

Algebra

It is now time to split mathematics into branches. First, algebra. A section on algebraic identities hones computational skills. It is followed naturally by inequalities. In general, any inequality can be reduced to the question of finding the minimum of a function. But this is a highly nontrivial matter, and that makes the subject exciting. We discuss the fact that squares are nonnegative, the Cauchy-Schwarz inequality, the triangle inequality, the arithmetic mean-geometric mean inequality, and also Sturm's method for proving inequalities.

Our treatment of algebra continues with polynomials. We focus on quadratic polynomials, the relations between zeros and coefficients, the properties of the derivative of a polynomial, problems about the location of the zeros in the complex plane or on the real axis, and methods for proving irreducibility of polynomials (such as the Eisenstein criterion). From all special polynomials we present the most important, the Chebyshev polynomials.

Linear algebra comes next. The first three sections, about operations with matrices, determinants, and the inverse of a matrix, insist on both the array structure of a matrix and the ring structure of the set of matrices. They are more elementary, as is the section on linear systems. The last three sections, about vector spaces and linear transformations, are more advanced, covering among other things the Cayley-Hamilton Theorem and the Perron-Frobenius Theorem.

The chapter concludes with a brief incursion into abstract algebra: binary operations, groups, and rings, really no further than the definition of a group or a ring.

2.1 Identities and Inequalities

2.1.1 Algebraic Identities

The scope of this section is to train algebraic skills. Our idea is to hide behind each problem an important algebraic identity. We commence with three examples, the first and the last written by the second author of the book, and the second given at a Soviet Union college entrance exam and suggested to us by A. Soifer.

Example. Solve in real numbers the system of equations

$$\begin{aligned}(3x + y)(x + 3y)\sqrt{xy} &= 14, \\ (x + y)(x^2 + 14xy + y^2) &= 36.\end{aligned}$$

Solution. By substituting $\sqrt{x} = u$, $\sqrt{y} = v$, we obtain the equivalent form

$$\begin{aligned}uv(3u^4 + 10u^2v^2 + 3v^4) &= 14, \\ u^6 + 15u^4v^2 + 14u^2v^4 + v^6 &= 36.\end{aligned}$$

Here we should recognize elements of the binomial expansion with exponent equal to 6. Based on this observation we find that

$$36 + 2 \cdot 14 = u^6 + 6u^5v + 15u^4v^2 + 20u^3v^3 + 15u^2v^4 + 6uv^5 + v^6$$

and

$$36 - 2 \cdot 14 = u^6 - 6u^5v + 15u^4v^2 - 20u^3v^3 + 15u^2v^4 - 6uv^5 + v^6.$$

Therefore, $(u + v)^6 = 64$ and $(u - v)^6 = 8$, which implies $u + v = 2$ and $u - v = \pm\sqrt{2}$ (recall that u and v have to be positive). So $u = 1 + \frac{\sqrt{2}}{2}$ and $v = 1 - \frac{\sqrt{2}}{2}$ or $u = 1 - \frac{\sqrt{2}}{2}$ and $v = 1 + \frac{\sqrt{2}}{2}$. The solutions to the system are

$$(x, y) = \left(\frac{3}{2} + \sqrt{2}, \frac{3}{2} - \sqrt{2}\right) \quad \text{and} \quad (x, y) = \left(\frac{3}{2} - \sqrt{2}, \frac{3}{2} + \sqrt{2}\right). \quad \square$$

Example. Given two segments of lengths a and b , construct with a straightedge and a compass a segment of length $\sqrt[4]{a^4 + b^4}$.

Solution. The solution is based on the following version of the Sophie Germain identity:

$$a^4 + b^4 = (a^2 + \sqrt{2}ab + b^2)(a^2 - \sqrt{2}ab + b^2).$$

Write

$$\sqrt[4]{a^4 + b^4} = \sqrt{\sqrt{a^2 + \sqrt{2}ab + b^2} \cdot \sqrt{a^2 - \sqrt{2}ab + b^2}}.$$

Applying the law of cosines, we can construct segments of lengths $\sqrt{a^2 \pm \sqrt{2}ab + b^2}$ using triangles of sides a and b with the angle between them 135° , respectively, 45° .

On the other hand, given two segments of lengths x , respectively y , we can construct a segment of length \sqrt{xy} (their geometric mean) as the altitude AD in a right triangle ABC ($\angle A = 90^\circ$) with $BD = x$ and $CD = y$. These two steps combined give the method for constructing $\sqrt[4]{a^4 + b^4}$. \square

Example. Let x, y, z be distinct real numbers. Prove that

$$\sqrt[3]{x-y} + \sqrt[3]{y-z} + \sqrt[3]{z-x} \neq 0.$$

Solution. The solution is based on the identity

$$a^3 + b^3 + c^3 - 3abc = (a + b + c)(a^2 + b^2 + c^2 - ab - bc - ca).$$

This identity arises from computing the circulant determinant

$$D = \begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$$

in two ways: first by expanding with Sarrus' rule, and second by adding up all columns to the first, factoring $(a + b + c)$, and then expanding the remaining determinant. Note that this identity can also be written as

$$a^3 + b^3 + c^3 - 3abc = \frac{1}{2}(a + b + c)[(a - b)^2 + (b - c)^2 + (c - a)^2].$$

Returning to the problem, let us assume the contrary, and set $\sqrt[3]{x-y} = a$, $\sqrt[3]{y-z} = b$, $\sqrt[3]{z-x} = c$. By assumption, $a + b + c = 0$, and so $a^3 + b^3 + c^3 = 3abc$. But this implies

$$0 = (x - y) + (y - z) + (z - x) = 3\sqrt[3]{x-y}\sqrt[3]{y-z}\sqrt[3]{z-x} \neq 0,$$

since the numbers are distinct. The contradiction we have reached proves that our assumption is false, and so the sum is nonzero. \square

And now the problems.

- 86.** Show that for no positive integer n can both $n + 3$ and $n^2 + 3n + 3$ be perfect cubes.
- 87.** Let A and B be two $n \times n$ matrices that commute and such that for some positive integers p and q , $A^p = \mathcal{I}_n$ and $B^q = \mathcal{O}_n$. Prove that $A + B$ is invertible, and find its inverse.
- 88.** Prove that any polynomial with real coefficients that takes only nonnegative values can be written as the sum of the squares of two polynomials.
- 89.** Prove that for any nonnegative integer n , the number $5^{5^{n+1}} + 5^{5^n} + 1$ is not prime.
- 90.** Show that for an odd integer $n \geq 5$,

$$\binom{n}{0}5^{n-1} - \binom{n}{1}5^{n-2} + \binom{n}{2}5^{n-3} - \cdots + \binom{n}{n-1}$$

is not a prime number.

- 91.** Factor $5^{1985} - 1$ into a product of three integers, each of which is greater than 5^{100} .

92. Prove that the number $\frac{5^{125} - 1}{5^{25} - 1}$ is not prime.
93. Let a and b be coprime integers greater than 1. Prove that for $n \geq 0$ is $a^{2n} + b^{2n}$ divisible by $a + b$.
94. Prove that any integer can be written as the sum of five perfect cubes.
95. Prove that

$$\sum_{k=1}^{31} \frac{1}{(k-1)^{4/5} - k^{4/5} + (k-1)^{4/5}} < \frac{3}{2} + \sum_{k=1}^{31} (k-1)^{1/5}.$$

96. Solve in real numbers the equation

$$\sqrt[3]{x-1} + \sqrt[3]{x} + \sqrt[3]{x+1} = 0.$$

97. Find all triples (x, y, z) of positive integers such that

$$x^3 + y^3 + z^3 - 3xyz = p,$$

where p is a prime number greater than 3.

98. Let a, b, c be distinct positive integers such that $ab + bc + ca \geq 3k^2 - 1$, where k is a positive integer. Prove that

$$a^3 + b^3 + c^3 \geq 3(abc + 3k).$$

99. Show that the expression

$$(x^2 - yz)^3 + (y^2 - zx)^3 + (z^2 - xy)^3 - 3(x^2 - yz)(y^2 - zx)(z^2 - xy)$$

is a perfect square.

100. Find all triples (m, n, p) of positive integers such that $m + n + p = 2002$ and the system of equations

$$\frac{x}{y} + \frac{y}{x} = m, \quad \frac{y}{z} + \frac{z}{y} = n, \quad \frac{z}{x} + \frac{x}{z} = p$$

has at least one solution in nonzero real numbers.

2.1.2 $x^2 \geq 0$

We now turn to inequalities. The simplest inequality in algebra says that the square of any real number is nonnegative, and it is equal to zero if and only if the number is zero. We illustrate how this inequality can be used with an example by the second author of the book.

Example. Find the minimum of the function $f : (0, \infty)^3 \rightarrow \mathbb{R}$,

$$f(x, y, z) = x^z + y^z - (xy)^{z/4}.$$

Solution. Rewrite the function as

$$f(x, y, z) = (x^{z/2} - y^{z/2})^2 + 2 \left[(xy)^{z/4} - \frac{1}{4} \right]^2 - \frac{1}{8}.$$

We now see that the minimum is $-\frac{1}{8}$, achieved if and only if

$$(x, y, z) = \left(a, a, \log_a \frac{1}{16} \right),$$

where $a \in (0, 1) \cup (1, \infty)$. □

We continue with a problem from the 2001 USA team selection test proposed also by the second author of the book.

Example. Let $(a_n)_{n \geq 0}$ be a sequence of real numbers such that

$$a_{n+1} \geq a_n^2 + \frac{1}{5}, \text{ for all } n \geq 0.$$

Prove that $\sqrt{a_{n+5}} \geq a_{n-5}^2$, for all $n \geq 5$.

Solution. It suffices to prove that $a_{n+5} \geq a_n^2$, for all $n \geq 0$. Let us write the inequality for five consecutive indices:

$$\begin{aligned} a_{n+1} &\geq a_n^2 + \frac{1}{5}, \\ a_{n+2} &\geq a_{n+1}^2 + \frac{1}{5}, \\ a_{n+3} &\geq a_{n+2}^2 + \frac{1}{5}, \\ a_{n+4} &\geq a_{n+3}^2 + \frac{1}{5}, \\ a_{n+5} &\geq a_{n+4}^2 + \frac{1}{5}. \end{aligned}$$

If we add these up, we obtain

$$\begin{aligned} a_{n+5} - a_n^2 &\geq (a_{n+1}^2 + a_{n+2}^2 + a_{n+3}^2 + a_{n+4}^2) - (a_{n+1} + a_{n+2} + a_{n+3} + a_{n+4}) + 5 \cdot \frac{1}{5} \\ &= \left(a_{n+1} - \frac{1}{2} \right)^2 + \left(a_{n+2} - \frac{1}{2} \right)^2 + \left(a_{n+3} - \frac{1}{2} \right)^2 + \left(a_{n+4} - \frac{1}{2} \right)^2 \geq 0. \end{aligned}$$

The conclusion follows. □

And finally a more challenging problem from the 64th W.L. Putnam Mathematics Competition.

Example. Let f be a continuous function on the unit square. Prove that

$$\begin{aligned} & \int_0^1 \left(\int_0^1 f(x, y) dx \right)^2 dx + \int_0^1 \left(\int_0^1 f(x, y) dy \right)^2 dy \\ & \leq \left(\int_0^1 \int_0^1 f(x, y) dx dy \right)^2 + \int_0^1 \int_0^1 f(x, y)^2 dx dy. \end{aligned}$$

Solution. To make this problem as simple as possible, we prove the inequality for a Riemann sum, and then pass to the limit. Divide the unit square into n^2 equal squares, then pick a point (x_i, y_j) in each such square and define $a_{ij} = f(x_i, y_j)$, $i, j = 1, 2, \dots, n$. Written for the Riemann sum, the inequality becomes

$$\frac{1}{n^3} \sum_i \left(\left(\sum_j a_{ij} \right)^2 + \left(\sum_j a_{ji} \right)^2 \right) \leq \frac{1}{n^4} \left(\sum_{ij} a_{ij} \right)^2 + \frac{1}{n^2} \left(\sum_{ij} a_{ij}^2 \right).$$

Multiply this by n^4 , then move everything to one side. After cancellations, the inequality becomes

$$(n-1)^2 \sum_{ij} a_{ij}^2 + \sum_{i \neq k, j \neq l} a_{ij} a_{kl} - (n-1) \sum_{ijk, j \neq k} (a_{ij} a_{ik} + a_{ji} a_{ki}) \geq 0.$$

Here we have a quadratic function in the a_{ij} 's that should always be nonnegative. In general, such a quadratic function can be expressed as an algebraic sum of squares, and it is nonnegative precisely when all squares appear with a positive sign. We are left with the problem of representing our expression as a sum of squares. To boost your intuition, look at the following tableau:

$$\begin{array}{cccccccc} a_{11} & \dots & \dots & \dots & \dots & \dots & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \dots & \dots & a_{ij} & \dots & a_{il} & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \dots & \dots & a_{kj} & \dots & a_{kl} & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & \dots & \dots & \dots & \dots & \dots & a_{nn} \end{array}$$

The expression

$$(a_{ij} + a_{kl} - a_{il} - a_{kj})^2$$

when expanded gives rise to the following terms:

$$a_{ij}^2 + a_{kl}^2 + a_{il}^2 + a_{kj}^2 + 2a_{ij}a_{kl} + 2a_{il}a_{kj} - 2a_{il}a_{ij} - 2a_{ij}a_{kj} - 2a_{kl}a_{il} - 2a_{kl}a_{kj}.$$

For a fixed pair (i, j) , the term a_{ij} appears in $(n-1)^2$ such expressions. The products $2a_{ij}a_{kl}$ and $2a_{il}a_{kj}$ appear just once, while the products $2a_{il}a_{ij}$, $2a_{ij}a_{kj}$, $2a_{kl}a_{il}$, $2a_{kl}a_{kj}$ appear $(n-1)$

times (once for each square of the form $(i, j), (i, l), (k, j), (k, l)$). It follows that the expression that we are trying to prove is nonnegative is nothing but

$$\sum_{ijkl} (a_{ij} + a_{kl} - a_{il} - a_{kj})^2,$$

which is of course nonnegative. This proves the inequality for all Riemann sums of the function f , and hence for f itself. \square

101. Find $\min_{a,b \in \mathbb{R}} \max(a^2 + b, b^2 + a)$.

102. Prove that for all real numbers x ,

$$2^x + 3^x - 4^x + 6^x - 9^x \leq 1.$$

103. Find all positive integers n for which the equation

$$nx^4 + 4x + 3 = 0$$

has a real root.

104. Find all triples (x, y, z) of real numbers that are solutions to the system of equations

$$\begin{aligned} \frac{4x^2}{4x^2 + 1} &= y, \\ \frac{4y^2}{4y^2 + 1} &= z, \\ \frac{4z^2}{4z^2 + 1} &= x. \end{aligned}$$

105. Find the minimum of

$$\log_{x_1} \left(x_2 - \frac{1}{4} \right) + \log_{x_2} \left(x_3 - \frac{1}{4} \right) + \cdots + \log_{x_n} \left(x_1 - \frac{1}{4} \right),$$

over all $x_1, x_2, \dots, x_n \in \left(\frac{1}{4}, 1 \right)$.

106. Let a and b be real numbers such that

$$9a^2 + 8ab + 7b^2 \leq 6.$$

Prove that $7a + 5b + 12ab \leq 9$.

107. Let a_1, a_2, \dots, a_n be real numbers such that $a_1 + a_2 + \cdots + a_n \geq n^2$ and $a_1^2 + a_2^2 + \cdots + a_n^2 \leq n^3 + 1$. Prove that $n - 1 \leq a_k \leq n + 1$ for all k .

108. Find all pairs (x, y) of real numbers that are solutions to the system

$$\begin{aligned} x^4 + 2x^3 - y &= -\frac{1}{4} + \sqrt{3}, \\ y^4 + 2y^3 - x &= -\frac{1}{4} - \sqrt{3}. \end{aligned}$$

109. Let n be an even positive integer. Prove that for any real number x there are at least $2^{n/2}$ choices of the signs $+$ and $-$ such that

$$\pm x^n \pm x^{n-1} \pm \cdots \pm x < \frac{1}{2}.$$

2.1.3 The Cauchy-Schwarz Inequality

A direct application of the discussion in the previous section is the proof of the Cauchy-Schwarz (or Cauchy-Bunyakovski-Schwarz) inequality

$$\sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2 \geq \left(\sum_{k=1}^n a_k b_k \right)^2,$$

where the equality holds if and only if the a_i 's and the b_i 's are proportional. The expression

$$\sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2 - \left(\sum_{k=1}^n a_k b_k \right)^2$$

is a quadratic function in the a_i 's and b_i 's. For it to have only nonnegative values, it should be a sum of squares. And this is true by the Lagrange identity

$$\sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2 - \left(\sum_{k=1}^n a_k b_k \right)^2 = \sum_{i < k} (a_i b_k - a_k b_i)^2.$$

Sadly, this proof works only in the finite-dimensional case, while the Cauchy-Schwarz inequality is true in far more generality, such as for square integrable functions. Its correct framework is that of a real or complex vector space, which could be finite or infinite dimensional, endowed with an inner product $\langle \cdot, \cdot \rangle$.

By definition, an inner product is subject to the following conditions:

- (i) $\langle x, x \rangle \geq 0$, with equality if and only if $x = 0$,
- (ii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, for any vectors x, y (here the bar stands for complex conjugation if the vector space is complex),
- (iii) $\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle$, for any vectors x_1, x_2, y and scalars λ_1 and λ_2 .

The quantity $\|x\| = \sqrt{\langle x, x \rangle}$ is called the norm of x . Examples of inner product spaces are \mathbb{R}^n with the usual dot product, \mathbb{C}^n with the inner product

$$\langle (z_1, z_2, \dots, z_n), (w_1, w_2, \dots, w_n) \rangle = z_1 \overline{w_1} + z_2 \overline{w_2} + \dots + z_n \overline{w_n},$$

but also the space of square integrable functions on an interval $[a, b]$ with the inner product

$$\langle f, g \rangle = \int_a^b f(t) \overline{g(t)} dt.$$

The Cauchy-Schwarz inequality. Let x, y be two vectors. Then

$$\|x\| \cdot \|y\| \geq |\langle x, y \rangle|,$$

with equality if and only if the vectors x and y are parallel and point in the same direction.

Proof. We have

$$0 \leq \langle \|y\|x - \|x\|y, \|y\|x - \|x\|y \rangle = 2\|x\|^2\|y\|^2 - \|x\| \cdot \|y\|(\langle x, y \rangle + \langle y, x \rangle),$$

hence $2\|x\| \cdot \|y\| \geq (\langle x, y \rangle + \langle y, x \rangle)$. Yet another trick: rotate y by $\langle x, y \rangle / |\langle x, y \rangle|$. The left-hand side does not change, but because of property (ii) the right-hand side becomes $\frac{1}{|\langle x, y \rangle|}(\langle x, y \rangle \overline{\langle x, y \rangle} + \overline{\langle x, y \rangle} \langle x, y \rangle)$, which is the same as $2|\langle x, y \rangle|$. It follows that

$$\|x\| \cdot \|y\| \geq |\langle x, y \rangle|,$$

which is the Cauchy-Schwarz inequality in its full generality. In our sequence of deductions, the only inequality that showed up holds with equality precisely when the vectors are parallel and point in the same direction. \square

As an example, if f and g are two complex-valued continuous functions on the interval $[a, b]$, or more generally two square integrable functions, then

$$\int_a^b |f(t)|^2 dt \int_a^b |g(t)|^2 dt \geq \left| \int_a^b f(t) \overline{g(t)} dt \right|^2.$$

Let us turn to more elementary problems.

Example. Find the maximum of the function $f(x, y, z) = 5x - 6y + 7z$ on the ellipsoid

$$2x^2 + 3y^2 + 4z^2 = 1.$$

Solution. For a point (x, y, z) on the ellipsoid,

$$\begin{aligned} (f(x, y, z))^2 &= (5x - 6y + 7z)^2 = \left(\frac{5}{\sqrt{2}} \cdot \sqrt{2}x - \frac{6}{\sqrt{3}} \cdot \sqrt{3}y + \frac{7}{2} \cdot 2z \right)^2 \\ &\leq \left(\left(\frac{5}{\sqrt{2}} \right)^2 + \left(-\frac{6}{\sqrt{3}} \right)^2 + \left(\frac{7}{2} \right)^2 \right) ((\sqrt{2}x)^2 + (\sqrt{3}y)^2 + (2z)^2) \\ &= \frac{147}{4} (2x^2 + 3y^2 + 4z^2) = \frac{147}{4}. \end{aligned}$$

Hence the maximum of f is $\sqrt{147}/2$, reached at the point (x, y, z) on the ellipsoid for which $x, z > 0, y < 0$, and $x : y : z = \frac{5}{\sqrt{2}} : -\frac{6}{\sqrt{3}} : \frac{7}{2}$. \square

The next problem was on the short list of the 1993 International Mathematical Olympiad, being proposed by the second author of the book.

Example. Prove that

$$\frac{a}{b+2c+3d} + \frac{b}{c+2d+3a} + \frac{c}{b+2a+3b} + \frac{d}{a+2b+3c} \geq \frac{2}{3},$$

for all $a, b, c, d > 0$.

Solution. Denote by E the expression on the left. Then

$$\begin{aligned}
 & 4(ab + ac + ad + bc + bd + cd)E \\
 &= (a(b + 2c + 3d) + b(c + 2d + 3a) + c(d + 2a + 3b) + d(a + 2b + 3c)) \\
 &\quad \times \left(\frac{a}{b + 2c + 3d} + \frac{b}{c + 2d + 3a} + \frac{c}{d + 2a + 3b} + \frac{d}{a + 2b + 3c} \right) \\
 &\geq (a + b + c + d)^2,
 \end{aligned}$$

where the last inequality is a well-disguised Cauchy-Schwarz. Finally,

$$3(a + b + c + d)^2 \geq 8(ab + ac + ad + bc + bd + cd),$$

because it reduces to

$$(a - b)^2 + (a - c)^2 + (a - d)^2 + (b - c)^2 + (b - d)^2 + (c - d)^2 \geq 0.$$

Combining these two and cancelling the factor $ab + ac + ad + bc + bd + cd$, we obtain the inequality from the statement. \square

And now a list of problems, all of which are to be solved using the Cauchy-Schwarz inequality.

110. If a, b, c are positive numbers, prove that

$$9a^2b^2c^2 \leq (a^2b + b^2c + c^2a)(ab^2 + bc^2 + ca^2).$$

111. If $a_1 + a_2 + \cdots + a_n = n$ prove that $a_1^4 + a_2^4 + \cdots + a_n^4 \geq n$.

112. Let a_1, a_2, \dots, a_n be distinct real numbers. Find the maximum of

$$a_1a_{\sigma(1)} + a_2a_{\sigma(2)} + \cdots + a_na_{\sigma(n)}$$

over all permutations of the set $\{1, 2, \dots, n\}$.

113. Let f_1, f_2, \dots, f_n be positive real numbers. Prove that for any real numbers x_1, x_2, \dots, x_n , the quantity

$$f_1x_1^2 + f_2x_2^2 + \cdots + f_nx_n^2 - \frac{(f_1x_1 + f_2x_2 + \cdots + f_nx_n)^2}{f_1 + f_2 + \cdots + f_n}$$

is nonnegative.

114. Find all positive integers n, k_1, \dots, k_n such that $k_1 + \cdots + k_n = 5n - 4$ and

$$\frac{1}{k_1} + \cdots + \frac{1}{k_n} = 1.$$

- 115.** Prove that the finite sequence a_0, a_1, \dots, a_n of positive real numbers is a geometric progression if and only if

$$(a_0a_1 + a_1a_2 + \dots + a_{n-1}a_n)^2 = (a_0^2 + a_1^2 + \dots + a_{n-1}^2)(a_1^2 + a_2^2 + \dots + a_n^2).$$

- 116.** Let $P(x)$ be a polynomial with positive real coefficients. Prove that

$$\sqrt{P(a)P(b)} \geq P(\sqrt{ab}),$$

for all positive real numbers a and b .

- 117.** Consider the real numbers $x_0 > x_1 > x_2 > \dots > x_n$. Prove that

$$x_0 + \frac{1}{x_0 - x_1} = \frac{1}{x_1 - x_2} + \dots + \frac{1}{x_{n-1} - x_n} \geq x_n + 2n.$$

When does equality hold?

- 118.** Prove that

$$\frac{\sin^3 a}{\sin b} + \frac{\cos^3 a}{\cos b} \geq \sec(a - b),$$

for all $a, b \in (0, \frac{\pi}{2})$.

- 119.** Prove that

$$\frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a} + \frac{1}{2\sqrt[3]{abc}} \geq \frac{(a+b+c+\sqrt[3]{abc})^2}{(a+b)(b+c)(c+a)},$$

for all $a, b, c > 0$.

2.1.4 The Triangle Inequality

In its most general form, the triangle inequality states that in a metric space X the distance function δ satisfies

$$\delta(x, y) \leq \delta(x, z) + \delta(z, y), \text{ for any } x, y, z \in X.$$

An equivalent form is

$$|\delta(x, y) - \delta(y, z)| \leq \delta(x, z).$$

Here are some familiar examples of distance functions: the distance between two real or complex numbers as the absolute value of their difference, the distance between two vectors in n -dimensional Euclidean space as the length of their difference $\|v - w\|$, the distance between two matrices as the norm of their difference, the distance between two continuous functions on the same interval as the supremum of the absolute value of their difference. In all these cases the triangle inequality holds.

Let us see how the triangle inequality can be used to solve a problem from T.B. Soulamis's book *Les olympiades de mathématiques: Réflexes et stratégies* (Ellipses, 1999).

Example. For positive numbers a, b, c prove the inequality

$$\sqrt{a^2 - ab + b^2} + \sqrt{b^2 - bc + c^2} \geq \sqrt{a^2 + ac + c^2}.$$

Solution. The inequality suggests the following geometric construction. With the same origin O , draw segments OA , OB , and OC of lengths a , b , respectively c , such that OB makes 60° angles with OA and OC (see Figure 12).

The law of cosines in the triangles OAB , OBC , and OAC gives $AB^2 = a^2 - ab + b^2$, $BC^2 = b^2 - bc + c^2$, and $AC^2 = a^2 + ac + c^2$. Plugging these formulas into the triangle inequality $AB + BC \geq AC$ produces the inequality from the statement. \square

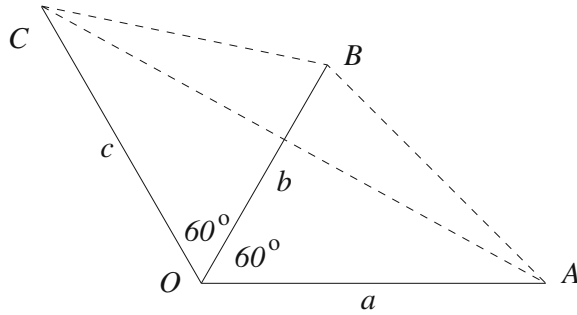


Figure 12

Example. Let $P(x)$ be a polynomial whose coefficients lie in the interval $[1, 2]$, and let $Q(x)$ and $R(x)$ be two nonconstant polynomials such that $P(x) = Q(x)R(x)$, with $Q(x)$ having the dominant coefficient equal to 1. Prove that $|Q(3)| > 1$.

Solution. Let $P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$. We claim that the zeros of $P(x)$ lie in the union of the half-plane $\operatorname{Re} z = 0$ and the disk $|z| < 2$.

Indeed, suppose that $P(x)$ has a zero z such that $\operatorname{Re} z > 0$ and $|z| = 2$. From $P(z) = 0$, we deduce that

$$a_n z^n + a_{n-1} z^{n-1} = -a_{n-2} z^{n-2} - a_{n-3} z^{n-3} - \dots - a_0.$$

Dividing through by z^n , which is not equal to 0, we obtain

$$a_n + \frac{a_{n-1}}{z} = -\frac{a_{n-2}}{z^2} - \frac{a_{n-3}}{z^3} - \dots - \frac{a_0}{z^n}.$$

Note that $\operatorname{Re} z > 0$ implies that $\operatorname{Re} \frac{1}{z} > 0$. Hence

$$\begin{aligned} 1 \leq a_n &\leq \operatorname{Re} \left(a_n + \frac{a_{n-1}}{z} \right) = \operatorname{Re} \left(-\frac{a_{n-2}}{z^2} - \frac{a_{n-3}}{z^3} - \dots - \frac{a_0}{z^n} \right) \\ &\leq \left| -\frac{a_{n-2}}{z^2} - \frac{a_{n-3}}{z^3} - \dots - \frac{a_0}{z^n} \right| \leq \frac{a_{n-2}}{|z|^2} + \frac{a_{n-3}}{|z|^3} + \dots + \frac{a_0}{|z|^n}, \end{aligned}$$

where for the last inequality we used the triangle inequality. Because the a_i 's are in the interval $[1, 2]$, this is strictly less than

$$2|z|^{-2}(1 + |z|^{-1} + |z|^{-2} + \cdots) = \frac{2|z|^{-2}}{1 - |z|^{-1}}.$$

The last quantity must therefore be greater than 1. But this cannot happen if $|z| \geq 2$, because the inequality reduces to $\left(\frac{2}{|z|} - 1\right)\left(\frac{1}{|z|} + 1\right) > 0$, impossible. This proves the claim.

Returning to the problem, $Q(x) = (x - z_1)(x - z_2) \cdots (x - z_k)$, where z_1, z_2, \dots, z_k are some of the zeros of $P(x)$. Then

$$|Q(3)| = |3 - z_1| \cdot |3 - z_2| \cdots |3 - z_k|.$$

If $\operatorname{Re} z_i \leq 0$, then $|3 - z_i| \geq 0$. On the other hand, if $|z_i| < 2$, then by the triangle inequality $|3 - z_i| \geq 3 - |z_i| > 1$. Hence $|Q(3)|$ is a product of terms greater than 1, and the conclusion follows. \square

More applications follow.

120. Let a, b, c be the side lengths of a triangle with the property that for any positive integer n , the numbers a^n, b^n, c^n can also be the side lengths of a triangle. Prove that the triangle is necessarily isosceles.

121. Given the vectors $\vec{a}, \vec{b}, \vec{c}$ in the plane, show that

$$\|\vec{a}\| + \|\vec{b}\| + \|\vec{c}\| + \|\vec{a} + \vec{b} + \vec{c}\| \geq \|\vec{a} + \vec{b}\| + \|\vec{a} + \vec{c}\| + \|\vec{b} + \vec{c}\|.$$

122. Let $P(z)$ be a polynomial with real coefficients whose roots can be covered by a disk of radius R . Prove that for any real number k , the roots of the polynomial $nP(z) - kP'(z)$ can be covered by a disk of radius $R + |k|$, where n is the degree of $P(z)$, and $P'(z)$ is the derivative.

123. Prove that the positive real numbers a, b, c are the side lengths of a triangle if and only if

$$a^2 + b^2 + c^2 < 2\sqrt{a^2b^2 + b^2c^2 + c^2a^2}.$$

124. Let $ABCD$ be a convex cyclic quadrilateral. Prove that

$$|AB - CD| + |AD - BC| \geq 2|AC - BD|.$$

125. Let V_1, V_2, \dots, V_m and W_1, W_2, \dots, W_m be isometries of \mathbb{R}^n (m, n positive integers). Assume that for all x with $\|x\| \leq 1$, $\|V_i x - W_i x\| \leq 1$, $i = 1, 2, \dots, m$. Prove that

$$\left\| \left(\prod_{i=1}^m V_i \right) x - \left(\prod_{i=1}^m W_i \right) x \right\| \leq m,$$

for all x with $\|x\| \leq 1$.

126. Given an equilateral triangle ABC and a point P that does not lie on the circumcircle of ABC , show that one can construct a triangle with sides the segments PA , PB , and PC . If P lies on the circumcircle, show that one of these segments is equal to the sum of the other two.

127. Let M be a point in the plane of the triangle ABC whose centroid is G . Prove that

$$MA^3 \cdot BC + MB^3 \cdot AC + MC^3 \cdot AB \geq 3MG \cdot AB \cdot BC \cdot CA.$$

2.1.5 The Arithmetic Mean-Geometric Mean Inequality

Jensen's inequality, which will be discussed in the section about convex functions, states that if f is a real-valued concave function, then

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n),$$

for any x_1, x_2, \dots, x_n in the domain of f and for any positive weights $\lambda_1, \lambda_2, \dots, \lambda_n$ with $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$. Moreover, if the function is nowhere linear (that is, if it is strictly concave) and the numbers $\lambda_1, \lambda_2, \dots, \lambda_n$ are nonzero, then equality holds if and only if $x_1 = x_2 = \cdots = x_n$.

Applying this to the concave function $f(x) = \ln x$, the positive numbers x_1, x_2, \dots, x_n , and the weights $\lambda_1 = \lambda_2 = \cdots = \lambda_n = \frac{1}{n}$, we obtain

$$\ln \frac{x_1 + x_2 + \cdots + x_n}{n} \geq \frac{\ln x_1 + \ln x_2 + \cdots + \ln x_n}{n}.$$

Exponentiation yields the following inequality.

The arithmetic mean-geometric mean inequality. *Let x_1, x_2, \dots, x_n be nonnegative real numbers. Then*

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \cdots x_n},$$

with equality if and only if all numbers are equal.

Proof. We will call this inequality AM-GM for short. We give it an alternative proof using derivatives, a proof by induction on n . For $n = 2$ the inequality is equivalent to the obvious $(\sqrt{a_1} - \sqrt{a_2})^2 \geq 0$. Next, assume that the inequality holds for any $n - 1$ positive numbers, meaning that

$$\frac{x_1 + x_2 + \cdots + x_{n-1}}{n-1} \geq \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}},$$

with equality only when $x_1 = x_2 = \cdots = x_{n-1}$. To show that the same is true for n numbers, consider the function $f : (0, \infty) \rightarrow \mathbb{R}$,

$$f(x) = \frac{x_1 + x_2 + \cdots + x_{n-1} + x}{n} - \sqrt[n]{x_1 x_2 \cdots x_{n-1} x}.$$

To find the minimum of this function we need the critical points. The derivative of f is

$$f'(x) = \frac{1}{n} - \frac{\sqrt[n]{x_1 x_2 \cdots x_{n-1}}}{n} x^{\frac{1}{n}-1} = \frac{x^{\frac{1}{n}-1}}{n} \left(x^{1-\frac{1}{n}} - \sqrt[n]{x_1 x_2 \cdots x_{n-1}} \right).$$

Setting this equal to zero, we find the unique critical point $x = \sqrt[n-1]{x_1 x_2 \cdots x_n}$, since in this case $x^{1-\frac{1}{n}} = \sqrt[n]{x_1 x_2 \cdots x_{n-1}}$. Moreover, the function $x^{1-\frac{1}{n}}$ is increasing on $(0, \infty)$; hence $f'(x) < 0$ for $x < \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}$, and $f'(x) > 0$ for $x > \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}$. We find that f has a global minimum at $x = \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}$, where it takes the value

$$\begin{aligned} f\left(\sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}\right) &= \frac{x_1 + x_2 + \cdots + x_{n-1} + \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}}{n} \\ &\quad - \sqrt[n]{x_1 x_2 \cdots x_{n-1}} \cdot \sqrt[n(n-1)]{x_1 x_2 \cdots x_{n-1}} \\ &= \frac{x_1 + x_2 + \cdots + x_{n-1} + \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}}{n} \\ &\quad - \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}} \\ &= \frac{x_1 + x_2 + \cdots + x_{n-1} - (n-1) \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}}}{n}. \end{aligned}$$

By the induction hypothesis, this minimum is nonnegative, and is equal to 0 if and only if $x_1 = x_2 = \cdots = x_{n-1}$. We conclude that $f(x_n) \geq 0$ with equality if and only if $x_1 = x_2 = \cdots = x_{n-1}$ and $x_n = \sqrt[n-1]{x_1 x_2 \cdots x_{n-1}} = x_1$. This completes the induction. \square

We apply the AM-GM inequality to solve two problems composed by the second author of the book.

Example. Find the global minimum of the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = 3^{x+y}(3^{x-1} + 3^{y-1} - 1).$$

Solution. The expression

$$3f(x, y) + 1 = 3^{2x+y} + 3^{x+2y} + 1 - 3 \cdot 3^{x+y}$$

is of the form $a^3 + b^3 + c^3 - 3abc$, where $a = \sqrt[3]{3^{2x+y}}$, $b = \sqrt[3]{3^{x+2y}}$, and $c = 1$, all of which are positive. By the AM-GM inequality, this expression is nonnegative. It is equal to zero only when $a = b = c$, that is, when $2x + y = x + 2y = 0$. We conclude that the minimum of f is $f(0, 0) = -\frac{1}{3}$. \square

Example. Let a, b, c, d be positive real numbers with $abcd = 1$. Prove that

$$\frac{a}{b+c+d+1} + \frac{b}{c+d+a+1} + \frac{c}{d+a+b+1} + \frac{d}{a+b+c+1} \geq 1.$$

Solution. A first idea is to homogenize this inequality, and for that we replace the 1 in each denominator by $\sqrt[4]{abcd}$, transforming the inequality into

$$\begin{aligned} \frac{a}{b+c+d+\sqrt[4]{abcd}} + \frac{b}{c+d+a+\sqrt[4]{abcd}} + \frac{c}{d+a+b+\sqrt[4]{abcd}} \\ + \frac{d}{a+b+c+\sqrt[4]{abcd}} \geq 1. \end{aligned}$$

Then we apply the AM-GM inequality to the last term in each denominator to obtain the stronger inequality

$$\frac{4a}{a+5(b+c+d)} + \frac{4b}{b+5(c+d+a)} + \frac{4c}{c+5(d+a+b)} + \frac{4d}{d+5(a+b+c)} \geq 1,$$

which we proceed to prove.

In order to simplify computations, it is better to denote the four denominators by $16x$, $16y$, $16z$, $16w$, respectively. Then $a+b+c+d = x+y+z+w$, and so $4a+16x = 4b+16y = 4c+16z = 4d+16w = 5(x+y+z+w)$. The inequality becomes

$$\begin{aligned} \frac{-11x+5(y+z+w)}{16x} + \frac{-11y+5(z+w+x)}{16y} + \frac{-11z+5(w+x+y)}{16z} \\ + \frac{-11w+5(x+y+z)}{16w} \geq 1, \end{aligned}$$

or

$$-4 \cdot 11 + 5 \left(\frac{y}{x} + \frac{z}{x} + \frac{w}{x} + \frac{z}{y} + \frac{w}{y} + \frac{x}{y} + \frac{w}{z} + \frac{x}{z} + \frac{y}{z} + \frac{x}{w} + \frac{y}{w} + \frac{z}{w} \right) \geq 16.$$

And this follows by applying the AM-GM inequality to the twelve summands in the parentheses. \square

We continue with a third example, which is an problem of A. Basyoni that was given in 2015 at a preliminary selection test for the team that represented the United States at the International Mathematical Olympiad in 2016.

Example. Let x, y, z be real numbers satisfying $x^4 + y^4 + z^4 + xyz = 4$. Show that

$$\sqrt{2-x} \geq \frac{y+z}{2}.$$

Solution. We have selected the problem for the book because of this elegant solution based on the AM-GM inequality found by the member of the Canadian team Zh.Q. (Alex) Song. It suffices to show that

$$\sqrt{2-x} \geq \left| \frac{y+z}{2} \right|.$$

This inequality and the fact that the square root is well defined follow simultaneously if we prove that

$$x + \left(\frac{y+z}{2} \right)^2 \leq 2.$$

Apply the AM-GM inequality three times:

$$\begin{aligned} \frac{x^4}{8} + \frac{y^4}{8} + \frac{y^4}{8} + \frac{1}{8} &\geq \frac{xy^2}{2} \\ \frac{x^4}{8} + \frac{z^4}{8} + \frac{z^4}{8} + \frac{1}{8} &\geq \frac{xz^2}{2} \\ \frac{3x^4}{4} + \frac{3}{4} &\geq \frac{3x^2}{2}. \end{aligned}$$

Then apply the power-mean inequality:

$$\frac{y^4 + z^4}{2} \geq \left(\frac{y + z}{2} \right)^4,$$

to write

$$\frac{3y^4}{4} + \frac{3z^4}{4} \geq \frac{3}{2} \left(\frac{y + z}{2} \right)^4.$$

Now add the four inequalities and use the relation from the statement to obtain

$$5 \geq \frac{3}{2}x^2 + \frac{3}{2} \left[\left(\frac{y + z}{2} \right)^2 \right]^2 + 2 \left[x \left(\frac{y + z}{2} \right) \right]^2.$$

Finally, noticing that the AM-GM inequality implies

$$\frac{1}{4} \left(x^2 + \left[\left(\frac{y + z}{2} \right)^2 \right]^2 \right) \geq \frac{1}{2} \left[x \left(\frac{y + z}{2} \right)^2 \right],$$

we obtain

$$5 \geq \frac{5}{4} \left[x + \left(\frac{y + z}{2} \right)^2 \right]^2,$$

and the conclusion follows. □

For completeness let us prove this particular case of the power mean inequality:

$$\frac{y^4 + z^4}{2} \geq \left(\frac{y^2 + z^2}{2} \right)^2 \geq \left[\left(\frac{y + z}{2} \right)^2 \right]^2 = \left(\frac{y + z}{2} \right)^4.$$

It becomes clear after expanding the square that the first inequality is a consequence of the AM-GM inequality. Taking the square root of the second inequality, we recognize that it is of the same type. So we are done.

Try your hand at the following problems.

128. Show that all real roots of the polynomial $P(x) = x^5 - 10x + 35$ are negative.

129. Find all real numbers that satisfy

$$x \cdot 2^{\frac{1}{x}} + \frac{1}{x} \cdot 2^x = 4.$$

130. Let a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_n be nonnegative numbers. Show that

$$(a_1 a_2 \cdots a_n)^{1/n} + (b_1 b_2 \cdots b_n)^{1/n} \leq ((a_1 + b_1)(a_2 + b_2) \cdots (a_n + b_n))^{1/n}.$$

131. Let a, b, c be the side lengths of a triangle with semiperimeter 1. Prove that

$$1 < ab + bc + ca - abc \leq \frac{28}{27}.$$

132. Which number is larger,

$$\prod_{n=1}^{25} \left(1 - \frac{n}{365}\right) \quad \text{or} \quad \frac{1}{2}?$$

133. On a sphere of radius 1 are given four points A, B, C, D such that

$$AB \cdot AC \cdot AD \cdot BC \cdot BD \cdot CD = \frac{2^9}{3^3}.$$

Prove that the tetrahedron $ABCD$ is regular.

134. Prove that

$$\frac{y^2 - x^2}{2x^2 + 1} + \frac{z^2 - y^2}{2y^2 + 1} + \frac{x^2 - z^2}{2z^2 + 1} \geq 0,$$

for all real numbers x, y, z .

135. Let a_1, a_2, \dots, a_n be positive real numbers such that $a_1 + a_2 + \cdots + a_n < 1$. Prove that

$$\frac{a_1 a_2 \cdots a_n (1 - (a_1 + a_2 + \cdots + a_n))}{(a_1 + a_2 + \cdots + a_n)(1 - a_1)(1 - a_2) \cdots (1 - a_n)} \leq \frac{1}{n^{n+1}}.$$

136. Consider the positive real numbers x_1, x_2, \dots, x_n with $x_1 x_2 \cdots x_n = 1$. Prove that

$$\frac{1}{n-1+x_1} + \frac{1}{n-1+x_2} + \cdots + \frac{1}{n-1+x_n} \leq 1.$$

2.1.6 Sturm's Principle

In this section we present a method for proving inequalities that is based on real analysis. It is based on a principle attributed to R. Sturm, phrased as follows.

Sturm's principle. *Given a function f defined on a set M and a point $x_0 \in M$, if*

(i) *f has a maximum (minimum) on M , and*

(ii) *if no other point x in M is a maximum (minimum) of f ,*

then x_0 is the maximum (minimum) of f .

But how to decide whether the function f has a maximum or a minimum? Two results from real analysis come in handy.

Theorem. *A continuous function on a compact set always attains its extrema.*

Theorem. *A closed and bounded subset of \mathbb{R}^n is compact.*

Let us see how Sturm's principle can be applied to a problem from the first Balkan Mathematical Olympiad in 1984.

Example. Let $\alpha_1, \alpha_2, \dots, \alpha_n$ be positive real numbers, $n \geq 2$, such that $\alpha_1 + \alpha_2 + \dots + \alpha_n = 1$. Prove that

$$\frac{\alpha_1}{1 + \alpha_2 + \dots + \alpha_n} + \frac{\alpha_2}{1 + \alpha_1 + \dots + \alpha_n} + \dots + \frac{\alpha_n}{1 + \alpha_1 + \dots + \alpha_{n-1}} \geq \frac{n}{2n-1}.$$

Solution. Rewrite the inequality as

$$\frac{\alpha_1}{2 - \alpha_1} + \frac{\alpha_2}{2 - \alpha_2} + \dots + \frac{\alpha_n}{2 - \alpha_n} \geq \frac{n}{2n-1},$$

and then define the function

$$f(\alpha_1, \alpha_2, \dots, \alpha_n) = \frac{\alpha_1}{2 - \alpha_1} + \frac{\alpha_2}{2 - \alpha_2} + \dots + \frac{\alpha_n}{2 - \alpha_n}.$$

As said in the statement, this function is defined on the subset of \mathbb{R}^n consisting of points whose coordinates are positive and add up to 1. We would like to show that on this set f is greater than or equal to $\frac{n}{2n-1}$.

Does f have a minimum? The domain of f is bounded but is not closed, being the interior of a tetrahedron. We can enlarge it, though, by adding the boundary, to the set

$$M = \{(\alpha_1, \alpha_2, \dots, \alpha_n) \mid \alpha_1 + \alpha_2 + \dots + \alpha_n = 1, \alpha_i \geq 0, i = 1, 2, \dots, n\}.$$

We now know that f has a minimum on M .

A look at the original inequality suggests that the minimum is attained when all the α_i 's are equal. So let us choose a point $(\alpha_1, \alpha_2, \dots, \alpha_n)$ for which $\alpha_i \neq \alpha_j$ for some indices i, j . Assume that $\alpha_i < \alpha_j$ and let us see what happens if we substitute $\alpha_i + x$ for α_i and $\alpha_j - x$

for α_j , with $0 < x < \alpha_j - \alpha_i$. In the defining expression of f , only the i th and j th terms change. Moreover,

$$\begin{aligned} & \frac{\alpha_i}{2 - \alpha_i} + \frac{\alpha_j}{2 - \alpha_j} - \frac{\alpha_i + x}{2 - \alpha_i - x} - \frac{\alpha_j - x}{2 - \alpha_j + x} \\ &= \frac{2x(\alpha_j - \alpha_i - x)(4 - \alpha_i - \alpha_j)}{(2 - \alpha_i)(2 - \alpha_j)(2 - \alpha_i - x)(2 - \alpha_j + x)} > 0, \end{aligned}$$

so on moving the numbers closer, the value of f decreases. It follows that the point that we picked was not a minimum. Hence the only possible minimum is $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$, in which case the value of f is $\frac{n}{2n-1}$. This proves the inequality. \square

However, in most situations, as is the case with this problem, we can bypass the use of real analysis and argue as follows. If the α_i 's were not all equal, then one of them must be less than $\frac{1}{n}$ and one of them must be greater. Take these two numbers and move them closer until one of them reaches $\frac{1}{n}$. Then stop and choose another pair. Continue the algorithm until all numbers become $\frac{1}{n}$. At this very moment, the value of the expression is

$$\frac{1}{n} \left(2 - \frac{1}{n}\right)^{-1} \cdot n = \frac{n}{2n-1}.$$

Since during the process the value of the expression kept decreasing, initially it must have been greater than or equal to $\frac{n}{2n-1}$. This proves the inequality.

Let us summarize the last idea. We want to maximize (or minimize) an n -variable function, and we have a candidate for the extremum. If we can move the variables one by one toward the maximum without decreasing (respectively, increasing) the value of the function, then the candidate is indeed the desired extremum. This approach is more elementary but can be more time consuming than the application of the principle itself.

Let us revisit the AM-GM inequality with a proof using Sturm's principle.

The arithmetic mean-geometric mean inequality. Let x_1, x_2, \dots, x_n be nonnegative real numbers. Then

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n}.$$

with equality if and only if $x_1 = x_2 = \dots = x_n$.

Proof. The inequality is homogeneous in the variables, so the general case follows if we check the inequality for a fixed value of the sum of the numbers, say $x_1 + x_2 + \dots + x_n = 1$. This amounts to checking that $\sqrt[n]{x_1 x_2 \dots x_n} \leq \frac{1}{n}$ if $x_1 + x_2 + \dots + x_n = 1$ with equality only when $x_1 = x_2 = \dots = x_n$, and this is equivalent to checking $x_1 x_2 \dots x_n \leq \frac{1}{n^n}$ with equality as specified.

The set

$$K = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid x_j \geq 0, x_1 + x_2 + \dots + x_n = 1\}$$

contains all its limit points, so it is closed. It also lies in the hypercube $[0, 1]^n$ so it is bounded, thus it is compact. The function

$$f : K \rightarrow \mathbb{R}, \quad f(x_1, x_2, \dots, x_n) = x_1 x_2 \dots x_n$$

is continuous, being a polynomial, so it attains its maximum. This maximum is not attained at a point where not all x_j are equal, because if $x_j < x_k$ and we replace x_j by $x_j + \varepsilon$ and x_k by $x_k - \varepsilon$, where $\varepsilon = \frac{x_k - x_j}{2}$, then the value of f increases to

$$\begin{aligned} \prod_{i \neq j, k} x_i (x_j + \varepsilon)(x_k - \varepsilon) &= \prod_{i \neq j, k} x_i (x_j x_k + \varepsilon(x_k - x_j) + \varepsilon^2) \\ &= \prod_{i \neq j, k} x_i (x_j x_k + \varepsilon^2) = f(x_1, x_2, \dots, x_n) + \varepsilon^2 \prod_{i \neq j, k} x_i. \end{aligned}$$

Thus the only candidate for the maximum is $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ and in this case the inequality holds with equality. \square

You can find more applications of Sturm's principle below.

137. Let a, b, c be nonnegative real numbers such that $a + b + c = 1$. Prove that

$$4(ab + bc + ac) - 9abc \leq 1.$$

138. Let $x_1, x_2, \dots, x_n, n \geq 2$, be positive numbers such that

$$x_1 + x_2 + \dots + x_n = 1.$$

Prove that

$$\left(1 + \frac{1}{x_1}\right) \left(1 + \frac{1}{x_2}\right) \cdots \left(1 + \frac{1}{x_n}\right) \geq (n+1)^n.$$

139. Prove that a necessary and sufficient condition that a triangle inscribed in an ellipse have maximum area is that its centroid coincide with the center of the ellipse.

140. Let $n > 2$ be an integer. A convex n -gon of area 1 is inscribed in a circle. What is the minimum that the radius of the circle can be?

141. Let $a, b, c > 0, a + b + c = 1$. Prove that

$$0 \leq ab + bc + ac - 2abc \leq \frac{7}{27}.$$

142. Let x_1, x_2, \dots, x_n be n real numbers such that $0 < x_j \leq \frac{1}{2}$, for $1 \leq j \leq n$. Prove the inequality

$$\frac{\prod_{j=1}^n x_j}{\left(\sum_{j=1}^n x_j\right)^n} \leq \frac{\prod_{j=1}^n (1 - x_j)}{\left(\sum_{j=1}^n (1 - x_j)\right)^n}.$$

- 143.** Let a, b, c , and d be nonnegative numbers such that $a \leq 1, a + b \leq 5, a + b + c \leq 14, a + b + c + d \leq 30$. Prove that

$$\sqrt{a} + \sqrt{b} + \sqrt{c} + \sqrt{d} \leq 10.$$

- 144.** What is the maximal value of the expression $\sum_{i < j} x_i x_j$ if x_1, x_2, \dots, x_n are nonnegative integers whose sum is equal to m ?

- 145.** Given the $n \times n$ array $(a_{ij})_{ij}$ with $a_{ij} = i + j - 1$, what is the smallest product of n elements of the array provided that no two lie on the same row or column?

- 146.** Given a positive integer n , find the minimum value of

$$\frac{x_1^3 + x_2^3 + \dots + x_n^3}{x_1 + x_2 + \dots + x_n}$$

subject to the condition that x_1, x_2, \dots, x_n be distinct positive integers.

2.1.7 Other Inequalities

We conclude with a section for the inequalities aficionado. Behind each problem hides a famous inequality.

- 147.** If x and y are positive real numbers, show that $x^y + y^x > 1$.

- 148.** Prove that for all $a, b, c \geq 0$,

$$(a^5 - a^2 + 3)(b^5 - b^2 + 3)(c^5 - c^2 + 3) \geq (a + b + c)^3.$$

- 149.** Assume that all the zeros of the polynomial $P(x) = x^n + a_1 x^{n-1} + \dots + a_n$ are real and positive. Show that if there exist $1 \leq m < p \leq n$ such that $a_m = (-1)^m \binom{n}{m}$ and $a_p = (-1)^p \binom{n}{p}$, then $P(x) = (x - 1)^n$.

- 150.** Let $n > 2$ be an integer, and let x_1, x_2, \dots, x_n be positive numbers with the sum equal to 1. Prove that

$$\prod_{i=1}^n \left(1 + \frac{1}{x_i}\right) \geq \prod_{i=1}^n \left(\frac{n - x_i}{1 - x_i}\right).$$

- 151.** Let $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ be real numbers such that

$$(a_1^2 + a_2^2 + \dots + a_n^2 - 1)(b_1^2 + b_2^2 + \dots + b_n^2 - 1) > (a_1 b_1 + a_2 b_2 + \dots + a_n b_n - 1)^2.$$

Prove that $a_1^2 + a_2^2 + \dots + a_n^2 > 1$ and $b_1^2 + b_2^2 + \dots + b_n^2 > 1$.

- 152.** Let a, b, c, d be positive numbers such that $abc = 1$. Prove that

$$\frac{1}{a^3(b+c)} + \frac{1}{b^3(c+a)} + \frac{1}{c^3(a+b)} \geq \frac{3}{2}.$$

2.2 Polynomials

2.2.1 A Warmup in One-Variable Polynomials

A polynomial is a sum of the form

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0,$$

where x is the variable, and a_n, a_{n-1}, \dots, a_0 are constant coefficients. If $a_n \neq 0$, the number n is called the degree, denoted by $\deg(P(x))$. If $a_n = 1$, the polynomial is called monic. The sets, which, in fact, are rings, of polynomials with integer, rational, real, or complex coefficients are denoted, respectively, by $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, $\mathbb{R}[x]$, and $\mathbb{C}[x]$. A number r such that $P(r) = 0$ is called a zero of $P(x)$, or a root of the equation $P(x) = 0$. By the Gauss-d'Alembert theorem, also called the fundamental theorem of algebra, every nonconstant polynomial with complex coefficients has at least one complex zero. Consequently, the number of zeros of a polynomial equals the degree, multiplicities counted. For a number α , $P(\alpha) = a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_0$ is called the value of the polynomial at α .

We begin the section on polynomials with an old problem from the 1943 competition of the *Mathematics Gazette*, Bucharest, proposed by Gh. Buicliu.

Example. Verify the equality

$$\sqrt[3]{20 + 14\sqrt{2}} + \sqrt[3]{20 - 14\sqrt{2}} = 4.$$

Solution. Apparently, this problem has nothing to do with polynomials. But let us denote the complicated irrational expression by x and analyze its properties. Because of the cube roots, it becomes natural to raise x to the third power:

$$\begin{aligned} x^3 &= 20 + 14\sqrt{2} + 20 - 14\sqrt{2} \\ &\quad + 3\sqrt[3]{(20 + 14\sqrt{2})(20 - 14\sqrt{2})} \left(\sqrt[3]{20 + 14\sqrt{2}} + \sqrt[3]{20 - 14\sqrt{2}} \right) \\ &= 40 + 3x\sqrt[3]{400 - 392} = 40 + 6x. \end{aligned}$$

And now we see that x satisfies the polynomial equation

$$x^3 - 6x - 40 = 0.$$

We have already been told that 4 is a root of this equation. The other two roots are complex, and hence x can only equal 4, the desired answer. \square

Of course, one can also recognize the quantities under the cube roots to be the cubes of $2 + \sqrt{2}$ and $2 - \sqrt{2}$, but that is just a lucky strike.

The second example is a problem from the Russian Journal *Kvant* (*Quantum*), proposed by A. Alexeev.

Example. Prove that for every odd positive integer n , there is a constant c_n such that

$$\frac{\tan x + \tan\left(x + \frac{\pi}{n}\right) + \cdots + \tan\left(x + \frac{(n-1)\pi}{n}\right)}{\tan x \tan\left(x + \frac{\pi}{n}\right) \cdots \tan\left(x + \frac{(n-1)\pi}{n}\right)} = c_n,$$

for all x for which the denominator is nonzero. Find the value of c_n .

Solution. Since the tangent function is periodic with period π , it suffices to look at $x \in [0, \pi]$. Consider the function $f : [0, \pi] \rightarrow \mathbb{R}$,

$$f(x) = \frac{\tan x + \tan\left(x + \frac{\pi}{n}\right) + \cdots + \tan\left(x + \frac{(n-1)\pi}{n}\right)}{\tan x \tan\left(x + \frac{\pi}{n}\right) \cdots \tan\left(x + \frac{(n-1)\pi}{n}\right)}.$$

Denote $\tan x = \xi$ and $\tan \frac{k\pi}{n} = t_k$, $k = 0, 1, \dots, n-1$. Then the numerator and the denominator are of the form

$$\begin{aligned} \frac{P_1(\xi)}{Q(\xi)} &= \xi + \frac{\xi + t_1}{1 - \xi t_1} + \cdots + \frac{\xi + t_{n-1}}{1 - \xi t_{n-1}} \\ \frac{P_2(\xi)}{Q(\xi)} &= \xi \cdot \frac{\xi + t_1}{1 - \xi t_1} \cdots \frac{\xi + t_{n-1}}{1 - \xi t_{n-1}} \end{aligned}$$

where $P_1(\xi)$, $P_2(\xi)$, $Q(\xi)$ are polynomials, and $Q(\xi) = (1 - \xi t_1)(1 - \xi t_2) \cdots (1 - \xi t_{n-1})$.

The polynomials $P_1(\xi)$, $P_2(\xi)$ have n th degree. Because of the fact that n is odd, and of the trigonometric identity $\tan(\pi - x) = -\tan x$, the roots of $P_1(\xi)$ must be $0, t_1, t_2, \dots, t_{n-1}$. Of course these are also the roots of $P_2(\xi)$. It follows that one of the polynomials is a constant multiple of the other. This proves the existence of the constant c_n .

To find c_n , note that it is equal to the ratio of the dominant coefficient of the polynomials $P_1(\xi)$ and $P_2(\xi)$. In the case of the first polynomial, this coefficient is

$$\tan \frac{\pi}{n} \tan \frac{2\pi}{n} \cdots \tan \frac{(n-1)\pi}{n} = (-1)^{\frac{n-1}{2}n} \quad (\text{See Problem 207}).$$

For the second polynomial this number is equal to 1. Hence $c_n = (-1)^{\frac{n-1}{2}n}$. □

And now the problems.

153. Find all solutions to the equation

$$(x+1)(x+2)(x+3)^2(x+4)(x+5) = 360.$$

154. Solve the polynomial equation

$$x^3 - (7 + 2\sqrt{5})x + \sqrt{5} + 1 = 0.$$

- 155.** Let a, b, c be real numbers. Prove that three roots of the equation

$$\frac{b+c}{x-a} + \frac{c+a}{x-b} + \frac{a+b}{x-c} = 3$$

are real.

- 156.** Find all polynomials satisfying the functional equation

$$(x+1)P(x) = (x-10)P(x+1).$$

- 157.** Let $n > 1$ be an integer and x, a_1, a_2, \dots, a_n be distinct real numbers. Show that

$$\begin{aligned} & \frac{(x-a_2)(x-a_3)\cdots(x-a_n)}{(a_1-a_2)(a_1-a_3)\cdots(a_1-a_n)} + \frac{(x-a_1)(x-a_3)\cdots(x-a_n)}{(a_2-a_1)(a_2-a_3)\cdots(a_2-a_n)} \\ & + \cdots + \frac{(x-a_1)(x-a_2)\cdots(x-a_{n-1})}{(a_n-a_1)(a_n-a_2)\cdots(a_n-a_{n-1})} = 1. \end{aligned}$$

- 158.** Let $P(x)$ be a polynomial of odd degree with real coefficients. Show that the equation $P(P(x)) = 0$ has at least as many real roots as the equation $P(x) = 0$, counted without multiplicities.

- 159.** Let $P(x) = x^2 + 2007x + 1$. Prove that for every positive integer n , $P^{(n)}(x) = 0$ has at least one real root, where $P^{(n)}$ denotes P composed with itself n times.

- 160.** Determine all polynomials $P(x)$ with real coefficients for which there exists a positive integer n such that for all x ,

$$P\left(x + \frac{1}{n}\right) + P\left(x - \frac{1}{n}\right) = 2P(x).$$

- 161.** Find a polynomial with integer coefficients that has the zero $\sqrt{2} + \sqrt[3]{3}$.

- 162.** Let $P(x)$ be a polynomial with real coefficients that satisfies the functional equation

$$(x-1)P(x+2) = (x+1)P(x-1) + 2, \text{ for all } x \in \mathbb{R}.$$

Compute $P(-1989)$.

- 163.** Consider the polynomial with real coefficients $P(x) = x^6 + ax^5 + bx^4 + cx^3 + bx^2 + ax + 1$, and let x_1, x_2, \dots, x_6 be its zeros. Prove that

$$\prod_{k=1}^6 (x_k^2 + 1) = (2a - c)^2.$$

- 164.** Let $P(z) = (z - z_1)(z - z_2)\cdots(z - z_n)$ with $|z_i| \geq 1, i = 1, 2, \dots, n$. Prove that if $0 < r < 1$, then for any z , with $|z| = 1$,

$$\left| \frac{P(z)}{P(rz)} \right| \leq \left(\frac{2}{1+r} \right)^n.$$

165. Let $P(x) = x^4 + ax^3 + bx^2 + cx + d$ and $Q(x) = x^2 + px + q$ be two polynomials with real coefficients. Suppose that there exists an interval (r, s) of length greater than 2 such that both $P(x)$ and $Q(x)$ are negative for $x \in (r, s)$ and both are positive for $x < r$ or $x > s$. Show that there is a real number x_0 such that $P(x_0) < Q(x_0)$.

166. Let $P(x)$ be a polynomial of degree n . Knowing that

$$P(k) = \frac{k}{k+1}, \quad k = 0, 1, \dots, n,$$

find $P(m)$ for $m > n$.

167. Consider the polynomials with complex coefficients

$$P(x) = x^n + a_1x^{n-1} + \dots + a_n$$

with zeros x_1, x_2, \dots, x_n and

$$Q(x) = x^n + b_1x^{n-1} + \dots + b_n$$

with zeros $x_1^2, x_2^2, \dots, x_n^2$. Prove that if $a_1 + a_3 + a_5 + \dots$ and $a_2 + a_4 + a_6 + \dots$ are both real numbers, then so is $b_1 + b_2 + \dots + b_n$.

168. Let $P(x)$ be a polynomial with complex coefficients. Prove that $P(x)$ is an even function if and only if there exists a polynomial $Q(x)$ with complex coefficients satisfying

$$P(x) = Q(x)Q(-x).$$

2.2.2 Polynomials in Several Variables

Let us switch to polynomials in several variables. The first example was published by the first author in *Mathematical Reflections*.

Example. Given that the real numbers x, y, z satisfy $x + y + z = 0$ and

$$\frac{x^4}{2x^2 + yz} + \frac{y^4}{2y^2 + xz} + \frac{z^4}{2z^2 + xy} = 1,$$

determine, with proof, all possible values of $x^4 + y^4 + z^4$.

Solution. First note that x, y, z have to be distinct, or else one of the denominators will be zero. We have

$$2x^2 + yz = x^2 + x^2 + yz = x^2 - (y + z)x + yz = (x - y)(x - z).$$

Similarly $2y^2 + xz = (y - z)(y - x)$ and $2z^2 + xy = (z - x)(z - y)$. Hence the second equation from the statement can be written as

$$\frac{x^4}{(x - y)(z - x)} + \frac{y^4}{(x - y)(y - z)} + \frac{z^4}{(z - x)(y - z)} = -1,$$

which gives the following equality

$$x^4(y-z) + y^4(z-x) + z^4(x-y) = -(x-y)(y-z)(z-x).$$

Viewing the left-hand side as a polynomial in x , the zeros of the polynomial are y and z , and its coefficients are divisible by $y-z$. Hence there is a quadratic homogeneous symmetric polynomial $Q(x, y, z)$ such that

$$x^4(y-z) + y^4(z-x) + z^4(x-y) = (x-y)(y-z)(z-x)Q(x, y, z).$$

Write $Q(x, y, z) = \alpha(x^2 + y^2 + z^2) + \beta(xy + xz + yz)$. Equating the coefficients of x^4 on both sides gives $\alpha = -1$. Equating the coefficients of x^3y^2 on both sides gives $0 = -1 - \beta$, hence $\beta = -1$. We conclude that $Q(x, y, z) = -(x^2 + y^2 + z^2 + xy + xz + yz)$. Hence

$$(x-y)(y-z)(z-x)(x^2 + y^2 + z^2 + xy + xz + yz) = -(x-y)(y-z)(z-x).$$

Given that $(x+y+z)^2 = 0$, we have $x^2 + y^2 + z^2 = -2xy - 2xz - 2yz$, and we obtain

$$xy + xz + yz = -1,$$

or

$$x^2 + y^2 + z^2 = 2.$$

Then

$$1 = (xy + xz + yz)^2 = x^2y^2 + x^2z^2 + y^2z^2 + 2xyz(x+y+z) = x^2y^2 + x^2z^2 + y^2z^2,$$

and hence

$$x^4 + y^4 + z^4 = (x^2 + y^2 + z^2)^2 - 2(x^2y^2 + x^2z^2 + y^2z^2) = 4 - 2 = 2.$$

We conclude that the answer to the question is 2. □

We continue with problems left to the reader.

169. Given the polynomial $P(x, y, z)$ prove that the polynomial

$$\begin{aligned} Q(x, y, z) = & P(x, y, z) + P(y, z, x) + P(z, x, y) \\ & - P(x, z, y) - P(y, x, z) - P(z, y, x) \end{aligned}$$

is divisible by $(x-y)(y-z)(z-x)$.

170. Let x, y, z be positive integers greater than 1. Prove that the expression

$$(x+y+z)^3 - (-x+y+z)^3 - (x-y+z)^3 - (x+y-z)^3$$

is the product of seven (not necessarily distinct) integers each of which is greater than one.

171. Factor completely the expression

$$(x + y + z)^5 - (-x + y + z)^5 - (x - y + z)^5 - (x + y - z)^5.$$

172. Factor the expression

$$E = a^3(b - c) + b^3(c - a) + c^3(a - b).$$

173. What conditions should the real numbers a, b, c, d satisfy in order for the equation

$$\frac{(x - b)(x - c)(x - d)}{(a - b)(a - c)(a - d)} + \frac{(x - a)(x - c)(x - d)}{(b - a)(b - c)(b - d)} + \frac{(x - a)(x - b)(x - d)}{(c - a)(c - b)(c - d)} + \frac{(x - a)(x - b)(x - c)}{(d - a)(d - b)(d - c)} = abcd$$

to admit real solutions.

174. Is there a polynomial $P(x, y, z)$ with integer coefficients such that $P(x, y, z)$ and $x + \sqrt[3]{2}y + \sqrt[3]{3}z$ have the same sign for all integers x, y, z ?

175. Let $f(x, y, z) = x^2 + y^2 + z^2 + xyz$. Let $p(x, y, z), q(x, y, z), r(x, y, z)$ be polynomials with real coefficients satisfying

$$f(p(x, y, z), q(x, y, z), r(x, y, z)) = f(x, y, z).$$

Prove or disprove the assertion that the sequence p, q, r consists of some permutation of $\pm x, \pm y, \pm z$ where the number of minus signs is 0 or 2.

176. Find all positive integers p, q , with $p > 2q$, and real numbers a such that the two-variable polynomial

$$x^p + ax^{p-q}y^q + ax^{p-2q}y^{2q} + y^p$$

is divisible by $(x + y)^2$.

177. Find all polynomials of two variables satisfying

$$P(a, b)P(c, d) = P(ac + bd, ad + bc)$$

for all real numbers a, b, c, d .

2.2.3 Quadratic Polynomials

We continue our discussion of polynomials with the case of polynomials of second degree. We start with the following problem due to I. Cucurezeanu, whose solution is based just on the formula for the roots of a quadratic equation.

Example. Let a, b, c be integer numbers that are the sides of a triangle. Show that if the equation

$$x^2 + (a + 1)x + b - c = 0$$

has integer roots, then the triangle is isosceles.

Solution. The quadratic equation has solutions

$$\frac{-(a+1) \pm \sqrt{(a+1)^2 - 4(b-c)}}{2}.$$

For it to admit integer roots, it is necessary that the discriminant is a rational number. But then the discriminant has to be an integer number. If $b > c$ then $(a+1)^2 - 4(b-c)$ is a perfect square $< (a+1)^2$ and of the same parity with this number. Hence

$$(a+1)^2 - 4(b-c) \leq (a-1)^2$$

We conclude that $a+c \leq b$, which contradicts the triangle inequality. The case $b < c$ is similar. So the only possibility is $b = c$; the triangle is isosceles. \square

Here is a problem that uses the sign of a quadratic function. Recall that a quadratic function changes sign only if it has two distinct real zeros, and in that case it has the sign of the dominant coefficient outside of the interval formed by the zeros and opposite sign between the zero. If it has a double zero, or complex zeros, than it always has the sign of the dominant coefficient.

Example. Let a, b, c be distinct real numbers. Show that there is a real number x such that

$$x^2 + 2(a+b+c)x + 3(ab+bc+ac)$$

is negative.

Solution. We compute the discriminant

$$\Delta = 4(a^2 + b^2 + c^2 - ab - bc - ac) = 2[(a-b)^2 + (b-c)^2 + (c-a)^2] > 0.$$

Hence the quadratic function has two distinct real zeros. Between the zeros this function is negative. \square

From the equality

$$(x - x_1)(x - x_2) = x^2 + ax + b,$$

we see that the for the quadratic equation $x^2 + ax + b$ the sum of the roots is $-a$ and the product of the roots is b . This is a particular case of Viète's relations, which will be studied in general in the next section. Here is a problem.

Example. Find all positive integers a, b, c such that the equations

$$x^2 - ax + b = 0, \quad x^2 - bx + c = 0, \quad x^2 - cx + a = 0$$

have integer roots.

Solution. The roots must also be positive. Write $x_1 + x_2 = a = x_5x_6$, $x_3 + x_4 = b = x_1x_2$, $x_5 + x_6 = c = x_3x_4$. Adding we obtain

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = x_1x_2 + x_3x_4 + x_5x_6.$$

This is equivalent to

$$(x_1 - 1)(x_2 - 1) + (x_3 - 1)(x_4 - 1) + (x_5 - 1)(x_6 - 1) = 3.$$

On the left there are only non-negative integers, so they can only be $(0, 0, 3)$, $(0, 1, 2)$, or $(1, 1, 1)$. In the first case, if say the third term is 3 then $\{x_5, x_6\} = \{4, 2\}$, so $a = 8$, $c = 6$. Also one of x_1, x_2 is 1, so the other is $a - 1 = 7$, and thus $b = 7$. We obtain $(a, b, c) = (8, 7, 6)$ and its circular permutations.

If, say, the second term is 1 and the third term is 2, then on the one hand $x_3 = x_4 = 2$, so $b = c = 4$, and on the other hand $\{x_5, x_6\} = \{2, 3\}$ and so $c = 5$, impossible. A similar argument rules out the case where the second term is 1 and the first term is 2.

Finally, if each term is 1, then $x_i = 2$, $i = 1, 2, 3$, and so we obtain the triple $(a, b, c) = (4, 4, 4)$. \square

178. Let $a > 2$ be a real number. Solve the equation

$$x^3 - 2ax^2 + (a^2 + 1)x + 2 - 2a = 0.$$

179. Does there exist a positive integer n such that the quadratic equation

$$(n^3 - n + 1)x^2 - (n^5 - n + 1)x - (n^7 - n + 1) = 0$$

has rational solutions?

180. Assume that the quadratic function $f(x) = x^2 + ax + b$ has integer zeros, and has the property that there is an integer number n such that $f(n) = 13$. Prove that either $f(n + 1)$ or $f(n - 1)$ is equal to 28.

181. Let a, b, c be integer numbers that are the sides of a triangle.

(a) Show that if the equation

$$x^2 + (2ab + 1)x + a^2 + b^2 = c^2$$

has integer roots, then the triangle is right.

(b) Show that if the equation

$$x^2 + (a^2 + b^2 + c^2 + 1)x + ab + bc + ac = 0$$

has integer roots, then the triangle is equilateral.

182. Let $a < b < c < d$ be nonzero real numbers. Show that the equations

$$ax^2 + (b + d)x + c = 0$$

$$bx^2 + (c + d)x + a = 0$$

$$cx^2 + (a + d)x + b = 0$$

have a common root if and only if $a + b + c + d = 0$.

183. Find all real numbers a such that for all $x, y \in \mathbb{R}$ one has

$$2a(x^2 + y^2) + 4axy - y^2 - 2xy - 2x + 1 \geq 0.$$

184. Show that if the equation $x^2 + ax + b = 0$ has real roots, then so does the equation $x^2 - (a^2 - 2b + 2)x + a^2 + b^2 + 1 = 0$.

185. Prove that

$$\log_2 3 + \log_3 4 + \log_4 5 + \log_5 6 > 5.$$

186. Let a, b be integer numbers. Decide when the equation

$$(ax - b)^2 + (bx - a)^2 = x$$

has an integer solution.

187. Prove that if the real numbers p_1, p_2, q_1, q_2 satisfy

$$(q_1 - q_2)^2 + (p_1 - p_2)(p_1q_2 - p_2q_1) < 0,$$

then the quadratic equations

$$x^2 + p_1x + q_1 = 0 \text{ and } x^2 + p_2x + q_2 = 0$$

have real roots and between the roots of one there is a root of the other.

188. Prove that if the inequality $a^2 + ab + ac < 0$ holds, then so does $b^2 - 4ac > 0$.

189. Let a and b be positive integers such that $a^2 + b^2$ is a prime number. Prove that the equation $x^2 + ax + b + 1 = 0$ does not have integer roots.

190. Find all positive integers a, b, c such that the equations

$$x^2 - ax + b = 0, \quad x^2 - bx + c = 0, \quad x^2 - cx + a = 0$$

have integer roots.

191. Let ABC be a triangle. Show that there exists a point D inside the segment BC such that $AD^2 = BD \cdot DC$ if and only if $b + c \leq \sqrt{2}a$.

192. Let $a_1 \leq a_2 \leq \dots \leq a_n$ and $b_1 \geq b_2 \geq \dots \geq b_n$ be real numbers such that

$$\sum_{i=1}^n (n-i)a_i b_i \text{ and } \sum_{j=1}^n (j-1)a_j b_j$$

are both positive. Prove the inequality

$$\left[\left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) - \left(\sum_{i=1}^n a_i b_i \right) \right]^2 \geq 4 \left(\sum_{i=1}^n (n-i)a_i b_i \right) \left(\sum_{j=1}^n (j-1)a_j b_j \right).$$

193. Let $a, a_1, a_2, \dots, a_{2n}, b, b_1, \dots, b_{2n}$ be real numbers such that

$$a^2 > 2 \max (a_1^2 + a_3^2 + \dots + a_{2n-1}^2, a_2^2 + a_4^2 + \dots + a_{2n}^2).$$

Show that $(ab - a_1b_1 - a_2b_2 - \dots - a_{2n}b_{2n})^2$ is greater than or equal to the smaller of the quantities $(a^2 - 2a_1^2 - 2a_3^2 - \dots - 2a_{2n-1}^2)(b^2 - 2b_1^2 - 2b_3^2 - \dots - 2b_{2n-1}^2)$ and $(a^2 - 2a_2^2 - 2a_4^2 - \dots - 2a_{2n}^2)(b^2 - 2b_2^2 - 2b_4^2 - \dots - 2b_{2n}^2)$.

194. A sphere is inscribed in a regular cone. Around the sphere a cylinder is circumscribed so that its base is in the same plane as the base of the cone. Let V_1 be the volume of the cone, and V_2 the volume of the cylinder.

(a) Prove that V_1 cannot equal V_2 .

(b) Find the smallest positive number k such that $V_1 = kV_2$.

2.2.4 Viète's Relations

From the Gauss-d'Alembert fundamental theorem of algebra it follows that a polynomial

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$$

can be factored over the complex numbers as

$$P(x) = a_n(x - x_1)(x - x_2) \cdots (x - x_n).$$

Equating the coefficients of x in the two expressions, we obtain

$$\begin{aligned} x_1 + x_2 + \dots + x_n &= -\frac{a_{n-1}}{a_n}, \\ x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n &= \frac{a_{n-2}}{a_n}, \\ &\dots \\ x_1x_2 \cdots x_n &= (-1)^n \frac{a_0}{a_n}. \end{aligned}$$

These relations carry the name of the French mathematician F. Viète. They combine two ways of looking at a polynomial: as a sum of monomials and as a product of linear factors. As a first application of these relations, we have selected a problem from a 1957 Chinese mathematical competition.

Example. If $x + y + z = 0$, prove that

$$\frac{x^2 + y^2 + z^2}{2} \cdot \frac{x^5 + y^5 + z^5}{5} = \frac{x^7 + y^7 + z^7}{7}.$$

Solution. Consider the polynomial $P(X) = X^3 + pX + q$, whose zeros are x, y, z . Then

$$x^2 + y^2 + z^2 = (x + y + z)^2 - 2(xy + xz + yz) = -2p.$$

Adding the relations $x^3 = -px - q$, $y^3 = -py - q$, and $z^3 = -pz - q$, which hold because x, y, z are zeros of $P(X)$, we obtain

$$x^3 + y^3 + z^3 = -3q.$$

Similarly,

$$x^4 + y^4 + z^4 = -p(x^2 + y^2 + z^2) - q(x + y + z) = 2p^2,$$

and therefore

$$\begin{aligned} x^5 + y^5 + z^5 &= -p(x^3 + y^3 + z^3) - q(x^2 + y^2 + z^2) = 5pq, \\ x^7 + y^7 + z^7 &= -p(x^5 + y^5 + z^5) - q(x^4 + y^4 + z^4) = -5p^2q - 2p^2q = -7p^2q. \end{aligned}$$

The relation from the statement reduces to the obvious

$$\frac{-2p}{2} \cdot \frac{5pq}{5} = \frac{-7p^2q}{7}. \quad \square$$

Viète's relations can be used to solve, or analyze, the roots of a polynomial equation when additional information about the roots is given, as the following problem of B. Enescu shows.

Example. Let $P(x) = x^3 + ax^2 + bx + c$ be a polynomial with rational coefficients, having the roots x_1, x_2, x_3 . Show that if $\frac{x_1}{x_2}$ is a rational number different from 0 and -1 , then x_1, x_2, x_3 are all rational.

Solution. Set $\frac{x_1}{x_2} = t$. Let us observe that if either x_1 or x_2 is rational, so is the other, and by Viète's relations x_3 is rational as well. Also, if x_3 is rational, then $x_1 + x_2 = x_2(1 + \frac{x_1}{x_2})$ is rational, so x_2 is rational, and x_1 is rational as well. Hence it suffices to show that $P(x)$ has a rational root.

Substituting $x_1 = tx_2$ in Viète's relations we obtain

$$\begin{aligned} (t + 1)x_2 + x_3 &= -a \\ x_2[tx_2 + (t + 1)x_3] &= b. \end{aligned}$$

Substituting x_3 from the first equation we obtain the quadratic equation in x_2 ,

$$(t^2 + t + 1)x_2^2 + (t + 1)ax_2 + b = 0.$$

Thus x_2 is a zero of the quadratic polynomial with rational coefficients $Q(x) = (t^2 + t + 1)x^2 + (t + 1)ax + b$. We deduce that the greatest common divisor of $P(x)$ and $Q(x)$ is a non-constant polynomial. Moreover, because both $P(x)$ and $Q(x)$ have rational coefficients their greatest common divisor must have rational coefficients as well. So $P(x)$ can be written as a product of two polynomials with rational coefficients. One of the factors must be a linear polynomial, showing that $P(x)$ has a rational zero. Hence the conclusion. \square

Next, a problem from the short list of the 2005 Ibero-American Mathematical Olympiad.

Example. Find the largest real number k with the property that for all fourth-degree polynomials $P(x) = x^4 + ax^3 + bx^2 + cx + d$ whose zeros are all real and positive, one has

$$(b - a - c)^2 \geq kd,$$

and determine when equality holds.

Solution. Let r_1, r_2, r_3, r_4 be the zeros of $P(x)$. Viète's relations read

$$\begin{aligned} a &= -(r_1 + r_2 + r_3 + r_4), \\ b &= r_1r_2 + r_1r_3 + r_1r_4 + r_2r_3 + r_2r_4 + r_3r_4, \\ c &= -(r_1r_2r_3 + r_1r_2r_4 + r_1r_3r_4 + r_2r_3r_4), \\ d &= r_1r_2r_3r_4. \end{aligned}$$

From here we obtain

$$\begin{aligned} b - a - c &= (r_1r_2 + r_1r_3 + r_1r_4 + r_2r_3 + r_2r_4 + r_3r_4) + (r_1 + r_2 + r_3 + r_4) \\ &\quad + (r_1r_2r_3 + r_1r_2r_4 + r_1r_3r_4 + r_2r_3r_4). \end{aligned}$$

By the AM-GM inequality this is greater than or equal to

$$14 \sqrt[14]{(r_1r_2r_3r_4)^7} = 14\sqrt{d}.$$

Since equality can hold in the AM-GM inequality, we conclude that $k = 196$ is the answer to the problem. Moreover, equality holds exactly when $r_1 = r_2 = r_3 = r_4 = 1$, that is, when $P(x) = x^4 - 4x^3 + 6x^2 - 4x + 1$. \square

And now a challenging problem from A. Krechmar's *Problem Book in Algebra* (Mir Publishers, 1974).

Example. Prove that

$$\sqrt[3]{\cos \frac{2\pi}{7}} + \sqrt[3]{\cos \frac{4\pi}{7}} + \sqrt[3]{\cos \frac{8\pi}{7}} = \sqrt[3]{\frac{1}{2}(5 - 3\sqrt[3]{7})}.$$

Solution. We would like to find a polynomial whose zeros are the three terms on the left. Let us simplify the problem and forget the cube roots for a moment. In this case we have to find a polynomial whose zeros are $\cos \frac{2\pi}{7}$, $\cos \frac{4\pi}{7}$, $\cos \frac{8\pi}{7}$. The seventh roots of unity come in handy. Except for $x = 1$, which we ignore, these are also roots of the equation $x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 = 0$, and are $\cos \frac{2k\pi}{7} + i \sin \frac{2k\pi}{7}$, $k = 1, 2, \dots, 6$. We see that the numbers $2 \cos \frac{2\pi}{7}$, $2 \cos \frac{4\pi}{7}$, and $2 \cos \frac{8\pi}{7}$ are of the form $x + \frac{1}{x}$, with x one of these roots.

If we define $y = x + \frac{1}{x}$, then $x^2 + \frac{1}{x^2} = y^2 - 2$ and $x^3 + \frac{1}{x^3} = y^3 - 3y$. Dividing the equation $x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 = 0$ through by x^3 and substituting y in it, we obtain the cubic equation

$$y^3 + y^2 - 2y - 1 = 0.$$

The numbers $2 \cos \frac{2\pi}{7}$, $2 \cos \frac{4\pi}{7}$, and $2 \cos \frac{8\pi}{7}$ are the three roots of this equation. The simpler task is fulfilled.

But the problem asks us to find the sum of the cube roots of these numbers. Looking at symmetric polynomials, we have

$$X^3 + Y^3 + Z^3 - 3XYZ = (X + Y + Z)^3 - 3(X + Y + Z)(XY + YZ + ZX)$$

and

$$\begin{aligned} X^3Y^3 + Y^3Z^3 + Z^3X^3 - 3(XYZ)^2 &= (XY + YZ + XZ)^3 \\ &\quad - 3XYZ(X + Y + Z)(XY + YZ + ZX). \end{aligned}$$

Because X^3, Y^3, Z^3 are the roots of the equation $y^3 + y^2 - 2y - 1 = 0$, by Viète's relations, $X^3Y^3Z^3 = 1$, so $XYZ = \sqrt[3]{1} = 1$, and also $X^3 + Y^3 + Z^3 = -1$ and $X^3Y^3 + X^3Z^3 + Y^3Z^3 = -2$. In the above two equalities we now know the left-hand sides. The equalities become a system of two equations in the unknowns $u = X + Y + Z$ and $v = XY + YZ + ZX$, namely

$$\begin{aligned} u^3 - 3uv &= -4, \\ v^3 - 3uv &= -5. \end{aligned}$$

Writing the two equations as $u^3 = 3uv - 4$ and $v^3 = 3uv - 5$ and multiplying them, we obtain $(uv)^3 = 9(uv)^2 - 27uv + 20$. With the substitution $m = uv$ this becomes $m^3 = 9m^2 + 27m - 20$ or $(m - 3)^3 + 7 = 0$. This equation has the unique solution $m = 3 - \sqrt[3]{7}$. Hence $u = \sqrt[3]{3m - 4} = \sqrt[3]{5 - 3\sqrt[3]{7}}$. We conclude that

$$\sqrt[3]{\cos \frac{2\pi}{7}} + \sqrt[3]{\cos \frac{4\pi}{7}} + \sqrt[3]{\cos \frac{8\pi}{7}} = X + Y + Z = \frac{1}{\sqrt[3]{2}}u = \sqrt[3]{\frac{1}{2}(5 - 3\sqrt[3]{7})},$$

as desired. □

All problems below can be solved using Viète's relations.

195. Find the zeros of the polynomial

$$P(x) = x^4 - 6x^3 + 18x^2 - 30x + 25$$

knowing that the sum of two of them is 4.

196. Let a, b, c be real numbers. Show that $a \geq 0$, $b \geq 0$, and $c \geq 0$ if and only if $a + b + c \geq 0$, $ab + bc + ca \geq 0$, and $abc \geq 0$.

197. Solve the system

$$\begin{aligned} x + y + z &= 1, \\ xyz &= 1, \end{aligned}$$

knowing that x, y, z are complex numbers of absolute value equal to 1.

- 198.** Let x_1, x_2, x_3 be the roots of the equation

$$x^3 - x^2 - 2x + 4 = 0,$$

with $|x_1| \geq |x_2| \geq |x_3|$. Find a polynomial with integer coefficients of minimal degree that has the root $x_1^5 + x_2^3 + x_3^2$.

- 199.** Find all real numbers r for which there is at least one triple (x, y, z) of nonzero real numbers such that

$$x^2y + y^2z + z^2x = xy^2 + yz^2 + zx^2 = rxyz.$$

- 200.** Let a, b, c, d be real numbers with $a + b + c + d = 0$. Prove that

$$a^3 + b^3 + c^3 + d^3 = 3(abc + bcd + cda + dab)$$

- 201.** Given the real numbers x, y, z, t such that

$$x + y + z + t = x^7 + y^7 + z^7 + t^7 = 0,$$

prove that

$$x(x + y)(x + z)(x + t) = 0.$$

- 202.** For five integers a, b, c, d, e we know that the sums $a + b + c + d + e$ and $a^2 + b^2 + c^2 + d^2 + e^2$ are divisible by an odd number n . Prove that the number $a^5 + b^5 + c^5 + d^5 + e^5 - 5abcde$ is also divisible by n .

- 203.** Find all polynomials whose coefficients are equal either to 1 or -1 and whose zeros are all real.

- 204.** Let $P(z) = az^4 + bz^3 + cz^2 + dz + e = a(z - r_1)(z - r_2)(z - r_3)(z - r_4)$, where a, b, c, d, e are integers, $a \neq 0$. Show that if $r_1 + r_2$ is a rational number, and if $r_1 + r_2 \neq r_3 + r_4$, then r_1r_2 is a rational number.

- 205.** Let $P(x) = x^3 + ax^2 + bx + c$ be a polynomial with rational coefficients, having the roots x_1, x_2, x_3 . Show that if $\frac{x_1}{x_2}$ is a rational number different from 0 and -1 , then x_1, x_2, x_3 are all rational.

- 206.** The zeros of the polynomial $P(x) = x^3 - 10x + 11$ are u, v , and w . Determine the value of $\arctan u + \arctan v + \arctan w$.

- 207.** Prove that for every positive integer n ,

$$\tan \frac{\pi}{2n+1} \tan \frac{2\pi}{2n+1} \cdots \tan \frac{n\pi}{2n+1} = \sqrt{2n+1}.$$

- 208.** Let $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a polynomial of degree $n \geq 3$. Knowing that $a_{n-1} = -\binom{n}{1}$, $a_{n-2} = \binom{n}{2}$, and that all roots are real, find the remaining coefficients.

- 209.** Determine the maximum value of λ such that whenever $P(x) = x^3 + ax^2 + bx + c$ is a cubic polynomial with all zeros real and nonnegative, then

$$P(x) \geq \lambda(x - a)^3$$

for all $x \geq 0$. Find the equality condition.

- 210.** Prove that there are unique positive integers a, n such that

$$a^{n+1} - (a + 1)^n = 2001.$$

2.2.5 The Derivative of a Polynomial

This section adds some elements of real analysis. We remind the reader that the derivative of a polynomial

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

is the polynomial

$$P'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1.$$

We also recall the product rule: $(P(x)Q(x))' = P'(x)Q(x) + P(x)Q'(x)$. If x_1, x_2, \dots, x_n are the zeros of $P(x)$, then by using the product rule we obtain

$$\frac{P'(x)}{P(x)} = \frac{1}{x - x_1} + \frac{1}{x - x_2} + \cdots + \frac{1}{x - x_n}.$$

If a zero of $P(x)$ has multiplicity greater than 1, then it is also a zero of $P'(x)$, and the converse is also true. By Rolle's theorem, if all zeros of $P(x)$ are real, then so are those of $P'(x)$. Let us discuss in detail two problems, the first of which belonging to the second author of the book, and the second to R. Gologan.

Example. Let $P(x)$ be a polynomial with real zeros and let $a < b$ be two real numbers that are smaller than any of the zeros of $P(x)$. Prove that

$$\exp\left(\int_a^b \frac{P'''(x)P(x)}{P'(x)^2} dx\right) < \left|\frac{P(a)^2 P'(b)^3}{P'(a)^3 P(b)^2}\right|.$$

Solution. Differentiate the identity

$$\frac{P'(x)}{P(x)} = \sum_{k=1}^n \frac{1}{x - x_k}$$

to obtain

$$\frac{P''(x)P(x) - P'(x)^2}{P(x)^2} = -\sum_{k=1}^n \frac{1}{(x - x_k)^2}.$$

Differentiate one more time and obtain

$$\frac{(P'''(x)P(x) - P''(x)P'(x))P(x)^2 - (P''(x)P(x) - P'(x)^2)2P(x)P'(x)}{(P(x))^4} = \sum_{k=1}^n \frac{2}{(x - x_k)^3}.$$

Notice that the right-hand side is negative for $a \leq x \leq b < \min(x_1, \dots, x_n)$. Hence

$$P'''(x)P(x)^3 - P''(x)P'(x)P(x)^2 - 2P''(x)P'(x)P(x)^2 + 2P'(x)^3P(x) < 0,$$

that is

$$P'''(x)P(x)^3 - 3P''(x)P'(x)P(x)^2 + 2P'(x)^3P(x) < 0.$$

Dividing by $P(x)^2P'(x)^2$, we obtain

$$\frac{P'''(x)P(x)}{P'(x)^2} < \frac{3P''(x)}{P'(x)} - \frac{2P'(x)}{P(x)}.$$

Integrating we obtain

$$\begin{aligned} \int_a^b \frac{P'''(x)P(x)}{P'(x)^2} dx &< 3 \ln |P'(b)| - \ln |P'(a)| - 2 \ln |P(b)| - \ln |P(a)| \\ &= \ln \left| \frac{P(a)^2 P'(b)^3}{P'(a)^3 P(b)^2} \right|. \end{aligned}$$

After exponentiation we obtain the inequality from the statement. \square

Example. Let $P(x) \in \mathbb{Z}[x]$ be a polynomial with n distinct integer zeros. Prove that the polynomial $(P(x))^2 + 1$ has a factor of degree at least $2 \lfloor \frac{n+1}{2} \rfloor$ that is irreducible over $\mathbb{Z}[x]$.

Solution. The statement apparently offers no clue about derivatives. The standard approach is to assume that

$$(P(x))^2 + 1 = P_1(x)P_2(x) \cdots P_k(x)$$

is a decomposition into factors that are irreducible over $\mathbb{Z}[x]$. Letting x_1, x_2, \dots, x_n be the integer zeros of $P(x)$, we find that

$$P_1(x_j)P_2(x_j) \cdots P_k(x_j) = 1, \text{ for } j = 1, 2, \dots, n.$$

Hence $P_i(x_j) = \pm 1$, which then implies $\frac{1}{P_i(x_j)} = P_i(x_j), i = 1, 2, \dots, k, j = 1, 2, \dots, n$.

Now let us see how derivatives come into play. The key observation is that the zeros x_j of $(P(x))^2$ appear with multiplicity greater than 1, and so they are zeros of the derivative. Differentiating with the product rule, we obtain

$$\sum_{i=1}^k P_1(x_j) \cdots P'_i(x_j) \cdots P_k(x_j) = 0, \text{ for } j = 1, 2, \dots, n.$$

This sum can be simplified by taking into account that $P_1(x_j)P_2(x_j)\cdots P_k(x_j) = 1$ and $\frac{1}{P_i(x_j)} = P_i(x_j)$ as

$$\sum_{i=1}^k P'_i(x_j)P_i(x_j) = 0, \text{ for } j = 1, 2, \dots, n.$$

It follows that x_j is a zero of the polynomial

$$\sum_{i=1}^k 2P'_i(x)P_i(x) = \left(\sum_{i=1}^k P_i^2(x) \right)'.$$

Let us remember that $P_i(x_j) = \pm 1$, which then implies $\sum_{i=1}^k P_i^2(x_j) - n = 0$ for $j = 1, 2, \dots, n$. The numbers x_j , $j = 1, 2, \dots, n$, are zeros of both $\sum_{i=1}^k P_i^2(x) - n$ and its derivative, so they are zeros of order at least 2 of this polynomial. Therefore,

$$\sum_{i=1}^k P_i^2(x) = (x - x_1)^2(x - x_2)^2 \cdots (x - x_n)^2 Q(x) + n,$$

for some polynomial $Q(x)$ with integer coefficients. We deduce that there exists an index i_0 such that the degree of $P_{i_0}(x)$ is greater than or equal to n . For n even, $n = 2 \lfloor \frac{n+1}{2} \rfloor$, and we are done. For n odd, since $(P(x))^2 + 1$ does not have real zeros, neither does $P_{i_0}(x)$, so this polynomial has even degree. Thus the degree of $P_{i_0}(x)$ is at least $n + 1 = 2 \lfloor \frac{n+1}{2} \rfloor$. This completes the solution. \square

- 211.** Find all polynomials $P(x)$ with integer coefficients satisfying $P(P'(x)) = P'(P(x))$ for all $x \in \mathbb{R}$.
- 212.** Determine all polynomials $P(x)$ with real coefficients satisfying $(P(x))^n = P(x^n)$ for all $x \in \mathbb{R}$, where $n > 1$ is a fixed integer.
- 213.** Let $P(x)$ and $Q(x)$ be polynomials with complex coefficients and let a be a nonzero complex number. Prove that if

$$P(x)^3 = Q(x)^2 + a,$$

for all $x \in \mathbb{C}$, then $P(x)$ and $Q(x)$ are constant polynomials.

- 214.** Let $P(z)$ and $Q(z)$ be polynomials with complex coefficients of degree greater than or equal to 1 with the property that $P(z) = 0$ if and only if $Q(z) = 0$ and $P(z) = 1$ if and only if $Q(z) = 1$. Prove that the polynomials are equal.
- 215.** Let $P(x)$ be a polynomial with all zeros real and distinct and such that none of its zeros is equal to 0. Prove that the polynomial $x^2 P''(x) + 3x P'(x) + P(x)$ also has all roots real and distinct.

- 216.** Let $P(x)$ be a polynomial of degree 5, with real coefficients, all of whose zeros are real. Prove that for each real number a that is not a zero of $P(x)$ or $P'(x)$, there is a real number b such that

$$b^2 P(a) + 4b P'(a) + 5 P''(a) = 0.$$

- 217.** Let $P_n(x) = (x^n - 1)(x^{n-1} - 1) \cdots (x - 1)$, $n \geq 1$. Prove that for $n \geq 2$, $P'_n(x)$ is divisible by $P_{\lfloor n/2 \rfloor}$ in the ring of polynomials with integer coefficients.
- 218.** The zeros of the n th-degree polynomial $P(x)$ are all real and distinct. Prove that the zeros of the polynomial $G(x) = nP(x)P''(x) - (n-1)(P'(x))^2$ are all complex.
- 219.** Let $P(x)$ be a polynomial of degree $n > 3$ whose zeros $x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n$ are real. Prove that

$$P' \left(\frac{x_1 + x_2}{2} \right) \cdot P' \left(\frac{x_{n-1} + x_n}{2} \right) \neq 0.$$

- 220.** A polynomial $P(x)$ with real coefficients is called a mirror polynomial if $|P(a)| = |P(-a)|$ for all real numbers a . Let $F(x)$ be a polynomial with real coefficients, and consider polynomials with real coefficients $P(x)$ and $Q(x)$ such that $P(x) - P'(x) = F(x)$ and $Q(x) + Q'(x) = F(x)$. Prove that $P(x) + Q(x)$ is a mirror polynomial if and only if $F(x)$ is a mirror polynomial.

2.2.6 The Location of the Zeros of a Polynomial

Since not all polynomial equations can be solved by radicals, methods of approximation are necessary. Results that allow you to localize the roots in certain regions of the real axis or complex plane are therefore useful.

The qualitative study of the position of the zeros of a polynomial has far-reaching applications. For example, the solutions of a homogeneous ordinary linear differential equation with constant coefficients are stable (under errors of measuring the coefficients) if and only if the roots of the characteristic equation lie in the open left half-plane (i.e., have negative real part). Stability is, in fact, an essential question in control theory, where one is usually interested in whether the zeros of a particular polynomial lie in the open left half-plane (Hurwitz stability) or in the open unit disk (Schur stability). Here is a famous result.

Lucas' theorem. *The zeros of the derivative $P'(z)$ of a polynomial $P(z)$ lie in the convex hull of the zeros of $P(z)$.*

Proof. Because any convex domain can be obtained as the intersection of half-planes, it suffices to show that if the zeros of $P(z)$ lie in an open half-plane, then the zeros of $P'(z)$ lie in that half-plane as well. Moreover, by rotating and translating the variable z we can further reduce the problem to the case in which the zeros of $P(z)$ lie in the upper half-plane $\text{Im } z > 0$. Here $\text{Im } z$ denotes the imaginary part.

So let z_1, z_2, \dots, z_n be the (not necessarily distinct) zeros of $P(z)$, which by hypothesis have positive imaginary part. If $\operatorname{Im} w \leq 0$, then $\operatorname{Im} \frac{1}{w - z_k} > 0$, for $k = 1, \dots, n$, and therefore

$$\operatorname{Im} \frac{P'(w)}{P(w)} = \sum_{k=1}^n \operatorname{Im} \frac{1}{w - z_k} > 0.$$

This shows that w is not a zero of $P'(z)$ and so all zeros of $P'(z)$ lie in the upper half-plane. The theorem is proved. \square

221. Let a_1, a_2, \dots, a_n be positive real numbers. Prove that the polynomial

$$P(x) = x^n - a_1 x^{n-1} - a_2 x^{n-2} - \dots - a_n$$

has a unique positive zero.

222. Prove that the zeros of the polynomial

$$P(z) = z^7 + 7z^4 + 4z + 1$$

lie inside the disk of radius 2 centered at the origin.

223. Prove that if the complex coefficients p, q of the quadratic equation $x^2 + px + q = 0$ satisfy $|p| + |q| < 1$, then the roots of this equation lie in the interior of the unit disk.

224. Let $P(x)$ be a polynomial with integer coefficients all of whose roots are real and lie in the interval $(0, 3)$. Prove that the roots of this polynomial lie in the set

$$\left\{ 1, 2, \frac{3 - \sqrt{5}}{2}, \frac{3 + \sqrt{5}}{2} \right\}.$$

225. For $a \neq 0$ a real number and $n > 2$ an integer, prove that every nonreal root z of the polynomial equation $x^n + ax + 1 = 0$ satisfies the inequality $|z| \geq \sqrt[n]{\frac{1}{n-1}}$.

226. Let $a \in \mathbb{C}$ and $n \geq 2$. Prove that the polynomial equation $ax^n + x + 1 = 0$ has a root of absolute value less than or equal to 2.

227. Let $P(z)$ be a polynomial of degree n , all of whose zeros have absolute value 1 in the complex plane. Set $g(z) = \frac{P(z)}{z^{n/2}}$. Show that all roots of the equation $g'(z) = 0$ have absolute value 1.

228. The polynomial $x^4 - 2x^2 + ax + b$ has four distinct real zeros. Show that the absolute value of each zero is smaller than $\sqrt{3}$.

229. Let $P_n(z)$, $n \geq 1$, be a sequence of monic k th-degree polynomials whose coefficients converge to the coefficients of a monic k th-degree polynomial $P(z)$. Prove that for any $\varepsilon > 0$ there is n_0 such that if $n \geq n_0$ then $|z_i(n) - z_i| < \varepsilon$, $i = 1, 2, \dots, k$, where $z_i(n)$ are the zeros of $P_n(z)$ and z_i are the zeros of $P(z)$, taken in the appropriate order.

- 230.** Let $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ be a polynomial with complex coefficients, with $a_0 \neq 0$, and with the property that there exists an m such that

$$\left| \frac{a_m}{a_0} \right| > \binom{n}{m}.$$

Prove that $P(x)$ has a zero of absolute value less than 1.

- 231.** For a polynomial $P(x) = (x - x_1)(x - x_2) \cdots (x - x_n)$, with distinct real zeros $x_1 < x_2 < \cdots < x_n$, we set $\delta(P(x)) = \min_i (x_{i+1} - x_i)$. Prove that for any real number k ,

$$\delta(P'(x) - kP(x)) > \delta(P(x)),$$

where $P'(x)$ is the derivative of $P(x)$. In particular, $\delta(P'(x)) > \delta(P(x))$.

2.2.7 Irreducible Polynomials

A polynomial is irreducible if it cannot be written as a product of two polynomials in a nontrivial manner. The question of irreducibility depends on the ring of coefficients. When the coefficients are complex numbers, only linear polynomials are irreducible. For real numbers some quadratic polynomials are irreducible as well. Both these cases are rather dull. The interesting situations occur when the coefficients are rational or integer, in which case there is an interplay between polynomials and arithmetic. The cases of rational and integer coefficients are more or less equivalent, with minor differences such as the fact that $2x + 2$ is irreducible over $\mathbb{Q}[x]$ but reducible over $\mathbb{Z}[x]$. For matters of elegance we focus on polynomials with integer coefficients. We will assume implicitly from now on that for any polynomial with integer coefficients, the greatest common divisor of its coefficients is 1.

Definition. A polynomial $P(x) \in \mathbb{Z}[x]$ is called irreducible over $\mathbb{Z}[x]$ if there do not exist polynomials $Q(x), R(x) \in \mathbb{Z}[x]$ different from ± 1 such that $P(x) = Q(x)R(x)$. Otherwise, $P(x)$ is called reducible.

We commence with an easy problem.

Example. Let $P(x)$ be an n th-degree polynomial with integer coefficients with the property that $|P(x)|$ is a prime number for $2n + 1$ distinct integer values of the variable x . Prove that $P(x)$ is irreducible over $\mathbb{Z}[x]$.

Solution. Assume the contrary and let $P(x) = Q(x)R(x)$ with $Q(x), R(x) \in \mathbb{Z}[x]$, $Q(x), R(x) \neq \pm 1$. Let $k = \deg(Q(x))$. Then $Q(x) = 1$ at most k times and $Q(x) = -1$ at most $n - k$ times. Also, $R(x) = 1$ at most $n - k$ times and $R(x) = -1$ at most k times. Consequently, the product $|Q(x)R(x)|$ is composite except for at most $k + (n - k) + (n - k) + k = 2n$ values of x . This contradicts the hypothesis. Hence $P(x)$ is irreducible. \square

The bound is sharp. For example, $P(x) = (x + 1)(x + 5)$ has $|P(-2)| = |P(-4)| = 3$, $P(0) = 5$, and $|P(-6)| = 7$.

Probably the most beautiful criterion of irreducibility of polynomials is that discovered independently by F.G.M. Eisenstein in 1850 and T. Schönemann in 1846. We present it below.

Eisenstein-Schönemann theorem. *Given a polynomial $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ with integer coefficients, suppose that there exists a prime number p such that a_n is not divisible by p , a_k is divisible by p for $k = 0, 1, \dots, n-1$, and a_0 is not divisible by p^2 . Then $P(x)$ is irreducible over $\mathbb{Z}[x]$.*

Proof. We argue by contradiction. Suppose that $P(x) = Q(x)R(x)$, with $Q(x)$ and $R(x)$ not identically equal to ± 1 . Let

$$\begin{aligned} Q(x) &= b_k x^k + b_{k-1} x^{k-1} + \cdots + b_0, \\ R(x) &= c_{n-k} x^{n-k} + c_{n-k-1} x^{n-k-1} + \cdots + c_0. \end{aligned}$$

Let us look closely at the equalities

$$\sum_{j=0}^i b_j c_{i-j} = a_i, \quad i = 0, 1, \dots, n,$$

obtained by identifying the coefficients in the equality $P(x) = Q(x)R(x)$. From the first of them, $b_0 c_0 = a_0$, and because a_0 is divisible by p but not by p^2 it follows that exactly one of b_0 and c_0 is divisible by p . Assume that b_0 is divisible by p and take the next equality $b_0 c_1 + b_1 c_0 = a_1$. The right-hand side is divisible by p , and the first term on the left is also divisible by p . Hence $b_1 c_0$ is divisible by p , and since c_0 is not, b_1 must be divisible by p .

This reasoning can be repeated to prove that all the b_i 's are divisible by p . It is important that both $Q(x)$ and $R(x)$ have degrees greater than or equal to 1, for the fact that b_k is divisible by p follows from

$$b_k c_0 + b_{k-1} c_1 + \cdots = a_k,$$

where a_k is divisible by p for $k < n$. The contradiction arises in the equality $a_n = b_k c_{n-k}$, since the right-hand side is divisible by p , while the left-hand side is not. This proves the theorem.

The first three problems listed below use this result, while the others apply similar ideas.

232. Prove that the polynomial

$$P(x) = x^{101} + 101x^{100} + 102$$

is irreducible over $\mathbb{Z}[x]$.

233. Prove that for every prime number p , the polynomial

$$P(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$$

is irreducible over $\mathbb{Z}[x]$.

234. Prove that for every positive integer n , the polynomial $P(x) = x^{2^n} + 1$ is irreducible over $\mathbb{Z}[x]$.

235. Prove that for any distinct integers a_1, a_2, \dots, a_n the polynomial

$$P(x) = (x - a_1)(x - a_2) \cdots (x - a_n) - 1$$

cannot be written as a product of two nonconstant polynomials with integer coefficients.

236. Prove that for any distinct integers a_1, a_2, \dots, a_n the polynomial

$$P(x) = (x - a_1)^2(x - a_2)^2 \cdots (x - a_n)^2 + 1$$

cannot be written as a product of two nonconstant polynomials with integer coefficients.

237. Associate to a prime the polynomial whose coefficients are the decimal digits of the prime (for example, for the prime 7043 the polynomial is $P(z) = 7z^3 + 4z + 3$). Prove that this polynomial is always irreducible over $\mathbb{Z}[x]$.

238. Let p be a prime number of the form $4k + 3$, k an integer. Prove that for any positive integer n , the polynomial $(x^2 + 1)^n + p$ is irreducible in the ring $\mathbb{Z}[x]$.

239. Let p be a prime number. Prove that the polynomial

$$P(x) = x^{p-1} + 2x^{p-2} + 3x^{p-3} + \cdots + (p-1)x + p$$

is irreducible in $\mathbb{Z}[x]$.

240. Let $P(x)$ be a monic polynomial in $\mathbb{Z}[x]$, irreducible over this ring, and such that $|P(0)|$ is not the square of an integer. Prove that the polynomial $Q(x)$ defined by $Q(x) = P(x^2)$ is also irreducible over $\mathbb{Z}[x]$.

2.2.8 Chebyshev Polynomials

The n th Chebyshev polynomial $T_n(x)$ expresses $\cos n\theta$ as a polynomial in $\cos \theta$. This means that $T_n(x) = \cos(n \arccos x)$, for $n \geq 0$. These polynomials satisfy the recurrence

$$T_0(x) = 1, T_1(x) = x, T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \text{ for } n \geq 1.$$

For example, $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, $T_4(x) = 8x^4 - 8x^2 + 1$.

One usually calls these the Chebyshev polynomials of the first kind, to distinguish them from the Chebyshev polynomials of the second kind $U_n(x)$ defined by

$$U_0(x) = 1, U_1(x) = 2x, U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \text{ for } n \geq 1$$

(same recurrence relation but different initial condition). Alternatively, $U_n(x)$ can be defined by the equality

$$U_n(\cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}.$$

Chebyshev's theorem. For fixed $n \geq 1$, the polynomial $2^{-n+1}T_n(x)$ is the unique monic n th-degree polynomial satisfying

$$\max_{-1 \leq x \leq 1} |2^{-n+1}T_n(x)| \leq \max_{-1 \leq x \leq 1} |P(x)|,$$

for any other monic n th-degree polynomial $P(x)$.

One says that among all monic n th-degree polynomials, $2^{-n+1}T_n(x)$ has the smallest variation away from zero on $[-1, 1]$. This variation is $\frac{1}{2^{n-1}}$. Let us see how Chebyshev's theorem applies to a problem from *Challenging Mathematical Problems with Elementary Solutions* by A.M. Yaglom and I.M. Yaglom.

Example. Let A_1, A_2, \dots, A_n be points in the plane. Prove that on any segment of length l there is a point M such that

$$MA_1 \cdot MA_2 \cdots MA_n \geq 2 \left(\frac{l}{4} \right)^n.$$

Solution. Rescaling, we can assume that $l = 2$. Associate complex coordinates to points in such a way that the segment coincides with the interval $[-1, 1]$. Then

$$MA_1 \cdot MA_2 \cdots MA_n = |z - z_1| \cdot |z - z_2| \cdots |z - z_n| = |P(z)|,$$

where $P(z)$ is a monic polynomial with complex coefficients, and $z \in [-1, 1]$. Write $P(z) = R(z) + iQ(z)$, where $R(z)$ is the real part and $Q(z)$ is the imaginary part of the polynomial. Since z is real, we have $|P(z)| \geq |R(z)|$. The polynomial $R(z)$ is monic, so on the interval $[-1, 1]$ it varies away from zero at least as much as the Chebyshev polynomial. Thus we can find z in this interval such that $|R(z)| \geq \frac{1}{2^{n-1}}$. This implies $|P(z)| \geq 2 \cdot \frac{1}{2^n}$, and rescaling back we deduce the existence in the general case of a point M satisfying the inequality from the statement. \square

Stepping aside from the classical picture, let us also consider the families of polynomials $\mathcal{T}_n(x)$ and $\mathcal{U}_n(x)$ defined by $\mathcal{T}_0(x) = 2$, $\mathcal{T}_1(x) = x$, $\mathcal{T}_{n+1}(x) = x\mathcal{T}_n(x) - \mathcal{T}_{n-1}(x)$, and $\mathcal{U}_0(x) = 1$, $\mathcal{U}_1(x) = x$, $\mathcal{U}_{n+1}(x) = x\mathcal{U}_n(x) - \mathcal{U}_{n-1}(x)$. These polynomials are determined by the equalities

$$\mathcal{T}_n\left(z + \frac{1}{z}\right) = z^n + \frac{1}{z^n} \quad \text{and} \quad \mathcal{U}_n\left(z + \frac{1}{z}\right) = \left(z^{n+1} - \frac{1}{z^{n+1}}\right) / \left(z - \frac{1}{z}\right).$$

Also, $\mathcal{T}_n(x) = \frac{1}{2}\mathcal{T}_n(2x)$ and $\mathcal{U}_n(x) = \mathcal{U}_n(2x)$. Here is a quickie that uses $\mathcal{T}_n(x)$.

Example. Let a be a real number such that $a + a^{-1}$ is an integer. Prove that for any $n \geq 1$, the number $a^n + a^{-n}$ is an integer.

Solution. An inductive argument based on the recurrence relation shows that $\mathcal{T}_n(x)$ is a polynomial with integer coefficients. And since $a^n + a^{-n} = \mathcal{T}_n(a + a^{-1})$, it follows that this number is an integer. \square

241. Prove that for $n \geq 1$,

$$\begin{aligned} \mathcal{T}_{n+1}(x) &= x\mathcal{T}_n(x) - (1 - x^2)\mathcal{U}_{n-1}(x), \\ \mathcal{U}_n(x) &= x\mathcal{U}_{n-1}(x) + \mathcal{T}_n(x). \end{aligned}$$

242. Compute the $n \times n$ determinants

$$\begin{vmatrix} x & 1 & 0 & 0 & \dots & 0 \\ 1 & 2x & 1 & 0 & \dots & 0 \\ 0 & 1 & 2x & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 2x \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} 2x & 1 & 0 & 0 & \dots & 0 \\ 1 & 2x & 1 & 0 & \dots & 0 \\ 0 & 1 & 2x & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 2x \end{vmatrix}.$$

243. Prove Chebyshev's theorem for $n = 4$: namely, show that for any monic fourth-degree polynomial $P(x)$,

$$\max_{-1 \leq x \leq 1} |P(x)| \geq \max_{-1 \leq x \leq 1} |2^{-3}T_4(x)|,$$

with equality if and only if $P(x) = 2^{-3}T_4(x)$.

244. Let r be a positive real number such that $\sqrt[6]{r} + \frac{1}{\sqrt[6]{r}} = 6$. Find the maximum value of $\sqrt[4]{r} - \frac{1}{\sqrt[4]{r}}$.

245. Let $\alpha = \frac{2\pi}{n}$. Prove that the matrix

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \cos \alpha & \cos 2\alpha & \dots & \cos n\alpha \\ \cos 2\alpha & \cos 4\alpha & \dots & \cos 2n\alpha \\ \vdots & \vdots & \ddots & \vdots \\ \cos(n-1)\alpha & \cos 2(n-1)\alpha & \dots & \cos(n-1)n\alpha \end{pmatrix}$$

is invertible.

246. Find all quintuples (x, y, z, v, w) with $x, y, z, v, w \in [-2, 2]$ satisfying the system of equations

$$\begin{aligned} x + y + z + v + w &= 0, \\ x^3 + y^3 + z^3 + v^3 + w^3 &= 0, \\ x^5 + y^5 + z^5 + v^5 + w^5 &= -10. \end{aligned}$$

247. Let $x_1, x_2, \dots, x_n, n \geq 2$, be distinct real numbers in the interval $[-1, 1]$. Prove that

$$\frac{1}{t_1} + \frac{1}{t_2} + \dots + \frac{1}{t_n} \geq 2^{n-2},$$

where $t_k = \prod_{j \neq k} |x_j - x_k|, k = 1, 2, \dots, n$.

248. Let $n \geq 3$ be an odd integer. Evaluate

$$\sum_{k=1}^{\frac{n-1}{2}} \sec \frac{2k\pi}{n}.$$

249. For $n \geq 1$, prove the following identities:

$$\frac{T_n(x)}{\sqrt{1-x^2}} = \frac{(-1)^n}{1 \cdot 3 \cdot 5 \cdots (2n-1)} \frac{d^n}{dx^n} (1-x^2)^{n-\frac{1}{2}},$$

$$U_n(x) \sqrt{1-x^2} = \frac{(-1)^n (n+1)}{1 \cdot 3 \cdot 5 \cdots (2n+1)} \frac{d^n}{dx^n} (1-x^2)^{n+\frac{1}{2}}.$$

2.3 Linear Algebra

2.3.1 Operations with Matrices

An $m \times n$ matrix is an array with m rows and n columns. The standard notation is $A = (a_{ij})_{i,j}$, where a_{ij} is the entry (element) in the i th row and j th column. We denote by \mathcal{I}_n the $n \times n$ identity matrix (for which $a_{ij} = 1$ if $i = j$, and 0 otherwise) and by \mathcal{O}_n the $n \times n$ zero matrix (for which $a_{ij} = 0$ for all i, j).

Given the matrix $A = (a_{ij})_{i,j}$, A^t denotes the transpose of A , in which the i, j entry is a_{ji} , and \bar{A} denotes the complex conjugate, whose entries are the complex conjugates of the entries of A . Also, $\text{tr } A$ is the trace of A , namely the sum of the elements on the main diagonal: $a_{11} + a_{22} + \cdots + a_{nn}$.

We illustrate how matrix multiplication can be used to prove an identity satisfied by the Fibonacci sequence ($F_0 = 0$, $F_1 = 1$, $F_{n+1} = F_n + F_{n-1}$, $n \geq 1$). The identity we have in mind has already been discussed in the introductory chapter in the solution to Problem 27; we put it here in a new perspective.

Example. Prove that

$$F_{m+n+1} = F_{m+1}F_{n+1} + F_mF_n, \text{ for } m, n \geq 0.$$

Solution. Consider the matrix

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

An easy induction shows that for $n \geq 1$,

$$M^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

The equality $M^{m+n} = M^m M^n$ written in explicit form is

$$\begin{pmatrix} F_{m+n+1} & F_{m+n} \\ F_{m+n} & F_{m+n-1} \end{pmatrix} = \begin{pmatrix} F_{m+1} & F_m \\ F_m & F_{m-1} \end{pmatrix} \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

We obtain the identity by setting the upper left corners of both sides equal. □

Here are some problems for the reader.

250. Let M be an $n \times n$ complex matrix. Prove that there exist Hermitian matrices A and B such that $M = A + iB$. (A matrix X is called Hermitian if $\bar{X}^t = X$).

- 251.** Do there exist $n \times n$ matrices A and B such that $AB - BA = \mathcal{I}_n$?
- 252.** Let A and B be 2×2 matrices with real entries satisfying $(AB - BA)^n = \mathcal{I}_2$ for some positive integer n . Prove that n is even and $(AB - BA)^4 = \mathcal{I}_2$.
- 253.** Let A and B be two $n \times n$ matrices that do not commute and for which there exist nonzero real numbers p, q, r such that $pAB + qBA = \mathcal{I}_n$ and $A^2 = rB^2$. Prove that $p = q$.
- 254.** Let a, b, c, d be real numbers such that $c \neq 0$ and $ad - bc = 1$. Prove that there exist u and v such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & -v \\ 0 & 1 \end{pmatrix}.$$

- 255.** Compute the n th power of the $m \times m$ matrix

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}, \lambda \in \mathbb{C}.$$

- 256.** Let A and B be $n \times n$ matrices with real entries satisfying

$$\operatorname{tr}(AA^t + BB^t) = \operatorname{tr}(AB + A^t B^t).$$

Prove that $A = B^t$.

2.3.2 Determinants

The determinant of an $n \times n$ matrix $A = (a_{ij})_{i,j}$, denoted by $\det A$ or $|a_{ij}|$, is the volume taken with sign of the n -dimensional parallelepiped determined by the row (or column) vectors of A . Formally, the determinant can be introduced as follows. Let $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 1)$ be the canonical basis of \mathbb{R}^n . The exterior algebra of \mathbb{R}^n is the vector space spanned by products of the form $e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_k}$, where the multiplication \wedge is distributive with respect to sums and is subject to the noncommutativity rule $e_i \wedge e_j = -e_j \wedge e_i$ for all i, j (which then implies $e_i \wedge e_i = 0$, for all i). If the row vectors of the matrix A are r_1, r_2, \dots, r_n , then the determinant is defined by the equality

$$r_1 \wedge r_2 \wedge \dots \wedge r_n = (\det A) e_1 \wedge e_2 \wedge \dots \wedge e_n.$$

The explicit formula is

$$\det A = \sum_{\sigma} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

with the sum taken over all permutations σ of $\{1, 2, \dots, n\}$.

To compute the determinant of a matrix, one applies repeatedly the row operation that adds to one row a multiple of another until the matrix either becomes diagonal or has a row of zeros. In the first case this transforms the parallelepiped determined by the row vectors into a right parallelepiped in standard position without changing its volume, as suggested in Figure 13.

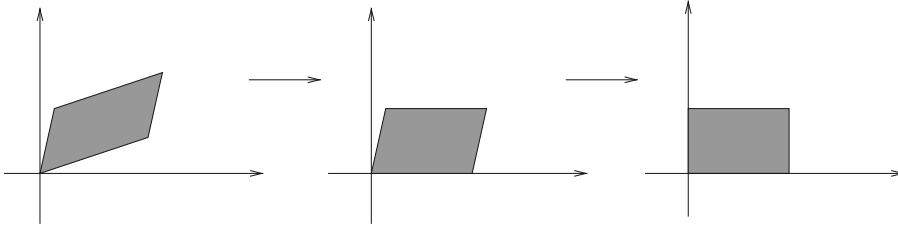


Figure 13

But it is not our purpose to teach the basics. We insist only on nonstandard tricks and methods. A famous example is the computation of the Vandermonde determinant.

Example. Let x_1, x_2, \dots, x_n be arbitrary numbers ($n \geq 1$). Compute the determinant

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}.$$

Solution. The key idea is to view x_n as a variable and think of the determinant as an $(n-1)$ st-degree polynomial in x_n . The leading coefficient is itself a Vandermonde determinant of order $n-1$, while the $n-1$ roots are obviously x_2, x_3, \dots, x_{n-1} . The determinant is therefore equal to

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \end{vmatrix} (x_n - x_1)(x_n - x_2) \cdots (x_n - x_{n-1}).$$

Now we can induct on n to prove that the Vandermonde determinant is equal to

$$\prod_{i>j} (x_i - x_j).$$

This determinant is equal to zero if and only if two of the x_i 's are equal. □

We continue with a problem of D. Andrica.

Example. (a) Consider the real numbers $a_{ij}, i = 1, 2, \dots, n-2, j = 1, 2, \dots, n, n \geq 3$, and

the determinants

$$A_k = \begin{vmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ a_{11} & \dots & a_{1,k-1} & a_{1,k+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n-2,1} & \dots & a_{n-2,k-1} & a_{n-2,k+1} & \dots & a_{n-2,n} \end{vmatrix}, \quad k \geq 1.$$

Prove that

$$A_1 + A_3 + A_5 + \dots = A_2 + A_4 + A_6 + \dots$$

(b) Define

$$p_k = \prod_{i=0}^{n-(k+1)} (x_{n-i} - x_k), \quad q_k = \prod_{i=1}^{k-1} (x_k - x_i),$$

where $x_i, i = 1, 2, \dots, n$, are some distinct real numbers. Prove that

$$\sum_{k=1}^n \frac{(-1)^k}{p_k q_k} = 0.$$

(c) Prove that for any positive integer $n \geq 3$ the following identity holds:

$$\sum_{k=1}^n \frac{(-1)^k k^2}{(n-k)!(n+k)!} = 0.$$

Solution. We have

$$\begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \\ a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-2,1} & a_{n-2,2} & \dots & a_{n-2,n-1} & a_{n-2,n} \end{vmatrix} = 0.$$

Expanding by the first row, we obtain

$$A_1 - A_2 + A_3 - A_4 + \dots = 0.$$

This implies

$$A_1 + A_3 + A_5 + \dots = A_2 + A_4 + A_6 + \dots,$$

and (a) is proved.

For (b), we substitute $a_{ij} = x_i^j, i = 1, 2, \dots, n-2, j = 1, 2, \dots, n$. Then

$$A_k = \begin{vmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_{k-1} & x_{k+1} & \dots & x_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-2} & \dots & x_{k-1}^{n-2} & x_{k+1}^{n-2} & \dots & x_n^{n-2} \end{vmatrix},$$

which is a Vandermonde determinant. Its value is equal to

$$\prod_{\substack{i>j \\ i,j \neq k}} (x_j - x_i) = \frac{\prod_{i>j} (x_j - x_i)}{p_k q_k}.$$

The equality proved in (a) becomes, in this particular case,

$$\sum_{k=1}^n \frac{(-1)^k}{p_k q_k} = 0,$$

as desired.

Finally, if in this we let $x_k = k^2$, then we obtain the identity from part (c), and the problem is solved. \square

And here comes a set of problems for the reader.

257. Prove that

$$\begin{vmatrix} (x^2 + 1)^2 & (xy + 1)^2 & (xz + 1)^2 \\ (xy + 1)^2 & (y^2 + 1)^2 & (yz + 1)^2 \\ (xz + 1)^2 & (yz + 1)^2 & (z^2 + 1)^2 \end{vmatrix} = 2(y - z)^2(z - x)^2(x - y)^2.$$

258. Let $(F_n)_n$ be the Fibonacci sequence. Using determinants, prove the identity

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n, \text{ for all } n \geq 1.$$

259. Let $p < m$ be two positive integers. Prove that

$$\begin{vmatrix} \binom{m}{0} & \binom{m}{1} & \cdots & \binom{m}{p} \\ \binom{m+1}{0} & \binom{m+1}{1} & \cdots & \binom{m+1}{p} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{m+p}{0} & \binom{m+p}{1} & \cdots & \binom{m+p}{p} \end{vmatrix} = 1.$$

260. Given distinct integers x_1, x_2, \dots, x_n , prove that $\prod_{i>j} (x_i - x_j)$ is divisible by $1!2! \cdots (n-1)!$.

261. Find all numbers in the interval $[-2015, 2015]$ that can be equal to the determinant of an 11×11 matrix with entries equal to 1 or -1 .

262. Prove the formula for the determinant of a circulant matrix

$$\begin{vmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_3 & x_4 & x_5 & \cdots & x_2 \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{vmatrix} = (-1)^{n-1} \prod_{j=0}^{n-1} \left(\sum_{k=1}^n \zeta^{jk} x_k \right),$$

where $\zeta = e^{2\pi i/n}$.

263. Let a and b be integers such that $a + b = 2014$. Prove that the determinant

$$\begin{vmatrix} a^3 & b^3 & 3ab & -1 \\ -1 & a^2 & b^2 & 2ab \\ 2b & -1 & a^2 & -b^2 \\ 0 & b & -1 & a \end{vmatrix}$$

is a multiple of 61.

264. Compute the determinant of the $n \times n$ matrix $A = (a_{ij})_{ij}$, where

$$a_{ij} = \begin{cases} (-1)^{|i-j|} & \text{if } i \neq j, \\ 2 & \text{if } i = j. \end{cases}$$

265. Prove that for any integers x_1, x_2, \dots, x_n and positive integers k_1, k_2, \dots, k_n , the determinant

$$\begin{vmatrix} x_1^{k_1} & x_2^{k_1} & \dots & x_n^{k_1} \\ x_1^{k_2} & x_2^{k_2} & \dots & x_n^{k_2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{k_n} & x_2^{k_n} & \dots & x_n^{k_n} \end{vmatrix}$$

is divisible by $n!$.

266. Let A and B be 3×3 matrices with real elements such that

$$\det A = \det B = \det(A + B) = \det(A - B) = 0.$$

Prove that $\det(xA + yB) = 0$ for any real numbers x and y .

Sometimes it is more convenient to work with blocks instead of entries. For that we recall the rule of Laplace, which is the direct generalization of the row or column expansion. The determinant is computed by expanding over all $k \times k$ minors of some k rows or columns. Explicitly, given $A = (a_{ij})_{i,j=1}^n$, when expanding by the rows i_1, i_2, \dots, i_k , the determinant is given by

$$\det A = \sum_{j_1 < j_2 < \dots < j_k} (-1)^{i_1 + \dots + i_k + j_1 + \dots + j_k} M_k N_k,$$

where M_k is the determinant of the $k \times k$ matrix whose entries are a_{ij} , with $i \in \{i_1, i_2, \dots, i_k\}$ and $j \in \{j_1, j_2, \dots, j_k\}$, while N_k is the determinant of the $(n - k) \times (n - k)$ matrix whose entries are a_{ij} with $i \notin \{i_1, i_2, \dots, i_k\}$ and $j \notin \{j_1, j_2, \dots, j_k\}$. We exemplify this rule with a problem from the 4th International Competition in Mathematics for University Students (1997).

Example. Let M be an invertible $2n \times 2n$ matrix, represented in block form as

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{and} \quad M^{-1} = \begin{pmatrix} E & F \\ G & H \end{pmatrix}.$$

Show that $\det M \cdot \det H = \det A$.

Solution. The idea of the solution is that the relation between determinants should come from a relation between matrices. To this end, we would like to find three matrices X, Y, Z such that $XY = Z$, while $\det X = \det M$, $\det Y = \det H$, and $\det Z = \det A$. Since among M, H , and A , the matrix M has the largest dimension, we might try to set $X = M$ and find $2n \times 2n$ matrices Y and Z . The equality $M \cdot M^{-1} = \mathcal{I}_{2n}$ yields two relations involving H , namely $AF + BH = 0$ and $CF + DH = \mathcal{I}_n$. This suggests that we should use both F and H in the definition of Y . So we need an equality of the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} * & F \\ * & H \end{pmatrix} = \begin{pmatrix} * & 0 \\ * & \mathcal{I}_n \end{pmatrix}.$$

We can try

$$Y = \begin{pmatrix} \mathcal{I}_n & F \\ 0 & H \end{pmatrix}.$$

The latter has determinant equal to $\det H$, as desired. Also,

$$Z = \begin{pmatrix} A & 0 \\ C & \mathcal{I}_n \end{pmatrix}.$$

According to the rule of Laplace, the determinant of Z can be computed by expanding along the $n \times n$ minors from the top n rows, and all of them are zero except for the first. Hence $\det Z = \det A \cdot \det \mathcal{I}_n = \det A$, and so the matrices X, Y, Z solve the problem. \square

267. Show that if

$$x = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{and} \quad x' = \begin{vmatrix} a' & b' \\ c' & d' \end{vmatrix},$$

then

$$(xx')^2 = \begin{vmatrix} ab' & cb' & ba' & da' \\ ad' & cd' & bc' & dc' \\ bb' & db' & aa' & ca' \\ bd' & dd' & ac' & cc' \end{vmatrix}.$$

268. Let A, B, C, D be $n \times n$ matrices such that $AC = CA$. Prove that

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB).$$

269. Let X and Y be $n \times n$ matrices. Prove that

$$\det(\mathcal{I}_n - XY) = \det(\mathcal{I}_n - YX).$$

A property exploited often in Romanian mathematics competitions states that for any $n \times n$ matrix A with real entries,

$$\det(\mathcal{I}_n + A^2) \geq 0.$$

The proof is straightforward:

$$\begin{aligned}\det(\mathcal{I}_n + A^2) &= \det((\mathcal{I}_n + iA)(\mathcal{I}_n - iA)) = \det(\mathcal{I}_n + iA) \det(\mathcal{I}_n - iA) \\ &= \det(\mathcal{I}_n + iA) \det(\overline{\mathcal{I}_n + iA}) = \det(\mathcal{I}_n + iA) \overline{\det(\mathcal{I}_n + iA)}.\end{aligned}$$

In this computation the bar denotes the complex conjugate, and the last equality follows from the fact that the determinant is a polynomial in the entries. The final expression is nonnegative, being equal to $|\det(\mathcal{I}_n + iA)|^2$.

Use this property to solve the following problems, while assuming that all matrices have real entries.

270. Let A and B be $n \times n$ matrices that commute. Prove that if $\det(A + B) = 0$, then $\det(A^k + B^k) \geq 0$ for all $k \geq 1$.

271. Let A be an $n \times n$ matrix such that $A + A^t = \mathcal{O}_n$. Prove that

$$\det(\mathcal{I}_n + \lambda A^2) \geq 0,$$

for all $\lambda \in \mathbb{R}$.

272. Let $P(t)$ be a polynomial of even degree with real coefficients. Prove that the function $f(X) = P(X)$ defined on the set of $n \times n$ matrices is not onto.

273. Let n be an odd positive integer and A an $n \times n$ matrix with the property that $A^2 = \mathcal{O}_n$ or $A^2 = \mathcal{I}_n$. Prove that $\det(A + \mathcal{I}_n) \geq \det(A - \mathcal{I}_n)$.

2.3.3 The Inverse of a Matrix

An $n \times n$ matrix A is called invertible if there exists an $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = \mathcal{I}_n$. The inverse of a matrix can be found either by using the adjoint matrix, which amounts to computing several determinants, or by performing row and column operations. We illustrate how the latter method can be applied to a problem from the first International Competition in Mathematics for University Students (1994).

Example. (a) Let A be an $n \times n$ symmetric invertible matrix with positive real entries, $n \geq 2$. Show that A^{-1} has at most $n^2 - 2n$ entries equal to zero.

(b) How many entries are equal to zero in the inverse of the $n \times n$ matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & 2 & \dots & 2 \\ 1 & 2 & 1 & 1 & \dots & 1 \\ 1 & 2 & 1 & 2 & \dots & 2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 1 & 2 & \dots & 1 \end{pmatrix}?$$

Solution. Denote by a_{ij} and b_{ij} the entries of A , respectively, A^{-1} . Then we have $\sum_{i=0}^n a_{mi} b_{im} = 1$, so for fixed m not all the b_{im} 's are equal to zero. For $k \neq m$ we have $\sum_{i=0}^n a_{ki} b_{im} = 0$, and from the positivity of the a_{ki} 's we conclude that at least one b_{im} is negative, and at least one is positive. Hence every column of A^{-1} contains at least two nonzero elements. This proves part (a).

To compute the inverse of the matrix in part (b), we consider the extended matrix $(A\mathcal{I}_n)$, and using row operations we transform it into the matrix $(\mathcal{I}_n A^{-1})$. We start with

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 2 & 2 & 2 & \dots & 2 & 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 1 & 1 & \dots & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & 2 & \dots & 2 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 1 & 2 & \dots & \dots & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Subtracting the first row from each of the others, then the second row from the first, we obtain

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & 1 & \dots & 1 & -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & -1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & \dots & 1 & -1 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & 1 & \dots & \dots & -1 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

We continue as follows. First, we subtract the second row from the third, fourth, and so on. Then we add the third row to the second. Finally, we multiply all rows, beginning with the third, by -1 . This way we obtain

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & -1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 1 & \dots & 1 & 0 & 1 & -1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 & 1 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & 0 & \dots & \dots & 1 & 0 & 0 & 0 & \dots & -1 \end{pmatrix}.$$

Now the inductive pattern is clear. At each step we subtract the k th row from the rows below, then subtract the $(k+1)$ st from the k th, and finally multiply all rows starting with the $(k+1)$ st by -1 . In the end we find that the entries of A^{-1} are $b_{1,1} = 2$, $b_{n,n} = (-1)^n$, $b_{i,i+1} = b_{i+1,i} = (-1)^i$, and $b_{ij} = 0$, for $|i-j| \geq 2$. This example shows that equality can hold in part (a). \square

274. For distinct numbers x_1, x_2, \dots, x_n , consider the matrix

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{pmatrix}.$$

It is known that $\det A$ is the Vandermonde determinant

$$\Delta(x_1, x_2, \dots, x_n) = \prod_{i>j} (x_i - x_j).$$

Prove that the inverse of A is $B = (b_{km})_{1 \leq k, m \leq n}$, where

$$b_{km} = (-1)^{k+m} \Delta(x_1, x_2, \dots, x_n)^{-1} \Delta(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \\ \times S_{n-1}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n).$$

Here S_{n-1} denotes the $(n-1)$ st symmetric polynomial in $n-1$ variables.

- 275.** Let A and B be 2×2 matrices with integer entries such that $A, A+B, A+2B, A+3B$, and $A+4B$ are all invertible matrices whose inverses have integer entries. Prove that $A+5B$ is invertible and that its inverse has integer entries.
- 276.** Determine the matrix A knowing that its adjoint matrix (the one used in the computation of the inverse) is

$$A^* = \begin{pmatrix} m^2 - 1 & 1 - m & 1 - m \\ 1 - m & m^2 - 1 & 1 - m \\ 1 - m & 1 - m & m^2 - 1 \end{pmatrix}, \quad m \neq 1, -2.$$

- 277.** Let $A = (a_{ij})_{ij}$ be an $n \times n$ matrix such that $\sum_{j=1}^n |a_{ij}| < 1$ for each i . Prove that $\mathcal{I}_n - A$ is invertible.
- 278.** Let $\alpha = \frac{\pi}{n+1}$, $n > 2$. Prove that the $n \times n$ matrix

$$\begin{pmatrix} \sin \alpha & \sin 2\alpha & \dots & \sin n\alpha \\ \sin 2\alpha & \sin 4\alpha & \dots & \sin 2n\alpha \\ \vdots & \vdots & \ddots & \vdots \\ \sin n\alpha & \sin 2n\alpha & \dots & \sin n^2\alpha \end{pmatrix}$$

is invertible.

- 279.** Assume that A and B are invertible complex $n \times n$ matrices such that $i(A^\dagger B - B^\dagger A)$ is positive semidefinite, where $X^\dagger = \overline{X}^t$, the transpose conjugate of X . Prove that $A + iB$ is invertible. (A matrix T is positive semidefinite if $\langle Tv, v \rangle \geq 0$ for all vectors v , where $\langle v, w \rangle = v^t \overline{w}$ the complex inner product.)

We continue with problems that exploit the ring structure of the set of $n \times n$ matrices. There are some special properties of matrices that do not hold in arbitrary rings. For example, an $n \times n$ matrix A is either a zero divisor (there exist nonzero matrices B and C such that $AB = CA = \mathcal{O}_n$), or it is invertible. Also, if a matrix has a left (or right) inverse, then the matrix is invertible, which means that if $AB = \mathcal{I}_n$ then also $BA = \mathcal{I}_n$.

A good example is a problem of I.V. Maftai that appeared in the 1982 Romanian Mathematical Olympiad.

Example. Let A, B, C be $n \times n$ matrices, $n \geq 1$, satisfying

$$ABC + AB + BC + AC + A + B + C = \mathcal{O}_n.$$

Prove that A and $B + C$ commute if and only if A and BC commute.

Solution. If we add \mathcal{I}_n to the left-hand side of the identity from the statement, we recognize this expression to be the polynomial $P(X) = (X + A)(X + B)(X + C)$ evaluated at the identity matrix. This means that

$$(\mathcal{I}_n + A)(\mathcal{I}_n + B)(\mathcal{I}_n + C) = \mathcal{I}_n.$$

This shows that $\mathcal{I}_n + A$ is invertible, and its inverse is $(\mathcal{I}_n + B)(\mathcal{I}_n + C)$. It follows that

$$(\mathcal{I}_n + B)(\mathcal{I}_n + C)(\mathcal{I}_n + A) = \mathcal{I}_n,$$

or

$$BCA + BC + BA + CA + A + B + C = \mathcal{O}_n.$$

Subtracting this relation from the one in the statement and grouping the terms appropriately, we obtain

$$ABC - BCA = (B + C)A - A(B + C).$$

The conclusion follows. □

Here are other examples.

280. Let A be an $n \times n$ matrix such that there exists a positive integer k for which

$$kA^{k+1} = (k + 1)A^k.$$

Prove that the matrix $A - \mathcal{I}_n$ is invertible and find its inverse.

281. Let A be an invertible $n \times n$ matrix, and let $B = XY$, where X and Y are $1 \times n$, respectively, $n \times 1$ matrices. Prove that the matrix $A + B$ is invertible if and only if $\alpha = YA^{-1}X \neq -1$, and in this case its inverse is given by

$$(A + B)^{-1} = A^{-1} - \frac{1}{\alpha + 1}A^{-1}BA^{-1}.$$

282. Given two $n \times n$ matrices A and B for which there exist nonzero real numbers a and b such that $AB = aA + bB$, prove that A and B commute.

283. Let A and B be $n \times n$ matrices, $n \geq 1$, satisfying $AB - B^2A^2 = \mathcal{I}_n$ and $A^3 + B^3 = \mathcal{O}_n$. Prove that $BA - A^2B^2 = \mathcal{I}_n$.

2.3.4 Systems of Linear Equations

A system of m linear equations with n unknowns can be written as

$$Ax = b,$$

where A is an $m \times n$ matrix called the coefficient matrix, and b is an m -dimensional vector. If $m = n$, the system has a unique solution if and only if the coefficient matrix A is invertible. If A is not invertible, the system can have either infinitely many solutions or none at all. If additionally $b = 0$, then the system does have infinitely many solutions and the codimension of the space of solutions is equal to the rank of A .

We illustrate this section with two problems that apparently have nothing to do with the topic. The first was published in *Mathematics Gazette, Bucharest*, by L. Pîrșan.

Example. Consider the matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad C = \begin{pmatrix} a\alpha & b\alpha & a\gamma & b\gamma \\ a\beta & b\beta & a\delta & b\delta \\ c\alpha & d\alpha & c\gamma & d\gamma \\ c\beta & d\beta & c\delta & d\delta \end{pmatrix},$$

where $a, b, c, d, \alpha, \beta, \gamma, \delta$ are real numbers. Prove that if A and B are invertible, then C is invertible as well.

Solution. Let us consider the matrix equation $AXB = D$, where

$$X = \begin{pmatrix} x & z \\ y & t \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} m & n \\ p & q \end{pmatrix}.$$

Solving it for X gives $X = A^{-1}DB^{-1}$, and so X is uniquely determined by A , B , and D . Multiplying out the matrices in this equation,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & z \\ y & t \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} m & n \\ p & q \end{pmatrix},$$

we obtain

$$\begin{pmatrix} a\alpha x + b\alpha y + a\gamma z + b\gamma t & a\beta x + b\beta y + a\delta z + b\delta t \\ c\alpha x + d\alpha y + c\gamma z + d\gamma t & c\beta x + d\beta y + c\delta z + d\delta t \end{pmatrix} = \begin{pmatrix} m & n \\ p & q \end{pmatrix}.$$

This is a system in the unknowns x, y, z, t :

$$\begin{aligned} a\alpha x + b\alpha y + a\gamma z + b\gamma t &= m, \\ a\beta x + b\beta y + a\delta z + b\delta t &= n, \\ c\alpha x + d\alpha y + c\gamma z + d\gamma t &= p, \\ c\beta x + d\beta y + c\delta z + d\delta t &= q. \end{aligned}$$

We saw above that this system has a unique solution, which implies that its coefficient matrix is invertible. This coefficient matrix is C . \square

The second problem we found in an old textbook on differential and integral calculus.

Example. Given the distinct real numbers a_1, a_2, a_3 , let x_1, x_2, x_3 be the three roots of the equation

$$\frac{u_1}{a_1 + t} + \frac{u_2}{a_2 + t} + \frac{u_3}{a_3 + t} = 1,$$

where u_1, u_2, u_3 are real parameters. Prove that u_1, u_2, u_3 are smooth functions of x_1, x_2, x_3 and that

$$\det \left(\frac{\partial u_i}{\partial x_j} \right) = - \frac{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}{(a_1 - a_2)(a_2 - a_3)(a_3 - a_1)}.$$

Solution. After eliminating the denominators, the equation from the statement becomes a cubic equation in t , so x_1, x_2, x_3 are well defined. The parameters u_1, u_2, u_3 satisfy the system of equations

$$\begin{aligned} \frac{1}{a_1 + x_1} u_1 + \frac{1}{a_2 + x_1} u_2 + \frac{1}{a_3 + x_1} u_3 &= 1, \\ \frac{1}{a_1 + x_2} u_1 + \frac{1}{a_2 + x_2} u_2 + \frac{1}{a_3 + x_2} u_3 &= 1, \\ \frac{1}{a_1 + x_3} u_1 + \frac{1}{a_2 + x_3} u_2 + \frac{1}{a_3 + x_3} u_3 &= 1. \end{aligned}$$

When solving this system, we might end up entangled in algebraic computations. Thus it is better instead to take a look at the two-variable situation. Solving the system

$$\begin{aligned} \frac{1}{a_1 + x_1} u_1 + \frac{1}{a_2 + x_1} u_2 &= 1, \\ \frac{1}{a_1 + x_2} u_1 + \frac{1}{a_2 + x_2} u_2 &= 1, \end{aligned}$$

with Cramer's rule we obtain

$$u_1 = \frac{(a_1 + x_1)(a_1 + x_2)}{(a_1 - a_2)} \quad \text{and} \quad u_2 = \frac{(a_2 + x_1)(a_2 + x_2)}{(a_2 - a_1)}.$$

Now we can extrapolate to the three-dimensional situation and guess that

$$u_i = \frac{\prod_{k=1}^3 (a_i + x_k)}{\prod_{k \neq i} (a_i - a_k)}, \quad i = 1, 2, 3.$$

It is not hard to check that these satisfy the system of equations. Observe that

$$\frac{\partial u_i}{\partial x_j} = \frac{\prod_{k \neq j} (a_i + x_k)}{\prod_{j \neq i} (a_i - a_j)}, \quad \text{and so} \quad \frac{\partial u_i}{\partial x_j} = \frac{1}{a_i + x_j} u_i, \quad i, j = 1, 2, 3.$$

The determinant in question looks again difficult to compute. Some tricks simplify the task. An observation is that the sum of the columns is 1. Indeed, these sums are

$$\frac{\partial u_1}{\partial x_i} + \frac{\partial u_2}{\partial x_i} + \frac{\partial u_3}{\partial x_i}, \quad i = 1, 2, 3,$$

which we should recognize as the left-hand sides of the linear system. So the determinant becomes much simpler if we add the first and second rows to the last. Another observation is that the determinant is a 3-variable polynomial in x_1, x_2, x_3 . Its total degree is 3, and it becomes zero if $x_i = x_j$ for some $i \neq j$. Consequently, the determinant is a number not depending on x_1, x_2, x_3 times $(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)$. This number can be determined by looking just at the coefficient of $x_2^2 x_3$. And an easy computation shows that this coefficient is equal to $\frac{1}{(a_1 - a_2)(a_2 - a_3)(a_3 - a_1)}$. \square

From the very many practical applications of the theory of systems of linear equations, let us mention the Global Positioning System (GPS). The principle behind the GPS is the measurement of the distances between the receiver and 24 satellites (in practice some of these satellites might have to be ignored in order to avoid errors due to atmospheric phenomena). This yields 24 quadratic equations $d(P, S_i)^2 = r_i^2$, $i = 1, 2, \dots, 24$, in the three spatial coordinates of the receiver. Subtracting the first of the equations from the others cancels the quadratic terms and gives rise to an overdetermined system of 23 linear equations in three unknowns. Determining the location of the receiver is therefore a linear algebra problem.

284. Solve the system of linear equations

$$\begin{aligned} x_1 + x_2 + x_3 &= 0, \\ x_2 + x_3 + x_4 &= 0, \\ &\dots \\ x_{99} + x_{100} + x_1 &= 0, \\ x_{100} + x_1 + x_2 &= 0. \end{aligned}$$

285. Find the solutions x_1, x_2, x_3, x_4, x_5 to the system of equations

$$\begin{aligned} x_5 + x_2 &= yx_1, & x_1 + x_3 &= yx_2, & x_2 + x_4 &= yx_3, \\ x_3 + x_5 &= yx_1, & x_4 + x_1 &= yx_5, \end{aligned}$$

where y is a parameter.

286. Let a, b, c, d be positive numbers different from 1, and x, y, z, t real numbers satisfying $a^x = bcd, b^y = cda, c^z = dab, d^t = abc$. Prove that

$$\begin{vmatrix} -x & 1 & 1 & 1 \\ 1 & -y & 1 & 1 \\ 1 & 1 & -z & 1 \\ 1 & 1 & 1 & -t \end{vmatrix} = 0.$$

287. Given the system of linear equations

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0, \\a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= 0,\end{aligned}$$

whose coefficients satisfy the conditions

- (i) a_{11}, a_{22}, a_{33} are positive,
- (ii) all other coefficients are negative,
- (iii) in each equation, the sum of the coefficients is positive,

prove that the system has the unique solution $x_1 = x_2 = x_3 = 0$.

288. Let $P(x) = x^n + x^{n-1} + \cdots + x + 1$. Find the remainder obtained when $P(x^{n+1})$ is divided by $P(x)$.

289. Find all functions $f : \mathbb{R} \setminus \{-1, 1\} \rightarrow \mathbb{R}$ satisfying

$$f\left(\frac{x-3}{x+1}\right) + f\left(\frac{3+x}{1-x}\right) = x \text{ for all } x \neq \pm 1.$$

290. Find all positive integer solutions (x, y, z, t) to the Diophantine equation

$$(x+y)(y+z)(z+x) = txyz$$

such that $\gcd(x, y) = \gcd(y, z) = \gcd(z, x) = 1$.

291. We have n coins of unknown masses and a balance. We are allowed to place some of the coins on one side of the balance and an equal number of coins on the other side. After thus distributing the coins, the balance gives a comparison of the total mass of each side, either by indicating that the two masses are equal or by indicating that a particular side is the more massive of the two. Show that at least $n - 1$ such comparisons are required to determine whether all of the coins are of equal mass.

292. Let $a_0 = 0, a_1, \dots, a_n, a_{n+1} = 0$ be a sequence of real numbers that satisfy

$$|a_{k-1} - 2a_k + a_{k+1}| \leq 1 \text{ for } k = 1, 2, \dots, n-1.$$

Prove that

$$|a_k| \leq \frac{k(n-k+1)}{2} \text{ for } k = 1, 2, \dots, n-1.$$

293. Prove that the Hilbert matrix

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{pmatrix}$$

is invertible. Prove also that the sum of the entries of the inverse matrix is n^2 .

2.3.5 Vector Spaces, Linear Combinations of Vectors, Bases

In general, a vector space V over a field of scalars (which in our book will be only \mathbb{C} , \mathbb{R} , or \mathbb{Q}) is a set endowed with a commutative addition and a scalar multiplication that have the same properties as those for vectors in Euclidean space.

A linear combination of the vectors v_1, v_2, \dots, v_m is a sum $c_1v_1 + c_2v_2 + \dots + c_mv_m$ with scalar coefficients. The vectors are called linearly independent if a combination of these vectors is equal to zero only when all coefficients are zero. Otherwise, the vectors are called linearly dependent. If v_1, v_2, \dots, v_n are linearly independent and if every vector in V is a linear combination of these vectors, then v_1, v_2, \dots, v_n is called a basis of V . The number of elements of a basis of a vector space depends only on the vector space, and is called the dimension of the vector space. We will be concerned only with finite-dimensional vector spaces. We also point out that if in a vector space there are given more vectors than the dimension, then these vectors must be linearly dependent.

The rank of a matrix is the dimension of its row vectors, which is the same as the dimension of the column vectors. A square matrix is invertible if and only if its rank equals its size.

Let us see some examples. The first appeared in the Soviet University Student Mathematical Competition in 1977.

Example. Let X and B_0 be $n \times n$ matrices, $n \geq 1$. Define $B_i = B_{i-1}X - XB_{i-1}$, for $i \geq 1$. Prove that if $X = B_{n^2}$, then $X = \mathcal{O}_n$.

Solution. Because the space of $n \times n$ matrices is n^2 -dimensional, B_0, B_1, \dots, B_{n^2} must be linearly dependent, so there exist scalars c_0, c_1, \dots, c_{n^2} such that

$$c_0B_0 + c_1B_1 + \dots + c_{n^2}B_{n^2} = \mathcal{O}_n.$$

Let k be the smallest index for which $c_k \neq 0$. Then

$$B_k = a_1B_{k+1} + a_2B_{k+2} + \dots + a_{n^2-k}B_{n^2},$$

where $a_j = -\frac{c_{k+j}}{c_k}$. Computing $B_{k+1} = B_kX - XB_k$, we obtain

$$B_{k+1} = a_1B_{k+2} + a_2B_{k+3} + \dots + a_{n^2-k}B_{n^2+1},$$

and inductively

$$B_{k+j} = a_1B_{k+j+1} + a_2B_{k+j+2} + \dots + a_{n^2-k}B_{n^2+j}, \text{ for } j \geq 1.$$

In particular,

$$B_{n^2} = a_1B_{n^2+1} + a_2B_{n^2+2} + \dots + a_{n^2-k}B_{n^2+k}.$$

But $B_{n^2+1} = B_{n^2}X - XB_{n^2} = X^2 - X^2 = \mathcal{O}_n$, and hence $B_{n^2+j} = \mathcal{O}_n$, for $j \geq 1$. It follows that X , which is a linear combination of $B_{n^2+1}, B_{n^2+2}, \dots, B_{n^2+k}$ is the zero matrix. And we are done. \square

The second example was given at the 67th W.L. Putnam Mathematical Competition in 2006, and the solution that we present was posted by C. Zara on the Internet.

Example. Let Z denote the set of points in \mathbb{R}^n whose coordinates are 0 or 1. (Thus Z has 2^n elements, which are the vertices of a unit hypercube in \mathbb{R}^n .) Let k be given, $0 \leq k \leq n$. Find the maximum, over all vector subspaces $V \subseteq \mathbb{R}^n$ of dimension k , of the number of points in $Z \cap V$.

Solution. Let us consider the matrix whose rows are the elements of $V \cap Z$. By construction it has row rank at most k . It thus also has column rank at most k ; in particular, there are k columns such that any other column is a linear combination of these k . It means that the coordinates of each point of $V \cap Z$ are determined by the k coordinates that lie in these k columns. Since each such coordinate can have only two values, $V \cap Z$ can have at most 2^k elements.

This upper bound is reached for the vectors that have all possible choices of 0 and 1 for the first k entries, and 0 for the remaining entries. \square

294. Prove that every odd polynomial function of degree equal to $2m - 1$ can be written as

$$P(x) = c_1 \binom{x}{1} + c_2 \binom{x+1}{3} + c_3 \binom{x+2}{5} + \dots + c_m \binom{x+m-1}{2m-1},$$

$$\text{where } \binom{x}{m} = \frac{x(x-1) \cdots (x-m+1)}{m!}.$$

295. Let n be a positive integer and $P(x)$ an n th-degree polynomial with complex coefficients such that $P(0), P(1), \dots, P(n)$ are all integers. Prove that the polynomial $n!P(x)$ has integer coefficients.

296. Let A be the $n \times n$ matrix whose i, j entry is $i + j$ for all $i, j = 1, 2, \dots, n$. What is the rank of A ?

297. For integers $n \geq 2$ and $0 \leq k \leq n - 2$, compute the determinant

$$\begin{vmatrix} 1^k & 2^k & 3^k & \dots & n^k \\ 2^k & 3^k & 4^k & \dots & (n+1)^k \\ 3^k & 4^k & 5^k & \dots & (n+2)^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n^k & (n+1)^k & (n+2)^k & \dots & (2n-1)^k \end{vmatrix}.$$

298. Let V be a vector space and let f, f_1, f_2, \dots, f_n be linear maps from V to \mathbb{R} . Suppose that $f(x) = 0$ whenever $f_1(x) = f_2(x) = \dots = f_n(x) = 0$. Prove that f is a linear combination of f_1, f_2, \dots, f_n .

299. Given a set S of $2n - 1$ different irrational numbers, $n \geq 1$, prove that there exist n distinct elements $x_1, x_2, \dots, x_n \in S$ such that for all nonnegative rational numbers a_1, a_2, \dots, a_n with $a_1 + a_2 + \dots + a_n > 0$, the number $a_1x_1 + a_2x_2 + \dots + a_nx_n$ is irrational.

300. There are given $2n + 1$ real numbers, $n \geq 1$, with the property that whenever one of them is removed, the remaining $2n$ can be split into two sets of n elements that have the same sum of elements. Prove that all the numbers are equal.

- 301.** Let V be an infinite set of vectors in \mathbb{R}^n containing n linearly independent vectors. A finite subset $S \subset V$ is called crucial if the set $V \setminus S$ contains no n linearly independent vectors, but every set $V \setminus T$, with T a subset of S does. Prove there are only finitely many crucial subsets of V .

2.3.6 Linear Transformations, Eigenvalues, Eigenvectors

A linear transformation between vector spaces is a map $T : V \rightarrow W$ that satisfies $T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2)$ for any scalars α_1, α_2 and vectors v_1, v_2 . A matrix A defines a linear transformation by $v \rightarrow Av$, and any linear transformation between finite-dimensional vector spaces with specified bases is of this form. An eigenvalue of a matrix A is a zero of the characteristic polynomial $P_A(\lambda) = \det(\lambda \mathcal{I}_n - A)$. Alternatively, it is a scalar λ for which the equation $Av = \lambda v$ has a nontrivial solution v . In this case v is called an eigenvector of the eigenvalue λ . If $\lambda_1, \lambda_2, \dots, \lambda_m$ are distinct eigenvalues and v_1, v_2, \dots, v_m are corresponding eigenvectors, then v_1, v_2, \dots, v_m are linearly independent. Moreover, if the matrix A is Hermitian, meaning that A is equal to its transpose conjugate, then v_1, v_2, \dots, v_m may be chosen to be pairwise orthogonal.

The set of eigenvalues of a matrix is called its spectrum. The reason for this name is that in quantum mechanics observable quantities are modelled by matrices. Physical spectra, such as the emission spectrum of the hydrogen atom, become spectra of matrices. Among all results in spectral theory we stopped at the spectral mapping theorem, mainly because we want to bring to your attention the method used in the proof.

The spectral mapping theorem. *Let A be an $n \times n$ matrix with not necessarily distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, and let $P(x)$ be a polynomial. Then the eigenvalues of the matrix $P(A)$ are $P(\lambda_1), P(\lambda_2), \dots, P(\lambda_n)$.*

Proof. To prove this result we will apply a widely used idea (see for example the splitting principle in algebraic topology). We will first assume that the eigenvalues of A are all distinct. Then A can be diagonalized by eigenvectors as

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

and in the basis formed by the eigenvectors of A , the matrix $P(A)$ assumes the form

$$\begin{pmatrix} P(\lambda_1) & 0 & \dots & 0 \\ 0 & P(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P(\lambda_n) \end{pmatrix}.$$

The conclusion is now straightforward. In general, the characteristic polynomial of a matrix depends continuously on the entries. Problem 229 in Section 2.2.6 proved that the roots of

a polynomial depend continuously on the coefficients. Hence the eigenvalues of a matrix depend continuously on the entries.

The set of matrices with distinct eigenvalues is dense in the set of all matrices. To prove this claim we need the notion of the discriminant of a polynomial. By definition, if the zeros of a polynomial are x_1, x_2, \dots, x_n , the discriminant is $\prod_{i < j} (x_i - x_j)^2$. It is equal to zero if and only if the polynomial has multiple zeros. Being a symmetric polynomial in the x_i 's, the discriminant is a polynomial in the coefficients. Therefore, the condition that the eigenvalues of a matrix be not all distinct can be expressed as a polynomial equation in the entries. By slightly varying the entries, we can violate this condition. Therefore, arbitrarily close to any matrix there are matrices with distinct eigenvalues.

The conclusion of the spectral mapping theorem for an arbitrary matrix now follows by a limiting argument. \square

We continue with two more elementary examples.

Example. Let $A : V \rightarrow W$ and $B : W \rightarrow V$ be linear maps between finite-dimensional vector spaces. Prove that the linear maps AB and BA have the same set of nonzero eigenvalues, counted with multiplicities.

Solution. Choose a basis that identifies V with \mathbb{R}^m and W with \mathbb{R}^n . Associate to A and B their matrices, denoted by the same letters. The problem is solved if we prove the equality

$$\det(\lambda \mathcal{I}_n - AB) = \lambda^k \det(\lambda \mathcal{I}_m - BA),$$

where k is of course $n - m$. The relation being symmetric, we may assume that $n \geq m$. In this case, complete the two matrices with zeros to obtain two $n \times n$ matrices A' and B' . Because $\det(\lambda \mathcal{I}_n - A'B') = \det(\lambda \mathcal{I}_n - AB)$ and $\det(\lambda \mathcal{I}_n - B'A') = \lambda^{n-m} \det(\lambda \mathcal{I}_n - BA)$, the problem reduces to proving that $\det(\lambda \mathcal{I}_n - A'B') = \det(\lambda \mathcal{I}_n - B'A')$. And this is true for arbitrary $n \times n$ matrices A' and B' . For a proof of this fact we refer the reader to problem 269 in Section 2.3.2. \square

If $B = A^\dagger$, the transpose conjugate of A , then this example shows that AA^\dagger and $A^\dagger A$ have the same nonzero eigenvalues. The square roots of these eigenvalues are called the singular values of A . The second example comes from the first International Mathematics Competition (for university students), 1994.

Example. Let α be a nonzero real number and n a positive integer. Suppose that F and G are linear maps from \mathbb{R}^n into \mathbb{R}^n satisfying $F \circ G - G \circ F = \alpha F$.

- (a) Show that for all $k \geq 1$ one has $F^k \circ G - G \circ F^k = \alpha k F^k$.
- (b) Show that there exists $k \geq 1$ such that $F^k = \mathcal{O}_n$.

Here $F \circ G$ denotes F composed with G , and F^k denotes F composed with itself k times.

Solution. Expand $F^k \circ G - G \circ F^k$ using a telescopic sum as follows:

$$\begin{aligned}
 F^k \circ G - G \circ F^k &= \sum_{i=1}^k (F^{k-i+1} \circ G \circ F^{i-1} - F^{k-i} \circ G \circ F^i) \\
 &= \sum_{i=1}^k F^{k-i} \circ (F \circ G - G \circ F) \circ F^{i-1} \\
 &= \sum_{i=1}^k F^{k-i} \circ \alpha F \circ F^{i-1} = \alpha k F^k.
 \end{aligned}$$

This proves (a). For (b), consider the linear map $L(F) = F \circ G - G \circ F$ acting on all $n \times n$ matrices F . Assuming $F^k \neq \mathcal{O}_n$ for all k , we deduce from (a) that αk is an eigenvalue of L for all k . This is impossible since the linear map L acts on an n^2 -dimensional space, so it can have at most n^2 eigenvalues. This contradiction proves (b). \square

- 302.** Let A be a 2×2 matrix with complex entries and let $C(A)$ denote the set of 2×2 matrices that commute with A . Prove that $|\det(A + B)| \geq |\det B|$ for all $B \in C(A)$ if and only if $A^2 = \mathcal{O}_2$.
- 303.** Let A, B be 2×2 matrices with integer entries, such that $AB = BA$ and $\det B = 1$. Prove that if $\det(A^3 + B^3) = 1$, then $A^2 = \mathcal{O}_2$.
- 304.** Consider the $n \times n$ matrix $A = (a_{ij})$ with $a_{ij} = 1$ if $j - i \equiv 1 \pmod{n}$ and $a_{ij} = 0$ otherwise. For real numbers a and b find the eigenvalues of $aA + bA^t$.
- 305.** Let A be an $n \times n$ matrix such that $\det A = 1$ and $A^t A = \mathcal{I}_n$. Show that 1 is an eigenvalue of A .
- 306.** Let A be an $n \times n$ matrix that has zeros on the main diagonal and all other entries from the set $\{-1, 1\}$. Is it possible that $\det A = 0$ for $n = 2007$? What about for $n = 2008$?
- 307.** Let A be an $n \times n$ skew-symmetric matrix (meaning that for all i, j , $a_{ij} = -a_{ji}$) with real entries. Prove that

$$\det(A + x\mathcal{I}_n) \cdot \det(A + y\mathcal{I}_n) \geq \det(A + \sqrt{xy}\mathcal{I}_n)^2,$$

for all $x, y \in [0, \infty)$.

- 308.** Let A be an $n \times n$ matrix. Prove that there exists an $n \times n$ matrix B such that $ABA = A$.
- 309.** Consider the angle formed by two half-lines in three-dimensional space. Prove that the average of the measure of the projection of the angle onto all possible planes in the space is equal to the angle.
- 310.** A linear map A on the n -dimensional vector space V is called an involution if $A^2 = \mathcal{I}$.
- (a) Prove that for every involution A on V there exists a basis of V consisting of eigenvectors of A .
- (b) Find the maximal number of distinct pairwise commuting involutions.

- 311.** Let A be a 3×3 real matrix such that the vectors Au and u are orthogonal for each column vector $u \in \mathbb{R}^3$. Prove that
- (a) $A^t = -A$, where A^t denotes the transpose of the matrix A ;
 - (b) there exists a vector $v \in \mathbb{R}^3$ such that $Au = v \times u$ for every $u \in \mathbb{R}^3$.
- 312.** Denote by $M_n(\mathbb{R})$ the set of $n \times n$ matrices with real entries and let $f : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ be a linear function. Prove that there exists a unique matrix $C \in M_n(\mathbb{R})$ such that $f(A) = \text{tr}(AC)$ for all $A \in M_n(\mathbb{R})$. In addition, if $f(AB) = f(BA)$ for all matrices A and B , prove that there exists $\lambda \in \mathbb{R}$ such that $f(A) = \lambda \text{tr} A$ for any matrix A .
- 313.** Let U and V be isometric linear transformations of \mathbb{R}^n , $n \geq 1$, with the property that $\|Ux - x\| \leq \frac{1}{2}$ and $\|Vx - x\| \leq \frac{1}{2}$ for all $x \in \mathbb{R}^n$ with $\|x\| = 1$. Prove that

$$\|UVU^{-1}V^{-1}x - x\| \leq \frac{1}{2},$$

for all $x \in \mathbb{R}^n$ with $\|x\| = 1$.

- 314.** For an $n \times n$ matrix A denote by $\phi_k(A)$ the symmetric polynomial in the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A ,

$$\phi_k(A) = \sum_{i_1 i_2 \dots i_k} \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_k}, \quad k = 1, 2, \dots, n.$$

For example, $\phi_1(A)$ is the trace and $\phi_n(A)$ is the determinant. Prove that for two $n \times n$ matrices A and B , $\phi_k(AB) = \phi_k(BA)$ for all $k = 1, 2, \dots, n$.

2.3.7 The Cayley-Hamilton and Perron-Frobenius Theorems

We devote this section to two more advanced results, which seem to be relevant to mathematics competitions. All matrices below are assumed to have complex entries.

The Cayley-Hamilton Theorem. Any $n \times n$ matrix A satisfies its characteristic equation, which means that if $P_A(\lambda) = \det(\lambda \mathcal{I}_n - A)$, then $P_A(A) = \mathcal{O}_n$.

Proof. Let $P_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$. Denote by $(\lambda \mathcal{I}_n - A)^*$ the adjoint of $(\lambda \mathcal{I}_n - A)$ (the one used in the computation of the inverse). Then

$$(\lambda \mathcal{I}_n - A)(\lambda \mathcal{I}_n - A)^* = \det(\lambda \mathcal{I}_n - A) \mathcal{I}_n.$$

The entries of the adjoint matrix $(\lambda \mathcal{I}_n - A)^*$ are polynomials in λ of degree at most $n - 1$. Splitting the matrix by the powers of λ , we can write

$$(\lambda \mathcal{I}_n - A)^* = B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \dots + B_0.$$

Equating the coefficients of λ on both sides of

$$(\lambda \mathcal{I}_n - A)(B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \dots + B_0) = \det(\lambda \mathcal{I}_n - A) \mathcal{I}_n,$$

we obtain the equations

$$\begin{aligned}
 B_{n-1} &= \mathcal{I}_n, \\
 -AB_{n-1} + B_{n-2} &= a_{n-1}\mathcal{I}_n, \\
 -AB_{n-2} + B_{n-3} &= a_{n-2}\mathcal{I}_n, \\
 &\dots \\
 -AB_0 &= a_0\mathcal{I}_n.
 \end{aligned}$$

Multiply the first equation by A^n , the second by A^{n-1} , the third by A^{n-2} , and so on, then add the $n + 1$ equations to obtain

$$\mathcal{O}_n = A^n + a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \dots + a_0\mathcal{I}_n.$$

This equality is just the desired $P_A(A) = \mathcal{O}_n$. □

As a corollary we prove the trace identity for $SL(2, \mathbb{C})$ matrices. This identity is important in the study of characters of group representations.

Example. Let A and B be 2×2 matrices with determinant equal to 1. Prove that

$$\text{tr}(AB) - (\text{tr}A)(\text{tr}B) + \text{tr}(AB^{-1}) = 0.$$

Solution. By the Cayley-Hamilton Theorem,

$$B^2 - (\text{tr}B)B + \mathcal{I}_2 = \mathcal{O}_2.$$

Multiply on the left by AB^{-1} to obtain

$$AB - (\text{tr}B)A + AB^{-1} = \mathcal{O}_2,$$

and then take the trace to obtain the identity from the statement. □

Five more examples are left to the reader.

315. Let A be a 2×2 matrix. Show that if for some complex numbers u and v the matrix $u\mathcal{I}_2 + vA$ is invertible, then its inverse is of the form $u'\mathcal{I}_2 + v'A$ for some complex numbers u' and v' .

316. Find the 2×2 matrices X with real entries that satisfy the equation

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

317. Let A, B, C, D be 2×2 matrices. Prove that the matrix $[A, B] \cdot [C, D] + [C, D] \cdot [A, B]$ is a multiple of the identity matrix (here $[A, B] = AB - BA$, the commutator of A and B).

318. Let A and B be two 2×2 matrices that do not commute. Assume that there is a nonconstant polynomial $P(x)$ with real coefficients such that $P(AB) = P(BA)$. Prove that there exists a real number a such that $P(AB) = aI_2$.

319. Let A and B be 3×3 matrices. Prove that

$$\det(AB - BA) = \frac{\operatorname{tr}((AB - BA)^3)}{3}.$$

320. Show that there do not exist real 2×2 matrices A and B such that their commutator is nonzero and commutes with both A and B .

Here is the simplest version of the other result that we had in mind.

The Perron-Frobenius theorem. *Any square matrix with positive entries has a unique eigenvector with positive entries (up to a multiplication by a positive scalar), and the corresponding eigenvalue has multiplicity one and is strictly greater than the absolute value of any other eigenvalue.*

Proof. The proof uses real analysis. Let $A = (a_{ij})_{i,j=1}^n$, $n \geq 1$. We want to show that there is a unique $v \in [0, \infty)^n$, $v \neq 0$, such that $Av = \lambda v$ for some λ . Of course, since A has positive entries and v has positive coordinates, λ has to be a positive number. Denote by K the intersection of $[0, \infty)^n$ with the $n - 1$ -dimensional unit sphere. Reformulating the problem, we want to show that the function $f : K \rightarrow K$, $f(v) = \frac{Av}{\|Av\|}$ has a fixed point.

Now, there is a rather general result that states that a contractive function on a compact metric space has a unique fixed point (see Section 3.2.3). Recall that a metric space is a set X endowed with a function $\delta : X \times X \rightarrow [0, \infty)$ satisfying

- (i) $\delta(x, y) = 0$ if and only if $x = y$,
- (ii) $\delta(x, y) = \delta(y, x)$ for all $x, y \in X$,
- (iii) $\delta(x, y) + \delta(y, z) \geq \delta(x, z)$ for all $x, y, z \in X$.

We use the property in the case of a compact set in \mathbb{R}^n , where compact sets are characterized by being closed and bounded. A function $f : X \rightarrow X$ is contractive if

$$\delta(f(x), f(y)) < \delta(x, y), \text{ for every } x \neq y.$$

With this in mind, we want to find a distance on the set K that makes the function f defined above contractive. This is the Hilbert metric defined by the formula

$$\delta(v, w) = \ln \left(\max_i \left\{ \frac{v_i}{w_i} \right\} / \min_i \left\{ \frac{v_i}{w_i} \right\} \right),$$

for $v = (v_1, v_2, \dots, v_n)$ and $w = (w_1, w_2, \dots, w_n) \in K$. That this satisfies the triangle inequality $\delta(u, w) + \delta(w, u) \geq \delta(v, w)$ is a consequence of the inequalities

$$\max_i \left\{ \frac{v_i}{w_i} \right\} \cdot \max_i \left\{ \frac{w_i}{u_i} \right\} \geq \max_i \left\{ \frac{v_i}{u_i} \right\},$$

$$\min_i \left\{ \frac{v_i}{w_i} \right\} \cdot \min_i \left\{ \frac{w_i}{u_i} \right\} \geq \min_i \left\{ \frac{v_i}{w_i} \right\}.$$

Let us show that f is contractive. If $v = (v_1, v_2, \dots, v_n)$ and $w = (w_1, w_2, \dots, w_n)$ are in K , $v \neq w$, and if $\alpha_i > 0$, $i = 1, 2, \dots, n$, then

$$\min_i \left\{ \frac{v_i}{w_i} \right\} < \frac{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n}{\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_n w_n} < \max_i \left\{ \frac{v_i}{w_i} \right\}.$$

Indeed, to prove the first inequality, add the obvious inequalities

$$\alpha_j w_j \min_i \left\{ \frac{v_i}{w_i} \right\} \leq \alpha_j v_j, \quad j = 1, 2, \dots, n.$$

Because $v \neq w$ and both vectors are on the unit sphere, at least one inequality is strict. The second inequality follows from

$$\alpha_j w_j \max_i \left\{ \frac{v_i}{w_i} \right\} \geq \alpha_j v_j, \quad j = 1, 2, \dots, n,$$

where again at least one inequality is strict.

Using this fact, we obtain for all j , $1 \leq j \leq n$,

$$\frac{\frac{a_{j1}v_1 + \dots + a_{jn}v_n}{a_{j1}w_1 + \dots + a_{jn}w_n}}{\max_i \left\{ \frac{v_i}{w_i} \right\}} < 1 < \frac{\frac{a_{j1}v_1 + \dots + a_{jn}v_n}{a_{j1}w_1 + \dots + a_{jn}w_n}}{\min_i \left\{ \frac{v_i}{w_i} \right\}}.$$

Therefore,

$$\frac{\max_j \left\{ \frac{a_{j1}v_1 + \dots + a_{jn}v_n}{a_{j1}w_1 + \dots + a_{jn}w_n} \right\}}{\max_i \left\{ \frac{v_i}{w_i} \right\}} < \frac{\min_j \left\{ \frac{a_{j1}v_1 + \dots + a_{jn}v_n}{a_{j1}w_1 + \dots + a_{jn}w_n} \right\}}{\min_i \left\{ \frac{v_i}{w_i} \right\}}.$$

It follows that for $v, w \in K$, $v \neq w$, $\delta(f(v), f(w)) < \delta(v, w)$.

Now, K is closed and but is not bounded in the Hilbert metric; some points are infinitely far apart. But even if K is not bounded in the Hilbert metric, $f(K)$ is (prove it!). If we denote by K_0 the closure of $f(K)$ in the Hilbert metric, then this space is closed and bounded. On K_0 , f is contractive, and so it has a unique fixed point. Note that all fixed points of f are necessarily in K_0 (because if $f(v) = v$, then $v = f(v) \in f(K)$).

We are done with the first half of the proof. Now let us show that the eigenvalue of this positive vector is larger than the absolute value of any other eigenvalue. Let $r(A)$ be the largest of the absolute values of the eigenvalues of A and let λ be an eigenvalue with $|\lambda| = r(A)$. In general, for a vector v we denote by $|v|$ the vector whose coordinates are the absolute values of the coordinates of v . Also, for two vectors v, w we write $v \geq w$ if each coordinate of v is greater than the corresponding coordinate of w . If v is an eigenvector of A corresponding to the eigenvalue λ , then $|Av| = |\lambda| \cdot |v|$. The triangle inequality implies $A|v| \geq |Av| = r(A)|v|$. It follows that the set

$$K_1 = \{v \mid \|v\| = 1, v \geq 0, Av \geq r(A)v\},$$

is nonempty. Because A has positive entries, $A(Av - r(A)v) \geq 0$ for $v \in K_1$. So $A(Av) \geq r(A)(Av)$, for $v \in K_1$, proving that $f(K_1) \subset K_1$. Again K_1 is closed and $f(K_1)$ is bounded, so we can reason as above to prove that f restricted to K_1 has a fixed point, and because $K_1 \subset K$, this is the fixed point that we detected before. Thus $r(A)$ is the unique positive eigenvalue.

There cannot exist another eigenvalue λ with $|\lambda| = r(A)$, for otherwise, for a small $\varepsilon > 0$ the matrix $A - \varepsilon I_n$ would still have positive entries, but its positive eigenvalue $r(A) - \varepsilon$ would be smaller than the absolute value of the other eigenvalue contradicting what we just proved. This concludes the proof of the theorem. \square

Nowhere in the book are more appropriate the words of Sir Arthur Eddington: “Proof is an idol before which the mathematician tortures himself.”

The conclusion of the theorem still holds in the more general setting of irreducible matrices with nonnegative entries (*irreducible* means that there is no reordering of the rows and columns that makes it block upper triangular). This more general form of the Perron-Frobenius Theorem is currently used by the Internet browser Google to sort the entries of a search. The idea is the following: Write the adjacency matrix of the Internet with a link highlighted if it is related to the subject. Then multiply each nonzero entry by a larger or smaller number that takes into account how important the subject is in that page. The Perron-Frobenius vector of this new matrix assigns a positive weight to each site on the Internet. The Internet browser then lists the sites in decreasing order of their weights.

We now challenge you with some problems.

- 321.** Let A be a square matrix whose off-diagonal entries are positive. Prove that the rightmost eigenvalue of A in the complex plane is real and all other eigenvalues are strictly to its left in the complex plane.
- 322.** Let a_{ij} , $i, j = 1, 2, 3$, be real numbers such that a_{ij} is positive for $i = j$ and negative for $i \neq j$. Prove that there exist positive real numbers c_1, c_2, c_3 such that the numbers

$$a_{11}c_1 + a_{12}c_2 + a_{13}c_3, \quad a_{21}c_1 + a_{22}c_2 + a_{23}c_3, \quad a_{31}c_1 + a_{32}c_2 + a_{33}c_3$$

are all negative, all positive, or all zero.

- 323.** Let x_1, x_2, \dots, x_n be differentiable (real-valued) functions of a single variable t that satisfy

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ &\dots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, \end{aligned}$$

for some constants $a_{ij} > 0$. Suppose that for all i , $x_i(t) \rightarrow 0$ as $t \rightarrow \infty$. Are the functions x_1, x_2, \dots, x_n necessarily linearly independent?

- 324.** For a positive integer n and any real number c , define $(x_k)_{k \geq 0}$ recursively by $x_0 = 0$, $x_1 = 1$, and for $k \geq 0$,

$$x_{k+2} = \frac{cx_{k+1} - (n-k)x_k}{k+1}.$$

Fix n and then take c to be the largest value for which $x_{n+1} = 0$. Find x_k in terms of n and k , $1 \leq k \leq n$.

2.4 Abstract Algebra

2.4.1 Binary Operations

A binary operation $*$ on a set S associates to each pair $(a, b) \in S \times S$ an element $a * b \in S$. The operation is called associative if $a * (b * c) = (a * b) * c$ for all $a, b, c \in S$, and commutative if $a * b = b * a$ for all $a, b \in S$. If there exists an element e such that $a * e = e * a = a$ for all $a \in S$, then e is called an identity element. If an identity exists, it is unique. In this case, if for an element $a \in S$ there exists $b \in S$ such that $a * b = b * a = e$, then b is called the inverse of a and is denoted by a^{-1} . If an element has an inverse, the inverse is unique.

Just as a warmup, we present a problem from the 62nd W.L. Putnam Competition, 2001.

Example. Consider a set S and a binary operation $*$ on S . Assume that $(a * b) * a = b$ for all $a, b \in S$. Prove that $a * (b * a) = b$ for all $a, b \in S$.

Solution. Substituting $b * a$ for a , we obtain

$$((b * a) * b) * (b * a) = b.$$

The expression in the first set of parentheses is a . Therefore,

$$a * (b * a) = b,$$

as desired. □

Often, problems about binary operations look like innocent puzzles, yet they can have profound implications. This is the case with the following example.

Example. For three-dimensional vectors $X = (p, q, t)$ and $Y = (p', q', t')$ define the operations $(p, q, t) * (p', q', t') = (0, 0, pq' - qp')$, and $X \circ Y = X + Y + \frac{1}{2}X * Y$, where $+$ denotes the addition in \mathbb{R}^3 .

(a) Prove that (\mathbb{R}^3, \circ) is a group.

(b) Let $\alpha : (\mathbb{R}^3, \circ) \rightarrow (\mathbb{R}^3, \circ)$ be a continuous map satisfying $\alpha(X \circ Y) = \alpha(X) \circ \alpha(Y)$ for all X, Y (which means that α is a homomorphism). Prove that

$$\alpha(X + Y) = \alpha(X) + \alpha(Y) \quad \text{and} \quad \alpha(X * Y) = \alpha(X) * \alpha(Y).$$

Solution. (a) Associativity can be verified easily, the identity element is $(0, 0, 0)$, and the inverse of (p, q, t) is $(-p, -q, -t)$.

(b) First, note that $X * Y = -Y * X$. Therefore, if X is a scalar multiple of Y , then $X * Y = Y * X = 0$. In general, if $X * Y = 0$, then $X \circ Y = X + Y = Y \circ X$. Hence in this case,

$$\alpha(X + Y) = \alpha(X \circ Y) = \alpha(X) \circ \alpha(Y) = \alpha(X) + \alpha(Y) + \frac{1}{2}\alpha(X) * \alpha(Y)$$

on the one hand, and

$$\alpha(X + Y) = \alpha(Y \circ X) = \alpha(Y) \circ \alpha(X) = \alpha(Y) + \alpha(X) + \frac{1}{2}\alpha(Y) * \alpha(X).$$

Because $\alpha(X) * \alpha(Y) = -\alpha(Y) * \alpha(X)$, this implies that $\alpha(X) * \alpha(Y) = 0$. and consequently $\alpha(X + Y) = \alpha(X) + \alpha(Y)$. In particular, α is additive on every one-dimensional space, whence $\alpha(rX) = r\alpha(X)$, for every rational number r . But α is continuous, so $\alpha(sX) = s\alpha(X)$ for every real number s . Applying this property we find that for any $X, Y \in \mathbb{R}^3$ and $s \in \mathbb{R}$,

$$\begin{aligned} s\alpha\left(X + Y + \frac{1}{2}sX * Y\right) &= \alpha\left(sX + sY + \frac{1}{2}s^2X * Y\right) = \alpha((sX) \circ (sY)) \\ &= \alpha(sX) \circ \alpha(sY) = (s\alpha(X)) \circ (s\alpha(Y)) \\ &= s\alpha(X) + s\alpha(Y) + \frac{1}{2}s^2\alpha(X) * \alpha(Y). \end{aligned}$$

Dividing both sides by s , we obtain

$$\alpha\left(X + Y + \frac{1}{2}sX * Y\right) = \alpha(X) + \alpha(Y) + \frac{1}{2}s\alpha(X) * \alpha(Y).$$

In this equality if we let $s \rightarrow 0$, we obtain $\alpha(X + Y) = \alpha(X) + \alpha(Y)$. Also, if we let $s = 1$ and use the additivity we just proved, we obtain $\alpha(X * Y) = \alpha(X) * \alpha(Y)$. The problem is solved. \square

Traditionally, $X * Y$ is denoted by $[X, Y]$ and \mathbb{R}^3 endowed with this operation is called the Heisenberg Lie algebra. Also, \mathbb{R}^3 endowed with \circ is called the Heisenberg group. And we just proved a famous theorem showing that a continuous automorphism of the Heisenberg group is also an automorphism of the Heisenberg Lie algebra. The Heisenberg group and algebra are fundamental concepts of quantum mechanics.

- 325.** With the aid of a calculator that can add, subtract, and determine the inverse of a nonzero number, find the product of two nonzero numbers using at most 20 operations.
- 326.** Invent a binary operation from which $+$, $-$, \times , and $/$ can be derived.
- 327.** A finite set S with at least four elements is endowed with an associative binary operation $*$ that satisfies

$$(a * a) * b = b * (a * a) = b \text{ for all } a, b \in S.$$

Prove that the set of all elements of the form $a * (b * c)$ with a, b, c distinct elements of S coincides with S .

- 328.** Let S be the smallest set of rational functions containing $f(x, y) = x$ and $g(x, y) = y$ and closed under subtraction and taking reciprocals. Show that S does not contain the nonzero constant functions.
- 329.** Let $*$ and \circ be two binary operations on the set M , with identity elements e , respectively, e' , and with the property that for every $x, y, u, v \in M$,

$$(x * y) \circ (u * v) = (x \circ u) * (y \circ v).$$

Prove that

- (a) $e = e'$;
 - (b) $x * y = x \circ y$, for every $x, y \in M$;
 - (c) $x * y = y * x$, for every $x, y \in M$.
- 330.** Consider a set S and a binary operation $*$ on S such that $x * (y * x) = y$ for all x, y in S . Prove that each of the equations $a * x = b$ and $x * a = b$ has a unique solution in S .
- 331.** On a set M an operation $*$ is given satisfying the properties
- (i) there exists an element $e \in M$ such that $x * e = x$ for all $x \in M$;
 - (ii) $(x * y) * z = (z * x) * y$ for all $x, y, z \in M$.

Prove that the operation $*$ is both associative and commutative.

- 332.** Prove or disprove the following statement: If F is a finite set with two or more elements, then there exists a binary operation $*$ on F such that for all $x, y, z \in F$,
- (i) $x * z = y * z$ implies $x = y$ (right cancellation holds), and
 - (ii) $x * (y * z) \neq (x * y) * z$ (no case of associativity holds).
- 333.** Let $*$ be an associative binary operation on a set S satisfying $a * b = b * a$ only if $a = b$. Prove that $a * (b * c) = a * c$ for all $a, b, c \in S$. Give an example of such an operation.
- 334.** Let S be a set and $*$ a binary operation on S satisfying the laws
- (i) $x * (x * y) = y$ for all $x, y \in S$,
 - (ii) $(y * x) * x = y$ for all $x, y \in S$.

Show that $*$ is commutative but not necessarily associative.

- 335.** Let $*$ be a binary operation on the set \mathbb{Q} of rational numbers that is associative and commutative and satisfies $0 * 0 = 0$ and $(a + c) * (b + c) = a * b + c$ for all $a, b, c \in \mathbb{Q}$. Prove that either $a * b = \max(a, b)$ for all $a, b \in \mathbb{Q}$, or $a * b = \min(a, b)$ for all $a, b \in \mathbb{Q}$.

2.4.2 Groups

Definition. A group is a set of transformations (of some space) that contains the identity transformation and is closed under composition and under the operation of taking the inverse.

The isometries of the plane, the permutations of a set, the continuous bijections on a closed bounded interval all form groups.

There is a more abstract, and apparently more general definition, which calls a group a set G endowed with a binary operation \cdot that satisfies

- (i) (associativity) $x(yz) = (xy)z$ for all $x, y, z \in G$;
- (ii) (identity element) there is $e \in G$ such that for any $x \in G$, $ex = xe = x$;
- (iii) (existence of the inverse) for every $x \in G$ there is $x^{-1} \in G$ such that

$$xx^{-1} = x^{-1}x = e.$$

But Cayley observed the following fact.

Theorem. *Any group is a group of transformations.*

Proof. Indeed, any group G acts on itself on the left. Specifically, $x \in G$ acts as a transformation of G by $y \rightarrow xy$, $y \in G$. \square

A group G is called Abelian (after N. Abel) if the operation is commutative, that is, if $xy = yx$ for all $x, y \in G$. An example of an Abelian group is the Klein four-group, introduced abstractly as $K = \{a, b, c, e \mid a^2 = b^2 = c^2 = e, ab = ac, ac = b, bc = a\}$, or concretely as the group of the symmetries of a rectangle (depicted in Figure 14).

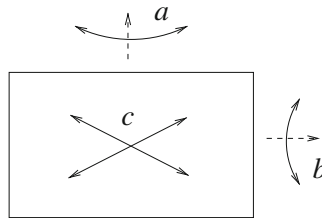


Figure 14

A group is called cyclic if it is generated by a single element, that is, if it consists of the identity element and the powers of some element.

Let us turn to problems and start with one published by L. Daia in the *Mathematics Gazette, Bucharest*.

Example. A certain multiplicative operation on a nonempty set G is associative and allows cancellations on the left, and there exists $a \in G$ such that $x^3 = axa$ for all $x \in G$. Prove that G endowed with this operation is an Abelian group.

Solution. Replacing x by ax in the given relation, we obtain $axaxax = a^2xa$. Cancelling a on the left, we obtain $x(axa)x = axa$. Because $axa = x^3$, it follows that $x^5 = x^3$, and cancelling an x^2 , we obtain

$$x^3 = x \text{ for all } x \in G.$$

In particular, $a^3 = a$, and hence $a^3x = ax$ for all $x \in G$. Cancel a on the left to find that

$$a^2x = x \text{ for all } x \in G.$$

Substituting x by xa , we obtain $a^2xa = xa$, or $ax^3 = xa$, and since $x^3 = x$, it follows that a commutes with all elements in G . We can therefore write

$$a^2x = a(ax) = a(xa) = (xa)a = xa^2,$$

whence $xa^2 = a^2x = x$. This shows that a^2 is the identity element of the multiplicative operation; we denote it by e . The relation from the statement implies $x^3 = axa = xa^2 = xe$; cancelling x , we obtain $x^2 = e$; hence for all $x \in G$, $x^{-1} = x$. It follows that G is a group. It is Abelian by the well-known computation

$$xy = (xy)^{-1} = y^{-1}x^{-1} = yx. \quad \square$$

Here are more examples of the kind.

336. Prove that in order for a set G endowed with an associative operation to be a group, it suffices for it to have a left identity, and for each element to have a left inverse. This means that there should exist $e \in G$ such that $ex = x$ for all $x \in G$, and for each $x \in G$, there should exist $x' \in G$ such that $x'x = e$. The same conclusion holds if “left” is replaced by “right”.

337. Let (G, \perp) and $(G, *)$ be two group structures defined on the same set G . Assume that the two groups have the same identity element and that their binary operations satisfy

$$a * b = (a \perp a) \perp (a \perp b),$$

for all $a, b \in G$. Prove that the binary operations coincide and the group they define is Abelian.

338. Let r, s, t be positive integers that are pairwise relatively prime. If the elements a and b of an Abelian group with identity element e satisfy $a^r = b^s = (ab)^t = e$, prove that $a = b = e$. Does the same conclusion hold if a and b are elements of an arbitrary nonAbelian group?

339. A is a subset of a finite group G which contains more than one half of the elements of G . Prove that every element of G is the product of two elements of A .

340. On the set $M = \mathbb{R} \setminus \{3\}$ the following binary operation is defined:

$$x * y = 3(xy - 3x - 3y) + m,$$

where $m \in \mathbb{R}$. Find all possible values of m for which $(M, *)$ is a group.

341. Assume that a and b are elements of a group with identity element e satisfying $(aba^{-1})^n = e$ for some positive integer n . Prove that $b^n = e$.

342. Let G be a group with the following properties:

- (i) G has no element of order 2,
- (ii) $(xy)^2 = (yx)^2$, for all $x, y \in G$.

Prove that G is Abelian.

343. A multiplicative operation on a set M satisfies

- (i) $a^2 = b^2$, (ii) $ab^2 = a$, (iii) $a^2(bc) = cb$, (iv) $(ac)(bc) = ab$, for all $a, b, c \in M$.

Define on M the operation

$$a * b = a(b^2b).$$

Prove that $(M, *)$ is a group.

We would like to point out the following property of the set of real numbers.

Kronecker's theorem. *A nontrivial subgroup of the additive group of real numbers is either cyclic or it is dense in the set of real numbers.*

Proof. Denote the group by G . It is either discrete, or it has an accumulation point on the real axis. If it is discrete, let a be its smallest positive element. Then any other element is of the form $b = ka + \alpha$ with $0 \leq \alpha < a$. But b and ka are both in G ; hence α is in G as well. By the minimality of a , α can only be equal to 0, and hence the group is cyclic.

If there is a nonconstant sequence $(x_n)_n$ in G converging to some real number, then $\pm(x_n - x_m)$ approaches zero as $n, m \rightarrow \infty$. Choosing the indices m and n appropriately, we can find a sequence of positive elements in G that converges to 0. Thus for any $\varepsilon > 0$ there is an element $c \in G$ with $0 < c < \varepsilon$. For some integer k , the distance between kc and $(k+1)c$ is less than ε ; hence any interval of length ε contains some multiple of c . Varying ε , we conclude that G is dense in the real axis. \square

Try to use this result to solve the following problems.

344. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function satisfying

$$f(x) + f(x + \sqrt{2}) = f(x + \sqrt{3}) \text{ for all } x.$$

Prove that f is constant.

345. Prove that the sequence $(\sin n)_{n \geq 1}$ is dense in the interval $[-1, 1]$.

346. Show that infinitely many powers of 2 start with the digit 7.

347. Given a rectangle, we are allowed to fold it in two or in three, parallel to one side or the other, in order to form a smaller rectangle. Prove that for any $\varepsilon > 0$ there are finitely many such operations that produce a rectangle with the ratio of the sides lying in the interval $(1 - \varepsilon, 1 + \varepsilon)$ (which means that we can get arbitrarily close to a square).

348. A set of points in the plane is invariant under the reflections across the sides of some given regular pentagon. Prove that the set is dense in the plane.

We continue with problems about groups of matrices.

- 349.** Prove that the group of invertible 4×4 matrices with rational entries has no elements of order 7.
- 350.** Given Γ a finite multiplicative group of invertible matrices with complex entries, denote by M the sum of the matrices in Γ . Prove that $\det M$ and $\operatorname{tr} M$ are integers.
- 351.** Let n be a positive integer. What is the size of the largest multiplicative group of invertible $n \times n$ matrices with integer entries such that for every matrix A in the group all the entries of $A - I_n$ are even?
- 352.** For an $n \times n$ matrix with complex entries, A , we define its norm to be

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\|,$$

where $\|x\|$ denotes the usual norm on \mathbb{C}^n (the square root of the sum of the squares of the absolute values of the coordinates). Let $a < 2$, and let G be a multiplicative group of invertible $n \times n$ matrices such that

$$\|A - I_n\| \leq a \text{ for all } A \in G.$$

Prove that G is finite.

“There is no certainty in sciences where one of the mathematical sciences cannot be applied, or which are not in relation with this mathematics.” This thought of Leonardo da Vinci motivated us to include an example of how groups show up in natural sciences.

The groups of symmetries of three-dimensional space play an important role in chemistry and crystallography. In chemistry, the symmetries of molecules give rise to physical properties such as optical activity. The point groups of symmetries of molecules were classified by A. Schönflies as follows:

- C_s : a reflection with respect to a plane, isomorphic to \mathbb{Z}_2 ,
- C_i : a reflection with respect to a point, isomorphic to \mathbb{Z}_2 ,
- C_n : the rotations by multiples of $\frac{2\pi}{n}$ about an axis, isomorphic to \mathbb{Z}_n ,
- C_{nv} : generated by a C_n and a C_s with the reflection plane containing the axis of rotation; in mathematics this is called the dihedral group,
- C_{nh} : generated by a C_n and a C_s with the reflection plane perpendicular to the axis of rotation, isomorphic to $C_n \times C_2$,
- D_n : generated by a C_n and a C_2 , with the rotation axes perpendicular to each other, isomorphic to the dihedral group,
- D_{nd} : generated by a C_n and a C_2 , together with a reflection across a plane that divides the angle between the two rotation axes,
- D_{nh} : generated by a C_n and a C_2 with perpendicular rotation axes, together with a reflection with respect to a plane perpendicular to the first rotation axis,

- S_n : improper rotations by multiples of $\frac{2\pi}{n}$, i.e., the group generated by the element that is the composition of the rotation by $\frac{2\pi}{n}$ and the reflection with respect to a plane perpendicular to the rotation axis,
- Special point groups: $C_{\infty v}$'s and $D_{\infty h}$'s (same as C_{nv} and D_{nh} but with all rotations about the axis allowed), together with the symmetry groups of the five Platonic solids.

When drawing a molecule, we use the convention that all segments represent bonds in the plane of the paper, all bold arrows represent bonds with the tip of the arrow below the tail of the arrow. The molecules from Figure 15 have respective symmetry point groups the octahedral group and C_{3h} .



Figure 15

353. Find the symmetry groups of the molecules depicted in Figure 16.

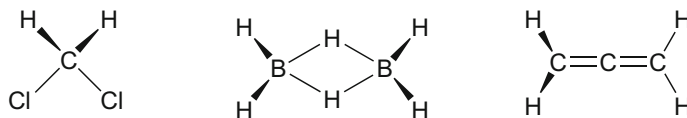


Figure 16

2.4.3 Rings

Rings mimic in the abstract setting the properties of the sets of integers, polynomials, or matrices.

Definition. A ring is a set R endowed with two operations $+$ and \cdot (addition and multiplication) such that $(R, +)$ is an Abelian group with identity element 0 and the multiplication satisfies

- (associativity) $x(yz) = (xy)z$ for all $x, y, z \in R$, and
- (distributivity) $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$ for all $x, y, z \in R$.

A ring is called commutative if the multiplication is commutative. It is said to have identity if there exists $1 \in R$ such that $1 \cdot x = x \cdot 1 = x$ for all $x \in R$. An element $x \in R$ is called invertible if there exists $x^{-1} \in R$ such that $xx^{-1} = x^{-1}x = 1$.

We consider two examples, the second of which appeared many years ago in the Balkan Mathematics Competition for university students.

Example. Let x and y be elements in a ring with identity. Prove that if $1 - xy$ is invertible, then so is $1 - yx$.

Solution. If we naively use the expansion $(1 - x)^{-1} = 1 + x + x^2 + x^3 + \cdots$ to write

$$\begin{aligned}(1 - xy)^{-1} &= 1 + xy + xyxy + xyxyxy + \cdots \\ (1 - yx)^{-1} &= 1 + yx + yxyx + yxyxyx + \cdots,\end{aligned}$$

we can rearrange the second as

$$(1 - yx)^{-1} = 1 + y(1 + xy + xyxy + xyxyxy + \cdots)x$$

So we can guess that if v be the inverse of $1 - xy$ then $1 + yvx$ is the inverse of $1 - yx$. We have $v(1 - xy) = (1 - xy)v = 1$; hence $vxy = xyv = v - 1$. We compute

$$(1 + yvx)(1 - yx) = 1 - yx + yvx - yvxyx = 1 - yx + yvx - y(v - 1)x = 1.$$

A similar verification shows that $(1 - yx)(1 + yvx) = 1$. It follows that $1 - yx$ is invertible and its inverse is $1 + yvx$. \square

Example. Prove that if in a ring R (not necessarily with identity element) $x^3 = x$ for all $x \in R$, then the ring is commutative.

Solution. For $x, y \in R$, we have

$$\begin{aligned}xy^2 - y^2xy^2 &= (xy^2 - y^2xy^2)^3 = xy^2xy^2xy^2 - xy^2xy^2y^2xy^2 - xy^2y^2xy^2xy^2 \\ &\quad - y^2xy^2xy^2xy^2 + y^2xy^2xy^2y^2xy^2 + y^2xy^2y^2xy^2xy^2 \\ &\quad - y^2xy^2y^2xy^2y^2xy^2 + xy^2y^2xy^2y^2xy^2.\end{aligned}$$

Using the fact that $y^4 = y^2$, we see that this is equal to zero, and hence $xy^2 - y^2xy^2 = 0$, that is, $xy^2 = y^2xy^2$. A similar argument shows that $y^2x = y^2xy^2$, and so $xy^2 = y^2x$ for all $x, y \in R$.

Using this we obtain

$$xy = xyxyxy = xy(xy)^2 = x(xy)^2y = x^2yxy^2 = y^3x^3 = yx.$$

This proves that the ring is commutative, as desired. \square

We remark that both this and the third problem below are particular cases of the following result by N. Jacobson:

Jacobson theorem. If a ring (with or without identity) has the property that for every element x there exists an integer $n(x) > 1$ such that $x^{n(x)} = x$, then the ring is commutative.

Try your hand at the following problems.

354. Let a, b, c be elements of a ring with identity.

- (a) Show that if $I_n - abc$ is invertible, then $I_n - cab$ is invertible.
 - (b) Can it happen that $I_n - abc$ is invertible but $I_n - cba$ is not?
- 355.** Let R be a nontrivial ring with identity, and $M = \{x \in R \mid x = x^2\}$ the set of its idempotents. Prove that if M is finite, then it has an even number of elements.
- 356.** Let R be a ring with identity such that $x^6 = x$ for all $x \in R$. Prove that $x^2 = x$ for all $x \in R$. Prove that any such ring is commutative.
- 357.** Let R be a ring with identity with the property that $(xy)^2 = x^2y^2$ for all $x, y \in R$. Show that R is commutative.
- 358.** Let R be a finite ring with unit, having n elements and such that the equation $x^n = 1$ has the unique solution $x = 1$ in R . Prove that
- (a) 0 is the unique nilpotent element of R ;
 - (b) there is a positive integer $k \geq 2$ such that the equation x^k has n solutions in R .
- ($x \in R$ is called nilpotent if there is a positive integer m such that $x^m = 0$.)
- 359.** Let R be a finite ring such that $1 + 1 = 0$. Prove that the number of solutions to the equation $x^2 = 0$ is equal to the number of solutions to the equation $x^2 = 1$.
- 360.** Let x and y be elements in a ring with identity and n a positive integer. Prove that if $1 - (xy)^n$ is invertible, then so is $1 - (yx)^n$.
- 361.** Let R be a ring with the property that if $x \in R$ and $x^2 = 0$, then $x = 0$.
- (a) Prove that if $x, z \in R$ and $z^2 = z$, then $zxz - xz = 0$.
 - (b) Prove that any idempotent of R belongs to the center of R (the center of a ring consists of those elements that commute with all elements of the ring).
- 362.** Show that if a ring R with identity has three elements a, b, c such that
- (i) $ab = ba, bc = cb$;
 - (ii) for any $x, y \in R, bx = by$ implies $x = y$;
 - (iii) $ca = b$ but $ac \neq b$,
- then the ring cannot be finite.

Putnam and Beyond

Gelca, R.; Andreescu, T.

2017, XVIII, 850 p. 297 illus., Softcover

ISBN: 978-3-319-58986-2