Number Theory

Definition: Let $a, b$ be integers, $a \neq 0$. We say that $a$ divides $b$, (or $b$ is a multiple of $a$), if there is an integer $k$ for which $b = ka$. Notation: $a | b$.

Definition: The integer $|p| > 1$ is a prime, if $p = ab$, then either $p = a$ or $p = b$ (i.e. $p$ has no proper divisors, only 1 and itself).

Proposition: For an integer $p$, $p$ is a prime if and only if it has the following property: if $p$ divides $ab$, then either $p$ divides $a$ or $p$ divides $b$.

Theorem (Fundamental Theorem of Algebra): Every integer $|n| > 1$ can be written as a product of primes in a unique way (up to the order and the signs of the primes).

Corollary: Every integer $n > 1$ has a unique canonical form, $n = p_1^{e_1} \cdot p_2^{e_2} \cdots p_r^{e_r}$, where $p_1 < p_2 < \cdots < p_r$ are primes, and $e_i > 0$ for all $i = 1, 2, \ldots, r$.

Proposition: If $n = p_1^{e_1} \cdot p_2^{e_2} \cdots \cdot p_r^{e_r}$ and $m = p_1^{f_1} \cdot p_2^{f_2} \cdots \cdot p_r^{f_r}$ with $e_i, f_i \geq 0$, then
1. $m|n$ if and only if $f_i \leq e_i$ for all $i = 1, 2, \ldots, r$.
2. $\operatorname{gcd}(m, n) = p_1^{\min(e_i, f_i)} \cdot p_2^{\min(e_2, f_2)} \cdots p_r^{\min(e_r, f_r)}$.
3. $\operatorname{lcm}(m, n) = p_1^{\max(e_1, f_1)} \cdot p_2^{\max(e_2, f_2)} \cdots p_r^{\max(e_r, f_r)}$.

Definition: $m$ and $n$ are relatively prime, if $\operatorname{gcd}(m, n) = 1$.

Definition: For an integer $n > 1$, $d(n)$ is the number of divisors of $n$.

Proposition: 1. $d(n) \geq 2$ and $d(n) = 2$ if and only if $n$ is a prime.
2. If $n = p_1^{e_1} \cdot p_2^{e_2} \cdots \cdot p_r^{e_r}$ then $d(n) = (e_1 + 1)(e_2 + 1)\cdots(e_r + 1)$.
3. The $d(n)$ function is multiplicative, i.e. if $\operatorname{gcd}(m, n) = 1$, then $d(mn) = d(m) \cdot d(n)$.

Definition: For an integer $n > 1$, $\varphi(n)$, the value of the Euler’s $\varphi$ function, is the number of positive integers less than $n$ which are relatively prime to $n$.

Proposition: 1. $\varphi(n) \leq n - 1$ and $\varphi(n) = n - 1$ if and only if $n$ is a prime.
2. If $n = p_1^{e_1} \cdot p_2^{e_2} \cdots \cdot p_r^{e_r}$ then $\varphi(n) = (p_1^{e_1} - p_1^{e_1 - 1}) \cdot (p_2^{e_2} - p_2^{e_2 - 1}) \cdots (p_r^{e_r} - p_r^{e_r - 1}) = n \cdot (1 - \frac{1}{p_1}) \cdot (1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_r})$.
3. The $\varphi(n)$ function is multiplicative, i.e. if $\operatorname{gcd}(m, n) = 1$, then $\varphi(mn) = \varphi(m) \cdot \varphi(n)$.

Theorems about primes

Theorem (Euclid): There are infinitely many primes.

Theorem: There are arbitrarily large gaps between consecutive primes.

Theorem (prime number theorem): if $\pi(n)$ is the number of primes less than $n$, then $\pi(n) \sim n/\ln n$, i.e. $\pi(n)/(n/\ln n) \to 1$, as $n \to \infty$.

Theorem (Dirichlet): If $\operatorname{gcd}(a, b) = 1$, then there are infinitely many primes of the form $ak + b$, where $k$ is an integer.

Congruences

Definition: For $m > 1$, and $a, b \in \mathbb{Z}$ we say that $a$ is congruent to $b$ modulo $m$, if $m$ divides $a - b$. Notation: $a \equiv b \pmod{m}$. $m$ is the modulus of the congruence.

Proposition: The congruence mod $m$ is compatible with the usual operations on integers (addition, subtraction, multiplication, exponentiation), i.e. if $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then
1. $a + c \equiv b + d \pmod{m}$.
2. $a - c \equiv b - d \pmod{m}$.
3. $ac \equiv bd \pmod{m}$
4. $a^k \equiv b^k \pmod{m}$ for every $k \in \mathbb{N}$.

Proposition (cancellation in congruences): $ac \equiv bc \pmod{m}$ if and only if $a \equiv b \pmod{\frac{m}{\operatorname{gcd}(c, m)}}$.

Definition: A congruence $ax \equiv b \pmod{m}$ with unknown $x$ is called a linear congruence.
Theorem: 1. If \( \text{g.c.d.}(a, m) \nmid b \), then the linear congruence \( ax \equiv b \pmod{m} \) has no solutions.
2. If \( \text{g.c.d.}(a, m) \mid b \), then the linear congruence \( ax \equiv b \pmod{m} \) has \( \text{g.c.d.}(a, m) \) solutions mod \( m \).

Definition: \( x \equiv a_1 \pmod{m_1}, x \equiv a_2 \pmod{m_2} \) is a simultaneous congruence system.

Theorem: The simultaneous congruence system \( x \equiv a_1 \pmod{m_1}, x \equiv a_2 \pmod{m_2} \) has a solution if and only if \( \text{g.c.d.}(m_1, m_2) / |a_1 - a_2| \), and in this case it has a unique solution mod l.c.m. \((m_1, m_2)\).

Definition: A reduced residue system mod \( m \) is a set of \( \varphi(m) \) pairwise non-congruent integers mod \( m \), each relatively prime to \( m \), i.e. a set of integers \( a_1, a_2, \ldots, a_k \), s.t.
1. \( \text{g.c.d.}(a_i, m) = 1 \) for all \( i = 1, 2, \ldots, k \),
2. \( a_i \neq a_j \pmod{m} \), if \( i \neq j \), \( i, j = 1, 2, \ldots, k \),
3. \( k = \varphi(m) \).

Lemma: If \( \text{g.c.d.}(a, m) = 1 \) and \( a_1, a_2, \ldots, a_\varphi(m) \) is a reduced residue system mod \( m \), then \( a_1 a, a_2 a, \ldots, a_\varphi(m) a \) is also a reduced residue system mod \( m \).

Theorem (Euler-Fermat): If \( \text{g.c.d.}(a, m) = 1 \) then \( a^{\varphi(m)} \equiv 1 \pmod{m} \).

Theorem (Fermat): 1. If \( p \nmid a \) then \( a^{p-1} \equiv 1 \pmod{p} \).
2. \( a^p \equiv a \pmod{p} \) for every integer \( a \).

Euclidean algorithm for determining the g.c.d. of \( m \) and \( n, m > n \):
Let \( m = k_1 \cdot n + r_1, 0 < r_1 < n \),
\( n = k_2 \cdot r_1 + r_2, 0 < r_2 < r_1 \),
\( r_1 = k_3 \cdot r_2 + r_3, 0 < r_3 < r_2 \),
\( \vdots \)
\( r_{i-1} = k_i \cdot r_i + r_{i+1}, 0 < r_i < r_{i-1} \),
\( r_i = k_{i+1} \cdot r_{i+1} + 0, \)
then g.c.d. \((m, n) = r_{i+1}.

Geometry of 3-space

Points and vectors in 3D have 3 coordinates.

Equation of a plane
A plane is determined by its normal vector \( \mathbf{n} = (A, B, C) \) and one of its points \( P_0(x_0, y_0, z_0) \).
If \( P(x, y, z) \) is a point on the plane, then \( \mathbf{n} \cdot \overrightarrow{P_0P} = 0 \).
The equation of the plane (with coordinates): \( A(x-x_0) + B(y-y_0) + C(z-z_0) = 0 \), or \( Ax + By + Cz = D \).

System of equations of a line
A line is determined by its direction vector \( \mathbf{v} = (a, b, c) \) and one of its points \( P_0(x_0, y_0, z_0) \).
If \( P(x, y, z) \) is a point on the line, then \( \overrightarrow{P_0P} = t \cdot \mathbf{v} \) for some \( t \in \mathbb{R} \).
The system of equations of the line (with coordinates): \( x = x_0 + ta, y = y_0 + tb, z = z_0 + tc, t \in \mathbb{R} \), or \( x = x_0, y = y_0, z = z_0 \) if none of \( a, b, c \) is 0. If \( a = 0 \), then \( x = x_0, y = y_0, -\frac{x_0 - x}{b} = -\frac{y - y_0}{c} \), and if \( a = b = 0 \), then \( x = x_0, y = y_0, z \in \mathbb{R} \).

The vector space \( \mathbb{R}^n \)

Definition 1: \( \mathbb{R}^n \) is the set of all column vectors with \( n \) real numbers (\( n \) coordinates).
The operations on \( \mathbb{R}^n \) are the (coordinatewise) addition of vectors and the (coordinatewise) multiplication by a scalar (=real number).

Proposition: If \( \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R} \), then
1. \( \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \) (addition is commutative),
2. \( (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) \) (addition is associative),
3. \( \lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v} \),
4. \( (\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u} \),
5. \( \lambda(\mu\mathbf{u}) = \lambda\mu\mathbf{u} \).

Definition 2: \( W \subseteq \mathbb{R}^n \), \( W \neq \emptyset \) is a subspace of \( \mathbb{R}^n \), if it satisfies
1. if \( \mathbf{u}, \mathbf{v} \in W \) then \( \mathbf{u} + \mathbf{v} \in W \) (i.e. \( W \) is closed under addition), and
2. if \( \mathbf{u} \in W, \lambda \in \mathbb{R} \), then \( \lambda\mathbf{u} \in W \) (i.e. \( W \) is closed under multiplication by a scalar).
Definition 3: Let \( u_1, \ldots, u_k \in \mathbb{R}^n \). Then for some \( \lambda_1, \ldots, \lambda_k \in \mathbb{R} \), the vector \( v = \lambda_1 u_1 + \ldots + \lambda_k u_k \) is a linear combination of \( u_1, \ldots, u_k \).

(In other words, \( v \) can be expressed using \( u_1, \ldots, u_k \).)

Definition 4: Let \( u_1, \ldots, u_k \in \mathbb{R}^n \). Then \( W \), which is the set of all the linear combinations of \( u_1, \ldots, u_k \), is the subspace spanned (or generated) by \( u_1, \ldots, u_k \).

We say that \( u_1, \ldots, u_k \) is a generating system for the subspace \( W \).

Notation: \( W = \langle u_1, \ldots, u_k \rangle \)

Proposition: \( W \) above is really a subspace of \( \mathbb{R}^n \) (according to Definition 2).

Definition 5: Let \( u_1, \ldots, u_k \in \mathbb{R}^n \). \( u_1, \ldots, u_k \) are linearly independent, if none of \( u_1, \ldots, u_k \) is a linear combination of the remaining vectors.

\( u_1, \ldots, u_k \) are linearly dependent, if they are not linearly independent, i.e. at least one of the vectors \( u_1, \ldots, u_k \) is a linear combination of the remaining vectors.

Definition 6: \( u_1, \ldots, u_k \) are linearly independent, if the zero vector can be expressed using them only in the trivial way (i.e. when all the coefficients are 0).

Proposition: Definitions 5 and 6 are equivalent.

Lemma: If the vectors \( u_1, \ldots, u_k \) are linearly independent, but \( u_1, \ldots, u_k, u_{k+1} \) are linearly dependent, then \( u_{k+1} \in \langle u_1, \ldots, u_k \rangle \).

Lemma: (Exchange theorem) Let \( W \subseteq \mathbb{R}^n \) be a subspace, \( u_1, \ldots, u_k \in W \) linearly independent, and \( v_1, \ldots, v_m \in W \) a generating system of \( W \). Then for each \( 1 \leq i \leq k \) there exists a \( 1 \leq j \leq m \) such that \( u_1, \ldots, u_{i-1}, v_j, u_{i+1}, \ldots, u_k \in W \) are also linearly independent.

Theorem: (I-G inequality) Let \( W \subseteq \mathbb{R}^n \) be a subspace, \( u_1, \ldots, u_k \in W \) linearly independent, and \( v_1, \ldots, v_m \in W \) a generating system of \( W \). Then \( k \leq m \).

Definition 7: Let \( W \subseteq \mathbb{R}^n \) be a subspace. \( B = \{ b_1, \ldots, b_k \} \) is a basis in \( W \), if it is linearly independent and a generating system in \( W \).

Theorem: Let \( W \subseteq \mathbb{R}^n \) be a subspace. If \( b_1, \ldots, b_k \) and \( c_1, \ldots, c_m \) are both bases in \( W \), then \( k = m \).

Definition 8: Let \( W \subseteq \mathbb{R}^n \) be a subspace. If \( b_1, \ldots, b_k \) is a basis in \( W \), then the dimension of \( W \) is \( k \).

Notation: \( \dim(W) = k \).

Remark: The dimension of \( W \) is well-defined because of the previous theorem.

Proposition: \( \{ u_1 = (1, 0, \ldots, 0)^T, u_2 = (0, 1, \ldots, 0)^T, \ldots, u_n = (0, 0, \ldots, 1)^T \} \) is a basis in \( \mathbb{R}^n \), the standard basis. Therefore \( \dim(\mathbb{R}^n) = n \).

Theorem: \( B = \{ b_1, \ldots, b_k \} \) is a basis in \( W \) if and only if each vector in \( W \) is a linear combination of \( b_1, \ldots, b_k \) in a unique way.

Definition: The coordinate vector of \( v \in W \) in a given basis \( B = \{ b_1, \ldots, b_k \} \) in \( W \) is \((\lambda_1, \lambda_2, \ldots, \lambda_k)^T\), if \( v = \lambda_1 b_1 + \ldots + \lambda_k b_k \).

Notation: \( [v]_B = (\lambda_1, \lambda_2, \ldots, \lambda_k)^T \).

Theorem: Let \( W \subseteq \mathbb{R}^n \) be a subspace. If \( u_1, \ldots, u_k \) are linearly independent vectors in \( W \), then \( u_1, \ldots, u_k \) can be extended to a basis in \( W \) (with finitely many vectors, maybe 0).

Corollary 1: Every subspace of \( \mathbb{R}^n \) has a basis (and a dimension).

Corollary 2: Let \( W \subseteq \mathbb{R}^n \) be a subspace of dimension \( k \). If \( u_1, \ldots, u_k \) are \( k \) linearly independent vectors in \( W \) then they are a basis in \( W \).

Corollary 3: If \( V \subseteq W \), \( V \neq W \) are subspaces in \( \mathbb{R}^n \), then \( \dim(V) < \dim(W) \).

Proposition: Let \( W \subseteq \mathbb{R}^n \) be a subspace. If \( u_1, \ldots, u_k \) is a generating system in \( W \), then there is a subset of \( u_1, \ldots, u_k \) which is a basis in \( W \).

Corollary: Let \( W \subseteq \mathbb{R}^n \) be a subspace of dimension \( k \). If \( u_1, \ldots, u_k \) is a generating system in \( W \) consisting of \( k \) vectors then they are a basis in \( W \).

**Linear mappings**
Definition: A mapping $f : \mathbb{R}^n \to \mathbb{R}^k$ is a linear mapping, if there is a $k \times n$ matrix $A$ for which $f(x) = A \cdot x$ for every $x \in \mathbb{R}^n$.
If $n = k$, then $f$ is a linear transformation.

A is the matrix of $f$, notation: $A = [f]$.

Theorem: $f : \mathbb{R}^n \to \mathbb{R}^k$ is a linear mapping if and only if it preserves the addition and the multiplication by a scalar, i.e.,
1) $f(y + z) = f(y) + f(z)$ for all $y, z \in \mathbb{R}^n$, and
2) $f(\lambda y) = \lambda f(y)$ for all $\lambda \in \mathbb{R}$, $y \in \mathbb{R}^n$.
In this case the $i$th column of the matrix of $f$ is $f(e_i)$ for $i = 1, 2, \ldots, n$, where $e_1, \ldots, e_n$ is the standard basis in $\mathbb{R}^n$.

Definition: If $f : \mathbb{R}^n \to \mathbb{R}^k$ is a linear mapping, then the kernel of $f$ is the set of vectors in $\mathbb{R}^n$ whose image is the zero vector in $\mathbb{R}^k$.
The image of $f$ is the set of vectors in $\mathbb{R}^k$ which are images under $f$ of some vector in $\mathbb{R}^n$.

Notation: $\text{Ker} f$ and $\text{Im} f$.

Proposition: $\text{Ker} f$ is a subspace in $\mathbb{R}^n$ and $\text{Im} f$ is a subspace in $\mathbb{R}^k$.

Theorem: (Dimension theorem) If $f : \mathbb{R}^n \to \mathbb{R}^k$ is a linear mapping, then $\dim \text{Ker} f + \dim \text{Im} f = n$.

Definition: If $f : \mathbb{R}^n \to \mathbb{R}^k$ and $g : \mathbb{R}^k \to \mathbb{R}^m$ are linear mappings, then the product (or composition) of them is $g \circ f : \mathbb{R}^n \to \mathbb{R}^m$, for which $(g \circ f)(x) = g(f(x))$ for every $x \in \mathbb{R}^n$.

Theorem: If $f : \mathbb{R}^n \to \mathbb{R}^k$ and $g : \mathbb{R}^k \to \mathbb{R}^m$ are linear mappings, then $g \circ f : \mathbb{R}^n \to \mathbb{R}^m$ is also a linear mapping, and its matrix is $[g \circ f] = [g] \cdot [f]$.

Definition: The inverse of a mapping $f : A \to B$ is $g : B \to A$, if $f(x) = y \iff g(y) = x$.

Theorem: A linear transformation $f : \mathbb{R}^n \to \mathbb{R}^n$ is invertible if and only if $\det(f) \neq 0$. In this case $f^{-1}$ is also a linear transformation and $[f^{-1}] = [f]^{-1}$.

Remark: $f$ is invertible $\iff \text{Ker} f = \{0\} \iff \text{Im} f = \mathbb{R}^n$.

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, $B = \{b_1, \ldots, b_n\}$ a basis in $\mathbb{R}^n$, and $B$ the $n \times n$ matrix whose columns are $b_1, \ldots, b_n$. Then the mapping $g : [x]_B \to [f(x)]_B$ is also a linear transformation.

Definition: In this case we say that the matrix of $f$ in the basis $B$ is $[g]$.

Notation: $[g] = [f]_B$.

Theorem: With the above notations,
1) $[f]_B = B^{-1} \cdot [f] \cdot B$,
2) $[f(\bar{x})]_B = [f]_B \cdot \bar{x}_B$,
3) the $i$th column of $[f]_B$ is $[f(b_i)]_B$ for $i = 1, 2, \ldots, n$.

Definition: Let $A$ be an $n \times n$ matrix. If $A \cdot x = \lambda \cdot x$ holds for a nonzero vector $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ then $x$ is an eigenvector of $A$ and $\lambda$ is an eigenvalue of $A$.

Theorem: $\lambda \in \mathbb{R}$ is an eigenvalue of the square matrix $A$ if and only if $\det(A - \lambda I) = 0$, where $I$ is the identity matrix. In this case the eigenvectors belonging to $\lambda$ are the nontrivial solutions of the system of equations $(A - \lambda I)x = 0$.

Definition: The characteristic polynomial of the square matrix $A$ is $\det(A - \lambda I)$, where $\lambda$ is a variable.

Proposition: (Diagonalisation of the matrix of a linear transformation) Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, and $B = \{b_1, \ldots, b_n\}$ a basis in $\mathbb{R}^n$. Then $[f]_B$ is a diagonal matrix if and only if each vector $b_j$ in $B$ is an eigenvector of $[f]$.

Determinants

Definition: If $A$ is an $n \times n$ (square) matrix with entries $a_{i,j}$, $i, j = 1, 2, \ldots, n$, then
$$\det(A) = \sum_{\text{all permutations } \pi} (-1)^{I(\pi)} \cdot a_{1,\pi(1)} \cdot a_{2,\pi(2)} \cdots a_{n,\pi(n)},$$
with $I(\pi)$ is the number of inversions of the permutation $\pi(1), \pi(2), \ldots, \pi(n)$. 