Pigeonhole and Inclusion-Exclusion Principles

1 Pigeonhole principle

This is the observation that, if n objects have to be placed in less than n sets, at least one of the sets will contain two objects or more. More generally, if n objects have to be placed in k sets, at least one of the sets will contain $\lceil n/k \rceil$ objects or more. Here are some famous theorems where this principle is applied.

Theorem 1 (Dirichlet). If x is an irrational number, then there are infinitely many pairs (p, q) of integers such that $|x - p/q| < 1/q^2$.

Proof. Suppose on the contrary that the set S of $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that $|x - p/q| < 1/q^2$ is finite. Define

$$\varepsilon := \min_{(p,q)\in\mathbb{Z}\times\mathbb{N}} |x - p/q| > 0.$$

Now consider an integer $Q \geq 1/\varepsilon$. The $Q + 1$ fractional parts $\{qx\}$, $q \in [0 : Q]$, belong to $[0, 1)$, which is the union of the Q intervals $[\ell/Q, (\ell + 1)/Q)$, $\ell \in [0, Q - 1]$. Therefore, there exist $q', q'' \in [0 : Q]$ with $q' > q''$, say, such that $\{q'x\}$ and $\{q''x\}$ belong to the same interval, hence $|\{q'x\} - \{q''x\}| < 1/Q$. This reads

$$\big|q'x - \lfloor q'x \rfloor - (q''x - \lfloor q''x \rfloor)\big| = \big|(q' - q'')x - (\lfloor q'x \rfloor - \lfloor q''x \rfloor)\big| < \frac{1}{Q}.$$

Setting $q := q' - q'' \in [1 : Q]$ and $p := \lfloor q'x \rfloor - \lfloor q''x \rfloor \in \mathbb{Z}$, we derive

$$\left|x - \frac{p}{q}\right| < \frac{1}{qQ} \leq \begin{cases} 1/q^2, \\ 1/Q \leq \varepsilon. \end{cases}$$

This shows that $(p, q) \in S$ and that $|x - p/q| < \varepsilon$, which is a contradiction. □

Theorem 2 (Erdős–Szekeres). Every sequence of $(m - 1)(n - 1) + 1$ distinct real numbers admits either an increasing subsequence of length m or a decreasing subsequence of length n.

Proof. Suppose that the sequence — denote it by $(u_i)_{i\in[1:(m-1)(n-1)+1]}$ — has no increasing subsequence of length m. This means that

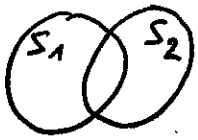
$$S_k := \{i \in [1 : (m-1)(n-1)+1] : \text{ the largest increasing subsequence starting at } i \text{ has length } k\}$$

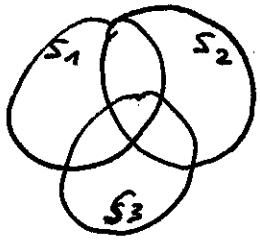
is nonempty only for $k = 1, 2, \ldots, m - 1$. Since every $i \in [1 : (m - 1)(n - 1) + 1]$ belongs to one of $S_1, S_2, \ldots, S_{m-1}$, there is an S_k with size $\geq n$. Consider $i_1 < i_2 < \cdots < i_n$ in S_k. For $j \in [1 : n]$, notice that an increasing subsequence of length k starts at i_{j+1} and that no increasing sequence of length $k + 1$ starts at i_j, therefore we must have $u_{i_j} > u_{i_{j+1}}$. This gives a decreasing subsequence of length n, namely $(u_{i_j})_{j\in[1:n]}$. □

2 Inclusion-exclusion principle

The inclusion-exclusion principle generalizes the formulas

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|,$$
$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + |S_1 \cap S_2 \cap S_3|.$$





Theorem 3. Given finite sets $S_1, S_2, \ldots, S_n$,

$$|S_1 \cup \cdots \cup S_n| = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \le i_1 < \cdots < i_r \le n} |S_{i_1} \cap \cdots \cap S_{i_r}|.$$

Remark. Finite sets can be replaced by measurable sets, say, with $|\cdot|$ meaning measure.

Proof. We could proceed by induction on n. Here is another instructive argument based on characteristic functions. Recall first that, for two sets A and B, one has $\chi_{A \cup B} = \chi_A + \chi_B - \chi_{A \cap B}$ and $\chi_{A \cap B} = \chi_A \chi_B$. It follows that $1 - \chi_{A \cap B} = (1 - \chi_A)(1 - \chi_B)$. We then derive

$$\begin{aligned} 1 - \chi_{S_1 \cup S_2 \cup \cdots \cup S_n}(x) &= (1 - \chi_{S_1}(x))(1 - \chi_{S_2}(x)) \cdots (1 - \chi_{S_n}(x)) \\ &= 1 + \sum_{r=1}^{n} (-1)^r \sum_{1 \le i_1 < \cdots < i_r \le n} \chi_{S_{i_1}}(x) \chi_{S_{i_2}}(x) \cdots \chi_{S_{i_r}}(x) \\ &= 1 + \sum_{r=1}^{n} (-1)^r \sum_{1 \le i_1 < \cdots < i_r \le n} \chi_{S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_r}}(x). \end{aligned}$$

It now suffices to sum over all possible x's. □

As an application, we count the derangements of $[1 : n]$, i.e., the permutations of $[1 : n]$ with no fixed points.

Theorem 4. The number of derangements of $[1 : n]$ is

$$D_n = n! \sum_{r=0}^{n} \frac{(-1)^r}{r!}.$$

Remark. The proportion of derangements among all permutations approaches $1/e \approx 0.3679$ when n grows, since $D_n/n! \to \sum_{r=0}^{\infty} (-1)^r / r! = e^{-1}$.

Proof. Let $\mathcal{P}$ be the set of permutations of $[1:n]$ and $\mathcal{D}$ the set of derangements of $[1:n]$. For each $i \in [1:n]$, we consider the set $\mathcal{P}_i$ of $\sigma \in \mathcal{P}$ such that $\sigma(i) = i$. Since a permutation is not a derangement iff it belongs to one of the $\mathcal{P}_i$, we have $\mathcal{P} \setminus \mathcal{D} = \bigcup_{i=1}^{n} \mathcal{P}_i$. Therefore,

$$|\mathcal{P} \setminus \mathcal{D}| = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \le i_1 < \cdots < i_r \le n} |\mathcal{P}_{i_1} \cap \cdots \cap \mathcal{P}_{i_r}|.$$

Note that $|\mathcal{P} \setminus \mathcal{D}| = |\mathcal{P}| - |\mathcal{D}| = n! - D_n$, that there are $\binom{n}{r}$ ways of choosing $(i_1 < \cdots < i_r)$ in $[1:r]$, and that $|\mathcal{P}_{i_1} \cap \cdots \cap \mathcal{P}_{i_r}| = (n-r)!$ since a permutation of $[1:n]$ that fixes $i_1, \ldots, i_r$ is equivalent to a permutation of $[1:n] \setminus \{i_1, \ldots, i_r\}$. This yields

$$n! - D_n = \sum_{r=1}^{n} (-1)^{r-1} \binom{n}{r} (n-r)! = \sum_{r=1}^{n} (-1)^{r-1} \frac{n!}{r!}.$$

It suffices to rearrange the latter to obtain the announced formula. □

3 Exercises

Ex.1: What is the maximum size of a subset of $\{3, 11, 19, 27, \ldots, 139, 147, 155\}$ for which no two elements add up to 158?

Ex.2: Prove that at any party there are two guests with the same number of friends present.

Ex.3: How many numbers between 1 and 2012 (inclusive) are not divisible by 2, 3, 5, and 7?

Ex.4: Let S be a subset of $[1:100]$ of size 10. Show that there are two subsets of S for which the sums of the elements are the same.

Ex.5: Calculate the generating function of the sequence $(D_n/n!)$, i.e., the formal power series

$$\sum_{n \ge 0} \frac{D_n}{n!} z^n.$$

Ex.6: Show that, if a collar made of n pearls has more than $(k-1)n/k$ white pearls, then there is a string of k consecutive white pearls.

Ex.7: Given real numbers $x_1, x_2, \ldots, x_n$, prove that

$$\max\{x_1, \ldots, x_n\} = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \le i_1 < \cdots < i_r \le n} \min\{x_{i_1}, \ldots, x_{i_r}\}.$$

Ex.8: Prove that, in a party with n guests, one can find two guests such that at least half of the remaining guests know either both or neither of them.

Ex.9: How many matrices in $\{0,1\}^{m \times n}$ have no row and no column consisting only of zeros?

Ex.10: Show that the number of permutations σ of $[1:n]$ with $\sigma(i+1) \neq \sigma(i)$ for all $i \in [1:n]$ equals $D_n + D_{n-1}$.

Sequences and Series

1 Some convergence criteria

For real-valued sequences,

- If a nondecreasing sequence is bounded above, then it is convergent;
- A sequence $(u_n)_{n\geq 0}$ is convergent if and only if it is a Cauchy sequence, i.e.,

$$\forall \varepsilon, \exists n_0 \geq 0 : \forall n \geq n_0, \forall p \geq 0, |u_{n+p} - u_n| < \varepsilon.$$

The counterparts for real-valued series are:

- If $u_k \geq 0$ for all k and if $\sum_{k=0}^{n} u_k \leq U$ for some $U > 0$ and all n, then $\sum_{k=0}^{\infty} u_k$ converges;
  [Consequence: if $0 \leq u_k \leq v_k$ and $\sum_k v_k$ converges, then $\sum_k u_k$ converges, too.]
- A series $\sum_{k=0}^{\infty} u_k$ converges if and only if

$$\forall \varepsilon > 0, \exists n_0 \geq 0 : \forall n \geq n_0, \forall p \geq 0, \left| \sum_{k=n}^{n+p} u_k \right| < \varepsilon.$$

  [Consequence: If $\sum_k |u_k|$ converges (absolute convergence), then $\sum_k u_k$ converges, too.]

The latter comparison criterion is useful when classical series are at our disposal, such as:

$$\text{(1)} \qquad \begin{array}{ll} \sum_n x^n \text{ converges iff } |x| < 1, & \sum_n \frac{1}{n^a} \text{ converges iff } a > 1, \\ & \sum_n \frac{1}{n \ln^a(n)} \text{ converges iff } a > 1. \end{array}$$

From there, we can deduce the ratio test: if a positive sequence satisfies $u_{n+1}/u_n \underset{n\to\infty}{\longrightarrow} \rho$ with $0 \leq \rho < 1$, then $\sum_k u_k$ converges. Once can also prove, for instance, that the series $\sum 1/(n(n+1))$ converges by remarking that $0 \leq 1/(n(n+1)) \leq 1/n^2$ and that $\sum 1/n^2$ converges, which implies that $\sum 1/(n(n+1))$ converges, too. To find the exact value of the sum, we interpret it as a telescoping sum, i.e.,

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1} \Longrightarrow \sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots = 1.$$

Another useful technique is the summation by parts, consisting of the following manipulation: given $(a_k)_{k\geq 1}$ and $(b_k)_{k\geq 1}$, define $A_0 = 0$, $A_n = \sum_{k=1}^{n} a_k$ for $n \geq 1$, and write

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} (A_k - A_{k-1}) b_k = \sum_{k=1}^{n} A_k b_k - \sum_{k=1}^{n} A_{k-1} b_k = \sum_{k=1}^{n} A_k b_k - \sum_{k=0}^{n-1} A_k b_{k+1}$$
$$= A_n b_n + \sum_{k=1}^{n-1} A_k (b_k - b_{k+1}).$$

2 Power series

A power series is a series of the type $\sum u_n z^n$. Here are a few classical ones, together with their regions of absolute convergence in the complex plane:

(2) $$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n = 1 + z + z^2 + z^3 \cdots, \qquad |z| < 1$$

(3) $$\frac{1}{(1-z)^2} = \sum_{n=0}^{\infty} (n+1) z^n = 1 + 2z + 3z^2 + 4z^3 \cdots, \qquad |z| < 1$$

(4) $$\ln(1-z) = \sum_{n=1}^{\infty} \frac{1}{n} z^n = -z - \frac{z^2}{2} - \frac{z^3}{3} - \frac{z^4}{4} \cdots, \qquad |z| < 1$$

(5) $$(1+z)^a = \sum_{n=0}^{\infty} \frac{a(a-1)\cdots(a-n+1)}{n!} z^n = 1 + az + \frac{a(a-1)}{2} z^2 + \frac{a(a-1)(a-2)}{6} z^3 \cdots, \qquad |z| < 1$$

(6) $$\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} z^n = 1 + z + \frac{1}{2} z^2 + \frac{1}{6} z^3 \cdots, \qquad \text{all } z$$

(7) $$\cosh(x) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} z^{2n} = 1 + \frac{1}{2} z^2 + \frac{1}{24} z^4 \cdots \qquad \text{all } z$$

(8) $$\sinh(x) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} z^{2n+1} = z + \frac{1}{6} z^3 + \frac{1}{120} z^5 \cdots \qquad \text{all } z$$

(9) $$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} = 1 - \frac{1}{2} z^2 + \frac{1}{24} z^4 \cdots \qquad \text{all } z$$

(10) $$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} = z - \frac{1}{6} z^3 + \frac{1}{120} z^5 \cdots \qquad \text{all } z$$

Some of these equalities may remain valid at particular points on the boundary of the region of absolute convergence. For instance, (4) holds for $z = -1$, and it then reads

$$\ln(2) = \sum_{n=1}^{\infty} \frac{1}{n} (-1)^n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \cdots.$$

3 Progression and sums

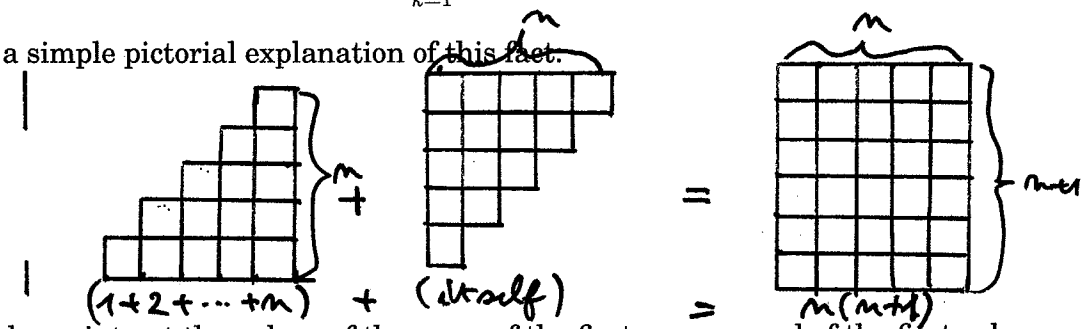
The fundamental identity (2) is a consequence of the general expression for the sum of a geometric progression. A geometric progression is a sequence $(u_k)_{k\geq 0}$ obeying the recurrence relation $u_{k+1} = ru_k$ for some $r \in \mathbb{C}$ and all $k \geq 0$. Alternatively, it can be viewed as the sequence defined by $u_k = r^k u_0$ for all $k \geq 0$. The sum of its first $n+1$ terms is derived from the identity

$$\sum_{k=0}^{n} r^k = \begin{cases} \dfrac{1-r^{n+1}}{1-r} & \text{if } r \neq 1, \\ n+1 & \text{if } r = 1. \end{cases}$$

An arithmetic progression is a sequence $(u_k)_{k\geq 0}$ obeying the recurrence relation $u_{k+1} = u_k + d$ for some $d \in \mathbb{C}$ and all $k \geq 0$. Alternatively, it can be viewed as the sequence defined by $u_k = u_0 + kd$ for all $k \geq 0$. The sum of its first $n+1$ terms is derived from the identity

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

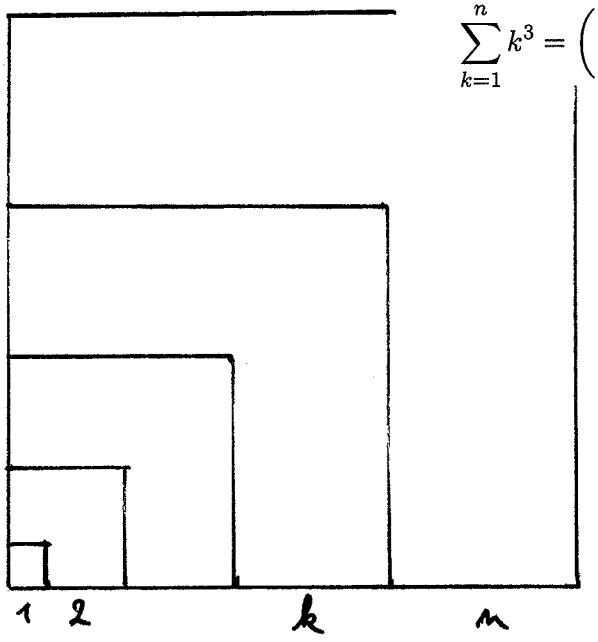
Here is a simple pictorial explanation of this fact.



Let us also point out the values of the sums of the first squares and of the first cubes:

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6},$$

$$\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2.$$



total area $= (1+2+\cdots+n)^2 = \left(\frac{n(n+1)}{2}\right)^2$

$= \sum_{k=1}^{n}$ kth instance

note that the kth instance is

$2 \times k \times (1+2+\cdots+k) - k^2$

$= k\,k(k+1) - k^2 = k^3$.

4 Exercises

Ex.1: Given $t > 0$, consider the sequence defined recursively by $x_0 > 0$ and

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{t}{x_n}\right), \qquad n \geq 0.$$

Does the sequence $(x_n)_{n\geq 0}$ converge, and if so what is the value of its limit?

Ex.2: Let $(u_n)_{n\geq 1}$ be a convergent sequence. Prove that the Cesaro mean $\frac{1}{n}\sum_{k=0}^{n} u_k$ converges to the same limit as the original sequence when $n \to \infty$.

Ex.3: Prove the assertion (1).

Ex.4: Given the value $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, find the values of $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$.

Ex.5: State the power series expansions of $1/(1-z)$, $1/(1+z)^2$, $\ln(1+z)$, and $\sqrt{1+z}$. Derive (3) and (4) from (2) by formal differentiation and integration. Show that (5) reduces to the binomial theorem when a is an integer. Derive (7), (8), (9), and (10) from (6) using the expressions of the (hyperbolic) trigonometric functions in terms of the exponential function.

Ex.6: Prove that the series $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}(k+1) + k\sqrt{k+1}}$ converges and find its value.

Ex.7: Evaluate the sums $\sum_{k=0}^{n} \cos\left(\frac{k\pi}{n}\right)$ and $\sum_{k=0}^{n} \sin\left(\frac{k\pi}{n}\right)$.

Ex.8: Consider a sequence $(u_k)_{k\geq 0}$ obeying the recurrence relation $u_{k+1} = ru_k + d$ for some $r, d \in \mathbb{C}$ and all $k \geq 0$. Find the explicit form for the general term u_k and deduce an expression for the sum $\sum_{k=0}^{n} u_k$.

Ex.9: Consider the power series expansion

$$\frac{1}{1-2z-z^2} = \sum_{n=0}^{\infty} a_n z^n.$$

Prove that, for all $n \geq 0$, there exists $m \geq 0$ such that $a_n^2 + a_{n+1}^2 = a_m$.

Ex.10: Consider partitioning the natural numbers into the sets $\{1\}$, $\{2,3\}$, $\{4,5,6\}$, $\{7,8,9,10\}$, etc. Find a simple expression for the sum of the integers in the nth set.

Ex.11: If $\sum a_n$ is a convergent series of nonnegative terms, prove that $\sum \sqrt[n]{a_1 a_2 \cdots a_n}$ is also convergent.

Induction and Recurrence

1 Mathematical induction

The principle of mathematical induction states that if an assertion $\mathcal{P}_n$ is true for an integer $n = n_0$ (the base case) and if $\mathcal{P}_{n+1}$ is true as soon as $\mathcal{P}_n$ is true (the inductive step), then the assertion $\mathcal{P}_n$ is true for all integers $n \geq n_0$. Equivalently, if an assertion $\mathcal{P}_n$ is true for an integer $n = n_0$ and if $\mathcal{P}_{n+1}$ is true as soon as $\mathcal{P}_n, \mathcal{P}_{n-1}, \ldots, \mathcal{P}_{n_0}$ are true, then the assertion $\mathcal{P}_n$ is true for all integers $n \geq n_0$. To see how the first version implies the second one, apply it to $\mathcal{P}'_n := (\mathcal{P}_{n_0}$ and $\mathcal{P}_{n_0+1}$ and $\ldots$ and $\mathcal{P}_n)$. It is advisable to systematically work with the second version. Sometimes, several base cases may need to be verified (see Section 2, for instance).

The validity of the principle of mathematical induction is a consequence of the well-ordering principle: any nonempty set of nonnegative integers has a minimal element. Indeed, consider the set $\mathcal{S} := \{n \geq n_0 : \mathcal{P}_n \text{ is wrong}\}$. If $\mathcal{S}$ was nonempty, it would have a minimal element n_1, and necessarily $n_1 > n_0$ (by the base case). The minimality implies that $n_1 - 1 \notin \mathcal{S}$, i.e., $\mathcal{P}_{n_1-1}$ is true. But then $\mathcal{P}_{n_1}$ is also true (by the inductive step), meaning that $n_1 \notin \mathcal{S}$. This is a contradiction. Hence $\mathcal{S}$ is empty, or in other words $\mathcal{P}_n$ is true for all $n \geq n_0$.

2 Recurrence relation

Mathematical induction is often used to establish rigorously a statement guessed from the first few cases (or given by the question). For instance, consider a sequence $(u_n)_{n\geq 1}$ given by the values $u_1, u_2, \ldots, u_p$ and the p-term recurrence relation $u_{n+p} = f(u_{n+p-1}, \ldots, u_{n+1}, u_n)$ for $n \geq 1$. One can compute in turn u_{p+1}, next u_{p+2}, then u_{p+3}, etc. If we see a pattern emerging for a closed-form formula, we can justify it using mathematical induction. We have already seen arithmetic and geometric progressions as examples of sequences defined by one-term recurrence relations. We now consider the particular case of a linear function f, i.e.,

$$u_{n+p} = c_{p-1}u_{n+p-1} + \cdots + c_1 u_{n+1} + c_0 u_n, \qquad c_0 \neq 0.$$

If the polynomial $p(z) := z^p - c_{p-1}z^{p-1} - \cdots - c_1 z - c_0$ has distinct roots $r_1, r_2, \ldots, r_n$ (the case of repeated roots can be treated, too), then it is proved below that the general term is given by

$$u_n = \alpha_1 r_1^n + \alpha_2 r_2^n + \cdots + \alpha_p r_p^n \qquad \text{for all } n \geq 1, \tag{1}$$

where the p coefficients $\alpha_1, \alpha_2, \ldots, \alpha_p$ are uniquely determined by the p values of $u_1, u_2, \ldots, u_p$. We proceed by induction on n. The p base cases hold since $\alpha_1, \alpha_2, \ldots, \alpha_p$ are determined precisely for this purpose. Assuming that (1) holds up to $n \geq 1$, let us now prove that it also holds for $n+1$. Using the recurrence relation, the induction hypothesis for $n-1, n-2, \ldots, n-p+1$,

and the fact that $r_j^p = c_{p-1}r_j^{p-1} + \cdots + c_1 r_j + c_0$ for all $j \in [1:p]$, we derive

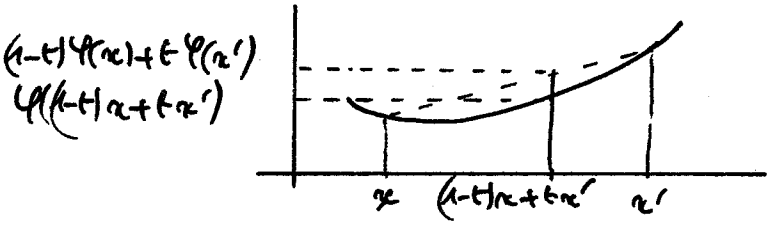
$$u_{n+1} = \sum_{k=1}^{p} c_{k-1}u_{n-p+k} = \sum_{k=1}^{p} c_{k-1} \sum_{j=1}^{p} \alpha_j r_j^{n-p+k} = \sum_{j=1}^{p} \alpha_j r_j^{n-p+1} \sum_{k=1}^{p} c_{k-1} r_j^{k-1} = \sum_{j=1}^{p} \alpha_j r_j^{n-p+1} r_j^p$$
$$= \sum_{j=1}^{p} \alpha_j r_j^{n+1}.$$

This shows that (1) holds for $n+1$. The principle of mathematical induction allows us to conclude that (1) holds for all $n \geq 1$.

3 An application: Jensen's inequality

A function φ defined on an interval I is called convex if

$$\varphi((1-t)x + tx') \leq (1-t)\varphi(x) + t\varphi(x') \qquad \text{for all } x, x' \in I \text{ and all } t \in [0,1].$$



Given a convex function φ on an interval I, Jensen's inequality says that the image of a convex combination is smaller than or equal to the convex combination of the images. Precisely, if $x_1, \ldots, x_n \in I$ and if $t_1, \ldots, t_n \geq 0$ satisfy $t_1 + \cdots + t_n$, then

$$\varphi\Big(\sum_{j=1}^{n} t_j x_j\Big) \leq \sum_{j=1}^{n} t_j \varphi(x_j). \tag{2}$$

This can be proved by induction on n. Indeed, in the base case $n = 1$, (2) holds with equality. Let us now assume that (2) holds up to an integer $n-1$, $n \geq 2$, and let us prove that it also holds for the integer n. To this end, consider $x_1, \ldots, x_n \in I$ and $t_1, \ldots, t_n \geq 0$ with $t_1 + \cdots + t_n = 1$. If $t_n = 1$, then all other t_j are zero, and (2) holds with equality. So we may assume that $t_n < 1$, and we set $t'_j := t_j/(1-t_n) \geq 0$ for $j \in [1:n-1]$. Notice that $\sum_{j=1}^{n-1} t'_j = (\sum_{j=1}^{n} t_j)/(1-t_n) = 1$. Applying the defining property of a convex function and then the induction hypothesis, we get

$$\varphi\Big(\sum_{j=1}^{n} t_j x_j\Big) = \varphi\Big(\sum_{j=1}^{n-1} t_j x_j + t_n x_n\Big) = \varphi\Big((1-t_n)\sum_{j=1}^{n-1} t'_j x_j + t_n x_n\Big)$$
$$\leq (1-t_n)\varphi\Big(\sum_{j=1}^{n-1} t'_j x_j\Big) + t_n \varphi(x_n) \leq (1-t_n)\sum_{j=1}^{n-1} t'_j \varphi(x_j) + t_n \varphi(x_n) = \sum_{j=1}^{n} t_j \varphi(x_j).$$

This shows that (2) holds for n. The principle of mathematical induction allows us to conclude that (1) holds for all $n \geq 1$.

4 Exercises

Ex.1: Prove that, for any integer $n \geq 1$,

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Ex.2: For $n \geq 1$, prove that $n(n-1)(n+1)(3n+2)$ is divisible by 24.

Ex.3: Find the number $R(n)$ of regions in which the plane can be divided by n straight lines.

Ex.4: The Fibonacci sequence is defined by $F_0 = 0$, $F_1 = 1$, and $F_{n+1} = F_n + F_{n-1}$ for $n \geq 1$. Find a closed-form formula for F_n involving the golden ration $\phi := (1+\sqrt{5})/2$.

Ex.5: Guess a formula for the sum $1^4 + 2^4 + \cdots + n^4$, and then provide a rigorous proof (Hint: the answer is a polynomial in n, it has degree 5, and it is divisible by $n(n+1)(2n+1)$).

Ex.6: Given $r \in \mathbb{R}$, consider the two-term recurrence relation $u_{n+2} = 2ru_{n+1} - r^2 u_n$ (note that the polynomial $p(z) = z^2 - 2rz + r^2$ has a double root at r). Prove that $u_n = \alpha r^n + \beta n r^n$ for all $n \geq 1$, where the coefficients α, β are uniquely determined by the values of u_1, u_2.

Ex.7: Find a formula for the general term of the sequence defined by $u_1 = 3$ and the recurrence relation $u_{n+1} = u_n(u_n + 2)$ for $n \geq 1$.

Ex.8: For $n \geq 2$, prove that any $2n$ points joined by at least $n^2 + 1$ segments contain at least one triangle. Show that this is not true if the number of segments is n^2.

Ex.9: For integers $n, d \geq 0$, prove the relation

$$\binom{n}{n} + \binom{n+1}{n} + \binom{n+2}{n} + \cdots + \binom{n+d}{n} = \binom{n+d+1}{n}.$$

Ex.10: Consider a $2^n \times 2^n$ checkerboard from which an arbitrary square has been removed. Can it be paved with polyominos, that is L-shaped tiles covering three squares each?

Ex.11: In Section 2, we have used the fact that if $r_1, \ldots, r_p$ are distinct nonzero numbers, then the system of p linear equations $\alpha_1^n r_1^n + \cdots + \alpha_p^n r_p^n = u_n$, $n \in [1:p]$, in the p unknowns $\alpha_1, \ldots, \alpha_p$ has a unique solution. In linear algebra terms, this condition is equivalent to the invertibility of the matrix whose (i,j)th entry is r_j^i. Establish this invertibility by proving the formula for the Vandermonde determinant, i.e.,

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ r_1 & r_2 & \cdots & r_p \\ \vdots & \cdots & \ddots & \vdots \\ r_1^{p-1} & r_2^{p-1} & \cdots & r_p^{p-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (r_j - r_i).$$

Ex.12: Given a convex function φ on an interval I and given $x_0 \in I$, prove that the slope $x \in I \mapsto (f(x) - f(x_0))/(x - x_0)$ is an increasing function of $x \in I$.

Generating Functions

1 Definition and first examples

Generating functions offer a convenient way to carry the totality of the information about a sequence in a condensed form. Precisely, the (ordinary) generating function of the sequence $(a_n)_{n\geq 0}$ is defined as the formal power series

$$\sum_{n=0}^{\infty} a_n z^n.$$

For instance, the power series of the constant sequence $(1)_{n\geq 0}$ is $\sum_{n=0}^{\infty} z^n = 1/(1-z)$. From there, k successive differentiations lead to the generating of the sequence $\left(\binom{n+k}{k}\right)_{n\geq 0}$:

$$\sum_{n=0}^{\infty} \binom{n+k}{k} z^n = \frac{1}{(1-z)^{k+1}}.$$

A more striking illustration concerns the number p_n of partitions of an integer n, i.e., the number of ways to write it as the sum of an nondecreasing sequence. For instance, $p_4 = 5$, since 4 can be written as $1+1+1+1 = 1+1+2 = 1+3 = 2+2 = 4$. Although there is no simple form for the sequence (p_n), its generating function admits a nice expression (uncovered by Euler), namely

$$\sum_{n=0}^{\infty} p_n z^n = \prod_{k=1}^{\infty} \frac{1}{1-z^k}.$$

This can be understood by looking at the coefficient of z^n in the right-hand side expressed as

$$(1+z+z^2+\cdots)(1+z^2+z^4+\cdots)(1+z^3+z^6+\cdots)\cdots.$$

Indeed, the coefficient of z^n is the number of ways to write

$$n = n_1 + 2n_2 + 3n_3 + \cdots = (1+\cdots+1) + (2+\cdots+2) + (3+\cdots+3) + \cdots,$$

which is precisely p_n.

2 Two classics: Fibonacci and Catalan

Sometimes, the cumbersome determination of the general term of a sequence can be shortcut by an argument exploiting generating functions. As a first example, consider the Fibonacci numbers defined by $F_0 = 1$, $F_1 = 1$, and

$$F_{n+2} = F_{n+1} + F_n \qquad \text{for } n \geq 0. \tag{1}$$

Let $f(z) := \sum_{n=0}^{\infty} F_n z^n$ denote the generating function of $(F_n)_{n\geq 0}$. Multiplying (1) by z^{n+2} and summing over all $n \geq 0$, we obtain

$$f(z) - F_0 - F_1 z = z f(z) - F_0 z + z^2 f(z), \qquad \text{i.e.,} \qquad f(z) = \frac{z}{1-z-z^2}.$$

Since $1 - z - z^2 = (1 - \phi z)(1 + z/\phi)$, where $\phi = (1 + \sqrt{5})/2$, we derive the partial fraction decomposition (remember to multiply through by $1 - \phi z$ and to take the value $z = 1/\phi$, next to multiply through by $1 + z/\phi$ and to take the value $z = -\phi$)

$$\frac{z}{1 - z - z^2} = \frac{1/\sqrt{5}}{1 - \phi z} - \frac{1/\sqrt{5}}{1 + z/\phi}.$$

Calling upon known power series expansions, we deduce

$$f(z) = \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty} (\phi z)^n - \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty} (-z/\phi)^n = \sum_{n=0}^{\infty} \frac{\phi^n - (-1/\phi)^n}{\sqrt{5}} z^n.$$

By identifying the coefficients of z^n, we conclude that

$$F_n = \frac{\phi^n - (-1/\phi)^n}{\sqrt{5}}.$$

As a second example, consider the Catalan numbers C_n defined (among many alternative definitions) as the numbers of binary trees that possess n branching nodes (hence $n+1$ leaves). Starting from $C_0 = 1$, they obey the recurrence relation

$$C_{n+1} = \sum_{i+j=n} C_i C_j, \qquad n \geq 0. \tag{2}$$

This translates the fact that a binary tree with $n + 1$ branching nodes is decomposed, when the root is removed, as two binary trees with i and j branching nodes satisfying $i + j = n$. Let $f(z) := \sum_{n=0}^{\infty} C_n z^n$ be the generating function of the Catalan numbers. Multiplying (2) by z^{n+1} and summing over all $n \geq 0$ leads to

$$f(z) - 1 = \sum_{n=0}^{\infty} \left(\sum_{i+j=n} C_i C_j \right) z^{n+1} = z \left(\sum_{i=0}^{\infty} C_i z^i \right) \left(\sum_{j=0}^{\infty} C_j z^j \right) = z f(z)^2.$$

Solving this quadratic equation in $f(z)$ gives (note that the second solution is rejected in view of its value at $z = 0$)

$$f(z) = \frac{1 - \sqrt{1 - 4z}}{2z}.$$

Calling upon known power series expansions, we deduce

$$f(z) = \frac{1}{2z} \left(-\sum_{n=1}^{\infty} \frac{(1/2)(-1/2)(-3/2) \cdots (1/2 - n + 1)}{n!} (-4z)^n \right) = \sum_{n=1}^{\infty} 2^{n-1} \frac{1 \cdot 3 \cdot \cdots \cdot (2n-3)}{n!} z^{n-1}$$

$$= \sum_{n=0}^{\infty} 2^n \frac{1 \cdot 3 \cdot \cdots \cdot (2n-1)}{(n+1)\, n!} z^n = \sum_{n=0}^{\infty} \frac{1 \cdot 2 \cdot 3 \cdot \cdots \cdot (2n-1) \cdot (2n)}{(n+1)\, n!\, n!} z^n = \sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} z^n.$$

By identifying the coefficients of z^n, we conclude that

$$C_n = \frac{1}{n+1} \binom{2n}{n}.$$

3 Stirling numbers

The Stirling numbers of the second kind, denoted $\left\{ {n \atop k} \right\}$, count the number of ways to partition $[1:n]$ into k nonempty blocks. For instance, $\left\{ {4 \atop 3} \right\} = 6$, since $\{1,2,3,4\}$ can be partitioned as $\{1\} \cup \{2\} \cup \{3,4\}$, $\{1\} \cup \{3\} \cup \{2,4\}$, $\{1\} \cup \{4\} \cup \{2,3\}$, $\{2\} \cup \{3\} \cup \{1,4\}$, $\{2\} \cup \{4\} \cup \{1,3\}$, and $\{3\} \cup \{4\} \cup \{1,2\}$. Note that $\left\{ {n \atop k} \right\} = 0$ for $k > n$ and for $k \le 0$ (unless $n = 0$, in which case the convention $\left\{ {0 \atop 0} \right\} = 1$ is used). The Stirling numbers of the second kind obey the recurrence relation

$$\left\{ {n+1 \atop k} \right\} = \left\{ {n \atop k-1} \right\} + k \left\{ {n \atop k} \right\}. \tag{3}$$

This translates the fact that, when partitioning $[1:n+1]$ into k blocks, the element $n+1$ either forms a block on its own, leading to $k-1$ blocks that partition $[1:n]$, or it joins one of k blocks that partition $[1:n]$. For $k \ge 0$, consider the generating function $f_k(z) := \sum_{n=0}^{\infty} \left\{ {n \atop k} \right\} z^n$. With $k \ge 1$, multiplying (3) by z^{n+1} and summing over all $n \ge 0$ leads to

$$f_k(z) = z f_{k-1}(z) + kz f_k(z), \qquad \text{i.e.,} \qquad f_k(z) = \frac{z}{1-kz} f_{k-1}(z).$$

In view of $f_0(z) = 1$, we obtain by immediate induction

$$f_k(z) = \frac{z^k}{(1-z)(1-2z)\cdots(1-kz)}.$$

The partial fraction decomposition of the latter is

$$f_k(z) = c_0 + \frac{c_1}{1-z} + \cdots + \frac{c_k}{1-kz}.$$

We have $c_0 = \lim_{n\to\infty} f_k(z) = (-1)^k/k!$ and, for $j \in [1:k]$,

$$c_j = [f_k(z)(1-jz)]_{|z=1/j} = \frac{(1/j)^k}{(1-1/j)\cdots(1-(j-1)/j)(1-(j+1)/j)\cdots(1-k/j)} = \frac{(-1)^{k-j}}{j!(k-j)!}.$$

In conjunction with

$$f_k(z) = \sum_{j=0}^{k} \frac{c_j}{1-jz} = \sum_{j=0}^{k} c_j \sum_{n=0}^{\infty} (jz)^n = \sum_{n=0}^{\infty} \sum_{j=0}^{k} c_j j^n z^n,$$

we conclude that

$$\left\{ {n \atop k} \right\} = \sum_{j=0}^{k} \frac{(-1)^{k-j}}{j!(k-j)!} j^n = \sum_{j=0}^{k} \frac{(-1)^{k-j}}{k!} \binom{k}{j} j^n.$$

The Stirling numbers of the first kind, denoted $\left[{n \atop k} \right]$, count the number of partitions of $[1:n]$ with k cycles. For instance, $\left[{3 \atop 2} \right] = 3$, since $(1)(32)$, $(2)(13)$, and $(3)(12)$ are the three permutations of $\{1,2,3\}$ with two cycles. Note that $\left[{n \atop k} \right] = 0$ for $k > n$ and for $k \le 0$ (unless $n = 0$, in

which case the convention $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$ is used). The Stirling numbers of the first kind obey the recurrence relation

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = \begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix}. \tag{4}$$

This translates the fact that, when considering a partition of $[1 : n+1]$ with k cycles, the element $n+1$ either forms a cycle on its own, leading to a permutation of $[1 : n]$ with $k-1$ cycles, or it incorporates (at one of n possible positions) one of k cycles making a partition of $[1 : n]$. For $k \geq 0$, consider now the exponential generating function of $\left(\begin{bmatrix} n \\ k \end{bmatrix}\right)_{n \geq 0}$ given by

$$f_k(z) := \sum_{n=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} \frac{z^n}{n!}.$$

Multiplying (4) by $z^n/n!$ and summing over all $n \geq 0$ leads to

$$f_k'(z) = f_{k-1}(z) + z f_k'(z), \qquad \text{i.e.,} \qquad f_k'(z) = \frac{f_{k-1}(z)}{1-z}.$$

In view of $f_0(z) = 1$, we obtain by immediate induction

$$f_k(z) = \frac{1}{k!} \ln^k \left(\frac{1}{1-z} \right). \tag{5}$$

4 Exercises

Ex.1: Find the sequence $(a_n)_{n \geq 0}$ given by $a_0 = 1$ and $a_n = \dfrac{1 - \sum_{k=1}^{n-1} a_k a_{n-k}}{2}$ for $n \geq 1$.

Ex.2: Find the number of different ways a convex polygon with $n+2$ sides can be cut into triangles by connecting vertices with straight lines.

Ex.3: Prove that the number of partitions of an integer into odd positive integers equals the number of its partitions into distinct positive integers.

Ex.4: It follows from the definition of the Stirling numbers of the first kind that $\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} = n!$. Recover this fact from the expression (5) of the exponential generating function.

Ex.5: The positive differences of the four numbers $0, 2, 5, 6$ are the numbers $1, 2, 3, 4, 5, 6$, each taken exactly once. Prove that this phenomenon cannot occur if there are more than four numbers.

Complex Analysis

1 The field of complex numbers

The set of complex numbers is denoted by $\mathbb{C}$. The cartesian representation of $z \in \mathbb{C}$ is $z = x+iy$ with $x, y \in \mathbb{R}$ and $i^2 = -1$. The real and imaginary parts of z are $\mathrm{Re}\,(z) = x$ and $\mathrm{Im}\,(z) = y$, respectively. Addition and multiplication of complex numbers (defined in a predictable way) satisfy all the properties we would have expected — meaning that $\mathbb{C}$ is a field. The polar representation of $z \in \mathbb{C}$ is $z = re^{i\theta}$ with $r \geq 0$ and $\theta \in \mathbb{R}$. We call $r = |z|$ the modulus of z and $\theta = \arg(z)$ — not necessarily unique — an argument of z. We have $r = \sqrt{x^2+y^2}$ and $\tan(\theta) = y/x$. De Moivre's theorem states that $(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$, or in simplified form, that $(e^{i\theta})^n = e^{in\theta}$ — this uses Euler formula $e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)$. Note also the identities $\cos(\theta) = (e^{i\theta} + e^{-i\theta})/2$ and $\sin(\theta) = (e^{i\theta} - e^{-i\theta})/(2i)$. In general, one has $\mathrm{Re}\,(z) = (z+\bar{z})/2$, $\mathrm{Im}\,(z) = (z-\bar{z})/(2i)$, and $|z|^2 = z\bar{z}$. Here $\bar{z} = x - iy = re^{-i\theta}$ is the complex conjugate of z. The fundamental theorem of algebra ensures that every nonconstant polynomial $p(z) = a_n z^n + \cdots + a_1 z + a_0$ has a complex roots (in turn, that every polynomial with complex coefficients has all its roots in $\mathbb{C}$, i.e., $\mathbb{C}$ is algebraically closed).
A possible argument goes along those lines: pick $z_0 \in \mathbb{C}$ such that $|p(z_0)| = \min_{z\in\mathbb{C}} |p(z)|$ and suppose $|p(z_0)| > 0$; write that p equals its Taylor polynomial at z_0, i.e., $p(z_0) + \sum_{j=k}^n b_j (z-z_0)^j$ where $b_k \neq 0$; note that $\sum_{j=k+1}^n |b_j|\rho^j < |b_k|\rho^k < |p(z_0)|$ for $\rho > 0$ sufficiently small; observe that $p(z_0) + b_k(z-z_0)^k$ describes k times the circle $\{|\zeta - p(z_0)| = |b_k|\rho^k\}$ when z describes the circle $\{|z-z_0| = \rho\}$, hence there exists z_1 with $|z_1 - z_0| = \rho$ such that $p(z_0) + b_k(z_1-z_0)^k$ lies between 0 and $p(z_0)$, so that $|p(z_0) + b_k(z_1-z_0)^k| = |p(z_0)| - |b_k|\rho^k$; derive a contradiction from

$$|p(z_1)| \leq |p(z_0) + b_k(z_1-z_0)^k| + |\sum_{j=k+1}^n b_j(z_1-z_0)^j| \leq |p(z_0)| - |b_k|\rho^k + \sum_{j=k+1}^n |b_j|\rho^j < |p(z_0)|.$$

Another possible argument involves Cauchy formula for holomorphic functions (see below): suppose that p does not vanish on $\mathbb{C}$, so that $q = 1/p$ is holomorphic on $\mathbb{C}$; for $R > 0$ sufficiently large to have $|p(z)| \geq (|a_n| - |a_{n-1}|/|z| - \cdots |a_0|/|z|^n)|z|^n \geq |a_n||z|^n/2$ whenever $|z| = R$, a contradiction follows from

$$0 < |q(0)| = \left|\frac{1}{2\pi i}\oint_{|z|=R} \frac{q(z)dz}{z}\right| \leq \frac{1}{2\pi}\oint_{|z|=R} \frac{dz}{|z||p(z)|} \leq \frac{1}{2\pi}\oint_{|z|=R} \frac{2\,dz}{|a_n|R^{n+1}} = \frac{2}{|a_n|R^n} \underset{R\to\infty}{\longrightarrow} 0.$$

2 Holomorphic functions

A function f defined on an open subset of $\mathbb{C}$ is differentiable at z_0 if one can make sense of

$$f'(z) = \lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

In particular, the limit is independent of how z_0 is approached. If the function f of the variable $z = x + iy$ is differentiable at $z_0 = x_0 + iy_0$, then it satisfies the Cauchy–Riemann equations

$$\frac{\partial \mathrm{Re}\, f}{\partial x}(x_0, y_0) = \frac{\partial \mathrm{Im}\, f}{\partial y}(x_0, y_0) \qquad \text{and} \qquad \frac{\partial \mathrm{Re}\, f}{\partial y}(x_0, y_0) = -\frac{\partial \mathrm{Im}\, f}{\partial x}(x_0, y_0).$$

A converse holds provided the first-order partial derivatives are continuous.

A function f is called holomorphic at z_0 if it is differentiable in some neighborhood of z_0 (i.e., whenever $|z - z_0| < r$ for some $r > 0$). Every power series $\sum_{n=0}^{\infty} c_n(z - z_0)^n$ with radius of convergence $R > 0$ defines a holomorphic function on $\{|z - z_0| < R\}$. Conversely, every holomorphic function is analytic, i.e., locally representable by powers series (hence holomorphic and analytic are synonymous terms for complex functions). This fact shows that holomorphic functions are infinitely differentiable and that their zeros are isolated (unless the function vanishes everywhere).

Let G be a simply connected open region, let γ be a simple closed path oriented counterclockwise and contained in G, and let $z_0 \in \mathbb{C}$ be inside γ. If f is holomorphic in G, then it satisfies Cauchy integral formulas

$$\text{(1)} \quad \oint_\gamma f(z)dz = 0, \qquad \oint_\gamma \frac{f(z)dz}{z - z_0} = 2\pi i f(z_0), \qquad \oint_\gamma \frac{f(z)dz}{(z - z_0)^n} = 2\pi i f^{(n)}(z_0) \ \text{ for all integer } n \geq 0.$$

Cauchy formula implies Liouville's theorem, which states that a function f holomorphic and bounded on $\mathbb{C}$ is constant. Indeed, if γ is the circular contour oriented counterclockwise with center 0 and radius R large enough so that $|z - z_0|, |z - z_1| \geq R/2$, then, for all $z_0, z_1 \in \mathbb{C}$,

$$\begin{aligned} |f(z_0) - f(z_1)| &= \left| \frac{1}{2\pi i} \oint_\gamma f(z) \left(\frac{1}{z - z_0} - \frac{1}{z - z_1} \right) dz \right| = \left| \frac{z_0 - z_1}{2\pi i} \oint_\gamma \frac{f(z)}{(z - z_0)(z - z_1)} dz \right| \\ &\leq \frac{|z_0 - z_1|}{2\pi} \oint_\gamma \frac{\max(|f|)}{(R/2)^2} dz = \frac{4|z_0 - z_1| \max(|f|)}{R} \underset{R \to \infty}{\longrightarrow} 0. \end{aligned}$$

Cauchy formula also implies the maximum principle, which sates that, if f is homomorphic on $\{|z - z_0| \leq r\}$, then

$$\max_{|z - z_0| \leq r} |f(z)| = \max_{|z - z_0| = r} |f(z)|.$$

3 Meromorphic functions

If a function is holomorphic on an annulus $A = \{r < |z - z_0| < R\}$ for some $R > r \geq 0$, then f has a unique Laurent expansion at z_0 of the form

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n, \qquad z \in A.$$

A function f holomorphic in some punctured neighborhood of z_0 (i.e., an annulus where $r = 0$) but not at z_0 is said to have an isolated singularity at z_0. These can be of three different kinds: removable singularity if $c_n = 0$ for all $n < 0$ (for instance $\sin(z)/z$ at $z_0 = 0$), poles if $c_{-m} \neq 0$ and $c_n = 0$ for all $n < -m$, in which case m is called the order of the pole (for instance rational functions at z_0 equal to a zero of the denominator), and essential singularities if $\inf\{n : c_n \neq 0\} = -\infty$. A function which is holomorphic in an open subset G of $\mathbb{C}$ except possibly for poles is said to be meromorphic in G.

Let G be a a simply connected open region and let γ be a simple closed path oriented counterclockwise and contained in G. Cauchy residue theorem states that, if f is meromorphic in G with all its poles $z_1, \ldots, z_N$ inside γ, then

$$\oint_\gamma f(z)dz = 2\pi i \sum_{k=1}^{N} \operatorname{Res}(f, z_k),$$

where the residue $\operatorname{Res}(f, z_k)$ of f at z_k is defined as the coefficient c_{-1} of $(z - z_k)^{-1}$ in the Laurent expansion of f at z_k. It follows that, if f is holomorphic on G and does not vanish on γ, then the number of zeros of f inside γ equals $\dfrac{1}{2\pi i}\oint_\gamma \dfrac{f'(z)}{f(z)}dz$. From here, we can deduce Rouché's theorem which states that, if f and g are holomorphic in G and if $|f(z)| > |g(z)|$ on γ, then f and $f + g$ have the same number of zeros (counting multiplicity) inside γ.

4 Exercises

Ex.1: Find the set of all $z \in \mathbb{C}^n$ such that $|z| + |z + 1| = 2$.

Ex.2: Prove the identity

$$\cos^n(\theta) = \frac{1}{2^n}\sum_{k=0}^{n} \binom{n}{k} \cos((n - 2k)\theta).$$

Ex.3: Prove the necessity of the Cauchy–Riemann equations.

Ex.4: Establish the fundamental integral

$$\oint_{\gamma(z_0,r)} (z - z_0)^n dz = \begin{cases} 0 & \text{if } n \neq -1, \\ 2\pi i & \text{if } n = -1, \end{cases}$$

where $\gamma(z_0, r)$ denotes the circular contour oriented counterclockwise with center z_0 and radius r. Derive (informally) formulas (1) with $\gamma = \gamma(z_0, r)$ for analytic functions.

Ex.5: Use the maximum principle to prove Schwarz lemma: if f is holomorphic on $\{|z| = 1\}$, if $M := \max\limits_{|\zeta|=1} |f(\zeta)|$, and if $f(0) = 0$, then $|f(z)| \leq M\,|z|$ whenever $|z| \leq 1$.

Ex.6: Use Cauchy residue theorem to evaluate

$$\oint_\gamma \frac{dz}{1 + z^4},$$

where γ is the semicircle $\{|z| = R, \operatorname{Im}(z) \geq 0\} \cup [-R, R]$ oriented counterclockwise. Deduce the value of the integral

$$\int_0^\infty \frac{dx}{1 + x^4}.$$

Classical Inequalities

Arithmetic-geometric means: The arithmetic mean $(a+b)/2$ of two nonnegative numbers a and b is always larger than or equal to its geometric mean $\sqrt{ab}$, with equality if and only if $a=b$. This can be seen from $a+b-2\sqrt{ab}=(\sqrt{a}-\sqrt{b})^2\geq 0$. The inequality generalizes to more than two numbers: for all $a_1,a_2,\ldots,a_n\geq 0$,

$$\frac{a_1+a_2+\cdots+a_n}{n}\geq \sqrt[n]{a_1a_2\cdots a_n},$$

with equality if and only if $a_1=a_2=\cdots=a_n$. A weighted version involves weights $w_1,w_2,\ldots,w_n$ not all equal to $1/n$. Namely, given $w_1,w_2,\ldots,w_n>0$ with $w_1+w_2+\cdots+w_n=1$, for all $a_1,a_2,\ldots,a_n\geq 0$,

$$\sum_{i=1}^n w_i a_i\geq \prod_{i=1}^n a_i^{w_i}, \tag{1}$$

with equality if and only if $a_1=a_2=\cdots=a_n$. This can be proved as follows.
Set $G:=a_1^{w_1}a_2^{w_2}\cdots a_n^{w_n}$ and $A:=w_1a_1+w_2a_2+\cdots+w_na_n$. Assume without loss of generality that $a_1\leq a_2\leq\cdots\leq a_n$. Since $a_1\leq G\leq a_n$, we consider the integer $k\in[1:n-1]$ such that $a_k\leq G\leq a_{k+1}$. Then one can write

$$\sum_{i=1}^k w_i\int_{a_i}^G\left(\frac{1}{x}-\frac{1}{G}\right)dx+\sum_{i=k+1}^n w_i\int_G^{a_i}\left(\frac{1}{G}-\frac{1}{x}\right)dx\geq 0. \tag{2}$$

It follows that

$$\sum_{i=1}^n w_i\int_G^{a_i}\frac{dx}{G}\geq\sum_{i=1}^n w_i\int_G^{a_i}\frac{dx}{x},\qquad\text{i.e.,}\quad \frac{A}{G}-1\geq\sum_{i=1}^n w_i\ln\frac{a_i}{G}=0,$$

as desired. Equality throughout means equality in (2), i.e., $a_1=\cdots=a_k=a_{k+1}=\cdots=a_n=G$.

Cauchy–Schwarz inequality: For all real numbers $a_1,\ldots,a_n,b_1,\ldots,b_n$,

$$\Big(\sum_{j=1}^n a_jb_j\Big)^2\leq\Big(\sum_{j=1}^n a_j^2\Big)\Big(\sum_{j=1}^n b_j^2\Big),$$

with equality if and only if $a_1=b_1,a_2=b_2,\ldots,a_n=b_n$. Cauchy–Schwarz inequality extends to other situations, for instance we can replace sums by integrals and obtain, for all real-valued functions f,g that are continuous on $[a,b]$,

$$\Big(\int_a^b f(x)g(x)dx\Big)^2\leq\Big(\int_a^b f(x)^2dx\Big)\Big(\int_a^b g(x)^2dx\Big),$$

with equality if and only if $f=g$.

Hölder inequality: This is a generalization of Cauchy–Schwarz inequality to all $p, q > 1$ satisfying $1/p + 1/q = 1$ rather than $p = q = 2$. It reads, for all real numbers $a_1, \dots, a_n, b_1, \dots, b_n$,

$$\sum_{j=1}^{n} a_j b_j \le \Big(\sum_{j=1}^{n} |a_j|^p \Big)^{1/p} \Big(\sum_{j=1}^{n} |b_j|^q \Big)^{1/q},$$

with equality if and only if $a_1 = b_1, a_2 = b_2, \dots, a_n = b_n$. The integral version reads, for all real-valued functions f, g that are continuous on $[a, b]$,

$$\int_a^b f(x) g(x) dx \le \Big(\int_a^b |f(x)|^p dx \Big)^{1/p} \Big(\int_a^b |g(x)|^q dx \Big)^{1/q},$$

with equality if and only if $f = g$. For the proof, set $u_j = |a_j|/A$ where $A := \big(\sum_{j=1}^n |a_j|^p \big)^{1/p}$ and $v_j = |b_j|/B$ where $B := \big(\sum_{j=1}^n |b_j|^q \big)^{1/q}$. Notice that is is enough to prove that $\sum_{j=1}^n u_j v_j \le 1$, knowing that $u_1, \dots, u_n, v_1, \dots, v_n \ge 0$, $\sum_{j=1}^n u_j^p = 1$, and $\sum_{j=1}^n v_j^q = 1$. In turn, it is enough to prove that $uv \le u^p/p + v^q/q$ for all $u, v \ge 0$ — this is known as Young's inequality. To justify the latter, rewrite it as $1 \le u^{p-1} v^{-1}/p + (p-1) u^{-1} v^{1/(p-1)}/p$, or, with $t := u^{-1} v^{1/(p-1)}$, as $t^{-(p-1)} + (p-1)t - 1 \ge 0$. This can now be seen by studying the variations of the function $f(x) := x^{-(p-1)} + (p-1)x - 1$ on $[0, \infty)$.

Jensen inequality: Let φ be a convex function on an interval I — if φ is twice differentiable, this means that $\varphi''(x) \ge 0$ for all $x \in I$. We have seen in 'Induction and Recurrence' that, if $x_1, \dots, x_n \in I$ and if $t_1, \dots, t_n \ge 0$ satisfy $t_1 + \cdots + t_n = 1$, then

$$\varphi\Big(\sum_{j=1}^{n} t_j x_j \Big) \le \sum_{j=1}^{n} t_j \varphi\big(x_j\big). \tag{3}$$

The integral version of Jensen inequality reads

$$\varphi\Big(\frac{1}{b-a} \int_a^b f(x) dx \Big) \le \frac{1}{b-a} \int_a^b \varphi(f(x)) dx \tag{4}$$

for any continuous fonction f on $[a, b]$.

Chebyshev inequality: If $a_1 \le a_2 \le \cdots \le a_n$ and $b_1 \le b_2 \le \cdots \le b_n$ or $a_1 \ge a_2 \ge \cdots \ge a_n$ and $b_1 \ge b_2 \ge \cdots \ge b_n$, then

$$\frac{1}{n} \sum_{j=1}^{n} a_j b_j \ge \Big(\frac{1}{n} \sum_{j=1}^{n} a_j \Big) \Big(\frac{1}{n} \sum_{j=1}^{n} b_j \Big). \tag{5}$$

An easy argument consists in rearranging the inequality $\sum_{i,j=1}^n (a_i - a_j)(b_i - b_j) \ge 0$. An integral version of Chebyshev inequality reads, for functions f, g both nondecreasing on $[a, b]$ or both nonincreasing on $[a, b]$,

$$\frac{1}{b-a} \int_a^b f(x) g(x) dx \ge \Big(\frac{1}{b-a} \int_a^b f(x) dx \Big) \Big(\frac{1}{b-a} \int_a^b g(x) dx \Big) \tag{6}$$

Rearrangement inequality: If $a_1 \le a_2 \le \cdots \le a_n$ and $b_1 \le b_2 \le \cdots \le b_n$ and if σ is a permutation of $[1:n]$, then

$$\sum_{j=1}^{n} a_j b_{n+1-j} \le \sum_{j=1}^{n} a_j b_{\sigma(j)} \le \sum_{j=1}^{n} a_j b_j. \tag{7}$$

One can use the technique of summation by parts for the proof of the rightmost inequality in (7), say. Setting $B_0 = 0$, $B'_0 = 0$, and

$$B_j = \sum_{i=1}^{j} b_{\sigma(i)}, \qquad B'_j = \sum_{i=1}^{n} b_i, \qquad j \in [1:n],$$

we have $B'_j \le B_j$ for $j \in [1:n-1]$ and $B'_n = B_n$. It follows that

$$\sum_{j=1}^{n} a_j b_{\sigma(j)} = \sum_{j=1}^{n} a_j B_j - \sum_{j=1}^{n} a_j B_{j-1} = a_n B_n + \sum_{j=1}^{n-1} \underbrace{(a_j - a_{j+1})}_{\le 0} \underbrace{B_j}_{\ge B'_j}$$

$$\le a_n B'_n + \sum_{j=1}^{n-1} (a_j - a_{j+1}) B'_j = \sum_{j=1}^{n} a_j b_j,$$

where the last equality is just the reversal of the summation by parts process.

1 Exercises

Ex.1: Prove the inequality between the geometric and harmonic means, namely

$$\frac{n}{1/a_1 + 1/a_2 + \cdots + 1/a_n} \le \sqrt[n]{a_1 a_2 \ldots a_n},$$

for all $a_1, a_2, \ldots, a_n > 0$.

Ex.2: For a continuous convex function φ on $[a,b]$, deduce (4) from (3).

Ex.3: For $a, b, c, d, \ldots \ge 0$, prove that

$$\sqrt{a+b+c+d+\cdots} + \sqrt{b+c+d+\cdots} + \sqrt{c+d+\cdots} + \cdots \ge \sqrt{a+4b+9c+16d+\cdots}.$$

Ex.4: Prove the leftmost inequality of (7).

Ex.5: Deduce (1) from Jensen inequality.

Ex.6: Prove Chebyshev inequality (5) using summation by parts.

Ex.7: Let $P(x)$ be a polynomial with positive coefficients. Prove that $P(1/x) \ge 1/P(x)$ for all $x > 0$, provided $P(1) \ge 1$.

Ex.8: If f is a continuous real-valued function on $[0,1]^2$, prove that

$$\int_0^1 \Big(\int_0^1 f(x,y)dx \Big)^2 dy + \int_0^1 \Big(\int_0^1 f(x,y)dy \Big)^2 dx \le \Big(\int_0^1 \int_0^1 f(x,y)dxdy \Big)^2 + \int_0^1 \int_0^1 f(x,y)^2 dxdy.$$

Group Theory

1 Definitions and first examples

A group $(G, *)$ is a set G equipped with an operation $(x, y) \in G \times G \mapsto x * y \in G$ satisfying the axioms of
G_1 – associativity: $\forall\, x, y, z \in G, x * (y * z) = (x * y) * z$,
G_2 – existence of an identity: $\exists e \in G : \forall\, x \in G, e * x = x * e = x$,
G_3 – existence of inverses: $\forall\, x \in G, \exists x' \in G : x * x' = x' * x = e$.
The axioms imply the uniqueness of an identity element and of inverses. One frequently uses either the additive notation with $+$ for $*$, 0 for the identity element, and $-x$ for the inverse (e.g. $G = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$) or the multiplicative notation with $\cdot$ (or nothing at all) for $*$, 1 for the identity element, and x^{-1} for the inverse (e.g. $G = \mathbb{Q} \setminus \{0\}, \mathbb{R} \setminus \{0\}, \mathbb{C} \setminus \{0\}$). Other examples include $(\mathbb{Z}_n, +)$ and $(\mathbb{Z}_p \setminus \{0\}, \cdot)$ when p is prime — see *Modular Arithmetic* . All the examples mentioned so far were commutative (aka abelian) groups, meaning that $x * y = y * x$ for all $x, y \in G$. The symmetric group S_n, i.e., the set of all permutations of $[1 : n]$ equipped with the operation of composition, is an example of a noncommutative group.

A subgroup of a group $(G, *)$ is a subset H of G which forms a group when equipped with the operation $*$. When H is a subset of G, it forms a subgroup of $(G, *)$ if and only if

$$xy^{-1} \in H \qquad \text{whenever } x, y \in H. \tag{1}$$

For a subset X of a group G, the smallest subgroup of G containing X, i.e., the intersection of all subgroups of G containing X, is called the subgroup generated by X. In particular, given $x \in G$, the subgroup generated by $\{x\}$ (or equivalently by $\{x^n, n \in \mathbb{Z}\}$) is called the cyclic group generated by x.

2 Finite Groups

Given a group $(G, *)$, if the set G is a finite, then its cardinality is called the order of G. The order of the cyclic group generated by $x \in G$ is called the order of x — it is the smallest positive integer m such that $x^m = 1$.

Lagrange theorem states that the order of any subgroup H of a group G divides the order of G (in particular, a group of prime order has no nontrivial subgroups). The argument consists in considering the sets $xH := \{xh, h \in H\}$: two sets xH and $x'H$ are either disjoint or equal, thus, they all have the same size m (which is the order of H), and if q is the number of those sets, one has $n = qm$.

Applying Lagrange theorem to cyclic subgroups generated by one element of a group G of order n, one derives in particular that $x^n = 1$ for every element $x \in G$. For instance, any permutation σ of $[1 : n]$ satisfies $\underbrace{\sigma \circ \sigma \circ \cdots \circ \sigma}_{n! \text{times}} = \mathrm{id}$, since the order of S_n is $n!$.

The product of groups $\{(G_i, *_i), i \in I\}$ is the set $\prod_{i\in I} G_i$ equipped with the operation $*$ as defined by

$$\prod_{i\in I} G_i := \{(x_i)_{i\in I}, , x_i \in G_i \text{ for each } i \in I\}, \qquad (x_i)_{i\in I} * (y_i)_{i\in I} := (x_i *_i y_i)_{i\in I}.$$

The structure theorem for finite abelian groups states that any finite abelian group is isomorphic to a product of cyclic groups of orders equal to powers of prime numbers. In other words, if G is a finite abelian group of order n, then it can be written as

$$G \cong \mathbb{Z}_{p_1^{k_1}} \times \mathbb{Z}_{p_2^{k_2}} \times \cdots \times \mathbb{Z}_{p_m^{k_m}},$$

where $p_1, \ldots, p_m$ are prime numbers, $k_1, \ldots, k_m$ are positive integers, and $p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} = n$. Saying that groups $(G, *)$ and $(G', \star)$ are isomorphic means that there is an isomorphism from G to G', i.e., f is a homomorphism from G to G' ($f(x * y) = f(x) \star f(y)$ for all $x, y \in G$) and that f is invertible.

3 Exercises

Ex.1: Verify that the axioms G_1, G_2, and G_3 imply the uniqueness of an identity element and of inverses. Verify also that a subset H of a group G forms an subgroup of G iff (1) holds.

Ex.2: Verify that, if f is a homomorphism from a group $(G, *)$ to another group $(G', \star)$, then $f(1_G) = 1_{G'}$ and $f(x^{-1}) = (f(x))^{-1}$ for all $x \in G$. Verify that, if f is in addition invertible, then its inverse f^{-1} is automatically an homomorphism from $(G', \star)$ to $(G, *)$.

Ex.3: Let m be the order of an element x in a group G. Prove that m divides any positive integer k such that $x^k = 1_G$.

Ex.4: Prove that the elements of order $\leq m$ in a group G form a subgroup of G.

Ex.5: Prove that

$$SL_n(\mathbb{Z}) := \{A \in \mathcal{M}_{n\times n}(\mathbb{Z}) : |\det(A)| = 1\}$$

of $n \times n$ matrices with integer entries and determinant equal to ± 1 is a group.

Ex.6: Let p and q be the order of two elements x and y in a group G. Suppose that x and y commute and that p and q are relatively prime. Prove that the order of xy equals pq.

Ex.7: For a subset E of a group G, prove that

$$N(E) := \{x \in G : xE = Ex\},$$
$$C(E) := \{x \in G : xy = yx \text{ for all } y \in E\}.$$

are subgroups of G. If E is a subgroup of G, prove that $N(E)$ is the largest subgroup of G containing E as a subgroup and such that $xE = Ex$ for all $x \in N(E)$.

Number Theory

1 The fundamental theorem of arithmetic

An integer $p > 1$ is a prime number if its only positive divisors are 1 and p (by convention, $p = 1$ is not considered a prime number). The prime numbers form an infinite set. Indeed, if there was finitely many prime numbers $p_1 < p_2 < \cdots < p_k$, then $q := p_1 p_2 \cdots p_k + 1 > p_k$ would not be prime, hence it would be divisible by some prime number p_i, but then $p_i | q - p_1 \cdots p_i \cdots p_k = 1$, which is absurd. In fact, the prime number theorem states that the number $\pi(n)$ of primes less than or equal to n behaves like $n/\ln(n)$ as $n \to \infty$.

The fundamental theorem of arithmetic states that every integer $n > 1$ can be written uniquely (up to the order of factors) as product of primes.

2 Euclid algorithm and its consequences

Two integers $n > 1$ and $m > 1$ are called coprime (or relatively prime) if they share no common prime factor. Stated differently, n and m are coprime if their greatest common divisor is 1. The notions of greatest common divisor and least common multiple are self-explanatory. With obvious notations, we have $\gcd(n, m) \cdot \mathrm{lcm}(n, m) = n \cdot m$. The greatest common divisor of n and m can be found via Euclid algorithm: with $n > m$, set $r_0 = n$, $r_1 = m$, and produce r_k inductively for $k \geq 2$ from the division of r_{k-2} by r_{k-1} as

$$r_{k-2} = q_{k-1} r_{k-1} + r_k, \qquad 0 \leq r_k < r_{k-1}.$$

Since the sequence of nonnegative numbers $(r_k)_{k \geq 0}$ is strictly decreasing, it eventually reaches $r_K = 0$, and $\gcd(n, m) = r_{K-1}$. This is the case because $\gcd$ is preserved at each iteration, i.e.,

$$\gcd(r_{k-2}, r_{k-1}) = \gcd(r_{k-1}, r_k), \qquad k \geq 2, \tag{1}$$

hence $\gcd(n, m) = \gcd(r_0, r_1) = \gcd(r_{K-2}, r_{K-1}) = r_{K-1}$ where the latter equality is due to the fact that r_{K-2} divides r_{K-1}. The set of integer combinations is also preserved at each iteration, i.e.,

$$\{p r_k + q r_{k-1}, (p, q) \in \mathbb{Z}\} = \{p r_{k+1} + q r_k, (p, q) \in \mathbb{Z}\}, \qquad k \geq 2, \tag{2}$$

so the equality between the first and last sets gives

$$\{pn + qm, (p, q) \in \mathbb{Z}\} = \gcd(n, m)\mathbb{Z}.$$

This implies in particular Bézout lemma, i.e.,

$$\gcd(n, m) = 1 \iff \exists p, q \in \mathbb{Z} \text{ such that } pn + qm = 1.$$

In turn, the latter is used to prove Euclid lemma (obvious with prime factor decompositions, but needed in the uniqueness part of the fundamental theorem of arithmetic) which says that

> if m divides nr and if m and n are coprime, then m divides r.

To see this, write $nr = dm$ and $pn + qm = 1$, so that $r = (pn + qm)r = pdm + qmr = (pd + qr)m$.

3 Euler totient function

Define the Euler function ϕ on the positive integers by

$$\phi(n) := \mathrm{card}\{k \in [1:n] \text{ such that } k \text{ and } n \text{ are coprime}\}.$$

Note that, if p is prime and if $s \geq 1$ is an integer, then $\phi(p^s) = p^s - p^{s-1} = p^s(1-1/p)$ (because there are p^{s-1} integers in $[1:p^s]$ that are not coprime with p^s, namely $p, 2p, 3p, \ldots, p^{s-1}p = p^s$). Note also that ϕ is multiplicative, meaning that $\phi(nm) = \phi(n)\phi(m)$ whenever n and m are coprime (this is a consequence of the Chinese remainder theorem, see *Modular Arithmetic*). Combining these two facts with the prime factor decomposition $n = p_1^{s_1} p_2^{s_2} \cdots p_\ell^{s_\ell}$ of a positive integer gives the formula

$$\phi(n) = n \prod_{p \,\mathrm{prime},\, p|n} \Big(1 - \frac{1}{p}\Big).$$

Multiplying out the right-hand side yields

$$\phi(n) = \sum_{d|n} \frac{n}{d}\, \mu(d) = \sum_{d|n} d\, \mu\Big(\frac{n}{d}\Big), \tag{3}$$

where μ is the Möbius function defined by $\mu(1) = 1$ and, for $m > 1$,

$$\mu(m) := \begin{cases} (-1)^\ell & \text{if } m = p_1 p_2 \cdots p_\ell \text{ is a product of } \ell \text{ distinct primes,} \\ 0 & \text{if } p^2 | m \text{ for some prime } p. \end{cases}$$

This can be concisely written as $\phi = \mathrm{id} * \mu$, where the Dirichlet convolution is the commutative operation defined, for two functions a, b on positive integers, by

$$(a * b)(n) = \sum_{ij=n} a(i) b(j).$$

This operation has an identity given by $e(1) = 1$ and $e(m) = 0$, $m > 1$, and is associative, since

$$[a * (b * c)](n) = \sum_{im=n} a(i)(b*c)(m) = \sum_{im=n} a(i) \sum_{jk=m} b(j)c(k) = \sum_{ijk=n} a(i)b(j)c(k),$$

$$[(a * b) * c](n) = \sum_{km=n} (a*b)(m)c(k) = \sum_{km=n} \sum_{ij=m} a(i)b(j)c(k) = \sum_{ijk=n} a(i)b(j)c(k).$$

Let us also notice that, for an integer $m > 1$ decomposed in prime factors as $m = p_1^{s_1} p_2^{s_2} \cdots p_\ell^{s_\ell}$,

$$\sum_{d|m} \mu(d) = \sum_{r_1,\ldots,r_\ell \in \{0,1\}} \mu(p_1^{r_1} p_2^{r_2} \cdots p_\ell^{r_\ell}) = \sum_{h=0}^{\ell} \binom{\ell}{h} (-1)^\ell = (1-1)^\ell, \qquad \text{i.e.,} \qquad \sum_{d|m} \mu(d) = 0.$$

Since the sum takes the value 1 for $m = 1$, we have $\mu * 1 = e$. Now, if $a = b * \mu$, then $a * 1 = b * \mu * 1 = b * e = b$, and conversely, if $a * 1 = b$, then $b * \mu = a * 1 * \mu = a * e = a$. Spelling out the convolutions leads to Möbius inversion formula: for functions a, b on positive integers,

$$a(n) = \sum_{d|n} b(d)\, \mu(n/d) \quad \text{for all } n \geq 1 \iff b(n) = \sum_{d|n} a(d) \quad \text{for all } n \geq 1.$$

Taking $a = \phi$ and $b = \mathrm{id}$ in the latter and using (3) gives Euler formula, that is

$$\sum_{d|n} \phi(d) = n.$$

4 Exercises

Ex.1: Verify the statements made in (1) and (2).

Ex.2: Prove that the distance between two consecutive prime numbers is unbounded.

Ex.3: Prove that the product of three consecutive integers is never a perfect power (i.e., not a perfect square, not a perfect cube, etc.).

Ex.4: For an integer $n \geq 1$, prove that $n^4 - 7n^2 + 1$ cannot be a perfect square.

Ex.5: If n is an integer with prime factor decomposition $n = p_1^{k_1} p_2^{k_2} \cdots p_\ell^{k_\ell}$, let $f(n) := \sum_{i=1}^{\ell} k_i p_i$ and $g(n) := \lim_{m \to \infty} \underbrace{f \circ \cdots \circ f}_{m \text{ times}}(n)$. Evaluate $g(100)$ and $g(10^{10})$. Find all odd integers $n > 1$ such that $n/2 < g(n) < n$.

Modular Arithmetic

1 Residue classes

Given an integer $n \geq 2$, we say that $a \in \mathbb{Z}$ is congruent to $b \in \mathbb{Z}$ modulo n if n divides $a - b$ — equivalently, if $a = b + kn$ for some $k \in \mathbb{Z}$, or if a and b have the same remainder in the division by n. In this case, we write $a \equiv b \pmod{n}$. Note that $\equiv$ is an equivalence relation on $\mathbb{Z}$ (reflexive: $a \equiv a \pmod{n}$; symmetric: if $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$; transitive: if $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$). Therefore, we can partition $\mathbb{Z}$ into the equivalence classes, called residue classes,

$$[a]_n = \{b \in \mathbb{Z} : b \equiv a \pmod{n}\} = \{a + kn, k \in \mathbb{Z}\}.$$

Each residue class has a unique representative in $\{0, 1, \ldots, n-1\}$ — the remainder of any element of the class in the division by n — and is often identified to this representative. Hence, the set $\mathbb{Z}_n$ of residue classes modulo n is identified to $\{0, 1, \ldots, n-1\}$. Defining an addition and a multiplication on $\mathbb{Z}_n$ by $[a]_n + [b]_n = [a+b]_n$ and $[a]_n \cdot [b]_n = [a \cdot b]_n$, it can be seen that $(\mathbb{Z}_n, +)$ is a group. With $\mathbb{Z}_n^* := \{a \in \mathbb{Z}_n : \exists\, b \in \mathbb{Z}_n : [a]_n \cdot [b]_n = [1]_n\}$ denoting the set of units (i.e., invertible elements) of $\mathbb{Z}_n$, it can be seen that $(\mathbb{Z}_n^*, \cdot)$ is also a group.

2 Euler theorem

Note that (the representative) of $a \in \mathbb{Z}$ is a unit of $\mathbb{Z}_n$ if and only if there exist $b \in \mathbb{Z}$ and $k \in \mathbb{Z}$ such that $ab + kn = 1$. By Bézout lemma, this means that $a \in \mathbb{Z}_n^*$ if and only if a and n are coprime. One consequence is that, if p is prime, then every nonzero element in $\mathbb{Z}_p$ is invertible — this makes $\mathbb{Z}_p$ a field, where usual calculation rules apply, for instance $ab \equiv 0 \pmod{n}$ implies $a \equiv 0 \pmod{n}$ or $b \equiv 0 \pmod{n}$. Another consequence is that

$$\operatorname{card}(\mathbb{Z}_n^*) = \operatorname{card}\{a \in [1 : n-1] \text{ such that } a \text{ and } n \text{ are coprime}\} = \phi(n),$$

where ϕ is Euler totient function. Thus, applying Lagrange theorem to the multiplicative group $\mathbb{Z}_n^*$ yields Euler theorem, that is

$$a^{\phi(n)} \equiv 1 \pmod{n} \qquad \text{whenever } \gcd(a, n) = 1.$$

When n is a prime number p, this becomes Fermat little theorem, that is

$$a^{p-1} \equiv 1 \pmod{p} \qquad \text{whenever } a \text{ is not a multiple of } p.$$

Euler theorem provides a way to compute the powers modulo n of an integer a coprime with n, i.e., $a^m \equiv a^{m \bmod \phi(n)} \pmod{n}$.

3 Chinese remainder theorem

Given integers $n_1, n_2, \ldots, n_k \geq 2$ that are pairwise coprime, the chinese remainder theorem says that the system of congruence

$$\begin{aligned} x &\equiv r_1 \pmod{n_1}, \\ x &\equiv r_2 \pmod{n_2}, \\ &\vdots \\ x &\equiv r_k \pmod{n_k}, \end{aligned}$$

has a unique solution modulo $N := n_1 n_2 \cdots n_k$. For the uniqueness, notice that, if x and x' are two solutions, then $n_1 | x - x'$, $n_2 | x - x'$, ..., $n_k | x - x'$, so $n_1 n_2 \cdots n_k | x - x'$ (because $n_1, n_2, \ldots, n_k$ are coprime), i.e., $x \equiv x' \pmod{N}$. For the existence, set $N_i := N/n_i$ and notice that N_i and n_i are coprime. Thus, we can consider the inverse m_i of N_i in $\mathbb{Z}_{n_i}$. It is now readily verified that $x := m_1 N_1 r_1 + m_2 N_2 r_2 + \cdots + m_k N_k r_k$ is a solution of the system of congruence. Stated differently, the theorem says that the map

$$x \in \mathbb{Z}_{n_1 n_2 \cdots n_k} \mapsto (x \bmod n_1, x \bmod n_2, \ldots, x \bmod n_k) \in \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \mathbb{Z}_{n_k}$$

is bijective when $n_1, n_2, \ldots, n_k$ are pairwise coprime.

We can now justify that Euler totient function is multiplicative, i.e., that $\phi(nm) = \phi(n)\phi(m)$ when n and m are coprime. Indeed, for $x \in \mathbb{Z}$, the fundamental theorem of arithmetic reveals

$$[\gcd(x, nm) = 1] \Leftrightarrow [\gcd(x, n) = 1,\ \gcd(x, m) = 1] \Leftrightarrow [\gcd(x \bmod n, n) = 1,\ \gcd(x \bmod m, m) = 1],$$

so that $x \in \mathbb{Z}_{nm}^* \mapsto (x \bmod n, x \bmod m) \in \mathbb{Z}_n^* \times \mathbb{Z}_m^*$ is also a bijective map. The equality between the cardinalities of $\mathbb{Z}_{nm}^*$ and of $\mathbb{Z}_n^* \times \mathbb{Z}_m^*$ gives the desired result.

4 Exercises

Ex.1: Verify that $(\mathbb{Z}_n, +)$ and $(\mathbb{Z}_n^*, \cdot)$ are groups.

Ex.2: We define a function f on positive integers by $f(1) = 3$ and $f(n+1) = 3f(n)$. What are the last two digits of $f(2012)$?

Ex.3: For any integer $n > 1$, prove that n does not divides $2^n - 1$.

Ex.4: What is the lowest degree monic polynomial which vanishes identically on the integers $\pmod{p}$ when p is prime? Same question $\pmod{100}$?

Ex.5: How many perfect squares are there $\pmod{2^n}$?

Linear Algebra

1 Range and null space

For $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, define its range

$$\operatorname{ran} A := \{Ax, x \in \mathbb{C}^m\},$$

and its null space

$$\ker A := \{x \in \mathbb{C}^n : Ax = 0\}.$$

These are linear subspaces of $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively. The rank and the nullity of A are defined by

$$\operatorname{rk} A := \dim(\operatorname{ran} A), \qquad \operatorname{nul} A := \dim(\ker A).$$

They are deduced form one another by the rank-nullity theorem

$$\operatorname{rk} A + \operatorname{nul} A = n.$$

Recall that $A \in \mathcal{M}_{m\times n}(\mathbb{C})$ is injective if $\ker A = \{0\}$, and surjective if $\operatorname{ran} A = \mathbb{C}^m$. Note that a *square* matrix A is injective (or surjective) iff it is both injective and surjective, i.e., iff it is bijective. Bijective matrices are also called invertible matrices, because they are characterized by the existence of a unique square matrix B (the inverse of A, denoted by A^{-1}) such that $AB = BA = I$.

2 Trace and determinant

The trace and determinants are functions taking *square* matrices and returning scalars. The trace of $A \in \mathcal{M}_n(\mathbb{C})$ is the sum of its diagonal elements, i.e.,

$$\operatorname{tr} A := \sum_{i=1}^{n} a_{i,i} \quad \text{where} \quad A = [a_{i,j}]_{i,j=1}^{n}.$$

Notice that the trace is linear (i.e., $\operatorname{tr}(\lambda A + \mu B) = \lambda \operatorname{tr}(A) + \mu \operatorname{tr}(B)$) and that

$$\operatorname{tr}(AB) = \operatorname{tr}(BA) \qquad \text{whenever } A \in \mathcal{M}_{m\times n}(\mathbb{C}) \text{ and } B \in \mathcal{M}_{n\times m}(\mathbb{C}).$$

As for the determinant, it can be defined in several equivalent ways:

1. As a function of the columns of a marix, it is the only function $f : \mathbb{C}^n \times \cdots \times \mathbb{C}^n \to \mathbb{C}$ that is linear with respect to each columns ($f(\ldots, \lambda x + \mu y, \ldots) = \lambda f(\ldots, x, \ldots) + \mu f(\ldots, y, \ldots)$), alternating ($f(\ldots, x, \ldots, y, \ldots) = -f(\ldots, y, \ldots, x, \ldots)$), and unit-normalized ($f(I) = 1$). This can be used to derive the identity

$$\det(AB) = \det(A)\det(B) \quad \text{for all } A, B \in \mathcal{M}_n(\mathbb{C}).$$

2. $\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$,
   where S_n is the set of $n!$ permutations of $[1 : n]$ and $\operatorname{sgn}(\sigma) = (-1)^s$, s = number of pairwise interchanges composing σ (hence the computation rules for 2×2 and 3×3 determinants). This can be used to prove that

$$\det A^\top = \det A \qquad \text{for all } A \in \mathcal{M}_n(\mathbb{C}).$$

3. Laplace expansion with respect to a row or a column, e.g. with respect to the ith row

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j},$$

   where $A_{i,j}$ is the submatrix of A obtained by deleting the ith row and the jth column. The matrix $B \in \mathcal{M}_n(\mathbb{C})$ with entries $b_{i,j} := (-1)^{i+j} \det A_{i,j}$ is called the comatrix of A — note that $B^\top$ is also called the adjoint of A (*classical* adjoint, not to be confused with *hermitian* adjoint). Laplace expansion can be used to prove that $AB^\top = (\det A)I$. In turn, it is deduced that $A \in \mathcal{M}_n(\mathbb{C})$ is invertible iff $\det A \neq 0$, in which case $\mathbf{A}^{-1} = (1/\det(A))B^\top$.

3 Eigenvalues and eigenvectors

Given a *square* matrix $A \in \mathcal{M}_n(\mathbb{C})$, if there exist $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}^n$, $x \neq 0$, such that

$$Ax = \lambda x,$$

then λ is called an eigenvalue of A and x is called an eigenvector corresponding to the eigenvalue λ. The set of all eigenvectors corresponding to an eigenvalue λ is called the eigenspace corresponding to the eigenvalue λ — it is indeed a linear space. Note that λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$, i.e., iff λ is a zero of the characteristic polynomial of A defined by

$$p_A(x) := \det(A - xI).$$

Observe that p_A is a polynomial of the form

$$p_A(x) = (-1)^n x^n + (-1)^{n-1} \operatorname{tr}(A)\, x^{n-1} + \cdots + \det(A).$$

Since this polynomial can also be written in factorized form as $(\lambda_1 - x) \cdots (\lambda_n - x)$, where $\{\lambda_1, \ldots, \lambda_n\}$ is the set of eigenvalues of A (complex and possibly repeated), we have

$$\operatorname{tr}(A) = \lambda_1 + \cdots + \lambda_n, \qquad \det(A) = \lambda_1 \cdots \lambda_n.$$

The existence of n linearly independent eigenvectors $v_1, \ldots, v_n \in \mathbb{C}^n$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$ of $A \in \mathcal{M}_n$ (which occurs in particular if A has n distinct eigenvalues) is

equivalent to the existence of a invertible matrix $V \in \mathcal{M}_n$ and of a diagonal matrix $D \in \mathcal{M}_n$ such that

$$A = VDV^{-1}.$$

The columns of V are the $v_i's$ and the diagonal entries of D are the λ_i's. In this case, we say that the matrix A is diagonalizable. More generally, two matrices A and B are called similar if there exists an invertible matrix V such that $A = VBV^{-1}$. Note that two similar matrices have the same characteristic polynomial, hence the same eigenvalues (counting multiplicities), and in particular the same trace and determinant.

4 Exercises

Ex.1: We recall that $\mathrm{rk}\ A^* = \mathrm{rk}\ A$, where $A^* \in \mathcal{M}_{n\times m}(\mathbb{C})$ denotes the conjugate transpose of a matrix $A \in \mathcal{M}_{m\times n}$. In general, is it true that $\mathrm{nul}\ A^* = \mathrm{nul}\ A$? Establish that $\ker A = \ker A^*A$, deduce that $\mathrm{nul}\ A = \mathrm{nul}\ A^*A$ and that $\mathrm{rk}\ A = \mathrm{rk}\ A^*A = \mathrm{rk}\ A^* = \mathrm{rk}\ AA^*$, and finally conclude that $\mathrm{ran}\ A = \mathrm{ran}\ AA^*$.

Ex.2: Calculate $\mathrm{tr}(A^*A)$ and observe that $A = 0$ iff $\mathrm{tr}(A^*A) = 0$.

Ex.3: For $A, B \in \mathcal{M}_n(\mathbb{C})$, prove that $AB = I$ implies $BA = I$. Is this true if A and B are not square?

Ex.4: Determine the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & t & \cdots & t \\ t & 1 & t & \vdots \\ t & \cdots & \ddots & t \\ t & \cdots & t & 1 \end{bmatrix},$$

and diagonalize it.

Ex.5: For $A \in \mathcal{M}_n(\mathbb{Z})$, suppose that there exists a prime number p dividing $\sum_{j=1}^n a_{i,j}$ for all $i \in [1:n]$. Prove that p divides $\det(A)$.

Ex.6: Determine if the following statement is true or false: there exists $A \in \mathcal{M}_n(\mathbb{R})$ such that $A^2 + 2A + 5I = 0$ if and only if n is even.