MATH COMPREHENSIVE EXAM NOTES

ALEXANDER LEE

Multivariable Calculus

Elementary Vector Analysis

Notation. The Cartesian coordinates: $(x, y) \in \mathbb{R}^2$ and $(x, y, z) \in \mathbb{R}^3$.

Notation. A vector $\vec{v}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$ is often represented by a directed line segment. In terms of coordinates, we write $\vec{v} = \langle a_1, a_2 \rangle$ in $\mathbb{R}^2$ and $\vec{v} = \langle a_1, a_2, a_2 \rangle$ in $\mathbb{R}^3$.

Notation. The standard basis vectors $\vec{\imath} = \langle 1, 0 \rangle$, $\vec{\jmath} = \langle 1, 0 \rangle$ in $\mathbb{R}^2$ and $\vec{\imath} = \langle 1, 0, 0 \rangle$, $\vec{\jmath} = \langle 0, 1, 0 \rangle$, $\vec{k} = \langle 0, 0, 1 \rangle$ in $\mathbb{R}^3$.

Notation. A vector $\vec{v}$ has length $|\vec{v}|$, sometimes denoted $||\vec{v}||$.

Fact. Nonzero vectors $\vec{v}$ and $\vec{u}$ are parallel if and only if each is a constant multiple of the other.

Definition. A point P in $\mathbb{R}^2$ or $\mathbb{R}^3$ gives a vector from the origin to P, called the position vector of P. This allows us to regard points as vectors and vice versa.

Definition. The length of a n-dimensional vector $\vec{v} = \langle a_1, \ldots, a_n \rangle$ is

$$|\vec{v}| = \sqrt{{a_1}^2 + \cdots + {a_n}^2}$$

Definition. In $\mathbb{R}^2$, the dot (scalar) product of $\vec{u} = \langle a_1, a_2 \rangle$ and $\vec{v} = \langle b_1, b_2 \rangle$ is $\vec{u} \cdot \vec{v} = a_1 b_1 + a_2 b_2$, and similarly in $\mathbb{R}^3$, the dot product of $\vec{u} = \langle a_1, a_2, a_3, \rangle$ and $\vec{v} = \langle b_1, b_2, b_3 \rangle$ is $\vec{u} \cdot \vec{v} = a_1 b_1 + a_2 b_2 + a_3 b_3$.

Facts. The linearity properties of the dot product:

(1) $\vec{u} \cdot \vec{u} = |\vec{u}|^2$.
(2) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.
(3) $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.
(4) $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{v} \cdot (c\vec{v})$.
(5) $\vec{0} \cdot \vec{u} = 0$.

Additional properties:

(1) $\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos\theta$, where θ is the angle between $\vec{u}$ and $\vec{v}$.
(2) $\vec{u} \cdot \vec{v} = 0$ if and only if $\vec{u}$ and $\vec{v}$ are perpendicular (orthogonal).

Definition. Given $\vec{u} = \langle a_1, a_2, a_3 \rangle$ and $\vec{v} = \langle b_1, b_2, b_3 \rangle$ in $\mathbb{R}^3$, their cross (vector) product is

$$\vec{u} \times \vec{v} = \det \begin{pmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = \det \begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} \vec{\imath} - \det \begin{pmatrix} a_1 & a_3 \\ b_1 & b_3 \end{pmatrix} \vec{\jmath} + \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \vec{k}.$$

Facts. The linearity properties of the cross product:

(1) $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$.
(2) $(c\vec{u}) \times \vec{v} = c(\vec{u} \times \vec{v}) = \vec{u} \times (c\vec{v})$.
(3) $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$.
(4) $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$.
(5) $\vec{u} \cdot (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$.
(6) $\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w}$.

Additional properties:

(1) $|\vec{u} \times \vec{v}| = |\vec{u}||\vec{v}| \sin\theta$, where θ is the angle between $\vec{u}$ and $\vec{v}$.

(2) $\vec{u} \times \vec{v} = \vec{0}$ if and only if $\vec{u}$ and $\vec{v}$ are parallel.
(3) $\vec{u} \times \vec{v}$ is perpendicular to both $\vec{u}$ and $\vec{v}$.

Definition. In $\mathbb{R}^2$ or $\mathbb{R}^3$, a point $\vec{r}_0$ and a nonzero vector $\vec{v}$ determine the line parametrized by

$$\vec{r}(t) = \vec{r}_0 + t\vec{v}.$$

The vector $\vec{v}$ is called a direction vector of the line. The parametric equations of a line for the coordinates $(x, y, z) \in \mathbb{R}^3$ are

$$x = x_0 + at \quad y = y_0 + bt \quad z = z_0 + ct,$$

where $\vec{v} = \langle a, b, c\rangle$, $\vec{r}_0 = \langle x_0, y_0, z_0\rangle$, and $t \in \mathbb{R}$. For $\mathbb{R}^2$, omit z.

Definition. A plan in $\mathbb{R}^3$ is defined by an equation of the form $ax + by + cz = d$ where $\langle a, b, c\rangle \neq \langle 0, 0, 0\rangle$. A more geometric way to write the equation uses a nonzero vector $\vec{n}$ perpendicular to the plane and point (x_0, y_0, z_0) in the plan. Then:

$$\begin{aligned}(x, y, z) \text{ is in the plane } &\iff \vec{n} \text{ is perpendicular to the vector from } (x, y, z) \text{ to } (x_0, y_0, z_0)\\ &\iff \vec{n} \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0.\end{aligned}$$

The vector $\vec{n}$ is called a normal vector to the plane. For a plane defined by $ax + by + cz = d$, a normal vector is given by $\vec{n} = \langle a, b, c\rangle$.

Definition. Given a curve parameterization $\vec{r}(t) = (x(t), y(t))$ in the plane, the tangent vector to the curve at the point $\vec{r}(t)$ is

$$\vec{r'}(t) = \langle x'(t), y'(t)\rangle.$$

The situation is similar on $\mathbb{R}^3$.

Functions of Several Variables

Definition. If f is a function of two variables, its partial derivatives are the functions f_x and f_y defined by

$$f_x(x, y) = \lim_{h\to 0} \frac{f(x + h, y) - f(x, y)}{h}$$
$$f_y(x, y) = \lim_{h\to 0} \frac{f(x, y + h) - f(x, y)}{h}$$

If f is a function of three variables x, y, and z, then its partial derivative with respect to x is defined as

$$f_x(x, y, z) = \lim_{h\to 0} \frac{f(x + h, y, z) - f(x, y, z)}{h}$$

Notation. The standard notation for the partial derivatives: $\frac{\partial f}{\partial x} = f_x(x, y)$, $\frac{\partial f}{\partial y} = f_y(x, y)$, $\frac{\partial^2 f}{\partial^2 x} = f_{xx}(x, y)$, $\frac{\partial^2 f}{\partial x \partial y} = f_{yx}(x, y)$, $\frac{\partial^2 f}{\partial^2 y} = f_{yy}(x, y)$ for $f(x, y)$, and similarly for $f(x, y, z)$.

Interpretation. If $z = f(x, y)$, then $\partial z/\partial x$ represents the rate of change of z with respect to x when y is fixed. Similarly, $\partial z/\partial y$ represents the rate of change of z with respect to y when x is fixed.

If $w = f(x, y, z)$, then $f_x = \partial w/\partial x$ can be interpreted as the rate of change of w with respect to x when y and z are held fixed.

Fact. Rule for find partial derivatives of $z = f(x, y)$.

(1) To find f_x, regard y as a constant and differentiate $f(x, y)$ with respect to x.
(2) To find f_y, regard x as a constant and differentiate $f(x, y)$ with respect to y.

Definition. A unit vector is a vector whose length is 1. In general, if $\vec{v} \neq \vec{0}$, then the unit vector that has the same direction as $\vec{v}$ is

$$\vec{u} = \frac{\vec{v}}{|\vec{v}|}.$$

Definition. The directional derivative of $f(x,y)$ in the direction of a unit vector $\vec{u} = \langle u_1, u_2 \rangle$ at the point (a,b) is

$$D_{\vec{u}} f(a,b) = \lim_{h \to 0} \frac{f(a+hu_1, b+hu_2) - f(a,b)}{h}$$

if this limit exists.

The directional derivative of $f(x,y,z)$ in the direction of a unit vector $\vec{u} = \langle u_1, u_2, u_3 \rangle$ at the point (a,b,c) is

$$D_{\vec{u}} f(a,b,c) = \lim_{h \to 0} \frac{f(a+hu_1, b+hu_2, c+hu_3) - f(a,b,c)}{h}$$

if this limit exists.

Interpretation. The directional derivative is the rate of change of a function of two or more variables in any direction.

Definition. The gradient of $f(x,y)$ at (a,b) is the vector

$$\nabla f(a,b) = \frac{\partial f}{\partial x}(a,b)\vec{\imath} + \frac{\partial f}{\partial y}(a,b)\vec{\jmath} = \left\langle \frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b) \right\rangle$$

and similarly for $f(x,y,z)$.

Fact. $\nabla f(a,b)$ is perpendicular to the level curve $f(x,y) = f(a,b)$ at the point (a,b). Similarly, $\nabla f(a,b,c)$ is perpendicular to the level surface $f(x,y,z) = f(a,b,c)$ at (a,b,c).

Theorem. If $f(a,b)$ is differentiable at (a,b) and $\vec{u}$ is a unit vector, then

$$D_{\vec{u}} f(a,b) = \nabla f(a,b) \cdot \vec{u},$$

and similarly for $f(x,y,z)$.

Theorem. When $\nabla f(a,b) \neq \vec{0}$, the unit vector $\nabla f(a,b)/|\nabla f(a,b)|$ gives the direction in which $f(x,y)$ is increasing most rapidly. Furthermore, the maximum rate of increase is $|\nabla f(a,b)|$. Similar results hold for $f(x,y,z)$.

Facts. Tangent planes arise in two situations:

- If $f(x,y)$ is differentiable at (x_0,y_0), then the tangent plane to the graph $z = f(x,y)$ at the point $(x_0,y_0,f(x_0,y_0))$ is defined by

$$z - f(x_0,y_0) = f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0). \tag{1}$$

- If $F(x,y,z)$ is differentiable at (x_0,y_0,z_0), then (x_0,y_0,z_0) lies on the level surface $F(x,y,z) = F(x_0,y_0,z_0)$, and the equation of the tangent plane to the surface at this point is defined by

$$\nabla F(x_0,y_0,z_0) \cdot (x-x_0, y-y_0, z-z_0) = 0, \tag{2}$$

  provided that the gradient $\nabla F(x_0,y_0,z_0)$ is nonzero. Written out, this is the equation

$$F_x(x_0,y_0,z_0)(x-x_0) + F_y(x_0,y_0,z_0)(y-y_0) + F_z(x_0,y_0,z_0)(z-z_0) = 0.$$

  The two situations are related since the graph $z = f(x,y)$ is the level surface $F(x,y,z) = f(x,y) - z = 0$. Since $\nabla F = f_x\vec{\imath} + f_y\vec{\jmath} - \vec{k}$, equation (2) reduces to equation (1) in this case.

Linear Algebra

Vector Spaces and Subspaces

Definition. A (real) *vector space* is a set V (whose elements are called *vectors*) together with

(1) an operation called *vector addition*, which for each pair of vector $\vec{x}, \vec{y} \in V$ produces another vector in V denoted $\vec{x} + \vec{y}$, and
(2) an operation called *multiplication by a scalar* (a real number), which for each vector $\vec{x} \in V$, and each scalar $c \in \mathbb{R}$ produces another vector in V denoted $c\vec{x}$.

Furthermore, the two operations must satisfy the follow *axioms*:

(1) For all vectors $\vec{x}, \vec{y}$, and $\vec{z} \in V$, $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$.
(2) For all vectors $\vec{x}$ and $\vec{y} \in V$, $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.

(3) There exists a vector $\vec{0} \in V$ with the property that $\vec{x} + \vec{0} = \vec{x}$ for all vectors $\vec{x} \in V$.
(4) For each vector $\vec{x} \in V$, there exists a vector denoted $-\vec{x}$ with the property that $\vec{x} + -\vec{x} = \vec{0}$.
(5) For all vectors $\vec{x}$ and $\vec{y} \in V$ and all scalars $c \in \mathbb{R}$, $c(\vec{x} + \vec{y}) = c\vec{x} + c\vec{y}$.
(6) For all vectors $\vec{x} \in V$, and all scalars c and $d \in \mathbb{R}$, $(c + d)\vec{x} = c\vec{x} + d\vec{x}$.
(7) For all vectors $\vec{x} \in V$, and all scalars c and $d \in \mathbb{R}$, $(cd)\vec{x} = c(d\vec{x})$.
(8) For all vectors $\vec{x} \in V$, $1\vec{x} = \vec{x}$.

Examples. Some simple vector spaces:

- $\mathbb{R}^n$ is the vector space of ordered n-tuples of real numbers. Note: $\dim(\mathbb{R}^n) = n$.
- $P_n(\mathbb{R})$ is the vector space of polynomials of degree *less than or equal to* n. Note: $\dim(P_n(\mathbb{R})) = n+1$.
- $M_{m\times n}(\mathbb{R})$ is the vector space of $m \times n$ matrices with real entries. Note: $\dim(M_{m\times n}(\mathbb{R})) = mn$.

Definition. Let V be a vector space and let $W \subseteq V$ be a subset. Then W is a (vector) *subspace* of V if W is a vector space itself under the operations of vector sum and scalar multiplication from V.

Notes. The empty set $\emptyset$ is not a vector space. Instead the smallest vector space is the trivial space, $\{\vec{0}\}$. Every vector space V has two obvious subspaces: the trivial subspace $\{\vec{0}\} \subseteq V$, and the improper subspace $V \subseteq V$.

Theorem (Subspace Theorem)**.** Let V be a vector space. A subset $W \subseteq V$ is a subspace if it satisfies the following properties:

(1) $W \neq \emptyset$
(2) For all $\vec{x}, \vec{y} \in W$ and all $c \in \mathbb{R}$, we have $c\vec{x} + \vec{y} \in W$.

Definition. Let V be a vector space, and let $S = \{\vec{v}_1, \ldots, \vec{v}_n\} \subseteq V$ be a finite set of vectors in V.

- A *linear combination* of elements of S is an expression $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$ for some scalars $a_1, \ldots, a_n \in \mathbb{R}$.
- The *span* of S, denoted $\mathrm{Span}(S)$, is the set of all linear combinations of elements of S. That is,

$$\mathrm{Span}(S) = \{a_1\vec{v}_1 + \cdots a_n\vec{v}_n \mid a_1, \ldots, a_n \in \mathbb{R}\}.$$

- We define $\mathrm{Span}(\emptyset) = \{\vec{0}\}$.
- If $\mathrm{Span}(S) = W$, we say that S spans W.

Fact. Let V be a vector space and let S be any subset of V. Then $\mathrm{Span}(S)$ is a subspace of V.

Fact. If W is a subspace and $S \subseteq W$, then $\mathrm{Span}(S) \subseteq W$.

Definition. The set S is *linearly dependent* if there exists scalars $a_1, \ldots, a_n \in \mathbb{R}$ that are not all zero such that $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}$. S is *linearly independent* if it is not linearly dependent. Equivalently, for any scalars $a_1, \ldots, a_n \in \mathbb{R}$ such that $a_1\vec{v}_1 + \cdots a_n\vec{v}_n = \vec{0}$, we must have $a_1 = \cdots = a_n = 0$.

Definition. The set $S \subseteq V$ is a basis for V if S is linearly independent and $\mathrm{Span}(S) = V$.

Definition. The *dimension* of V is the number $\dim(V)$ of elements in a basis for V. If V has no finite basis, we say $\dim(V) = \infty$.

Theorem. Any two bases of V have the same number of elements.

Fact. The three kinds of row reduction steps are

(1) Switching two rows.
(2) Multiplying a row by a nonzero scalar.
(3) Adding a multiple of one row to another.

Definition. A matrix is in *echelon form* if it satisfies all of the following conditions:

(1) If a row is not a zero row (i.e., all entries of that row are zeros), then the first nonzero entry is a 1 (and called the *pivot*).
(2) If a column contains a pivot, then all other entries in that column are 0.
(3) If a row contains a pivot, then each row above contains a pivot further to the left. This also implies that zero rows, if any, appear at the bottom.

Variables corresponding to the pivots are called *pivot variables*. All other variables are called *free variables*.

Definition. A *homogeneous* system of linear equations is when all the linear combinations equal 0. A system is *inhomogeneous* otherwise.

Definition. The *nullspace* of a matrix A is the solution set of its corresponding homogeneous system of equations. The basis of the nullspace of A is the set of vectors that the free variables end up multiplied by in the solution.

Definition. The *column space* of a matrix A is the span of its columns. If B is the echelon form of A, then the columns of A corresponding to the columns of B with pivots form a basis of the column space.

Linear Transformations

Definition. Let V and W be vector spaces, and let $T : V \to W$ be a function. We say T is a *linear transformation* (or a *linear map*, or simply that T is *linear*) if for all $\vec{x}, \vec{y} \in V$ and all $c \in \mathbb{R}$, we have

$$T(c\vec{x} + \vec{y}) = cT(\vec{x}) + T(\vec{y}).$$

That is, T *respects addition* and *scalar multiplication*.

Corollary. When T is linear,

$$T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) = a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n).$$

Theorem. Let $T : V \to W$ be a linear transformation. If $U : W \to X$ is another linear transformation, then the composition $U \circ T : V \to X$ is also linear. The composition $U \circ T$ is often denoted simply UT.

Theorem. If $A \in M_{m\times n}(\mathbb{R})$ is an $m \times n$ matrix, then the function $T : \mathbb{R}^n \to \mathbb{R}^m$ by $T(\vec{x}) = A\vec{x}$ is linear.

Theorem. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then there is a matrix $A \in M_{m\times n}(\mathbb{R})$ so that T is given by $T(\vec{x}) = A\vec{x}$.

Definition. Let $T : V \to W$ be a linear transformation.

- The *kernel* or *nullspace* of T is
$$\{\vec{v} \in V \mid T(\vec{v}) = \vec{0}_W\} \subseteq V.$$
It is usually denoted as $\mathrm{Ker}(T)$ or $N(T)$. Its dimension $\dim(\mathrm{Ker}(T))$ is called the *nullity* of T, sometimes denoted $\mathrm{nullity}(T)$.
- The *image* or *range* of T is
$$\{T(\vec{v}) \mid \vec{v} \in V\} = \{\vec{w} \in W \mid \exists \vec{v} \in V \text{ s.t. } T(\vec{v}) = \vec{w}\} \subseteq W.$$
It is usually denoted as $\mathrm{Im}(T)$ or $R(T)$. Its dimension $\dim(\mathrm{Im}(T))$ is called the *rank* of T, sometimes denoted $\mathrm{rank}(T)$. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by a matrix A, the range of T is sometimes called the *column space* of A, because it is precisely the span of the columns of A.

Facts. Let $T : V \to W$ be a linear transformation.

- $N(T)$ is a subspace of V, and $R(T)$ is a subspace of W.
- T is one-to-one if and only if $N(T) = \{\vec{0}_V\}$, i.e., iff $N(T)$ is as small as possible, i.e., iff $\mathrm{nullity}(T) = 0$.
- T is onto if and only if $R(T) = W$, i.e., iff $R(T)$ is as big as possible. If $\dim(W) < \infty$, this is equivalent to saying $\mathrm{rank}(T) = \dim(W)$.
- If $T : \mathbb{R}^n \to \mathbb{R}^m$ is represented by a matrix A, then
$$N(T) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\} \text{ and } R(T) = \text{column space of } A.$$
Furthermore, after finding an echelon form of A via row reduction, $\mathrm{rank}(T)$ is the number of columns *with* pivots (since the corresponding columns form a basis of $R(T)$), and $\mathrm{nullity}(T)$ is the number of columns *without* pivots (since the vectors that give the general solution form a basis of $N(T)$).

Fact. If $X \subseteq V$ is a subspace, then $\dim(X) \leq \dim(V)$. Moreover, if $\dim(V) < \infty$, then $\dim(X) = \dim(V)$ if and only if $X = V$.

Fact. Let V be a vector space with $\dim(V) = n$, and let $S \subseteq V$ be a set of m distinct vectors in V.

- If $m < n$, then S cannot span V.
- If $m > n$, then S cannot be linearly independent.

Theorem. Let V be a vector space, and let $S \subseteq V$ be a set of n distinct vectors in V. If any two of the following conditions hold, then all three hold (and S is a basis for V):

(1) S is linearly independent.
(2) S spans V.
(3) $\dim(V) = n$.

Theorem. Let $T : V \to W$ be a linear transformation, and suppose that at least one of V, W is finite-dimensional. If any two of the following conditions hold, then all three hold:

(1) T is one-to-one.
(2) T is onto.
(3) $\dim(V) = \dim(W)$.

In this case, T is *invertible*.

Theorem (Rank-Nullity/Dimension Theorem)**.** Let $T : V \to W$ be a linear transformation. Then

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

Abstract Algebra

Groups

Definition. Suppose that:

(1) G is a set and $*$ is a binary operation on G,
(2) $*$ is associative,
(3) there exists an element e of G such that $\forall x \in G$

$$x * e = e * x = x \text{ (identity element)}$$

(4) for each $x \in G$, there exists an element $y \in G$ such that

$$x * y = y * x = e$$

Then G, together with the binary operation $*$, is called a *group* and denoted $(G, *)$.

Theorem. If $(G, *)$ is a group, then there is only one identity element in G.

Theorem. If $(G, *)$ is a group and $x \in G$, then x has only one inverse.

Definition. We say a group G is *abelian* if the group operation is commutative, i.e., if $xy = yx$ for all $x, y \in G$.

Notation. For $g \in G$ and $n \in \mathbb{Z}$, $g^n = g \cdot g \cdots g$ (n-times).

Notation. If the operation of G is called $+$, then we write ng instead of g^n.

Notation. When G is finite, its order $|G|$ is the number of elements in the group.

Notation. When $g \in G$ has finite order, its order $o(g)$ is the smallest integer $m > 0$ with $g^m = e$.

Notation. If $g^m = e$ for some $m \in \mathbb{Z}$, then $o(g) \mid m$.

Subgroups

Definition. Let $(G, *)$ be a group and $H \subseteq G$. H is a *subgroup* of G if the elements of H from a group under $*$. I.e. $(H, *)$ is a group.

Theorem. Let H be a nonempty subset of group G. Then, H is a subgroup of G if and only if

(1) $\forall a, b \in H$, $ab \in H$ and
(2) $\forall a \in H$, $a^{-1} \in H$.

Terminology: If H has property 1 we say it is closed under multiplication. If H has property 2 we say it is closed under inverses.

Definition. An element $g \in G$ generates the *cyclic* subgroup $\langle g \rangle = \{g^m \mid m \in \mathbb{Z}\}$.

Theorem. An element $g \in G$ has finite order if and only if $\langle g \rangle$ is finite, in which case $o(g) = |\langle g \rangle|$.

Theorem (Lagrange's Theorem)**.** If H is a subgroup of a finite group G, then $|H| \mid |G|$.

Corollary. If $|G|$ is prime, the G is cyclic.

Corollary. For all $g \in G$, we have $o(g) \mid |G|$.

Corollary. For all $g \in G$, we have $g^{|G|} = e$.

Cosets

Definition. If H is a subgroup of G, then by *right coset* of H in G we mean a subset of the form Hg, where $g \in G$ and

$$Hg = \{hg \mid h \in H\}.$$

Definition. If H is a subgroup of G, then by *left coset* of H in G we mean a subset of the form gH, where $g \in G$ and

$$gH = \{gh \mid h \in H\}.$$

Theorem. Two right cosets Hx and Hy are either the same set or disjoint sets. (That is, if they share even one element, they are exactly the same set.) The same holds for left cosets. (On the other hand, a right coset and a left coset can intersect each other without being the same set.)

Corollary (Right coset relation)**.** $Hx = Hy \iff xy^{-1} \in H \iff x \in Hy$.

Corollary (Left coset relation)**.** $xH = yH \iff y^{-1}x \in H \iff x \in yH$.

Corollary. $Hx = H \iff x \in H \iff xH = H$.

Notation. If G is abelian and the group operation is written as addition, then the left and right cosets of $H \subseteq G$ coincide and are written

$$H + a = \{h + a \mid h \in H\}.$$

Here, the coset relation becomes

$$H + a = H + b \iff a - b \in H \iff a \in H + b.$$

Definition. When a group G is a union of finitely many left cosets of a subgroup H, we say that H has finite index in G and the *index* of H in G is defined to be

$$[G : H] = \text{number of distinct left cosets of } H \text{ in } G.$$

The same holds for right cosets. When G is finite, $[G : H] = |G|/|H|$, since all cosets of H have the same number of elements.

Normal Subgroups

Theorem. Given a subgroup $H \subseteq G$, H being normal in G is equivalent to any of the following conditions:

- $gH = Hg$ for all $g \in G$.
- $gHg^{-1} = H$ for all $g \in G$.
- $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$.

Theorem. When $N \subseteq G$ is a normal subgroup, every left coset is a right coset, and vice versa. The set of all cosets of N in G forms a group and is denoted G/N. The group operation is defined by $Na \cdot Nb = Nab$, which is well-defined since N is normal. When G is finite, G/N is also finite and $|G/N| = [G : N] = |G|/|N|$.

Group Homomorphisms

Definition. Let G and H be groups and let $\phi : G \to H$ be a function. We say that ϕ is a *homomorphism* if for all $a, b \in G$,

$$\phi(ab) = \phi(a)\phi(b).$$

Theorem. If $\phi : G \to H$ is a homomorphism, then

- $\phi(e_G) = e_H$.
- $\phi(g^n) = \phi(g)^n$ for all $g \in G$ and $n \in \mathbb{Z}$.

Definition. If $\phi : G \to H$ is a homomorphism, then

- The *kernel* of ϕ is $\mathrm{Ker}(\phi) = \{g \in G \mid \phi(g) = e_H\} \subseteq G$.
- The *image* of ϕ is $\mathrm{Im}(\phi) = \{\phi(g) \mid g \in G\} \subseteq H$.

Theorem. If $\phi : G \to H$ is a homomorphism, then $\mathrm{Ker}(\phi)$ is a normal subgroup of G.

Theorem. If $\phi : G \to H$ is a homomorphism, then $\mathrm{Im}(\phi)$ is a subgroup of H, but not necessarily normal.

Theorem. Given $\phi : G \to H$ is a homomorphism, ϕ is one-to-one if and only if $\mathrm{Ker}(\phi) = \{e_G\}$.

Theorem. Given a group homomorphism $\phi : G \to H$, ϕ being an isomorphism is equivalent to

- ϕ is one-to-one and onto.
- ϕ has an inverse function $\phi^{-1} : H \to G$ that is a group homomorphism.

Theorem (The Fundamental Theorem of Group Homomorphisms)**.** If $\phi : G \to H$ is a group homomorphism, then there is a group isomorphism $\widetilde{\phi} : G/\mathrm{Ker}(\phi) \simeq \mathrm{Im}(\phi)$ defined by $\widetilde{\phi}(g\mathrm{Ker}(\phi)) = \phi(g)$.