^2 + C^2 - (AB + BC + CA) = \left(X + \frac{1}{2}Y\right)^2 + \left(\frac{\sqrt{3}}{2}Y\right)^2.$$

Next use the first problem.

(c) Deduce that

$$\det(A^2 + B^2 + C^2) + 2\det(AB + BC + CA) \geq 0.$$

4. Let $A, B \in M_2(\mathbf{C})$ be matrices with $\mathrm{Tr}(AB) = 0$. Prove that $(AB)^2 = (BA)^2$. Hint: use the Cayley–Hamilton theorem.
5. Let A be a 2×2 matrix with rational entries with the property that

$$\det(A^2 - 2I_2) = 0.$$

Prove that $A^2 = 2I_2$ and $\det A = -2$. Hint: use the fact that $A^2 - 2I_2 = (A - \sqrt{2}I_2)(A + \sqrt{2}I_2)$ and consider the characteristic polynomial of A.
6. Let x be a positive real number and let $A \in M_2(\mathbf{R})$ be a matrix such that $\det(A^2 + xI_2) = 0$. Prove that

$$\det(A^2 + A + xI_2) = x.$$

7. Let $A, B \in M_2(\mathbf{R})$ be such that $\det(AB - BA) \leq 0$. Consider the polynomial

$$f(X) = \det(I_2 + (1 - X)AB + XBA).$$

(a) Prove that $f(0) = f(1)$.
(b) Deduce that

$$\det(I_2 + AB) \leq \det\left(I_2 + \frac{1}{2}(AB + BA)\right).$$

8. Let $n \geq 3$ be an integer. Let $X \in M_2(\mathbf{R})$ be such that

$$X^n + X^{n-2} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

(a) Prove that $\det X = 0$. Hint: show that $\det(X^2 + I_2) = 0$.

(b) Let t be the trace of X. Prove that

$$t^n + t^{n-2} = 2.$$

(c) Find all possible matrices X satisfying the original equation.

9. Let $n \geq 2$ be a positive integer and let $A, B \in M_2(\mathbf{C})$ be two matrices such that $AB \neq BA$ and $(AB)^n = (BA)^n$. Prove that $(AB)^n = \alpha I_2$ for some complex number α.
10. Let $A, B \in M_2(\mathbf{R})$ and let $n \geq 1$ be an integer such that $C^n = I_2$, where $C = AB - BA$. Prove that n is even and $C^4 = I_2$. Hint: use Problem 2.15.

2.3 The Powers of a Square Matrix of Order 2

In this section we will use the Cayley–Hamilton theorem to compute the powers of a given matrix $A \in M_2(\mathbf{C})$. Let λ_1 and λ_2 be the eigenvalues of A. The discussion and the final result will be very different according to whether λ_1 and λ_2 are different or not.

Let us start with the case $\lambda_1 = \lambda_2$ and consider the matrix $B = A - \lambda_1 I_2$. Then the Cayley–Hamilton theorem in the form of relation (2.2) yields $B^2 = O_2$, thus $B^k = O_2$ for $k \geq 2$. Using the binomial formula we obtain

$$A^n = (B + \lambda_1 I_2)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda_1^{n-k} B^k = \lambda_1^n I_2 + n\lambda_1^{n-1} B.$$

Let us assume now that $\lambda_1 \neq \lambda_2$ and consider the matrices

$$B = A - \lambda_1 I_2 \quad \text{and} \quad C = A - \lambda_2 I_2.$$

Relation (2.2) becomes $BC = O_2$, or equivalently $B(A - \lambda_2 I_2) = O_2$. Thus $BA = \lambda_2 B$, which yields $BA^2 = \lambda_2 BA = \lambda_2^2 B$ and by an immediate induction $BA^n = \lambda_2^n B$ for all n. Similarly, the relation $BC = O_2$ yields $CA^n = \lambda_1^n C$ for all n. Taking advantage of the relation $C - B = (\lambda_1 - \lambda_2) I_2$, we obtain

$$(\lambda_1 - \lambda_2) A^n = (C - B) A^n = CA^n - BA^n = \lambda_1^n C - \lambda_2^n B.$$

Thus

$$A^n = \frac{1}{\lambda_1 - \lambda_2} (\lambda_1^n C - \lambda_2^n B).$$

All in all, we proved the following useful result, in which we change notations:

Theorem 2.25. *Let $A \in M_2(\mathbf{C})$ and let λ_1, λ_2 be its eigenvalues.*

(a) If $\lambda_1 \neq \lambda_2$, then for all $n \geq 1$ we have $A^n = \lambda_1^n B + \lambda_2^n C$, where

$$B = \frac{1}{\lambda_1 - \lambda_2}(A - \lambda_2 I_2) \text{ and } C = \frac{1}{\lambda_2 - \lambda_1}(A - \lambda_1 I_2).$$

(b) If $\lambda_1 = \lambda_2$, then for all $n \geq 1$ we have $A^n = \lambda_1^n B + n\lambda_1^{n-1} C$, where $B = I_2$ and $C = A - \lambda_1 I_2$.

Problem 2.26. Compute A^n, where $A = \begin{bmatrix} 1 & 3 \\ -3 & -5 \end{bmatrix}$.

Solution. As $\mathrm{Tr}(A) = -4$ and $\det A = 4$, the eigenvalues of A are solutions of the equation $t^2 + 4t + 4 = 0$, thus $\lambda_1 = \lambda_2 = -2$ are the eigenvalues of A. Using the previous theorem, we conclude that for any $n \geq 1$ we have

$$A^n = (-2)^n I_2 + n(-2)^{n-1}(A + 2I_2) = (-2)^{n-1} \begin{bmatrix} 3n - 2 & 3n \\ -3n & -3n - 2 \end{bmatrix}.$$

□

Though the exact statement of the previous theorem is a little cumbersome, the basic idea is very simple. If one learns this idea, then one can compute A^n easily. Keep in mind that when computing powers of a 2×2 matrix, one starts by computing the eigenvalues of the matrix (this comes down to solving the quadratic equation $t^2 - \mathrm{Tr}(A)t + \det A = 0$). If the eigenvalues are equal, say both equal to λ, then $B := A - \lambda I_2$ satisfies $B^2 = O_2$ and so one computes A^n by writing $A = B + \lambda I_2$ and using the binomial formula. On the other hand, if the eigenvalues are different, say λ_1 and λ_2, then there are two matrices B, C such that for all n we have

$$A^n = \lambda_1^n B + \lambda_2^n C.$$

One can easily find these matrices without having to learn the formulae by heart: if the previous relation holds for all $n \geq 0$, then it certainly holds for $n = 0$ and $n = 1$. Thus

$$I_2 = B + C, \quad A = \lambda_1 B + \lambda_2 C.$$

This immediately yields the matrices B and C in terms of I_2, A and λ_1, λ_2. Moreover, we see that they are of the form $xI_2 + yA$ for some complex numbers x, y. Combining this observation with Theorem 2.25 yields the following useful

Corollary 2.27. *For any matrix $A \in M_2(\mathbf{C})$ there are sequences $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$ of complex numbers such that*

$$A^n = x_n A + y_n I_2$$

for all $n \geq 0$.

One has to be careful that the sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ in the previous corollary are definitely not always characterized by the equality $A^n = x_n A + y_n I_2$ (this is however the case if A is not scalar). On the other hand, Theorem 2.25 shows that we can take

$$x_n = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} \text{ and } y_n = \frac{\lambda_1\lambda_2^n - \lambda_2\lambda_1^n}{\lambda_1 - \lambda_2}$$

when $\lambda_1 \neq \lambda_2$, and, when $\lambda_1 = \lambda_2$

$$x_n = n\lambda_1^{n-1} \text{ and } y_n = -(n-1)\lambda_1^n.$$

Problem 2.28. Let m, n be positive integers and let $A, B \in M_2(\mathbf{C})$ be two matrices such that $A^m B^n = B^n A^m$. If A^m and B^n are not scalar, prove that $AB = BA$.

Solution. From Corollary 2.27 we have

$$A^k = x_k A + y_k I_2 \text{ and } B^k = u_k B + v_k I_2, \quad k \geq 0,$$

where $(x_k)_{k\geq 0}$, $(y_k)_{k\geq 0}$, $(u_k)_{k\geq 0}$, $(v_k)_{k\geq 0}$ are sequences of complex numbers. Since A^m and B^n are not scalar matrices it follows that $x_m \neq 0$ and $u_n \neq 0$. The relation $A^m B^n = B^n A^m$ is equivalent to

$$(x_m A + y_m I_2)(u_n B + v_n I_2) = (u_n B + v_n I_2)(x_m A + y_m I_2)$$

i.e.

$$x_m u_n (AB - BA) = O_2.$$

Hence $AB = BA$. □

Problem 2.29. Let $t \in \mathbf{R}$ and let

$$A_t = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

Compute A_t^n for $n \geq 1$.

Solution. We offer three ways to solve this problem. The first is to follow the usual procedure: compute the eigenvalues of A_t and then use the general Theorem 2.25. Here the eigenvalues are e^{it} and e^{-it} and it is not difficult to deduce that

$$A_t^n = \begin{bmatrix} \cos nt & -\sin nt \\ \sin nt & \cos nt \end{bmatrix}.$$

Another argument is as follows: an explicit computation shows that $A_{t_1+t_2} = A_{t_1}A_{t_2}$, thus

$$A_t^n = A_t \cdot A_t \cdot \ldots \cdot A_t = A_{t+t+\ldots+t} = A_{nt}.$$

Finally, one can also argue geometrically: A_t is the matrix of a rotation of angle t, thus A_t^n is the matrix of a rotation of angle nt. □

2.3.1 *Problems for Practice*

1. Consider the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}.$$

   (a) Let n be a positive integer. Prove the existence of a unique pair of integers (x_n, y_n) such that

$$A^n = x_n A + y_n I_2.$$

   (b) Compute $\lim_{n\to\infty} \frac{x_n}{y_n}$.

2. Given a positive integer n, compute the nth power of the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

3. Let a, b be real numbers and let n be a positive integer. Compute the nth power of the matrix $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$.

4. Let x be a real number and let

$$A = \begin{bmatrix} \cos x + \sin x & 2\sin x \\ -\sin x & \cos x - \sin x \end{bmatrix}.$$

   Compute A^n for all positive integers n.

2.4 Application to Linear Recurrences

In this section we present two classical applications of the theory developed in the previous section. Let a, b, c, d, x_0, y_0 be complex numbers and consider two sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ recursively defined by

$$\begin{cases} x_{n+1} = ax_n + by_n \\ y_{n+1} = cx_n + dy_n, \quad n \geq 0 \end{cases} \tag{2.3}$$

We would like to find the general terms of the two sequences in terms of the initial data a, b, c, d, x_0, y_0 and n.

The key observation is that the system can be written in matrix form as follows

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \text{ i.e. } \begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = A \begin{bmatrix} x_n \\ y_n \end{bmatrix}, \quad n \geq 0,$$

where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the matrix of coefficients. An immediate induction yields

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = A^n \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \quad n \geq 0, \tag{2.4}$$

and so the problem is reduced to the computation of A^n, which is solved by Theorem 2.25.

Let us consider a slightly different problem, also very classical. It concerns second order linear recurrences with constant coefficients. More precisely, we fix complex numbers a, b, x_0, x_1 and look for the general term of the sequence $(x_n)_{n \geq 0}$ defined recursively by

$$x_{n+1} = ax_n + bx_{n-1}, \quad n \geq 1, \tag{2.5}$$

We can easily reduce this problem to the previous one by denoting $y_n = x_{n-1}$ for $n \geq 1$ and $y_0 = \frac{1}{b}(x_1 - ax_0)$ if $b \neq 0$ (which we will assume from now on, since otherwise the problem is very simple from the very beginning). Indeed, relation (2.5) is equivalent to the following system

$$\begin{cases} x_{n+1} = ax_n + by_n \\ y_{n+1} = x_n \end{cases}, \quad n \geq 0.$$

As we have already seen, finding x_n and y_n (or equivalently x_n) comes down to computing the powers of the matrix of coefficients

$$A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}.$$

The characteristic equation of this matrix is $\lambda^2 - a\lambda - b = 0$. If λ_1 and λ_2 are the roots of this equation, then Theorem 2.25 yields the following:

• If $\lambda_1 \neq \lambda_2$, then we can find constants u, v such that

$$x_n = u\lambda_1^n + v\lambda_2^n$$

for all n. These two constants are easily determined by imposing

$$x_0 = u + v \quad \text{and} \quad x_1 = u\lambda_1 + v\lambda_2$$

and solving this linear system in the unknowns u, v.

• If $\lambda_1 = \lambda_2$, then we can find constants u, v such that for all $n \geq 0$

$$x_n = (un + v)\lambda_1^n,$$

and u and v are found from the initial conditions by solving $x_0 = v$ and $x_1 = (u + v)\lambda_1$.

Problem 2.30. Find the general terms of $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$ if

$$\begin{cases} x_{n+1} = x_n + 2y_n \\ y_{n+1} = -2x_n + 5y_n, \quad n \geq 0, \end{cases}$$

and $x_0 = 1$, $y_0 = 2$.

Solution. The matrix of coefficients is $A = \begin{bmatrix} 1 & 2 \\ -2 & 5 \end{bmatrix}$, with characteristic equation $\lambda^2 - 6\lambda + 9 = 0$ and solutions $\lambda_1 = \lambda_2 = 3$. Theorem 2.25 yields (after a simple computation)

$$A^n = 3^n I_2 + n3^{n-1}\begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} (3-2n)3^{n-1} & 2n3^{n-1} \\ -2n3^{n-1} & (3+2n)3^{n-1} \end{bmatrix}$$

Combined with $x_0 = 1$ and $y_0 = 2$, we obtain

$$x_n = (2n + 3)3^{n-1} \text{ and } y_n = 2(n + 3)3^{n-1}, \quad n \geq 0.$$

□

Problem 2.31. Find the limits of sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$, where

$$\begin{cases} x_{n+1} = (1-\alpha)x_n + \alpha y_n \\ y_{n+1} = \beta x_n + (1-\beta)y_n, \end{cases}$$

and α, β are complex numbers with $|1 - \alpha - \beta| < 1$.

Solution. The matrix of coefficients is

$$A = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

and one easily checks that its eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 1-\alpha-\beta$. Note that $|\lambda_2| < 1$, in particular $\lambda_2 \neq 1$. Letting

$$B = \frac{1}{\lambda_1 - \lambda_2}(A - \lambda_2 I_2) = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix},$$

Theorem 2.25 gives the existence of an explicit matrix C such that

$$A^n = \lambda_1^n B + \lambda_2^n C = B + \lambda_2^n C.$$

Since $|\lambda_2| < 1$, we have $\lim_{n\to\infty} \lambda_2^n = 0$ and the previous relation shows that $\lim_{n\to\infty} A^n = B$.

Since $\begin{bmatrix} x_n \\ y_n \end{bmatrix} = A^n \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, we conclude that x_n and y_n are convergent sequences, and if l_1, l_2 are their limits, then

$$\begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = B \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

Taking into account the explicit form of B, we obtain

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \frac{\beta x_0 + \alpha y_0}{\alpha+\beta}.$$

$\square$

2.4.1 Problems for Practice

1. Find the general term of the sequence $(x_n)_{n\geq 0}$ defined by $x_1 = 1$, $x_2 = 0$ and for all $n \geq 1$

$$x_{n+2} = 4x_{n+1} - x_n.$$

2. Consider the sequence $(x_n)_{n\geq 0}$ defined by $x_0 = 1$, $x_1 = 2$ and for all $n \geq 0$

$$x_{n+2} = x_{n+1} - x_n.$$

   Is this sequence periodical? If so, find its minimal period.

3. Find the general terms of the sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ satisfying $x_0 = y_0 = 1$, $x_1 = 1$, $y_1 = 2$ and

$$x_{n+1} = \frac{2x_n + 3y_n}{5}, \quad y_{n+1} = \frac{2y_n + 3x_n}{5}.$$

4. A sequence $(x_n)_{n\geq 0}$ satisfies $x_0 = 2$, $x_1 = 3$ and for all $n \geq 1$

$$x_{n+1} = \sqrt{x_{n-1}x_n}.$$

Find the general term of this sequence (hint: take the logarithm of the recurrence relation).

5. Consider a map $f : (0, \infty) \to (0, \infty)$ such that

$$f(f(x)) = 6x - f(x)$$

for all $x > 0$. Let $x > 0$ and define a sequence $(z_n)_{n\geq 0}$ by $z_0 = x$ and $z_{n+1} = f(z_n)$ for $n \geq 0$.

(a) Prove that

$$z_{n+2} + z_{n+1} - 6z_n = 0$$

for $n \geq 0$.

(b) Deduce the existence of real numbers a, b such that

$$z_n = a \cdot 2^n + b \cdot (-3)^n$$

for all $n \geq 0$.

(c) Using the fact that $z_n > 0$ for all n, prove that $b = 0$ and conclude that $f(x) = 2x$ for all $x > 0$.

2.5 Solving the Equation $X^n = A$

Consider a matrix $A \in M_2(\mathbf{C})$ and an integer $n > 1$. In this section we will explain how to solve the equation $X^n = A$, with $X \in M_2(\mathbf{C})$.

A first key observation is that for any solution X of the equation we have

$$AX = XA.$$

Indeed, this is simply a consequence of the fact that $X^n \cdot X = X \cdot X^n$. We will need the following very useful:

Proposition 2.32. *Let $A \in M_2(\mathbf{C})$ be a non-scalar matrix. If $X \in M_2(\mathbf{C})$ commutes with A, then $X = \alpha I_2 + \beta A$ for some complex numbers α, β.*

Proof. Write $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $X = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$. The equality $AX = XA$ is equivalent, after an elementary computation, to

$$bz = cy, \quad ay + bt = bx + dy, \quad cx + dz = az + ct,$$

or

$$bz = cy, \quad (a-d)y = b(x-t), \quad c(x-t) = z(a-d).$$

If $a \neq d$, set $\beta = \frac{x-t}{a-d}$. Then $z = c\beta$, $y = b\beta$ and $\beta a - x = \beta d - t$. We deduce that $X = \alpha I_2 + \beta A$, where $\alpha = -\beta a + x = -\beta d + t$.

Suppose that $a = d$. If $x \neq t$, the previous relations yield $b = c = 0$ and so A is scalar, a contradiction. Thus $x = t$ and $bz = cy$. Moreover, one of b, c is nonzero (as A is not scalar), say b (the argument when $c \neq 0$ is identical). Setting $\beta = \frac{y}{b}$ and $\alpha = x - \beta a$ yields $X = \alpha I_2 + \beta A$.

□

Let us come back to our original problem, solving the equation $X^n = A$. Let λ_1 and λ_2 be the eigenvalues of A. We will discuss several cases, each of them having a very different behavior.

Let us start with the case $\lambda_1 \neq \lambda_2$. By Problem 2.21, we can then write $A = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$ for some $P \in \mathrm{GL}_2(\mathbf{C})$. Since $AX = XA$ and A is not scalar, by Proposition 2.32 there are complex numbers a, b such that $X = aI_2 + bA$. Thus

$$X = P \begin{bmatrix} a + b\lambda_1 & 0 \\ 0 & a + b\lambda_2 \end{bmatrix} P^{-1}.$$

The equation $X^n = A$ is then equivalent to

$$\begin{bmatrix} (a + b\lambda_1)^n & 0 \\ 0 & (a + b\lambda_2)^n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

It follows that $a + b\lambda_1 = z_1$ and $a + b\lambda_2 = z_2$, where $z_1^n = \lambda_1$ and $z_2^n = \lambda_2$, and $X = P \begin{bmatrix} z_1 & 0 \\ 0 & z_2 \end{bmatrix} P^{-1}$. Hence

Proposition 2.33. *Let $A \in M_2(\mathbf{C})$ be a matrix with distinct eigenvalues λ_1, λ_2. Let $P \in \mathrm{GL}_2(\mathbf{C})$ be a matrix such that $A = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$. Then the solutions of the* *equation $X^n = A$ are given by $X = P \begin{bmatrix} z_1 & 0 \\ 0 & z_2 \end{bmatrix} P^{-1}$, where z_1 and z_2 are solutions of the equations $t^n = \lambda_1$ and $t^n = \lambda_2$ respectively.*

Let us deal now with the case in which A is not scalar, but has equal eigenvalues, say both eigenvalues equal λ. Then the matrix $B = A - \lambda I_2$ satisfies $B^2 = O_2$ (by the Cayley–Hamilton theorem) and we have $A = B + \lambda I_2$. Now, since $AX = XA$ and A is not scalar, we can write $X = cI_2 + dA$ for some complex numbers c, d (Proposition 2.32). Since $A = B + \lambda I_2$, it follows that we can also write $X = aI_2 + bB$ for some complex numbers a, b. Since $B^2 = O_2$, the binomial formula and the given equation yield

$$A = X^n = (aI_2 + bB)^n = a^n I_2 + na^{n-1}bB.$$

Since $A = B + \lambda I_2$, we obtain

$$B + \lambda I_2 = na^{n-1}bB + a^n I_2.$$

Since B is not scalar (as A itself is not scalar), the previous relation is equivalent to

$$1 = na^{n-1}b \quad \text{and} \quad \lambda = a^n.$$

This already shows that $\lambda \neq 0$ (as the first equation shows that $a \neq 0$), so if $\lambda = 0$ (which corresponds to $A^2 = O_2$) then the equation has no solution. On the other hand, if $\lambda \neq 0$, then the equation $a^n = \lambda$ has n complex solutions, and for each of them we obtain a unique value of b, namely $b = \frac{1}{na^{n-1}}$. We have just proved the following

Proposition 2.34. *Suppose that $A \in M_2(\mathbf{C})$ is not scalar, but both eigenvalues of A are equal to some complex number λ. Then*

(a) If $\lambda = 0$, the equation $X^n = A$ has no solutions for $n > 1$, and the only solution $X = A$ for $n = 1$.
(b) If $\lambda \neq 0$, then the solutions of the equation $X^n = A$ are given by

$$X = aI_2 + \frac{1}{na^{n-1}}(A - \lambda I_2),$$

where a runs over the n solutions of the equation $z^n = \lambda$.

Finally, let us deal with the case when A is scalar, say $A = cI_2$ for some complex number c. If $c = 0$, then $X^n = O_2$ has already been solved, so let us assume that $c \neq 0$. Consider a solution X of the equation $X^n = cI_2$ and let λ_1, λ_2 be the eigenvalues of X. Then $\lambda_1^n = \lambda_2^n = c$. We have two possibilities:

- Either $\lambda_1 \neq \lambda_2$, in which case X has distinct eigenvalues and so (Problem 2.21) we can write $X = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$ for some invertible matrix P. Then $X^n = P \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} P^{-1}$ and this equals cI_2 since $\lambda_1^n = \lambda_2^n = c$. The conclusion is that for each pair (λ_1, λ_2) of **distinct** solutions of the equation $t^n = c$ we obtain a whole family of solutions, namely the matrices $X = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$ for some invertible matrix P.
- Suppose that $\lambda_1 = \lambda_2$ and let $Y = X - \lambda_1 I_2$, then $Y^2 = O_2$ and the equation $X^n = cI_2$ is equivalent to $(Y + \lambda_1 I_2)^n = cI_2$. Using again the binomial formula and the equality $Y^2 = O_2$, we can rewrite this equation as

$$\lambda_1^n I_2 + n\lambda_1^{n-1} Y = cI_2.$$

  Since $\lambda_1^n = c$ and $\lambda_1 \neq 0$ (as $c \neq 0$), we deduce that necessarily $Y = O_2$ and so $X = \lambda_1 I_2$, with λ_1 one of the n complex solutions of the equation $t^n = c$. Thus we obtain n more solutions this way.

We can unify the previous two possibilities and obtain

Proposition 2.35. *If $c \neq 0$ is a complex number, the solutions in $M_2(\mathbf{C})$ of the equation $X^n = cI_2$ are given by*

$$X = P \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} P^{-1} \tag{2.6}$$

where x, y are solutions (not necessary distinct) of the equation $z^n = c$, and $P \in \mathrm{GL}_2(\mathbf{C})$ is arbitrary.

Problem 2.36. Let $t \in (0, \pi)$ be a real number and let $n > 1$ be an integer. Find all matrices $X \in M_2(\mathbf{R})$ such that

$$X^n = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

Solution. With the notations of Problem 2.29, we need to solve the equation $X^n = A_t$. Let X be a solution, then $XA_t = A_t X = X^{n+1}$. Writing $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the relation $XA_t = A_t X$ yields $b \sin t = -c \sin t$ and $-a \sin t = -d \sin t$, thus $a = d$ and $c = -b$. Hence $X = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Next, since $X^n = A_t$, we have

$$(\det X)^n = \det X^n = \det A_t = 1,$$

and since $\det X = a^2 + b^2 \geq 0$, we deduce that $a^2 + b^2 = 1$. Thus we can write $a = \cos x$ and $b = \sin x$ for some real number x. Then $X = A_x$ and the equation $X^n = A_t$ is equivalent (thanks to Problem 2.29) to $A_{nx} = A_t$. This is further equivalent to $nx = t + 2k\pi$ for some integer k. It is enough to restrict to $k \in \{0, 1, \ldots, n-1\}$. We conclude that the solutions of the problem are the matrices

$$X_k = \begin{bmatrix} \cos t_k & -\sin t_k \\ \sin t_k & \cos t_k \end{bmatrix},$$

where $t_k = \dfrac{t + 2k\pi}{n}$, $k = 0, 1, \ldots, n-1$. □

Problem 2.37. Let $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in M_2(\mathbf{R})$. Prove that the following statements are equivalent:

(1) $A^n = I_2$ for some positive integer n;
(2) $a = \cos r\pi$, $b = \sin r\pi$ for some rational number r.

Solution. If $a = \cos(\frac{k}{n}\pi)$ and $b = \sin(\frac{k}{n}\pi)$ for some $n \geq 1$ and $k \in \mathbf{Z}$, then Problem 2.29 yields $A^{2n} = I_2$, thus (2) implies (1).

Assume now that (1) holds. Then $(\det A)^n = \det A^n = 1$ and since $\det A = a^2 + b^2 \geq 0$, we must have $\det A = 1$, that is $a^2 + b^2 = 1$. Thus we can find $t \in \mathbf{R}$ such that $a = \cos t$ and $b = \sin t$. Then $A = A_t$ and by Problem 2.29 we have $I_2 = A^n = A_{nt}$. This forces $\cos(nt) = 1$ and so t is a rational multiple of π. The problem is solved. □

2.5.1 Problems for Practice

1. Let $n > 1$ be an integer. Prove that the equation

$$X^n = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

has no solutions in $M_2(\mathbf{C})$.

2. Solve in $M_2(\mathbf{C})$ the binomial equation

$$X^4 = \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}.$$

3. Let $n > 1$ be an integer. Prove that the equation

$$X^n = \begin{bmatrix} 3 & -1 \\ 0 & 0 \end{bmatrix}$$

has no solutions in $M_2(\mathbf{Q})$.

4. Find all matrices $X \in M_2(\mathbf{R})$ such that

$$X^3 = \begin{bmatrix} 4 & 3 \\ -3 & -2 \end{bmatrix}.$$

5. Find all matrices $A, B \in M_2(\mathbf{C})$ such that

$$AB = O_2 \quad \text{and} \quad A^5 + B^5 = O_2.$$

6. Solve in $M_2(\mathbf{R})$ the equation

$$X^n = \begin{bmatrix} 7 & -5 \\ -15 & 12 \end{bmatrix}.$$

7. Solve in $M_2(\mathbf{R})$ the equation

$$X^n = \begin{bmatrix} -6 & -2 \\ 21 & 7 \end{bmatrix}.$$

2.6 Application to Pell's Equations

Let $D > 1$ be an integer which is not a perfect square. The diophantine equation, called **Pell's equation**

$$x^2 - Dy^2 = 1 \tag{2.7}$$

has an obvious solution $(1, 0)$ in nonnegative integers. A well-known but nontrivial result (which we take for granted) is that this equation also has nontrivial solutions (i.e., different from $(0, 1)$).

In this section we explain how the theory developed so far allows finding all solutions of the Pell equation once we know the smallest nontrivial solution. Let S_D be the set of all solutions in **positive integers** to the Eq. (2.7) and let (x_1, y_1) be the **fundamental solution**, i.e., the solution in S_D for which the first component x_1 is minimal among the first components of the elements of S_D.

If x, y are positive integers, consider the matrix

$$A_{(x,y)} = \begin{bmatrix} x & Dy \\ y & x \end{bmatrix},$$

so that $(x, y) \in S_D$ if and only if $\det A_{(x,y)} = 1$. Elementary computations yield the fundamental relation

$$A_{(x,y)} \cdot A_{(u,v)} = A_{(xu+Dyv, xv+yu)} \tag{2.8}$$

Passing to determinants in (2.8) we obtain the **multiplication principle**:

$$\text{if} \quad (x, y), (u, v) \in S_D, \quad \text{then } (xu + Dyv, xv + yu) \in S_D.$$

It follows from the multiplication principle that if we write

$$A^n_{(x_1,y_1)} = \begin{bmatrix} x_n & Dy_n \\ y_n & x_n \end{bmatrix}, \quad n \geq 1,$$

then $(x_n, y_n) \in S_D$ for all n. The sequences x_n and y_n are described by the recursive system

$$\begin{cases} x_{n+1} = x_1 x_n + Dy_1 y_n \\ y_{n+1} = y_1 x_n + x_1 y_n, \quad n \geq 1 \end{cases} \tag{2.9}$$

consequence of the equality $A^{n+1}_{(x_1,y_1)} = A_{(x_1,y_1)} A^n_{(x_1,y_1)}$. Moreover, Theorem 2.25 gives explicit formulae for x_n and y_n in terms of x_1, y_1, n: the characteristic equation of matrix $A_{(x_1,y_1)}$ is

$$\lambda^2 - 2x_1\lambda + 1 = 0$$

with $\lambda_{1,2} = x_1 \pm \sqrt{x_1^2 - 1} = x_1 \pm y_1\sqrt{D}$, and Theorem 2.25 yields, after an elementary computation of the matrices B, C involved in that theorem

$$\begin{cases} x_n = \dfrac{1}{2}[(x_1 + y_1\sqrt{D})^n + (x_1 - y_1\sqrt{D})^n] \\ \\ y_n = \dfrac{1}{2\sqrt{D}}[(x_1 + y_1\sqrt{D})^n - (x_1 - y_1\sqrt{D})^n], \quad n \geq 1. \end{cases} \tag{2.10}$$

Note that relation (2.10) also makes sense for $n = 0$, in which case it gives the trivial solution $(x_0, y_0) = (1, 0)$.

Theorem 2.38. *All solutions in positive integers of the Pell equation $x^2 - Dy^2 = 1$ are described by the formula (2.10), where (x_1, y_1) is the fundamental solution of the equation.*

Proof. Suppose that there are elements in S_D which are not covered by formula (2.10), and among them choose one (x, y) for which x is minimal. Using the multiplication principle, we observe that the matrix $A_{(x,y)} A^{-1}_{(x_1,y_1)}$ generates a solution in integers (x', y'), where

$$\begin{cases} x' = x_1 x - Dy_1 y \\ y' = -y_1 x + x_1 y \end{cases}$$

We claim that x', y' are positive integers. This is clear for x', as $x > \sqrt{D}y$ and $x_1 > \sqrt{D}y_1$, thus $x_1 x > Dy_1 y$. Also, $x_1 y > y_1 x$ is equivalent to $x_1^2(x^2 - 1) > x^2(x_1^2 - 1)$

or $x > x_1$, which holds because (x_1, y_1) is a fundamental solution and (x, y) is not described by relation (2.10) (while (x_1, y_1) is described by this relation, with $n = 1$). Moreover, since $A_{(x',y')}A_{(x_1,y_1)} = A_{(x,y)}$, we have $x = x'x_1 + Dy'y_1 > x'$ and $y = x'y_1 + y'x_1 > y'$. By minimality, (x', y') must be of the form (2.10), i.e., $A_{(x,y)}A^{-1}_{(x_1,y_1)} = A^k_{(x_1,y_1)}$ for some positive integer k. Therefore $A_{(x,y)}=A^{k+1}_{(x_1,y_1)}$, i.e., (x, y) is of the form (2.10), a contradiction. □

Problem 2.39. Find all solutions in positive integers to Pell's equation

$$x^2 - 2y^2 = 1.$$

Solution. The fundamental solution is $(x_1, y_1) = (3, 2)$ and the associated matrix is

$$A_{(3,2)} = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}$$

The solutions $(x_n, y_n)_{n\geq 1}$ are given by $A^n_{(3,2)}$, i.e.

$$\begin{cases} x_n = \dfrac{1}{2}[(3 + 2\sqrt{2})^n + (3 - 2\sqrt{2})^n] \\[2ex] y_n = \dfrac{1}{2\sqrt{2}}[(3 + 2\sqrt{2})^n - (3 - 2\sqrt{2})^n]. \end{cases}$$

□

We can extend slightly the study of the Pell equation by considering the more general equation

$$ax^2 - by^2 = 1 \tag{2.11}$$

where we assume that ab is not a perfect square (it is not difficult to see that if ab is a square, then the equation has only trivial solutions). Contrary to the Pell equation, this Eq. (2.11) does not always have solutions (the reader can check that the equation $3x^2 - y^2 = 1$ has no solutions in integers by working modulo 3).

Define the **Pell resolvent** of (2.11) by

$$u^2 - abv^2 = 1 \tag{2.12}$$

and let $S_{a,b}$ be the set of solutions in positive integers of Eq. (2.11). Thus $S_{1,ab}$ is the set denoted S_{ab} when considering the Pell equation (it is the set of solutions of the Pell resolvent). If x, y, u, v are positive integers consider the matrices

$$B_{(x,y)} = \begin{bmatrix} x & by \\ y & ax \end{bmatrix}, \quad A_{u,v} = \begin{bmatrix} u & abv \\ v & u \end{bmatrix},$$

the second matrix being the matrix associated with the Pell resolvent equation.

An elementary computation shows that

$$B_{(x,y)}A_{(u,v)} = B_{(xu+byv,axv+yu)},$$

Passing to determinants in the above relation and noting that $(x, y) \in S_{a,b}$ if and only if $\det B_{(x,y)} = 1$, we obtain the multiplication principle:

$$\text{if} \quad (x, y) \in S_{a,b} \quad \text{and} \quad (u, v) \in S_{ab}, \quad \text{then} \quad (xu + byv, axv + yu) \in S_{a,b},$$

i.e., the product $B_{(x,y)}A_{(u,v)}$ generates the solution $(xu + byv, axv + yu)$ of (2.11). Using the previous theorem and the multiplication principle, one easily obtains the following result, whose formal proof is left to the reader.

Theorem 2.40. *Assume that Eq. (2.11) is solvable in positive integers, and let (x_0, y_0) be its minimal solution (i.e., x_0 is minimal). Let (u_1, v_1) be the fundamental solution of the resolvent Pell equation (2.14). Then all solutions (x_n, y_n) in positive integers of Eq. (2.11) are generated by*

$$B_{(x_n,y_n)} = B_{(x_0,y_0)}A^n_{(u_1,v_1)}, \quad n \geq 0 \tag{2.13}$$

It follows easily from (2.13) that

$$\begin{cases} x_n = x_0u_n + by_0v_n \\ y_n = y_0u_n + ax_0v_n, \quad n \geq 0 \end{cases} \tag{2.14}$$

where $(u_n, v_n)_{n\geq 1}$ is the general solution to the Pell resolvent equation.

Problem 2.41. Solve in positive integers the equation

$$6x^2 - 5y^2 = 1.$$

Solution. This equation is solvable and its minimal solution is $(x_0, y_0) = (1, 1)$. The Pell resolvent equation is $u^2 - 30v^2 = 1$, with fundamental solution $(u_1, v_1) = (11, 2)$. Using formulae (2.14) and then (2.10), we deduce that the solutions in positive integers are $(x_n, y_n)_{n\geq 1}$, where

$$\begin{cases} x_n = \dfrac{6+\sqrt{30}}{12}(11+2\sqrt{30})^n + \dfrac{6-\sqrt{30}}{12}(11-2\sqrt{30})^n \\[2ex] y_n = \dfrac{5+\sqrt{30}}{12}(11+2\sqrt{30})^n + \dfrac{5-\sqrt{30}}{12}(11-2\sqrt{30})^n. \end{cases}$$

$\square$

2.6.1 Problems for Practice

1. A triangular number is a number of the form $1 + 2 + \ldots + n$ for some positive integer n. Find all triangular numbers which are perfect squares.
2. Find all positive integers n such that $n + 1$ and $3n + 1$ are simultaneously perfect squares.
3. Find all integers a, b such that $a^2 + b^2 = 1 + 4ab$.
4. The difference of two consecutive cubes equals n^2 for some positive integer n. Prove that $2n - 1$ is a perfect square.
5. Find all triangles whose sidelengths are consecutive integers and whose area is an integer.

Springer

http://www.springer.com/978-0-8176-4360-7

Essential Linear Algebra with Applications
A Problem-Solving Approach
Andreescu, T.
2014, X, 491 p. 2 illus. in color., Hardcover
ISBN: 978-0-8176-4360-7
A product of Birkhäuser Basel