Mathematics

P. Tansuntorn

Last updated July 21, 2019

Contents

Part I

Part IA

Chapter 1

Numbers and Sets

1.1 Introduction to number systems and logic

1.2 Sets, relations and functions

1.2.1 Union, intersection and equality of sets

1.2.2 Indicator functions

1.2.3 Functions

1.2.4 Relations and equivalence relations

1.2.5 The Inclusion-Exclusion Principle

1.3 The integers

1.3.1 Natural numbers

1.4 Elementary number theory

1.4.1 Prime numbers

Definition 1.4.1. For two integers a and b, a *divides* b if there exists an integer k such that $b = ak$. We call a a *factor* of b and write $a \mid b$.

Definition 1.4.2. A number p is *prime* if its divisors are only 1 and itself. A number which is not prime is called a *composite* number.

Theorem 1.1. Every number greater than 1 has a prime factor.

Proof. We proceed by induction. Note that 2 obviously has a prime factor 2. Suppose that every number less than m has a prime factor, we need to show that m also has a prime factor. If m is prime then we are done. If m is not prime then there exists $a, b \in \mathbb{N}$ with $a \leq m$ such that $ab = m$ and $a \neq 1$. Then by the hypothesis, a has a prime factor. That prime factor must also divide m. Thus every number greater than 1 has a prime factor. $\square$

This proof of infinitude of prime is first described by Euclid.

Theorem 1.2. There are infinitely many prime numbers.

Proof. Suppose there are only finitely many prime numbers, denoted $p_1, \ldots, p_k$. Consider the number obtained by multiplying all primes in the list, and then adding one; $p_1 p_2 \cdots p_k + 1$. This number is obviously greater than 1, and so it must have a prime factor q. It then follows that q must be one of the finitely many primes in the list. But for all p_i with $1 \leq i \leq k$, $p_i \nmid p_1 p_2 \cdots p_k + 1$. This means that q is not equal to any of the prime in the list, a contradiction. $\square$

1.4.2 Euclid's algorithm

1.4.3 Solution in integers of $ax + by = c$.

1.4.4 Modular arithmetic

1.4.5 Chinese remainder theorem

1.4.6 Wilson's theorem

1.5 The real numbers

1.5.1 Least upper bounds

1.5.2 Sequences and series

1.5.3 Irrationality of $\sqrt{2}$ and e

What does it mean for a number to be rational? Recalls the definition of a rational number, which says that a number a is rational if it can be expressed in the form

$$a = \frac{p}{q}$$

for relatively prime integers p, q with $q \neq 0$.

We start by the classic proof of irrationality of $\sqrt{2}$.

Theorem 1.3. $\sqrt{2}$ is irrational.

Proof. Suppose $\sqrt{2}$ is rational, and $\sqrt{2} = \frac{p}{q}$ with $(p, q) = 1$ and $q \neq 0$. Then $(\sqrt{2})^2 = 2 = \frac{p^2}{q^2}$, so $2q^2 = p^2$. Therefore $2 \mid p^2$; it follows that p is even. But then $p = 2p_0$ for some integer p_0, which means that $q^2 = 2(p_0)^2$ and q is even. But this contradicts our assumption that p and q are relatively prime. $\square$

More generally,

Theorem 1.4. $\sqrt{p}$ is irrational if p is prime.

Proof. We provide another proof using unique factorisation of integers. Assume that $\sqrt{p}$ is a rational number, and that $\sqrt{p} = \frac{a}{b}$, with coprime a, b and $b \neq 0$. If $b = 1$, then p must divide a^2, then it divides a, which is absurd. Then there exists a prime q in the factorisation of b such that $q \nmid a$, or else they have a common factor.

Now consider $2 = \frac{a^2}{b^2}$. a^2 is factored into the product of primes of a, but squared. The prime factor of b^2 includes q^2. As so the fraction $\frac{a^2}{b^2}$ cannot be reduced to an integer, contradicting $2 = \frac{a^2}{b^2}$. $\square$

We can extend the result to the following theorem.

Theorem 1.5. $\sqrt{\frac{p}{q}}$ is rational if and only if p and q are perfect squares.

Even more generally,

Theorem 1.6. If an integer a is not an exact k-th power of another integer then $\sqrt[k]{a}$ is irrational.

We now provide a proof that e is irrational, starting with the definition of e.

Definition 1.5.1. The number e is defined as

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n.$$

We will show later on that the two definition is indeed equal. The proof of irrationality of e will use the fact that

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots.$$

Note that $2 = 1 + 1 < e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots < 1 + \left(1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots\right) = 3$, that is e is bounded between 2 and 3. Now we present the proof of irrationality of e, as presented by Joseph Fourier.

Theorem 1.7. e is irrational.

Proof. Suppose e is rational and with usual condition $(a, b) = 1$, $e = \frac{a}{b}$. Define

$$x = b!\left(e - \sum_{n=0}^{b} \frac{1}{n!}\right). \tag{1.1}$$

This renders x an integer, for if we substitute $e = \frac{a}{b}$,

$$x = b!\left(\frac{a}{b} - \sum_{n=0}^{b} \frac{1}{n!}\right) = a(b-1)! - \sum_{n=0}^{b} \frac{b!}{n!}.$$

For each of $0 \le n \le b$, $n!$ divides entirely into $b!$, so the sum is an integer.

Notice that we are using an idea that the difference between the fast-converging series expansion of e and $\sum_{n=0}^{b} \frac{1}{n!}$ multiplied by $b!$ is still less than 1, thus making x an integer between 0 and 1. This would give us a contradiction.

Let's bound x first by showing that it is indeed positive, since

$$x = b!\left(e - \sum_{n=0}^{b} \frac{1}{n!}\right) = b!\left(\sum_{n=0}^{\infty} \frac{1}{n!} - \sum_{n=0}^{b} \frac{1}{n!}\right) = \sum_{n=b+1}^{\infty} \frac{b!}{n!}, \tag{1.2}$$

and all of its terms is positive, so $x > 0$.

Consider $\frac{b!}{n!}$. For all term $n \ge b+1$,

$$\frac{b!}{n!} = \frac{1}{(b+1)(b+2)\cdots(b+(n-b))} < \frac{1}{(b+1)^{n-b}}.$$

The inequality is strict for $n > b+1$, we now have

$$x = \sum_{n=b+1}^{\infty} \frac{b!}{n!} < \sum_{n=b+1}^{\infty} \frac{1}{(b+1)^{n-b}} = \sum_{k=1}^{\infty} \frac{1}{(b+1)^k} = \frac{1}{b+1}\left(\frac{1}{1-\frac{1}{b+1}}\right) = \frac{1}{b} < 1. \tag{1.3}$$

A contradiction. $\square$

Later in the 19th century, e is proven to be transcendental, i.e. e is not a root of any polynomial with rational coefficient, by Charles Hermite. Furthermore, e^a is transcendental if a is rational and non-zero, by the result of the Lindemann-Weierstrass theorem. The same theorem also shows that π is transcendental.

1.5.4 Decimal expansions

1.5.5 Construction of a transcendental number

1.6 Countability and uncountability

Chapter 2

Groups

2.1 Examples of groups

2.1.1 Axioms for groups

Definition 2.1.1. A *group* is a set G, together with a binary operation $*$ on G with the following properties.

1. (Closure) for all g and h in G, $g * h \in G$;
2. (Associativity) for all f, g and h in G, $g * h \in G$, $f * (g * h) = (f * g) * h$;
3. (Existence of identity) there is a unique e in G such that for all g in G, $g * e = g = e * g$;
4. (Existence of inverse) if $g \in G$ there is some h in G such that $g * h = e = h * g$.

These results follow nicely.

Lemma 2.1.1. Let G be any group. Then, given $g \in G$, there is only one element h such that $g * h = e = h * g$. Particularly $(g^{-1})^{-1} = g$.

Lemma 2.1.2 (Cancellation law)**.** Suppose that a, b and x are in a group G. If $a * x = b * x$ then $a = b$.

Lemma 2.1.3. Suppose that a and b are in a group G. Then the equation $a * x = b$ has a unique solution $x = a^{-1} * b$.

Lemma 2.1.4. In any group G, e is the unique solution of $x * x = x$.

Notice that we do not include the familiar assumption that $f * g = g * f$ normally found in arithmetic. In fact, for some interesting groups this equality does not hold.

Definition 2.1.2. Let G be a group with respect to $*$. The elements f and g *commute* if $f * g = g * f$. We call G *abelian* if for all f and g in G, we have $f * g = g * f$.

We adopt the notation gh as equivalent to $g * h$ for simplicity.

2.1.2 Examples from geometry

In this section we examine the idea of group in geometry, using polygons.

2.1.3 Permutation on a set

In this section we will show that permutations of a non-empty set X, in fact, form a group, starting with the definition of permutations acting on a set, although only for finite sets, before developing the idea further into arbitrary sets.

Definition 2.1.3. A *permutation* $\alpha\colon X \to X$ is a bijection from X to itself. We say that α acts on the set X. The set of all permutations of X is denoted $\mathcal{P}(X)$.

This set is indeed a group.

Theorem 2.1. The set $\mathcal{P}(X)$ forms a group under composition of functions. We shall write $\alpha\beta(x)$ in place of $\alpha(\beta(x))$.

Proof. We will show that all group axioms are satisfied.

1. It is obvious that if α, β are permutations, then $\alpha\beta$ is also a permutation. Thus the set $\mathcal{P}(X)$ is closed under composition.

2. For any permutations α, β, γ, let $\mu = \alpha\beta$ and $\nu = \beta\gamma$. Then for every x in X,

$$\begin{aligned}(\alpha(\beta\gamma))(x) &= (\alpha\nu)(x)\\ &= \alpha(\nu(x))\\ &= \alpha(\beta(\gamma(x)))\\ &= \mu(\gamma(x))\\ &= (\mu\gamma)(x)\\ &= ((\alpha\beta)\gamma)(x).\end{aligned} \tag{2.1}$$

   Thus the permutations are commutative under composition.

3. The identity permutation $\iota(x) = x$ is the identity of $\mathcal{P}(X)$, since $\alpha\iota(x) = \alpha(x) = \iota\alpha(x)$.

4. For any element α of X, the inverse is simply its functional inverse α^{-1}. Direct verification shows that $\alpha\alpha^{-1} = \iota = \alpha^{-1}\alpha$.

□

The above proof lets us write $\alpha\beta\gamma$ for any composition of three or more permutations without any confusion.

Setting $X = \{1, \dots, n\}$, the study of permutation groups is simpler. We shall give a name for such group.

Definition 2.1.4. The *symmetric group* S_n is a set of permutations of $\{1, \dots, n\}$. We say that the group is of degree n.

Theorem 2.2. The order of S_n is $n!$.

Proof. Evidently, there are $n!$ permutations on a set with n elements. $\square$

We now introduce a customary notation for permutation $\rho(x)$ in the form

$$\rho = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \rho(1) & \rho(2) & \rho(3) & \cdots & \rho(n) \end{pmatrix},$$

which mean that the image of the permutation $\rho(i)$ is underneath i in the first row. For example, let α be a permutation on $\{1, 2, 3, 4\}$ with $\alpha(1) = 1, \alpha(2) = 4, \alpha(3) = 2$ and $\alpha(4) = 3$, then

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}.$$

Example 2.1.1. There are 6 permutations in S_3, they are

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Therefore S_3 is not abelian. More generally S_n is not abelian for $n \geq 3$. We will study permutations in more details later on.

2.1.4 Subgroups and homomorphisms

Definition 2.1.5. A *subgroup* of a group G is a subset of G which itself form a group under the operation taken from G.

Theorem 2.3. Let H be a subgroup of G, then the identity element of H is that of G.

A group G always at least admits two subgroup, namely G and the singleton $\{e\}$. We call $\{e\}$ the *trivial subgroup* of G, and we say that H is the *non-trivial subgroup* of G if $H \neq \{e\}$. We say that H is a *proper subgroup* of G if $H \neq G$.

We now give a test for a subset to be a subgroup.

Theorem 2.4 (A test for subgroup)**.** Let G be a group, and H be a non-empty subset of G. Then H is a subgroup of G if and only if

1. if $g \in H$ and $h \in H$, then $gh \in H$, and
2. if $g \in H$ then $g^{-1} \in H$.

Another test is similar and follows from the above theorem.

Theorem 2.5. Let G be a group, and H be a non-empty subset of G. Then H is a subgroup of G if and only if $xy^{-1} \in H$ whenever $x, y \in H$.

Example 2.1.2. The group $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{R}, +)$.

The following property of the class of subsets of G is important.

Theorem 2.6. Let G be any group, then the intersection of any collection of subgroups of G is itself a subgroup of G.

Proof. Note that the intersection $\cap_t H_t$ of the subgroups of G, defined as H_t for some t in the index set T, is not empty. Then for every elements $g \in \cap_t H_t$ and $h \in \cap_t H_t$, they also lie in H_t for every t. And thus $gh \in H_t$, so $gh \in \cap_t H_t$. Any element $g \in \cap_t H_t$ also has its inverse in every subgroup H_t. It then follows that $g^{-1} \in \cap_t H_t$. Therefore $\cap_t H_t$ forms a subgroup under the operation of G. □

As a consequence, we see that for any non-empty subset G_0 of G, we can consider the intersection of the collection of all subgroups H of G than contain G_0. The collection is not empty, since G is itself in the collection. It follows that the intersection is itself not empty, and is a subgroup of G that contain G_0. In fact, it is the *smallest subgroup* to contain G_0. This allows us to propose the next definition.

Definition 2.1.6. Let G_0 be a non-empty subset of a group G. The subgroup of G *generated by* G_0 is the smallest subgroup of G that contains G_0.

The idea of subgroup is expanded into the notion of a **coset**, which will be explored later.

Let's now turn to *homomorphism*, as a tool to study relationship between two groups.

Definition 2.1.7. Let G, G' be groups. A function $\phi \colon G \to G'$ is a *homomorphism* if it takes the action of G to that of H, namely

$$\phi(xy) = \phi(x)\phi(y),$$

for all $x, y \in G$.

Definition 2.1.8. A homomorphism ϕ is called an *isomorphism* if it is bijective.

Lemma 2.1.5. The homomorphism $\phi \colon G \to H$ sends the identity of G to that of H.

Proof. Let $x = y = e_G$. So $\phi(e_G) = \phi(e_G)\phi(e_G)$. This equation is satisfied only when $\phi(e_G) = e_H$. □

Lemma 2.1.6. $\phi(xy^{-1}) = \phi(x)\phi(y)^{-1}$.

Proof. This is clear from the fact that $\phi(y)\phi(xy^{-1}) = \phi(x)$. □

Lemma 2.1.7. $\phi(x^{-1}) = \phi(x)^{-1}$.

Lemma 2.1.8. If $\phi\colon G \to H$ and $\theta\colon H \to K$ are homomorphisms, then $\theta\phi\colon G \to K$ is also a homomorphism. Similarly, if $\phi\colon G \to H$ and $\theta\colon H \to K$ are isomorphisms, then $\theta\phi\colon G \to K$ is an isomorphism.

The idea of kernel, introduced for vector spaces, motivates us to find an analogy for homomorphisms between groups. As the kernel of a linear map is the set of vectors mapped to the identity elements of the image spaces, we naturally define kernel as follows.

Definition 2.1.9. The *kernel* $\ker\phi$ of a homomorphism $\phi\colon G \to H$ is the set of elements of g mapped to the identity of H, that is,

$$\{g \in G\colon \phi(g) = e_H\}.$$

Theorem 2.7. Let $\phi\colon G \to H$. Then $\ker\phi$ is a subgroup of G.

This result is similar to those of kernels of vector spaces.

2.1.5 Symmetry groups

2.2 The Möbius group

We first begin with the definition of Möbius transformation.

Definition 2.2.1. A *Möbius transformation* is a function f of a complex variable z in the form

$$f(z) = \frac{az+b}{cz+d},$$

for some complex numbers a, b, c and d, with the condition that $ad - bc \neq 0$.

The condition $ad - bc \neq 0$ might not be obvious, but it follows from the fact that

$$f(z) - f(w) = \frac{(ad-bc)(z-w)}{(cz+d)(cw+d)}.$$

If $ad - bc = 0$, then f is constant. This also shows that f is injective.

This definition of the Möbius transformation has two problems. First, a Möbius transformation f is not unique. As for example, the 4-tuples (a, b, c, d) and (ma, mb, mc, md) with $m \neq 0$ will all map a complex number z to a same number. Thus, given f, we *cannot* say what are the coefficients.

The second problem stems from the fact that, for example $1/(z - z_0)$ is not defined at the point z_0. This means that there is no subset of $\mathbb{C}$ on which all Möbius maps are defined.

Here is an example of this.

Example 2.2.1. Let $f(z) = (z+2)/z$ and $g(z) = (z+1)/(z-1)$. Then,

$$f(g(z)) = \frac{g(z)+2}{g(z)} = \frac{(z+1)+2(z-1)}{z+1} = \frac{3z-1}{z+1},$$

so that fg fixes the point 1. However, g *is not defined when* $z = 1$. What's worse is that, if $h(z) = 1/z$ then $hfg(z) = (z+1)/(3z-1)$, although g is not defined when $z = 1$, $fg(z)$ is not defined when $z = -1$, and $hfg(z)$ is not defined when $z = 1/3$. More generally, a composition $f_1 \cdots f_n$ of Möbius maps will not be defined at n distinct points in the complex plane.

The following theorem addresses the first problem.

Theorem 2.8. Suppose that $a, b, c, d, \alpha, \beta, \gamma$ and δ are complex numbers with $(ad - bc)(\alpha\delta - \beta\gamma) \neq 0$, and such that for at least three distinct values of z in $\mathbb{C}$, $cz + d \neq 0, \gamma z + \delta \neq 0$, and

$$\frac{az+b}{cz+d} = \frac{\alpha z + \beta}{\gamma z + \delta}.$$

Then there is some non-zero complex number λ such that

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{2.2}$$

Proof. Consider the quadratic polynomial

$$(az+b)(\gamma z + \delta) = (\alpha z + \beta)(cz + d).$$

The polynomial has three distinct roots, and so it must be a zero polynomial. Therefore, $a\gamma = c\alpha, b\gamma + a\delta = c\beta + d\alpha$ and $b\delta = d\beta$, which is equivalent to

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \mu & 0 \\ 0 & \mu \end{pmatrix},$$

where $\mu^2 = (ad - bc)(\alpha\delta - \beta\gamma) \neq 0$. We then have

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \frac{\mu}{ad - bc} \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

□

The first problem is then resolved by showing that the 4-tuple (a, b, c, d) determines f, up to non-zero multiple. The second problem will be resolved differently, by joining an extra point, which is called *the point at infinity* to $\mathbb{C}$. This point is denoted ∞.

Definition 2.2.2. The set of complex numbers joined with the set $\{\infty\}$ of the point at infinity is called an *extended complex plane*, and is denoted $\mathbb{C}_\infty$.

We already have our notion of the Möbius map approaching infinity, since we have

$$\lim_{z\to\infty} \frac{az+b}{cz+d} = \frac{a}{c}, \quad \lim_{z\to -d/c} \frac{az+b}{cz+d} = \infty$$

when $c \neq 0$. And if $c = 0$ then $\lim_{z\to\infty} f(z) = \infty$. So we naturally use them to assign value to $f(\infty)$.

Definition 2.2.3. For $c \neq 0$, define $f(\infty) = a/c$ and $f(-d/c) = \infty$. If $c = 0$ then $f(\infty) = \infty$.

This assignment of values is well-defined only because we have shown before that either $c \neq 0$ or $c = 0$, and if $c \neq 0$ then the value of a/c and $-d/c$ is always the same for any multiple of c. The main result of this definition is that, all Möbius transformation is now defined on the set $\mathbb{C}_\infty$ so that the composition of any two Möbius is defined. In fact,

Theorem 2.9. Every Möbius map is a bijection from $\mathbb{C}_\infty$ onto itself, and that they form the Möbius group $\mathcal{M}$ with respect to composition.

Theorem 2.10. Every Möbius map transformation can be expressed as the composition of at most four maps, which are

1. rotation and dilation of the form $z \mapsto az$,
2. translation of the form $z \mapsto z + b$; and
3. *complex inversion* of the form $z \mapsto 1/z$.

There is a connection between Möbius maps and 2×2 complex matrices. We have seen that, if M is a non-singular 2×2 matrix with complex entries, then we can find a corresponding Möbius map f. Indeed this mapping, explicitly stated

$$\phi\colon \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto f, \quad f(z) = \frac{az+b}{cz+d},$$

gives us a homomorphism between the group of 2×2 non-singular complex matrices $\mathrm{GL}(2, \mathbb{C})$ and $\mathcal{M}$.

Theorem 2.11. The mapping ϕ is a homomorphism from the group $\mathrm{GL}(2, \mathbb{C})$ onto the Möbius group $\mathcal{M}$.

Lemma 2.2.1. The kernel of ϕ is $\{\lambda I \colon \lambda \in \mathbb{C}\}$. where I is the identity matrix.

2.2.1 Fixed points and uniqueness

2.2.2 Cross-ratios

2.3 Lagrange's theorem

2.3.1 Cosets

We have introduced the idea of subgroup in the previous section. Now we come to the idea of constructing a subset of any group G from its subgroup. For example, we could define a new subset XY of G by

$$XY = \{xy \colon x \in X, y \in Y\}$$

for any subgroup X, Y of G. If X is a singleton, that is $X = \{x\}$, we shall adopt a notation $XY = xY$. Such constructions which we shall consider are of the form

$$gH = \{gh\colon h \in H\} \text{ or } Hg = \{hg\colon h \in H\}$$

for some $g \in G$, and H is a subgroup of G. The set gH is called the *left coset* of H with respect to g, similarly, Hg is the *right coset* of H with respect to g. Some constructions of this type might turn out to be the same set H. This is illustrated below.

Theorem 2.12. Let H be a subgroup of G, and $g \in G$. Then $g \in H$ if and only if $gH = H$ (or $Hg = H$).

Thus we concern ourselves to the study of gH when $g \notin H$. We will adopt an additive notation $g + H$ in place of gH when such subgroups employ addition. The next results show that a group can be divided into disjoint cosets. This is called the *coset decomposition of* G.

Theorem 2.13. Let H be a subgroup of a group G, then G is a union of its left (or right) cosets.

Proof. Clearly, for any $g \in G$, $g \in gH$. So g is contained in the union. □

Theorem 2.14. Let H be a subgroup of a group G, then any two left cosets of G are either equal or disjoint.

Proof. Let $f, g \in G$ and fH, gH are the two left cosets. Suppose that fH and gH are disjoint, that is, the set $fH \cap gH$ is not empty. Then there exists an element $x \in fH \cap gH$, and so $fy_1 = gy_2$ for some $y_1, y_2 \in H$. Thus $g^{-1}f = y_2y_1^{-1} \in H$ and so $g^{-1}fH = H$; hence $gH = gg^{-1}fH = fH$, hereby proving the theorem. □

Corollary 2.3.1. If $fH = gH$, then $g^{-1}f \in H$.

2.3.2 Lagrange's theorem

Recall the definition of an *order* of a group, denoted $|G|$. The next theorem shows the connection between the orders of a group and its subgroup.

Theorem 2.15 (Lagrange's theorem)**.** Let H be a subgroup of a finite group G. Then $|H|$ divides G, and $|G|/|H|$ is the number of distinct left (or right) cosets of H in G.

Proof. From the previous theorem we can write a group G as a union of the pairwise disjoint coset left of H. Therefore $G = g_1H \cup g_2H \cup \cdots \cup g_rH$. Consequently,

$$|G| = |g_1H| + |g_2H| + \cdots + |g_rH|.$$

It remains to show that $|g_1H| = |g_2H| = \cdots = |g_rH| = |H|$. Notice that the map $x \mapsto g_jx$ is a bijection from H to g_jH, and so $|g_1H| = |g_2H| = \cdots = |g_rH| = |H|$. Therefore $|G| = r|H|$ and the results follow. □

The corollaries of Lagrange's theorem are as follows.

Corollary 2.3.2. Let g be an element of a finite group G. Then the order of g divides the order of G.

Proof. Let d be the order of g. The subgroup $H = \{e, g, g^1, \ldots, g^{d-1}\}$ is a subgroup of order d. By Lagrange's Theorem, $|H| \mid |G|$. $\square$

Corollary 2.3.3. If the order of a group is prime, then it is cyclic.

Proof. Let G be a group with prime order p. Suppose $x \in G$, $x \neq e$ and $H = \langle x \rangle$ be its subgroup. Then $|H| \mid |G|$. But $|G|$ is prime, therefore $|H|$ must either be 1 or p. But H contains both x and e, therefore $|H| = p$, that is $H = G = \langle x \rangle$ as claimed. $\square$

2.3.3 Group of small order (up to 8)

Now we use the result from Lagrange's theorem to classify all groups with order less than 8.

2.3.4 Quaternions

2.3.5 Fermat-Euler theorem

2.4 Group actions

This section studies **group actions**.

2.4.1 Group actions

2.4.2 Orbit-stabilizer theorem

2.4.3 Cayley's theorem

2.4.4 Conjugacy classes

2.4.5 Cauchy's theorem

2.5 Quotient groups

2.5.1 Normal subgroups

2.5.2 Quotient groups

2.5.3 The isomorphism theorem

2.6 Matrix groups

2.6.1 The general and special linear groups

2.6.2 The orthogonal and special orthogonal groups

2.6.3 Basis change

2.7 Permutations

2.7.1 Permutations, Cycles and Transpositions

We have given the definition of permutations before. More importantly, we have show that, generally, S_n is not abelian, but some elements of S_n is abelian.

Example 2.7.1. Let $\alpha, \beta \in S_6$, and define

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 1 & 4 & 3 & 6 \end{pmatrix}, \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 3 & 4 & 5 & 2 \end{pmatrix}.$$

Then

$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 1 & 4 & 3 & 2 \end{pmatrix} = \beta\alpha.$$

We shall now provide a sufficient condition for two permutations to commute.

Definition 2.7.1. Any permutations α, β are said to be *disjoint* if, for every k in $\{1, 2, \ldots, n\}$, either $\alpha(k) = k$ or $\beta(k) = k$.

Theorem 2.16. Two permutations commute if they are disjoint.

Proof. Let the two permutations be α and β. For any $k \in \{1, \ldots, n\}$, suppose that α fixes k, the case for β can be argued similarly.

Let $\beta(k) = k'$. Then $\alpha\beta(k) = \alpha(k')$ and $\beta\alpha(k) = \beta(k) = k'$. We shall prove that indeed $\alpha(k') = k$.

If $\beta(k') \neq k'$ then we are done by the premise. So suppose $\beta(k') = k'$, but then $\beta(k') = k' = \beta(k)$. This implies $k = k'$ and so $\alpha(k') = \alpha(k) = k'$ as required. □

The conventional notation for permutations is unwieldy, especially for large n. We shall further simplify it, by introducing fixed points.

Definition 2.7.2. We call that k is a *fixed point* of α, and that α fixes k, if $\alpha(k) = k$.

And so, by convention, we shall left out any integers fixed by α. For example, the permutation

$$\alpha = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

interchanges 1 and 3, and fixes 2. This notation is still too cumbersome for large n, this drives us to find a new notation. Let us start by noticing that, if we repeatedly apply any permutation α, any elements in $\{1, 2, \ldots, n\}$ must eventually return. For example, let

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{pmatrix},$$

then $\alpha^2(1) = 1, \alpha^3(2) = 1, \alpha^3(3) = 3, \alpha^3(4) = 4$ and $\alpha^2(5) = 5$. This is easily proven using the pigeonhole principle. Notice that 1 and 5 form a *cycle* between each other, as α sends 1 to 5 and also send 5 to 1; this is also the case for 2, 3, 4. The permutation α sends 2 to 3, 3 to 4, and 4 to 2. This is the motivation to define *cycles*.

Definition 2.7.3. A *cycle* between $n_1, n_2, \ldots, n_q$ is the permutation

$$\begin{pmatrix} n_1 & n_2 & \cdots & n_q \\ n_2 & n_3 & \cdots & n_1 \end{pmatrix}.$$

It is denoted by $(n_1\, n_2\, \cdots\, n_q)$. The cycle is said to be of length q.

Definition 2.7.4. A *transposition* is a cycle of length 2.

The integers $n_1, n_2, \ldots, n_q$ need not be in an increasing order. By inspection, $\alpha = (1\,5)(2\,3\,4) = (2\,3\,4)(1\,5)$. We will show that any permutation can be written in this manner, as the compositions of cycles.

Theorem 2.17. Any permutation α in the symmetric group S_n can be written as a composition of disjoint cycles.

Proof. This will employ the similar strategy used above. For any integer $k \in \{1, \ldots, n\}$, we apply α repeatedly, and so we have the sequence $k, \alpha(k), \alpha^2(k), \ldots$, and so some elements of this sequence must coincide. Let the two such elements be $\alpha^p(k) = \alpha^q(k)$, with $p < q$. Thus $\alpha^{q-p}(k) = k$. Now there exists a smallest positive number u such that $\alpha^u(k) = k$.

The sequence $k, \alpha(k), \alpha^2(k), \ldots, \alpha^{u-1}(k)$ must be distinct. Construct the cycle

$$\gamma_k = (k \ \alpha(k)\ \alpha^2(k) \ \cdots \ \alpha^{u-1}(k)).$$

Now, two cycles are either disjoint or identical. For if $y = \alpha^d(x)$ for some integer d, then $\gamma_x = \gamma_y$, and we say that x and y belong to the same cycle. Continue doing this for all elements of $\{1, \ldots, n\}$, we will have a collection of cycles $\{\gamma_{k_1}, \gamma_{k_2}, \ldots, \gamma_{k_m}\}$, all of them are pairwise disjoint.

Now consider the composition $\gamma_{k_1}\gamma_{k_2}\cdots\gamma_{k_m}$. For any $x \in \{1, \ldots, n\}$, then $\gamma_{k_d}(x) = \alpha(x)$ if x and k_d belong to the same cycle; else $\gamma_{k_d}(x) = x$. And so $\alpha = \gamma_{k_1}\gamma_{k_2}\cdots\gamma_{k_m}$. □

The proof above use the idea of constructing the sequence $k, \alpha k, \alpha^2(k), \ldots, \alpha^{u-1}(k)$ of elements of a group. This will be studied further in the notion of **orbits**. This decomposition is also unique up to the order of y_{k_i}, and it is called the *standard representation* of α.

Let's try to decompose a permutation using the theorem. Consider

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 4 & 1 & 8 & 2 & 7 & 6 \end{pmatrix}$$

with $\alpha \in S_8$. The cycle formed by 1 is $\gamma_1 = (134)$. Continuing this, we have the collection $\{(134), (2586), (7)\}$, and the standard representation of α is $(1\,3\,4)(2\,5\,8\,6)(7)$. One can drop the single cycle (7) and so

$$\alpha = (1\,3\,4)(2\,5\,8\,6).$$

Finally, consider a cycle α of length n. Note that $\alpha^n = \iota$. Furthermore, for any positive integer d,

$$\alpha^d = (\gamma_{k_1}\gamma_{k_2}\cdots\gamma_{k_m})^d = \gamma_{k_1}^d\gamma_{k_2}^d\cdots\gamma_{k_m}^d,$$

since all cycles commute. It follows that if d is the least common multiple of $n_{k_1}, n_{k_2}, \ldots, n_{k_m}$, where n_{k_i} is the length of γ_{k_i}, then $\alpha^d = \iota$. The least common multiple is indeed the smallest positive integer with such property.

2.7.2 Sign of Permutations

2.7.3 Conjugacy in S_n and A_n

2.7.4 Simple Groups

Chapter 3

Vectors and Matrices

3.1 Complex Numbers

3.1.1 Complex logarithm

3.2 Vectors

3.2.1 Vector Algebra in $\mathbb{R}^3$

This section will review algebra of vectors in $\mathbb{R}^3$. They are usually regarded as an arrow, one with *dimension* and *length*. Starting from two points inside the space $\mathbb{R}^3$, namely $P(p_1, p_2, p_3)$ and $Q(q_1, q_2, q_3)$ we may draw a vector from P to Q, and is expressed by

$$\overrightarrow{PQ} = (q_1 - p_1, q_2 - p_2, q_3 - p_3).$$

Generally let $\mathbf{u} = (u_1, u_2, u_3)$. This vector can also be written as a sum of unit vectors laying on the axis. Those unit vectors are

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \text{ and } \mathbf{k} = (0, 0, 1).$$

And $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$. We can multiply vectors by a scalar, which is a real number, by

$$\mu\mathbf{u} = (\mu u_1, \mu u_2, \mu u_3).$$

The usual properties of vectors should be familiar, that is $\mu(\mathbf{u} + \mathbf{v}) = \mu\mathbf{u} + \mu\mathbf{v}$, $(\mu + \lambda)\mathbf{u} = \mu\mathbf{u} + \lambda\mathbf{u}$, and $(\mu\lambda)\mathbf{u} = \mu(\lambda\mathbf{u})$.

3.2.2 Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$

Let us consider vectors in $\mathbb{R}^n$, the natural generalisation of $\mathbb{R}^3$.

Definition 3.2.1. Using the standard basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$, if $x = \sum_j x_j e_j$ and $y = \sum_j y_j e_j$, we write

$$x \cdot y = \sum_{j=1}^{n} x_j y_j, \|x\|^2 = x \cdot x = \sum_{j=1}^{n} x_j^2,$$

and $x \perp y$ when $x \cdot y = 0$.

Note that $\|x\| = \|-x\|$. The distance $\|x - y\|$ between the points x and y is given by the natural extension of Pythagoras' theorem, and importantly, satisfies the *triangle inequality.*

$$\|x - z\| \leq \|x - y\| + \|y - z\|. \tag{3.1}$$

To prove this assertion, it is sufficient to show that $|x \cdot y| \leq \|x\|\|y\|$, so that we have $\|x + y\| \leq \|x\| + \|y\|$, which readily implies the triangle inequality. Thus we seek to prove

Theorem 3.1 (the Cauchy-Schwarz inequality)**.** For all $x, y \in \mathbb{R}^n$,

$$|x \cdot y| \leq \|x\|\|y\|. \tag{3.2}$$

The equality holds if and only if $\|x\|y = \pm\|y\|x$, i.e. one vector is a multiple of one another.

Proof. Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. The equation holds true when $x = 0$ and when $y = 0$. So we assume that $\|x\|\|y\| > 0$.

Consider the equation

$$0 \leq \sum_{j=1}^{n} \left(\|x\|y_j - \|y\|x_j\right)^2 = 2\|x\|\|y\| \left(\|x\|\|y\| - xy\right),$$

so $x \cdot y \leq \|x\|\|y\|$; similarly, put $-x$ as x and we have $-x \cdot y \leq \|x\|\|y\|$. Therefore $|x \cdot y| \leq \|x\|\|y\|$. Equality holds if $\sum_{j=1}^{n} \left(\|x\|y_j - \|y\|x_j\right)^2$ or $\sum_{j=1}^{n} \left(\|x\|y_j + \|y\|x_j\right)^2$ is equal to zero, which implies $\|x\|y = \pm\|y\|x$. □

Now we are sufficiently equipped with the tool to prove the triangle inequality for general $\mathbb{R}^n$ space.

Theorem 3.2 (The triangle inequality for $\mathbb{R}^n$)**.** For all x, y, z in $\mathbb{R}^n$,

$$\|x - z\| \leq \|x - y\| + \|y - z\|. \tag{3.3}$$

Proof. Set $a = x - y$ and $b = y - z$. The inequality is equivalent to $\|a + b\| \leq \|a\| + \|b\|$, which we seek to prove. Note that

$$\begin{aligned} \|a + b\|^2 = (a + b) \cdot (a + b) &= \|a\|^2 + \|b\|^2 + 2a \cdot b \\ &\leq \|a\|^2 + \|b\|^2 + 2\|a\|\|b\| = \left(\|a\| + \|b\|\right)^2. \end{aligned}$$

Taking square root on both sides we arrive at $\|a + b\| \leq \|a\| + \|b\|$. □

3.2.3 Concepts in linear algebra

3.2.4 Suffix notation

3.2.5 Vector product and triple product

3.2.6 Solution of linear vector equations

3.2.7 Applications

3.3 Matrices

Definition 3.3.1. An $n \times m$ matrix is an array of numbers of the form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

Sometimes it will be denoted by (a_{ij}), where a_{ij} is the general element of the matrix, where i is the *row* and j is the *column* of the element.

Definition 3.3.2. An $n \times n$ matrix is called a *square* matrix.

3.3.1 Algebra of matrices

3.3.2 Determinant and trace

Theorem 3.3. The *trace* of an $n \times n$ matrix $\mathbf{A}$, denoted $\operatorname{tr}\mathbf{A}$ is the sum of its diagonal entries, that is

$$\operatorname{tr}\mathbf{A} = \sum_{i=1}^{n} a_{ii.} = a_{11} + a_{22} + \cdots + a_{nn}.$$

It is evident that, for two square matrices $\mathbf{A}$ and $\mathbf{B}$ with same dimension,

$$\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr}\mathbf{A} + \operatorname{tr}\mathbf{B}.$$

But it is not so obvious that the following holds.

$$\operatorname{tr}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \operatorname{tr}\mathbf{A}$$

for any invertible $n \times n$ matrix $\mathbf{P}$ and any matrix $\mathbf{A}$. Less so of its importance which shall be used later.

3.3.3 Matrix as linear transformation

We start with the definition of linear transformations.

Definition 3.3.3. A map $\alpha\colon V \to W$ between vector spaces V and W is *linear* if, for all scalars $\lambda_1, \ldots, \lambda_n$, and all vectors $v_1, \ldots, v_n$,

$$\alpha\left(\lambda_1 v_1 + \cdots + \lambda_n v_n\right) = \lambda_1 \alpha(v_1) + \cdots + \lambda_n \alpha(v_n).$$

If α is linear we say that it is a *linear transformation*, or a *linear map*, if for all scalars λ and all vectors u and v, $\alpha(\lambda x) = \lambda\alpha(x)$ and $\alpha(x + y) = \alpha(x) + \alpha(y)$.

The two definitions are equivalent.

Theorem 3.4 (Rank-nullity theorem)**.** content...

3.3.4 Simultaneous linear equations

3.4 Eigenvalues and Eigenvectors

Chapter 4

Differential Equations

4.1 Basic Calculus

4.1.1 Differentiation

Definition 4.1.1. The derivative of a function $f(x)$ with respect to x, is the rate of change of $f(x)$ at x, is defined as

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}. \tag{4.1}$$

The function f is differentiable at x if the limit exists. We may write $\frac{df}{dx} = f'(x)$. And more generally, $\frac{d^n}{dx^n} f(x) = f^{(n)}(x)$ is the n-th derivative of f.

We shall adopt the convention that $f'(x)$ is the derivative with respect to the argument, or variable, of the function. For example, $f'(2x)$ is to be view as a derivative of f with respect to $2x$, that is, $f'(2x) = \frac{df}{d(2x)}$.

4.1.2 Big O and small o notation

Definition 4.1.2. We say that $f(x) = o(g(x))$ as $x \to x_0$ if $\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$. That is, $f(x)$ is much *smaller* than $g(x)$.

Definition 4.1.3. We say that $f(x) = O(g(x))$ as $x \to x_0$ if $\frac{f(x)}{g(x)}$ is bounded as $x \to x_0$. That is, $f(x)$ is as big as $g(x)$.

The definition of O does not requires that $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ exists; $\sin x = O(1)$ as $x \to \infty$ but $\lim_{x\to\infty} \sin x$ does not exists.

Theorem 4.1. Let f be a function differentiable at x_0, then

$$f(x_0+h) = f(x_0) + f'(x_0)h + o(h) \tag{4.2}$$

as $h \to 0$.

Proof. From the definition of differentiation and o,

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{o(h)}{h}. \tag{4.3}$$

The result follows. □

4.1.3 Rules of differentiation

Theorem 4.2 (Chain rule)**.** Let $f(x) = F(g(x))$, F is differentiable at $g(x)$ and g is differentiable at x, then

$$\frac{df}{dx} = \frac{dF}{dg}\frac{dg}{dx}.$$

Proof. We have

$$\begin{aligned}
\frac{df}{dx} &= \lim_{h\to 0} \frac{F(g(x+h)) - F(g(x))}{h} \\
&= \lim_{h\to 0} \frac{F(g(x) + hg'(x) + o(h)) - F(g(x))}{h} \\
&= \lim_{h\to 0} \frac{}{den}
\end{aligned}$$

□

4.2 First-order Linear Differential Equations

4.2.1 Equations with constant coefficients

4.2.2 Equations with non-constant coefficients

4.3 Nonlinear first-order equations

4.3.1 Separable equations

4.3.2 Exact equations

4.4 Higher-order Linear Differential Equations

4.5 Multivariate Functions

Chapter 5

Analysis I

A rigorous theory of mathematical analysis must take an axiomatic approach as its foundation. Thus it is preferable to start from the construction of real numbers, and then discover their properties, as not to take them for granted. This foundational rigour is, fortunately, available for us by Dedekind and his model for the real number.

What are the essential properties of $\mathbb{R}$? We have learnt that $\mathbb{R}$ is a field, with the usual addition and multiplication; the usual subtraction and division is also possible.

Secondly, there is a *total order* on $\mathbb{R}$, that is, if $x, y \in \mathbb{R}$ then either $x \leq y$ or $y \leq x$, and only $x = y$ when both condition are satisfied. Furthermore, if $x \leq y$ and $y \leq z$ then $x \leq z$. This means $\mathbb{R}$ is an *ordered field* and that is, if $x \leq y$ then $x + z \leq y + z$, and if $w \geq 0$ then $xw \leq yw$.

Of course, $\mathbb{Q}$ is also an ordered field, but it is not *complete*. This is the most important property of $\mathbb{R}$ to keep in mind. Let's start by a notion of an *upper bound*. If A is a non-empty subset of $\mathbb{R}$ and $b \in \mathbb{R}$, then b is an upper bound for A if $b \geq a$ for all $a \in A$. By saying that $\mathbb{R}$ is complete, this means that, if A is a non-empty set of $\mathbb{R}$ with an upper bound, then A has a *least upper bound*, or *supremum* $\sup A$. This translates to, for any upper bound b of a set $A \subset \mathbb{R}$, should it exist, we have $\sup A \leq b$.

Another central theme of analysis regards *absolute value*, that is the function

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x = 0 \\ -x & \text{if } x \leq 0 \end{cases} . \tag{5.1}$$

Note that $|x - y| = |y - x|$ and $|x| \geq 0$ for all $x \in \mathbb{R}$.

Theorem 5.1. For all $x, y \in \mathbb{R}$, $|x + y| \leq |x| + |y|$, with equality when $xy \geq 0$.

Proof. Trivial proof by case. □

Theorem 5.2. (Triangle Inequality) For all $x, y, z \in \mathbb{R}$, we have

$$|x - z| \leq |x - y| + |y - z|. \tag{5.2}$$

Proof. Simply substitute $x - y$ and $y - z$ in place of x and y, respectively. □

5.1 Limit and Convergences

Let's start with sequences.

5.1.1 Series and sequences in $\mathbb{R}$ and $\mathbb{C}$

Definition 5.1.1. A *sequence* is an ordered list of number, with a natural number n corresponding to the nth term in the sequence.

Definition 5.1.2. A sequence s_n is a *null sequence* if, to every positive number ϵ, there corresponds an integer N such that

$$|s_n| < \epsilon \text{ for all values of } n \text{ greater than } N.$$

We can adapt the definition to any sequence whose terms approach any number s.

Definition 5.1.3. A sequence s_n is said to tend to the limit s if, given any positive number ϵ, there is an integer N (depending on ϵ) such that

$$|s_n - s| < \epsilon \text{ for all } n > N.$$

We then write $\lim s_n = s$.

A more clear notation $\lim_{n\to\infty} s_n = s$ can be given.

Note.

1. Clearly, $\lim s_n = s$ if and only if $s_n - s$ is a null sequence.
2. The inequality $|s_n - s| < \epsilon$ is equivalent to the two inequalities

   $$s - \epsilon < s_n < s + \epsilon.$$

   This is clear that s_n is bounded.
3. A short notation $s_n \to s$ stands for $\lim s_n = s$. A further symbolism for the above definition may be given:

   $$s_n \to s \text{ if } \epsilon > 0; \quad \exists N. |s_n - s| < \epsilon \text{ for all } n > N.$$

   If limits exist, they are unique.

Theorem 5.3. If $a_n \to s$ as $n \to \infty$ and $a_n \to l$ as $n \to \infty$, then $s = l$.

Proof. We will prove this theorem by contradiction. Suppose $s \neq l$. Let $\epsilon = |s - l|/3 > 0$. There exists n_0 such that $|a_n - s| < \epsilon$ for $n \geq n_0$, and there exists m_0 such that $|a_n - l| < \epsilon$ for $n \geq m_0$. Let $N = \max(n_0, m_0)$. Then if $n \geq N$,

$$|l - s| \leq |a_n - l| + |a_n - s| < 2\epsilon = 2|l - s|/3,$$

a contradiction. □

We have discussed on upper bound and lower bound of a set, it is time to introduce a notion of *boundedness*, and expand it to those of sequences in general.

Definition 5.1.4. A subset A of $\mathbb{R}$ is *bounded* if it is bounded above and bounded below. A sequence s_n is bounded if the set $\{s_n : n \in \mathbb{Z}^+\}$ is bounded.

Theorem 5.4. If a sequence tends to a limit, then it is bounded.

Proof. Let the sequence a_n tends to the limit l. We choose an arbitrary ϵ so that for any $n \geq n_0$ the difference $|a_n - l|$ is less than ϵ.

Let $\epsilon = 1$, so that $|a_n - l| < 1$ for all $n \geq n_0$. Choose

$$M = \max\{|a_1|, |a_2|, \ldots, |a_{n_0}|, |l| + 1\}.$$

Then for all $n \geq n_0$ $|a_n| \leq |a_n - l| + |l| < 1 + |l|$. Clearly, $|a_n| \leq M$ and we are set. $\square$

Note that the converse of the theorem might not be true; if a sequence is bounded, then it *might not* tends to a limit. Consider the sequence $a_n = \cos n\pi$. It is bounded, but a_n does not tend to a limit.

Theorem 5.5. Suppose that a_n is an increasing sequence of real numbers. If it is bounded then $a_n \to \sup\{a_n : n \in \mathbb{Z}^+\}$ as $n \to \infty$; otherwise $a_n \to +\infty$.

Similarly, for any decreasing sequence a_n, if it is bounded, then $a_n \to \inf\{a_n : n \in \mathbb{Z}^+\}$; otherwise $a_n \to -\infty$.

One sequence worth considering is the sequence $a_n = r^n$. The convergence of the sequence depends on the value of r.

1. If $r = 1$, then $a_n \to 1$, and if $r = 0$ then $a_n \to 0$.

2. If $r > 1$, then $r = 1 + k$ for some $k > 0$, so we have

$$a_n = (1 + k)^n > 1 + kn$$

   by considering the first two terms in the binomial expansion. And so $a_n \to +\infty$.

3. If $0 < r < 1$, then $r^{-1} = 1 + l > 1$ with $l > 0$, thus

$$0 < a_n = \frac{1}{(1 + l)^n} < \frac{1}{1 + nl}.$$

   As $n \to \infty$, $1/(1 + nl) \to 0$ and therefore $a_n \to 0$.

4. If $-1 < r < 0$, set $s = -r$, so that $0 < s < 1$, it follows that $s^n \to 0$ as $n \to 0$, and therefore $a_n = (-s)^n = 0$.

5. If $r = -1$, then a_n takes the values -1 and 1 alternatively, and so it oscillates finitely.

6. If $r < -1$, set $s = -r$, then $s^n \to \infty$. And so $a_n = (-s)^n$ takes increasing values alternating between negative and positive. That is to say a_n oscillates infinitely.

Another proof of convergence of $a_n = r^n$ when $0 < r < 1$ can be given, as follows: the sequence r^n is decreasing and bounded (by 0), therefore it tends to $\inf\{r^n : n \in \mathbb{Z}^+\}$, which is 0.

5.1.2 Sums, products and quotients

We start with important theorem of sums and products of null sequence.

Theorem 5.6. If s_n and t_n are null sequences, so is $s_n + t_n$.

Theorem 5.7. If s_n is a null sequence and t_n is a bounded sequence, then $s_n t_n$ is a null sequence.

Corollary 5.1.1. If s_n is a null sequence and c is a constant, then cs_n is a null sequence.

We then now extend the results to general sequences.

Theorem 5.8. If $s_n \to s$ and $t_n \to t$, then

1. $s_n + t_n \to s + t$,

2. $s_n t_n \to st$.

Theorem 5.9. If $s_n \to s$ and $t_n \to t$ with $t \neq 0$, then

$$\frac{s_n}{t_n} \to \frac{s}{t}$$

Theorem 5.10. If $s_n \to s$ and $t_n \to t$ and $s_n \leq b_n$ for all n, then $a \leq b$.

Theorem 5.11. If $s_n \to s$ and s_{n_k} is a subsequence, then $s_{n_k} \to s$.

5.1.3 Absolute convergence

5.1.4 Bolzano-Weierstrass theorem

Theorem 5.12. (Bolzano-Weierstrass theorem) Suppose that a_n is a bounded sequence of real numbers. There there is a subsequence a_{n_k} which converges.

5.1.5 Comparison and ratio test

5.1.6 Alternating series test

5.2 Continuity

5.2.1 Continuity of real and complex function

5.2.2 The intermediate value theorem

5.3 Differentiability

5.3.1 Differentiability of functions from $\mathbb{R}$ to $\mathbb{R}$

5.3.2 Derivative of sums and products

5.4 Power series

Definition 5.4.1. An infinite series of the form

$$\sum_{n=0}^{\infty} a_n z^n = a_0 + a_1 z + a_2 z^2 + \cdots$$

composed of multiples of powers of z is called a *power series*. Both the variable z and the coefficients a_n might be real of complex.

There are three possibilities with convergence of a power series.

1. The series converges for all $z \in \mathbb{C}$.

5.5 Integration

5.5.1 Integrability of monotonic functions

Chapter 6

Probability

6.1 Basic concepts

6.2 Axiomatic approach

6.3 Discrete random variables

6.4 Continuous random variables

6.5 Inequalities and limits

6.5.1 Markov's and Chebyshev's inequality

6.5.2 Weak law of large numbers

6.5.3 Convexity and Jensen's inequality

6.5.4 AM-GM inequality

Chapter 7

Vector Calculus

7.1 Curves in $\mathbb{R}^3$

7.2 Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$

7.3 Vector operators

7.4 Integration theorems

7.5 Laplace's equation

7.6 Cartesian tensors in $\mathbb{R}^3$

Chapter 8

Mechanics

8.1 Kinematics of a single particle

8.2 Equilibrium of a single particle

8.3 Equilibrium of a rigid body

8.4 Dynamics of particles

8.5 Energy

8.6 Momentum

8.7 Springs, strings and SHM

8.8 Motion in a circle

Chapter 9

Dynamics and Relativity

9.1 Basic concepts

9.2 Newtonian dynamics of a single particle

9.3 Newtonian dynamics of systems of particles

9.4 Rigid bodies

9.5 Special relativity

Part II

Part IB

Chapter 10

Metric and Topological Spaces

10.1 Metrics

10.1.1 Definition and examples

We extend the notion of continuity, first to $\mathbb{R}^2$. But let's start from the case of continuity in $\mathbb{R}$. What does it mean for f to be continuous at a? One might start form the definition that

$$\forall \epsilon > 0; \quad \exists \delta > 0 . |f(x) - f(a)| < \epsilon \text{ for } |x - a| < \delta.$$

That is, we can make the distance between $f(x)$ and $f(a)$ as small as we want, by choosing x so that the distance $|x - a|$ between x and a is sufficiently small. What about $\mathbb{R}^2$?

Definition 10.1.1. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *continous* at a point $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$, if given $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(a)| < \epsilon$ for every $x = (x_1, x_2 \ldots, x_n)$ satisfying

$$\sqrt{\sum_{i=1}^{n} (x_i - a_i)^2} < \delta.$$

10.1.2 Limits and continuity

10.1.3 Open sets and neighbourhoods

10.1.4 Characterising limits and continuity

10.2 Topology

10.2.1 Definition

10.2.2 Metric topologies

10.2.3 Neighbourhoods

10.2.4 Hausdorff spaces

10.2.5 Homeomorphisms

10.2.6 Topological and non-topological properties

10.2.7 Completeness

10.2.8 Subspace, quotient and product topologies

10.3 Connectedness

10.3.1 Definition

10.3.2 Components

10.3.3 Path-connectedness

10.4 Compactness

Chapter 11

Variational Principles

Chapter 12

Linear Algebra

12.1 Vector Spaces

12.2 Linear maps

12.3 Determinant

12.4 Eigenvalues and Eigenvectors

12.5 Duals

12.6 Bilinear Forms

12.7 Inner Product Spaces

Chapter 13

Groups, Rings and Modules

13.1 Groups

We have gone into details of groups in Part IA.

13.1.1 Basics concepts

13.1.2 Normal subgroups

13.1.3 Sylow subgroups and Sylow theorems

13.2 Rings

13.2.1 Definition

Rings are abstraction of systems with addition and multiplication. The prototype of rings are the set $\mathbb{Z}$ of integers.

We define the general notion of ring in a similar way. We say that a set R with two operations, addition and multiplication, denoted $x + y$ and $x \cdot y$, respectively. We write $x \cdot y$ as xy for comprehensiveness.

Definition 13.2.1. A set R is a ring if the following properties are satisfied:

1. R forms an abelian group under addition.

2. R forms a monoid under multiplication.

3. The distributive laws hold true, i.e.

$$x(y + z) = xy + xz, (y + z)x = yx + zx.$$

13.2.2 Ideals

13.2.3 Fields

13.2.4 Factorisation in rings

13.2.5 Rings $\mathbb{Z}[a]$ of algebraic integers

13.3 Modules

13.3.1 Definition

13.3.2 Submodules

13.3.3 Equivalence of matrices

13.3.4 Finitely generated modules over Euclidean domains

Chapter 14

Analysis II

14.1 Uniform Convergence

14.2 Uniform Continuity and Integration

14.3 $\mathbb{R}^n$ as a Normed Space

14.4 Differentiation from $\mathbb{R}^m$ to $\mathbb{R}^n$

14.5 Metric Spaces

14.6 The Contraction Mapping Theorem

Chapter 15

Complex Analysis

15.1 Analytic Functions

15.2 Contour Integration and Cauchy's Theorem

15.3 Expansions and Singularities

15.4 The Residue Theorem

Chapter 16

Complex Methods

16.1 Analytic Functions

16.2 Contour Integration and Cauchy's Theorem

16.3 Residue Calculus

16.4 Fourier and Laplace Transforms

Chapter 17

Geometry

Part III

Part II

Chapter 18

Number Theory

18.1 Basics

Most fundamentals are covered in part IA.

18.2 Chinese Remainder Theorem

18.3 Law of Quadratic Reciprocity

18.4 Binary Quadratic Forms

18.5 Distribution of the Primes

18.6 Continued fractions and Pell's equation

18.7 Primality testing

18.8 Factorisation

Chapter 19

Topics in Analysis

Chapter 20

Coding and Cryptography

Chapter 21

Automata and Formal Languages

21.1 Register machines

21.2 Regular languages and finite-state automata

21.3 Pushdown automata and context-free languages

Chapter 22

Logic and Set Theory

22.1 Ordinals and Cardinals

22.1.1 Well-orderings and order-types

22.2 Posets and Zorn's Lemma

22.3 Propositional Logic

22.4 Predicate Logic

22.5 Set Theory

22.6 Consistency

Chapter 23

Graph Theory

23.1 Introduction

23.2 Connectivity and matchinhs

23.3 Extremal graph theory

23.4 Eigenvalue methods

23.5 Graph colouring

23.6 Ramsey theory

23.7 Probabilistic methods

Chapter 24

Galois Theory

24.1 Fields extensions

Chapter 25

Representation Theory

25.1 Representations of Finite Groups

25.1.1 Representations on vector spaces

25.2 Character Theory

25.3 Arithmetic Properties of Characters

25.4 Tensor Products

25.5 Representations of S^1 and SU_2

25.6 Further Worked Examples

Chapter 26

Number Fields

26.1 Algebraic Number Fields

26.2 Ideals

26.3 Units

26.4 Ideal classes

26.5 Dedekind's theorem on the factorisation of primes

Chapter 27

Algebraic Topology

27.1 The Fundamental Group

27.2 Covering Spaces

27.3 The Seifert-Van Kampen Theorem

27.4 Simplicial Complexes

27.5 Homology

27.6 Homology Calculations

Chapter 28

Linear Analysis

Chapter 29

Analysis of Functions

Chapter 30

Riemann Surfaces

Chapter 31

Algebraic Geometry

Chapter 32

Differential Geometry

Chapter 33

Probability and Measure

Index