Estimates: factorial and binomial coefficients

Proposition. For each natural number $n \geq 1$:

$$2^{n-1} \leq n! \leq n^n.$$

Theorem. For each $n \in \mathbb{N}$:

$$n^{n/2} \leq n! \leq \left(\frac{n+1}{2}\right)^n.$$

Lemma (AM-GM inequality)**.** For every pair of non-negative reals a, b:

$$\sqrt{ab} \leq \frac{a+b}{2}.$$

Theorem. For every $n \in \mathbb{N}$:

$$\mathrm{e}\left(\frac{n}{\mathrm{e}}\right)^n \leq n! \leq \mathrm{e}n\left(\frac{n}{\mathrm{e}}\right)^n.$$

Claim. For every real number x:

$$1 + x \leq \mathrm{e}^x.$$

Claim (Stirling formula)**.** $n! \sim \sqrt{2\pi n} \cdot (\frac{n}{\mathrm{e}})^n$, where $f \sim g$ means $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = 1$.

Theorem. For every $1 \leq k \leq n$:

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{\mathrm{e}n}{k}\right)^k.$$

Theorem. For every $m \in \mathbb{N}$:

$$\frac{2^{2m}}{2n+1} \leq \binom{2m}{m} \leq 2^{2m}.$$

Using Stirling formula, we can get a more precise approximation:

$$\binom{2m}{m} \sim \frac{2^{2m}}{\sqrt{\pi m}}.$$

Generating functions

Theorem. Let $a_0, a_1, a_2, \ldots$ be an infinite sequence of real numbers such that $|a_i| \leq k^i$ for some $k \in \mathbb{R}$ and all $i \geq 1$. Then for each $x \in (-k, k)$ the power series $\sum_{i=0}^{\infty} a_i x^i$ is convergent and it determines a real function $a(x) = \sum_{i=0}^{\infty} a_i x^i$.

Moreover, the function $a(x)$ is uniquely determined by the sequence on the interval $(-k, k)$ and $a_i = a^{(i)}(0)/i!$. We call $a(x)$ the *generating function* of the sequence $a_0, a_1, a_2, \ldots$.

Example:
sequence $1, 1, 1, \ldots \leftrightarrow$ power series $1 + x + x^2 + \ldots \leftrightarrow$ generating function $\frac{1}{1-x}$.

Operations with sequences and generating functions:

1. sum: $a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots \leftrightarrow a(x) + b(x)$
2. multiplication by $\alpha \in \mathbb{R}$: $\alpha a_0, \alpha a_1, \alpha a_2, \ldots \leftrightarrow \alpha a(x)$
3. substitution of αx for x: $a_0, \alpha a_1, \alpha^2 a_2, \ldots \leftrightarrow a(\alpha x)$
4. substitution of x^n for x: $a_0, 0, \ldots, 0, a_1, 0, \ldots, 0, a_2, \ldots \leftrightarrow a(x^n)$
5. move right: $0, a_0, a_1, a_2, \ldots \leftrightarrow x a(x)$
6. move left: $a_1, a_2, a_3, \ldots \leftrightarrow \frac{a(x) - a_0}{x}$
7. differentiation: $a_1, 2a_2, 3a_3, \ldots \leftrightarrow a'(x)$
8. integration: $0, a_0, \frac{1}{2}a_1, \frac{1}{3}a_2, \ldots \leftrightarrow \int_0^x a(t)\, \mathrm{d}t$
9. product of functions: $c_0, c_1, c_2, \ldots \leftrightarrow a(x)b(x)$, where $c_k = \sum_{i=0}^{k} a_i b_{k-i}$
10. prefix sums: $a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots \leftrightarrow a(x)/(1-x)$

Fibonacci numbers: Let $F_0 = 0, F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$. Then

$$F_n = \frac{1}{\sqrt{5}} \cdot \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right].$$

Theorem (Generalized Binomial theorem)**.** For $r \in \mathbb{R}, k \in \mathbb{N}$, we define the *generalized binomial coefficients*

$$\binom{r}{k} = \frac{r(r-1)(r-2)\ldots(r-k+1)}{k!}, \quad \binom{r}{0} = 1.$$

Then $(1+x)^r$ is the generating function of the sequence $\binom{r}{0}, \binom{r}{1}, \binom{r}{2}, \ldots$ (the sum $\sum_{i=0}^{\infty} \binom{r}{i} x^i$ is convergent for $x \in (-1, 1)$).

Lemma. For non-negative integers a, b, we have:

$$\binom{-a}{b} = (-1)^b \cdot \binom{a+b-1}{b}.$$

Catalan numbers: Let $b_0 = 1$ and $b_{n+1} = \sum_{i=0}^{n} b_i b_{n-i}$. Then

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

Example: There are exactly b_n binary trees on n vertices.

Theorem (A cookbook for linear recurrent relations)**.** Let

$$A_{n+k} = c_0 A_n + c_1 A_{n+1} + \ldots + c_{k-1} A_{n+k-1}$$

be a *homogeneous linear recurrence relation with constant coefficients* and *initial conditions* $A_0, \ldots, A_{k-1}$. Let further

$$R(x) = x^k - c_{k-1} x^{k-1} - \ldots - c_1 x^1 - c_0 x^0$$

be its *characteristic polynomial* and $\lambda_1, \ldots, \lambda_z \in \mathbb{C}$ pairwise different roots of this polynomial with multiplicities $k_1, \ldots, k_z$. Then there are constants $C_{ij} \in \mathbb{C}$ such that for each n:

$$A_n = \sum_{i=1}^{z} \sum_{j=0}^{k_i - 1} \left(C_{ij} \cdot \binom{n+j}{j} \cdot \lambda_i^n \right).$$

If R has no multiple roots, the formula for A_n can be written in a simple form:

$$A_n = \sum_{i=1}^{z} C_i \lambda_i^n.$$

Proof. Only for simple roots. □

Finite projective planes

Definition. Let X be a finite set and $\mathcal{L} \subseteq 2^X$ a set of subsets of X. Then $(X, \mathcal{L})$ is called a *finite projective plane* if it satisfies:

(P0) There exists $F \subseteq X$ with $|F| = 4$ and $|F \cap L| \leq 2$ for each $L \in \mathcal{L}$.

(P1) For all distinct $L_1, L_2 \in \mathcal{L}$: $|L_1 \cap L_2| = 1$.

(P2) For all distinct $x_1, x_2 \in X$, there is a unique $L \in \mathcal{L}$ such that $x_1 \in L$ and $x_2 \in L$.

We will call the elements of X *points* of the projective plane and the elements of $\mathcal{L}$ its *lines.*

Lemma. For every line $L \in \mathcal{L}$, there exists a point $x \in X \setminus L$.

Proposition. Let $L_1, L_2 \in \mathcal{L}$ be two lines of the finite projective plane $(X, \mathcal{L})$, then $|L_1| = |L_2|$.

Definition. The *order* of a finite projective plane $(X, \mathcal{L})$ is $|L| - 1$, where $L \in \mathcal{L}$.

Theorem. Let $(X, \mathcal{L})$ be a finite projective plane of order n. Then:

(i) For all $x \in X$ we have $|\{L \in \mathcal{L} \mid x \in L\}| = n + 1$,

(ii) $|X| = n^2 + n + 1$,

(iii) $|\mathcal{L}| = n^2 + n + 1$.

Definition. A (finite) *set system* is a pair $(X, \mathcal{L})$, where X is a finite set and $\mathcal{L}$ is a multi-set of subsets of X. (Formally, you can avoid multi-sets by considering a sequence of subsets instead. This way, the set system would be a triple $(X, I, \mathcal{L})$, where I is an *index set* and $\mathcal{L}$ is a mapping from I to 2^X.)
The *incidence graph* of a set system is a bipartite graph with parts X and $\mathcal{L}$ and edges $\{x, L\}$ for all $x \in L \in \mathcal{L}$.

Observation. A set system is uniquely determined by its incidence graph.

Definition. A *dual* of a set system $(X, \mathcal{L})$ is defined by its incidence graph, which is obtained by taking the incidence graph of $(X, \mathcal{L})$ and exchaning the roles of its parts.

Theorem. A dual set system of a finite projective plane is a finite projective plane. (Roles of points and lines are exchanged by the duality.)

Theorem. If n is a prime power, then there exists a finite projective plane of order n.

Latin squares

Definition. A *Latin square* of order n is a matrix A of order $n \times n$ with entries from $\{1, 2, \ldots, n\}$ such that $a_{ij} \neq a_{ij'}$ for $j \neq j'$ and $a_{ij} \neq a_{i'j}$ for $i \neq i'$.
Two Latin squares A, B of order n are called *orthogonal* if $(a_{ij}, b_{ij}) = (a_{rs}, b_{rs})$ implies $(i, j) = (r, s)$.

Proposition. Let $A_1, A_2, \ldots, A_t$ be a collection of mutually orthogonal Latin squares of order n. Then $t \leq n - 1$.

Theorem. For $n \geq 2$, a finite projective plane of order n exists if and only if there exists a collection of $n - 1$ mutually orthogonal Latin squares of order n.

Hall's theorem and bipartite matching

Definition. Let $(X, \mathcal{L})$ be a set system. A function $f : \mathcal{L} \to X$ is called its *system of distinct representatives* if it is injective and $f(S) \in S$ for all $S \in \mathcal{L}$.

Theorem (Hall's theorem, set version)**.** A set system $(X, \mathcal{L})$ has a system of distinct representatives if and only if $|\bigcup_{\mathcal{K}}| \geq |\mathcal{K}|$ holds for all sub-systems $\mathcal{K} \subseteq \mathcal{L}$. (This is called the *Hall's condition.*)

Definition. A *matching* in a graph $G = (V, E)$ is a set of edges $F \subseteq E$ such that no two edges in F share a common vertex. A matching is called *perfect* if its edges contain all vertices of G. In a bipartite graph with parts L and R, we can define *L-perfect* and *R-perfetct* matchings similarly.

Observation. Systems of distinct representatives of a set system $(X, \mathcal{L})$ are in one-to-one correspondence with $\mathcal{L}$-perfect matchings in the incidence graph.

Theorem (Hall's theorem, graph version)**.** Let $G = (V, E)$ be a bipartite graph with parts L and R. Then G has a L-perfect matching iff $|\Gamma(K)| \geq |K|$ for each $K \subseteq L$, where $\Gamma(K) = \{v \in V \mid \exists w \in K : \{v, w\} \in E\}$ is the *neighborhood* of K.

Corollary. Every regular bipartite graph has a perfect matching.

Definition. A matrix $B \in \mathbb{R}^{m \times n}$ is *bistochastic* if all its entries are non-negative and every row/column sums to 1. In particular, a *permutation matrix* contains exactly one 1 in each row/column and zeroes everywhere else.

Observation. Every bistochastic matrix is square.

Theorem (Birkhoff)**.** Every bistochastic matrix is a convex linear combination of some permutation matrices. That is, for every bistochastic matrix $B \in \mathbb{R}^{n \times n}$ there exist permutation matrices $P_1, \ldots, P_k \in \{0, 1\}^{n \times n}$ and positive real numbers $\alpha_1, \ldots, \alpha_k$ such that $B = \sum_i \alpha_i P_i$ and $\sum_i \alpha_i = 1$.

Flows in networks

Definition. A *network* is a directed graph (V, E) with two designated vertices s (the *source*) and t (the *target*) and *capacities* on edges given by a function $c : E \to \mathbb{R}_0^+$. Without loss of generality, we can assume that $uv \in E$ implies $vu \in E$ (missing edges can be added with zero capacity).

Definition. For a function $f : E \to \mathbb{R}$ on a network, we define functions f^+ (*inflow*), f^- (*outflow*), and f^Δ (*excess*) from V to $\mathbb{R}$ by:

$$f^+(v) = \sum_{uv \in E} f(uv), \quad f^-(v) = \sum_{vw \in E} f(vw), \quad f^\Delta(v) = f^+(v) - f^-(v).$$

Definition. A function $f : E \to \mathbb{R}$ is a *flow* in a given network if it satisfies the following conditions:

1. *Capacity constraints:* $0 \le f(e) \le c(e)$ for all $e \in E$,
2. *Flow conservation:* $f^\Delta(v) = 0$ for all $v \in V \setminus \{s, t\}$
   (this is also known as the *Kirchhoff's law*).

The *value* of the flow is defined by $|f| = f^\Delta(t)$.

Observation. Equivalently, $|f| = -f^\Delta(s)$.

Observation. In every network, there is at least one flow: the everywhere-zero flow. A more interesting problem is finding a *maximum* flow, that is a flow with the maximum possible value. (Does it always exist?)

Definition. For a given network and a flow f, we define *residual capacities* $r : E \to \mathbb{R}$ as $r(uv) = c(uv) - f(uv) + f(vu)$. (Intuitively, it tells how much extra flow we can send from u to v either by adding to the flow on uv, or by subtracting from flow on vu.)

Definition. An *augmenting path* is a directed path from s to t whose all edges have non-zero residual capacities.

If there is an augmenting path, the flow can be improved along this path. Repeating this process yields the following algorithm.

Algorithm (Ford-Fulkerson maximum flow)**.**

1. Let $f(e) \leftarrow 0$ for every edge e.
2. While there exists an augmenting path P:
3. $\quad\varepsilon \leftarrow \min_{e \in P} r(e)$
4. $\quad$For all edges $uv \in P$:
5. $\quad\quad\delta \leftarrow \min(\varepsilon, c(uv) - f(uv))$
6. $\quad\quad f(uv) \leftarrow f(uv) + \delta$
7. $\quad\quad f(vu) \leftarrow f(vu) - (\varepsilon - \delta)$

Definition. For any two disjoint sets $A, B \subset V$, we define $E(A,B) = \{ab \in E \mid a \in A, b \in B\}$. This set of edges is called an *(elementary) cut* if $s \in A$ and $t \in B$.
When $E(A,B)$ is a cut and g is a real-valued function on edges, we define $g(A,B) = \sum_{e \in E(A,B)} f(e)$. In particular, $c(A,B)$ is called the *capacity of the cut.*

Observation. When f is a flow and $E(A,B)$ is a cut, then $|f| = f(A,B) - f(B,A)$. Since $f(A,B) \leq c(A,B)$, this implies $|f| \leq c(A,B)$. Hence if $|f| = c(A,B)$, then f is maximum and $E(A,B)$ minimum (it has the lowest possible capacity over all cuts).

Theorem. The Ford-Fulkerson algorithm has the following properties:

- During the whole computation, f is a flow.
- When the algorithm stops, f is a maximum flow.
- If the capacities are integers, the algorithm stops. Furthermore, it produces an integral maximum flow.
- If the capacities are rationals, the algorithm stops.
- For some real capacities, the computation can run forever.

Theorem (Edmonds-Karp algorithm)**.** When the Ford-Fulkerson algorithm always selects the shortest possible augmenting path, it stops within $\mathcal{O}(|V| \cdot |E|)$ iterations.

Corollary. Every network has a maximum flow.

Corollary. If all capacities are integers, there exists at least one maximum flow using only integers.

Corollary (Ford-Fulkerson min-max theorem)**.** For every network, the value of the maximum flow equals the capacity of the minimum cut.

Bipartite matchings

For any bipartite graph $(L \cup R, E)$, we can define an auxiliary network with vertices $L \cup R \cup \{s,t\}$, edges $\{su \mid u \in L\} \cup \{uv \mid u \in L, v \in R, \{u,v\} \in E\} \cup \{vt \mid v \in R\}$ and all capacities set to 1.

Observation. Integral flows in this network correspond to matchings, cuts correspond to vertex covers (sets of vertices which intersect every edge). This implies the Hall's theorem. By the min-max theorem, we also get:

Corollary (König's theorem)**.** In every bipartite graph, the size of a maximum matching equals the size of a minimum vertex cover.

Higher connectivity

Definition. Let $G = (V, E)$ be an undirected graph. A subset $F \subseteq E$ is an *edge cut* of G if $G - F$ is disconnected. For an integer k, the graph G is called *k-edge-connected,* if it has no edge cut of size smaller than k.
Similarly, a *vertex cut* of G is a subset $U \subseteq V$ such that $G - U$ is disconnected. The graph G is *k-vertex-connected*, if $|V| \geq k + 1$ and G has no vertex cut of size smaller than k.

Definition. The *edge connectivity function* $k_e(G)$ is defined as the minimum size of an edge cut of a graph G (alternatively, the maximum k such that G is k-edge-connected).
Similarly, the *vertex connectivity function* $k_v(G)$ gives the size of the smallest vertex cut of a non-complete graph G (i.e., the maximum k such that G is k-vertex-connected). For complete graphs, we define $k_v(K_n) = n - 1$.

Lemma. Let $G = (V, E)$ be a graph and e an arbitrary edge of G. Then

$$k_e(G) - 1 \leq k_e(G - e) \leq k_e(G)$$

and

$$k_v(G) - 1 \leq k_v(G - e) \leq k_v(G).$$

Theorem (Menger, edge version)**.** Let G be a graph and k a positive integer. Then G is k-edge-connected if and only if for every pair $u, v \in V$ of distinct vertices of G, there exists a system of k edge-disjoint paths between u and v.

Theorem (Menger, vertex version)**.** Let G be a graph and k a positive integer. Then G is k-vertex-connected if and only if for every pair $u, v \in V$ of distinct vertices of G there exists a system of k paths between u and v such that every two paths are vertex-disjoint except for u and v.

Corollary. For every graph G, we have $k_v(G) \leq k_e(G) \leq \delta(G)$.

Definition. An ear-decomposition of a graph $G = (V, E)$ is a sequence $G_0, G_1, \ldots, G_k$ of subgraphs of G satisfying

- G_0 is a cycle,
- for $i = 1, \ldots, k$, the graph G_i is obtained from G_{i-1} by adding a path P_i sharing exactly its endpoints with the graph G_{i-1} (and no edges).

Theorem. The following properties of a graph G are equivalent:

(i) G is 2-vertex-connected.

(ii) G has an ear-decomposition.

(iii) G can be obtained from K_3 by a sequence of edge additions and edge subdivisions.

Counting spanning trees

Definition. Let $\kappa(G)$ denote the number of distinct spanning trees of a graph G.

Proposition (Basic properties of κ)**.**

- $\kappa(C_n) = n$.
- $\kappa(G) = 0$ iff G is disconnected.
- $\kappa(G) = 1$ iff G is a tree.
- $\kappa(G \cup H) = \kappa(G) \cdot \kappa(H)$ if G and H are (multi)graphs with exactly one edge or exactly one vertex in common.

Theorem (Cayley's formula)**.** $\kappa(K_n) = n^{n-2}$ for every $n \geq 2$.

Theorem (Deletion-contraction formula)**.** Let G be a multigraph and e its edge. Then $\kappa(G) = \kappa(G - e) + \kappa(G/e)$, where G/e is multigraph contraction producing parallel edges, but no loops.

Definition. The *Laplace matrix* of a graph $G = (V, E)$, $V = \{v_1, \ldots, v_n\}$ is an $n \times n$ matrix with entries:

$$
\begin{aligned}
q_{ii} &= \deg(v_i) \\
q_{ij} &= \begin{cases} -1 & \text{if } \{v_i, v_j\} \in E \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Theorem. For every graph G, $\kappa(G) = \det Q_{11}$, where Q_{ij} denotes the matrix obtained from Q by deleting the i-th row and j-th column.

Extremal combinatorics

Theorem. Maximum number of edges of a graph on n vertices, containing no K_3 as a subgraph, is $\lceil n^2/4 \rceil$. Furthermore, all graphs achieving the maximum number of edges are isomorphic to $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Theorem. Let G be a graph on n vertices with m edges, containing no C_4 as a subgraph. Then $m \leq \frac{1}{2}(n^{3/2} + n)$.

Ramsey theory

Definition. The *clique number* $\omega(G)$ of a graph G is the maximum number of vertices in a complete subgraph. Similarly, the *independence number* $\alpha(G)$ is the maximum number of vertices in an independent set (that is, a set inducing a subgraph with no edges).

Theorem (Ramsey theorem on graphs)**.** Let $k, \ell \in \mathbb{N}$ and let $G = (V, E)$ be a graph with $|V| \geq \binom{k+\ell-2}{k-1}$. Then G contains a clique of order k or an independent set of order ℓ. (That is, $\omega(G) \geq k$ or $\alpha(G) \geq \ell$.)

Definition. For a given $k, \ell \in \mathbb{N}$, we define the *Ramsey number* $r(k, \ell)$ to be the minimal n such that every graph with at least n vertices contains a clique of order k or an independent set of order ℓ.

Theorem (Lower bound on Ramsey numbers)**.** $r(k, k) \geq 2^{k/2}$ for all $k \geq 3$.

Definition. $[n]$ will denote the set $\{1, \ldots, n\}$.

Theorem (The Pigeonhole principle)**.** Let k and t be positive integers and $n > (k-1) \cdot t$. Then for every function $c : [n] \to [t]$, there exists a k-element subset $A \subseteq [n]$ on which the function c is constant. (Intuitively: for every coloring of $[n]$ by t colors, there is a k-element monochromatic subset.)

Theorem (Ramsey for colored graphs)**.** For all integers $k > 0$ (required clique size) and $t > 0$ (the number of colors), there exists n (minimum graph size) such that for every function $c : \binom{[n]}{2} \to [t]$ (a coloring of edges of K_n) there is $A \in \binom{[n]}{k}$ such that c is constant on $\binom{A}{2}$ (a monochromatic copy of K_k).

Theorem (Infinite Pigeonhole principle)**.** For every $c : \mathbb{N} \to [t]$ (a coloring of natural numbers by t colors), there exists an infinite set $A \subseteq \mathbb{N}$ on which c is constant.

Theorem (Infinite Ramsey theorem)**.** For every $c : \binom{\mathbb{N}}{2} \to [t]$ (a coloring of an infinite complete graph by t colors), there exists an infinite set $A \subseteq \mathbb{N}$ such that c is constant on $\binom{A}{2}$ (an infinite monochromatic complete subgraph).

Theorem (Infinite Ramsey theorem for p-tuples)**.** For every $c : \binom{\mathbb{N}}{p} \to [t]$ (a coloring of p-tuples of natural numbers by t colors), there exists an infinite set $A \subseteq \mathbb{N}$ such that c is constant on $\binom{A}{p}$.

Claim (Finite Ramsey theorem for p-tuples)**.** For all integers $k > 0$ (required subset size), $t > 0$ (the number of colors) and $p > 0$ (tuple size), there exists n (minimum set size) such that for every function $c : \binom{[n]}{p} \to [t]$ (a coloring of p-tuples of $[n]$ by t colors), there exists $A \in \binom{[n]}{k}$ such that c is constant on $\binom{A}{p}$ (a monochromatic subsystem).

Theorem (Schur)**.** $\forall t \in \mathbb{N}\ \exists n \in \mathbb{N}\ \forall c : [n] \to [t]\ \exists x, y, z \in [n]$ such that $c(x) = c(y) = c(z)$ and $x + y = z$.

Theorem (Erdős-Szekeres)**.** For each $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that any n-element set of points in the plane in general position (no three on a line) contains k points forming a convex k-gon.

Error-correcting codes

Definition. Let Σ be a q-element set called the *alphabet*. Elements of Σ^n are called *words* of length n over Σ. A *code* of length n over Σ is a subset $C \subseteq \Sigma^n$. A *binary code* is a code over the alphabet $\{0, 1\}$. For a code C we define its *size* as $k = \log_q |C|$, and its *rate* as $\alpha(C) = k/n$.

Definition. The *Hamming distance* of words $x = (x_1, \ldots, x_n)$ and $y = (y_i, \ldots, y_n)$ in Σ^n is defined by $d(x, y) = |\{i \in \{1, \ldots, n\} \mid x_i \neq y_i\}|$. The *minimal distance* of a code C is $d(C) = \min d(x, y)$ over all distinct words $x, y \in C$. A code of length n and size k with minimal distance d is called a $(n, k, d)_q$*-code.* If q is clear from the context, it is usually omitted.

Example:

- The *total code* Σ^n contains all possible words. It is a $(n, n, 1)$-code.
- The *repetition code* $\{(x_1, \ldots, x_n) \mid x_1 = \ldots = x_n \in \Sigma\}$ is a $(n, 1, n)$-code.
- The *parity code* $\{(x_1, \ldots, x_{n-1}, x_1 + \ldots + x_{n-1})\}$ over the alphabet $\mathbb{Z}_t$ is a $(n, n-1, 2)$-code.

Theorem. A code *detects* up to e errors iff $d \geq e + 1$. A code *corrects* up to e errors iff $d \geq 2e + 1$.

Definition. A code C is *linear* if its alphabet is some finite field $\mathbb{F}_q$ and C is a subspace of the vector space $\mathbb{F}_q^n$. That is, codewords are closed under addition and multiplication by an element of $\mathbb{F}_q$. Parameters of linear codes are usually written in brackets: $[n, k, d]_q$.

Observation. The dimension of the subspace is equal to the size of the code. A linear code is completely described by the basis of the subspace, or by the corresponding *generator matrix* G, which is a $k \times n$ matrix whose rows are the vectors of the basis. The *dual code* $C^\perp$ is the orthogonal complement of C. Therefore, it has dimension $n - k$. Its generator matrix has size $(n - k) \times n$ and it is called the *parity check matrix* P of the code C.

Lemma. $x \in C \Leftrightarrow Px^{\mathrm{T}} = \mathbf{0}$.

Observation. Hamming distance in linear codes is invariant with respect to translation:

$$d(x, y) = d(x + z, y + z).$$

Therefore $d(C) = \min_{x \in C} w(x)$, where $w(x) = d(\mathbf{0}, x)$ is the *Hamming weight* of x.

Corollary. For a linear code, d is equal to the minimum non-zero number of columns of the parity-check matrix, which are linearly dependent.

Definition. The family of *binary Hamming codes* contains for every ℓ a linear code with a parity check matrix of shape $\ell \times (2^\ell - 1)$, whose columns contain binary expansions of all numbers $1, \ldots, 2^\ell - 1$.

Observation. For a given ℓ, the corresponding Hamming code is a $[2^\ell - 1, 2^\ell - \ell - 1, 3]$-code.

Theorem (Singleton's bound)**.** If there exists a (n, k, d)-code, then $k \leq n - d + 1$.

Definition. Let $x \in \{0,1\}^n$ and $0 \leq r \leq n$. A *combinatorial ball* with center x and radius r is the set

$$B(x, r) = \{z \in \{0,1\}^n \mid d(x, z) \leq r\}.$$

Lemma. The volume of the combinatorial ball $B(x, r)$ (in the space of dimension n) is

$$V(n, r) = \sum_{i=0}^{r} \binom{n}{i}.$$

Theorem (Hamming's bound)**.** Let C be a binary code with minimal distance $d(C) \geq 2r + 1$, then

$$|C| \leq \frac{2^n}{V(n, r)}.$$

Definition. A binary code C of length n and minimal distance $d(C) = 2r + 1$ is called *perfect* if $|C| = 2^n / V(n, r)$.

Corollary. All Hamming codes are perfect.