Mathematics

P. Tansuntorn

Last updated July 16, 2020.

Contents

I Part IA 1

1 Numbers and Sets 3

1 Introduction to number systems and logic . . . 3

2 Sets, relations and functions . . . 3

Union, intersection and equality of sets; Indicator functions; Functions; Relations and equivalence relations; The Inclusion-Exclusion Principle

3 The integers . . . 3

Natural numbers

4 Elementary number theory . . . 3

Prime numbers; Euclid's algorithm; Solution in integers of $ax+by=c$.; Modular arithmetic; Chinese remainder theorem; Wilson's theorem

5 The real numbers . . . 5

Least upper bounds; Sequences and series; Irrationality of $\sqrt{2}$ and e; Decimal expansions; Construction of a transcendental number

6 Countability and uncountability . . . 7

2 Groups 9

1 Examples of groups . . . 9

Axioms for groups; Examples from geometry; Permutation on a set; Subgroups and homomorphisms; Symmetry groups

2 The Möbius group . . . 14

Fixed points and uniqueness; Cross-ratios; Preservation of circles; Conjugation; Fixed points of Möbius maps and iteration

3 Lagrange's theorem . . . 17

Cosets; Lagrange's theorem; Group of small order (up to 8); Quaternions; Fermat-Euler theorem

4 Group actions . . . 20

Orbit-stabilizer theorem; Cayley's theorem; Conjugacy classes; Cauchy's theorem

5 Quotient groups . . . 24

Normal subgroups; Quotient groups; The isomorphism theorem

6 Matrix groups . . . 24

The general and special linear groups; The orthogonal and special orthogonal groups; Basis change

7 Permutations . . . 24

Permutations, Cycles and Transpositions; Sign of Permutations; Conjugacy in S_n and A_n; Simple Groups

3 Vectors and Matrices **27**
1 Complex Numbers . . . 27
Complex logarithm
2 Vectors . . . 27
Vector Algebra in $\mathbb{R}^3$; Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$; Concepts in linear algebra; Suffix notation; Vector product and triple product; Solution of linear vector equations; Applications
3 Matrices . . . 30
Algebra of matrices; Determinant and trace; Matrix as linear transformation; Simultaneous linear equations
4 Eigenvalues and Eigenvectors . . . 31

4 Differential Equations **33**
1 Basic Calculus . . . 33
Differentiation; Big O and small o notation; Rules of differentiation
2 1st-order LDEs . . . 35
Equations with constant coefficients; Equations with non-constant coefficients
3 Nonlinear first-order equations . . . 35
Separable equations; Exact equations
4 Higher-order LDEs . . . 35
5 Multivariate Functions . . . 35

5 Analysis I **37**
1 Limit and Convergences . . . 38
Series and sequences in $\mathbb{R}$ and $\mathbb{C}$; Sums, products and quotients; Bolzano-Weierstrass theorem; Absolute convergence; Comparison and ratio test; Alternating series test
2 Continuity . . . 41
Continuity of real and complex function; The intermediate value theorem
3 Differentiability . . . 42
Differentiability of functions from $\mathbb{R}$ to $\mathbb{R}$; Rolle's theorem; Derivative of sums and products; Taylor's theorem
4 Power series . . . 42
Differentiability of power series
5 Integration . . . 43
Integrability of functions; Fundamental theorem of calculus; Improper integrals

6 Probability **45**
1 Basic concepts . . . 45
2 Axiomatic approach . . . 45
3 Discrete random variables . . . 45
4 Continuous random variables . . . 45

5 Inequalities and limits . . . 45
Markov's and Chebyshev's inequality; Weak law of large numbers; Convexity and Jensen's inequality; AM-GM inequality

7 Vector Calculus 47
1 Curves in $\mathbb{R}^3$ . . . 47
Parameterised curves
2 Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$ . . . 47
3 Vector operators . . . 47
Directional derivatives
4 Integration theorems . . . 47
Divergence theorem
5 Laplace's equation . . . 48
Solution of Poisson's equation
6 Cartesian tensors in $\mathbb{R}^3$ . . . 48

II Part IB 49

8 Analysis and Topology 51
1 Uniform convergence and uniform continuity . . . 51
2 Metric spaces . . . 51
3 Topological spaces . . . 51
Metric spaces as topological spaces; Further examples; Concepts in topological spaces
4 Connectedness . . . 53
5 Compactness . . . 53
6 Differentiation from $\mathbb{R}^m$ to $\mathbb{R}^n$ . . . 54

9 Variational Principles 55

10 Linear Algebra 57
1 Vector Spaces . . . 57
Linear independence
2 Linear maps . . . 57
3 Determinant . . . 57
4 Eigenvalues and Eigenvectors . . . 57
5 Duals . . . 57
6 Bilinear Forms . . . 57
7 Inner Product Spaces . . . 57

11 Groups, Rings and Modules 59
1 Groups . . . 59
Basics concepts; Normal subgroups; Sylow subgroups and Sylow theorems
2 Rings . . . 59

Definition; Ideals; Fields; Factorisation in rings; Rings $\mathbb{Z}[a]$ of algebraic integers
3 Modules . . . 61
Definition; Submodules; Equivalence of matrices; Finitely generated modules over Euclidean domains

12 Complex Analysis 63
1 Analytic Functions . . . 63
Complex differentiation; Conformal mappings
2 Contour Integration and Cauchy's theorem . . . 64
Maximum modulus theorem
3 Expansions and singularities . . . 64
4 The residue theorem . . . 64
Open mapping theorem

13 Complex Methods 65
1 Analytic Functions . . . 65
2 Contour Integration and Cauchy's Theorem . . . 65
3 Residue Calculus . . . 65
4 Fourier and Laplace Transforms . . . 65

14 Geometry 67
1 Surfaces . . . 67

15 Methods 69
1 Self-adjoint ODEs . . . 69
2 PDEs on bounded domains: separation of variables . . . 69
3 Inhomogeneous ODEs: Green's functions . . . 69
4 Fourier transforms . . . 69
5 PDEs on unbounded domains . . . 69

16 Statistics 71
1 Estimation . . . 71
2 Hypothesis testing . . . 71
3 Linear models . . . 71

17 Markov chains 73

III Part II 75

18 Number Theory 77
1 Basics . . . 77
2 Chinese Remainder Theorem . . . 77
3 Law of quadratic reciprocity . . . 77
4 Binary quadratic forms . . . 77

5 Distribution of the primes 77
6 Continued fractions and Pell's equation 77
7 Primality testing 77
8 Factorisation 77

19 Topics in Analysis **79**

20 Coding and Cryptography **81**

21 Automata and Formal Languages **83**
1 Register machines 83
2 Regular languages and finite-state automata 83
3 Pushdown automata and context-free languages 83

22 Logic and Set Theory **85**
1 Ordinals and Cardinals 85
Well-orderings and order-types
2 Posets and Zorn's Lemma 85
3 Propositional Logic 85
Deduction and completeness theorem
4 Predicate Logic 85
Löwenheim-Skolem theorems
5 Set Theory 85
6 Consistency 85

23 Graph Theory **87**
1 Introduction 87
2 Connectivity and matchings 87
3 Extremal graph theory 87
4 Eigenvalue methods 87
5 Graph colouring 87
6 Ramsey theory 87
7 Probabilistic methods 87

24 Galois Theory **89**
1 Fields extensions 89
2 Separability 89

25 Representation Theory **91**
1 Representations of Finite Groups 91
Representations on vector spaces
2 Character Theory 91
3 Arithmetic Properties of Characters 91
4 Tensor Products 91
5 Representations of S^1 and SU_2 91

6 Further Worked Examples 91

26 Number Fields 93
1 Algebraic Number Fields 93
2 Ideals 93
3 Units 93
4 Ideal classes 93
5 Dedekind's theorem on the factorisation of primes 93

27 Algebraic Topology 95
1 The Fundamental Group 95
2 Covering Spaces 95
3 The Seifert-Van Kampen Theorem 95
4 Simplicial Complexes 95
5 Homology 95
6 Homology Calculations 95

28 Linear Analysis 97

29 Analysis of Functions 99
1 Lebesgue integration theory 99

30 Riemann Surfaces 101

31 Algebraic Geometry 103

32 Differential Geometry 105

33 Probability and Measure 107

IV Part III 109

34 Topics in Algebra 111

35 Topics in Set Theory 113
1 Model theory of set theory 113
Models of set theory; Transitive models of set theory; Inaccessible cardinals; Absoluteness; Simple independence results; Reflection principles
2 Inner models 113
Definability; Ordinal definability; Constructibility
3 Forcing 113
Generic extensions

36 Category Theory **115**
1 Categories, functors and natural transformations 115
2 Locally small categories 115

Yoneda lemma

3 Adjunctions 115
4 Limits 115
5 Monads 115
6 Filtered colimits 115
7 Regular categories 115
8 Abelian categories 115
9 Monoidal categories 115

37 Model Theory **117**

38 Modular Forms and L-Functions **119**

39 Algebraic Number Theory **121**

40 Elliptic Curves **123**

41 Analytic Number Theory **125**

Part IA

Chapter 1

Numbers and Sets

1 Introduction to number systems and logic

2 Sets, relations and functions

2.1 Union, intersection and equality of sets

2.2 Indicator functions

2.3 Functions

2.4 Relations and equivalence relations

2.5 The Inclusion-Exclusion Principle

3 The integers

3.1 Natural numbers

4 Elementary number theory

4.1 Prime numbers

Definition 1.1. 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 1.2. 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, 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. ☐

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. ☐

4.2 Euclid's algorithm

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

4.4 Modular arithmetic

4.5 Chinese remainder theorem

4.6 Wilson's theorem

5 The real numbers

5.1 Least upper bounds

5.2 Sequences and series

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. ☐

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 $2 = \frac{a^2}{b^2}$. ☐

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 1.3. 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 $0 \leq n \leq b$, $n!$ divides entirely into $b!$, and 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 the value of 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 $b!/n!$. For all term $n \geq b+1$,

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

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. □

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, the result of the Lindemann-Weierstrass theorem indicates that e^a is transcendental if a is rational and non-zero. The same theorem also shows that π is transcendental.

5.4 Decimal expansions

5.5 Construction of a transcendental number

6 Countability and uncountability

Chapter 2

Groups

1 Examples of groups

1.1 Axioms for groups

Definition 2.1. 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. 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 2.2 (Cancellation law)**.** Suppose that a, b and x are in a group G. If $a * x = b * x$ then $a = b$.

Lemma 2.3. 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 2.4. 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 2.2. 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.

1.2 Examples from geometry

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

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

Definition 2.4. The *symmetric group* S_n is a set of permutations of $\{1, \ldots, n\}$. 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. □

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. 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.

1.4 Subgroups and homomorphisms

Definition 2.5. 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 2.2. 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. $\square$

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 2.6. 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 2.7. Let G, H 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 2.8. A homomorphism ϕ is called an *isomorphism* if it is bijective.

Lemma 2.5. 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 2.6. $\phi(xy^{-1}) = \phi(x)\phi(y)^{-1}$.

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

Lemma 2.7. $\phi(x^{-1}) = \phi(x)^{-1}$.

Lemma 2.8. 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 2.9. 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.

1.5 Symmetry groups

2 The Möbius group

We first begin with the definition of Möbius transformations.

Definition 2.10. 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 2.3. 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 2.11. 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 2.12. 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 transformations are now defined on the set $\mathbb{C}_\infty$ so that the composition of any two Möbius maps 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 2.9. The kernel of ϕ is $\{\lambda I \colon \lambda \in \mathbb{C}\}$. where I is the identity matrix.

2.1 Fixed points and uniqueness

2.2 Cross-ratios

2.3 Preservation of circles

2.4 Conjugation

2.5 Fixed points of Möbius maps and iteration

3 Lagrange's theorem

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

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 2.2. 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 $H = \{e, g, g^1, \dots, g^{d-1}\}$ is a subgroup of order d. By Lagrange's Theorem, $|H| \mid |G|$. ☐

Corollary 2.3. 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 $|H| \mid |G|$. Since $|G|$ is prime, $|H|$ 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. ☐

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.

3.4 Quaternions

3.5 Fermat-Euler theorem

In this section we shall show some applications of group theory in the study of arithmetic, notably to prove Fermat's theorem and Euler's theorem using tools from group theory. Recall the definition of $\mathbb{Z}_n$,

$$\mathbb{Z}_n = \{0, 1, \ldots, n-1\},$$

that is, the set of all remainders after any integer is divided by n. We are interested in the group structure of $\mathbb{Z}_n$ with respect to multiplication. A group axiom states that all elements of $\mathbb{Z}_n$ must have an inverse, that is, there exists an element x of $\mathbb{Z}_n$ such that

$$ax \equiv 1 \pmod{n}$$

for all $a \in \mathbb{Z}_n$. But not all integer n satisfied this.

Example 2.4. Consider $2 \in \mathbb{Z}_8$. It does not has an inverse in $\mathbb{Z}_8$. Similary 4 and 6 all do not has an inverse. But $3 \cdot 3 \equiv 1 \pmod{8}$ and $5 \cdot 5 \equiv 1 \pmod{8}$, and so they each have inverses mod 8.

Such observation leads to an important result for general set $\mathbb{Z}_n$.

Theorem 2.16. An element a of a set $\mathbb{Z}_n$ has an inverse if and only if it is coprime to n.

Proof. Suppose that a has an inverse a' in $\mathbb{Z}_n$, then it satisfies the equation

$$aa' \equiv 1 \pmod{n}.$$

Therefore there exists an integer k such that $aa' - 1 = kn$. It follows from Bézout's lemma that $(a, n) = 1$.

Conversely, let $(a, n) = 1$, then there exists an integer k, l such that $ak + nl = 1$. Thus $ak \equiv 1 \pmod{n}$, and so a has an inverse in $\mathbb{Z}_n$. $\square$

Consequently, the set of all integers coprime to n forms a group under multiplication. We shall denote them as $\mathbb{Z}/n\mathbb{Z}$, using the quotient notation. This motivates the following definition.

Definition 2.13. The numbers of positive integers up to n, and are also coprime to n is equal to $\phi(n)$. This is called *Euler's totient function.*

Corollary 2.4. The order of the group $\mathbb{Z}/n\mathbb{Z}$ is equal to $\phi(n)$.

Lemma 2.10. For any prime p, we have $\phi(p) = p - 1$.

Now we can prove Euler's theorem using Lagrange's theorem.

Theorem 2.17 (Euler's theorem)**.** If a and n are coprime then

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

Proof. We have seen that the set $\mathbb{Z}/n\mathbb{Z}$ forms a group and that $|\mathbb{Z}/n\mathbb{Z}| = \phi(n)$. Let d be the order of $a \in \mathbb{Z}/n\mathbb{Z}$. Now the set $\{1, a, a^2, \ldots, a^{d-1}\}$ is a subgroup of $\mathbb{Z}/n\mathbb{Z}$. By Lagrange's theorem, $d \mid \phi(n)$. It then follows that $a^{\phi(n)} = a^{dk} \equiv 1 \pmod{n}$, as claimed. ☐

The Fermat's theorem is a direct result of Euler's theorem.

Corollary 2.5 (Fermat's theorem)**.** For any prime p and any integer a,

$$a^p \equiv a \pmod{p}.$$

4 Group actions

This section studies group actions, which involves a ”product” between a group and an arbitrary set X and returns an element of X. We refer to this by saying that G *acts* on X.

A more precise formulation of the above can be given as following:

Definition 2.14. Let X be any set and G be any group, we say that G acts on X on the left by the product $*$, if for each pair (g, x) with $g \in G, x \in X$, an element $g * x \in X$ is defined, such that for all $g, g' \in G$ and all $x \in X$ the following axioms hold:

1. $e * x = x$, and,
2. $g * (g' * x) = (gg') * x$.

An alternate formulation when G acts on X on the right can also be defined. Our study of group action will rely on this notion and some geometric ideas in the study. Later on we shall apply them to the study of the symmetry groups of regular solids.

Definition 2.15. Suppose that G acts on X. We say that x is a *fixed point* of g in G if $g * x = x$, and the set of fixed points of g is denoted by $\mathcal{F}(g)$.

Definition 2.16. Given x in X, the set $\{g \in G \colon g * x = x\}$ of elements of G that fix x is called the *stabilizer* G_x of x.

This set is indeed a group.

Theorem 2.18. The stabilizer G_x is a subgroup of G.

Definition 2.17. Suppose that G acts on X. Then the subset $\{g * x \colon g \in G\}$ of X is called the *orbit* O_x of x under G. As we could see, it is the set of all images of x varying g to every elements of G.

Definition 2.18. The action of G on X is called *transitive* if for each pair x, y in X there exists a g in G such that $g * x = y$.

Lemma 2.11. The group G acts transitively on X if and only if $O_x = X$ for one x in X.

Note that any two orbits of two elements of X are either disjoint or equal. Thus the orbits partition X into equivalence classes.

4.1 Orbit-stabilizer theorem

The most important result in the section is a variant of Lagrange's theorem.

Theorem 2.19 (Orbit-stabilizer theorem)**.** If a finite group G acts on a set X, then for any $x \in X$ the order of G is given by

$$|G| = |O_x||G_x|,$$

where O_x is the orbit of x under G and G_x is the stabilizer of x.

We present two proofs of the theorem, illustrating the power of cosets in group theory.

1st Proof. In this proof we let G acts on the left of X, the case when G acts on the right can be argued similarly. We find, from Lagrange's theorem, that $|G|/|G_x|$ is the number of distinct left cosets of G_x in G. It remain to show that, indeed, $|O_x|$ is equal to $|G|/|G_x|$, by finding a bijection between the two sets.

Let g be any element of G, and define $\theta(gG_x) = g * x$. Thus θ is surjective over O_x by definition of O_x. Now if $\theta(gG_x) = \theta(hG_x)$, then $g * x = h * x$ which implies $g^{-1}h \in G_x$, that is $g^{-1}hG_x = G_x$, and so $gG_x = hG_x$. Thus θ is injective, and this proves the theorem. ☐

2nd Proof. Take any x in X and let the orbit O_x of x be $\{g_1 * x, \dots, g_r * x\}$, with all of the $g_i * x$ are distinct. Consider the set g_iG_x, the left coset of G_x with respect to g_i. We will show that the set g_iG_x and g_jG_x is disjoint for $i \neq j$, and that G can be decomposed into the union of left cosets of G_x with respect to the elements of O_x.

Suppose g_iG_x and g_jG_x share at least one element, i.e. $g_iG_x \cap g_jG_x \neq \emptyset$, therefore $g_ih_i = g_jh_j$ for some $h_i, h_j \in G_x$. But then

$$\begin{gathered}(g_ih_i) * x = (g_jh_j) * x \\ g_i * (h_i * x) = g_j * (h_j * x) \\ g_i * x = g_j * x,\end{gathered}$$

implying $i = j$. Now let g be any element of G, we need to find g_i such that $g \in g_iG_x$. Let $g * x = y$, then y is in the orbit of x, and so there is one g_i such that $g_i * x = g * x = y$. But then we have $(g_i^{-1}g) * x = x$, i.e. $g_i^{-1}g \in G_x$, which implies $g_i(g_i^{-1}g) = g \in g_iG_x$, as claimed.

Thus we can write G as union of the left cosets of G_x as $G = g_1G_x \cup \cdots \cup g_rG_x$, this gives

$$|G| = |g_1G_x| + \cdots + |g_rG_x|.$$

The map $h \mapsto g_ih$ is obviously a bijection from G_x to g_iG_x, and so $|g_iG_x| = |G_x|$. Finally we get

$$|G| = r|G_x| = |O_x||G_x|,$$

as needed. □

The above proof gives an alluding idea for the following theorem, which shows that the map $g * x = y$ can be written in terms of g and a map that either fixes x or y.

Theorem 2.20. Suppose that G acts on X and that $g * x = y$, where $x, y \in X$ and $g \in G$. Then

$$gG_x = \{h \in G \colon h * x = y\} = G_yg.$$

The following theorem is an important result of orbit-stabilizer theorem in combinatorics.

Theorem 2.21 (Burnside's lemma)**.** Let G be a finite group acting on a finite set X. Then there are N orbits, where

$$N = \frac{1}{|G|}\sum_{g\in G}|\mathcal{F}(g)| = \frac{1}{|G|}\sum_{x\in X}|G_x|.$$

In particular, N is the averge number of fixed points that an element of G has.

4.2 Cayley's theorem

In this section we prove the result by Cayley, which show that, even abstract group is indeed not so abstract.

Theorem 2.22 (Cayley's theorem)**.** Every finite group is isomorphic to a subgroup of a symmetric group

Proof. Let g be any element of a group G, and define the map $\lambda_g \colon G \to G$ by the rule $\lambda_g(a) = ga$, for all $a \in G$. The map is surjective; take any element of h we have $g^{-1}h \in G$ and so $\lambda_g(g^{-1}h) = h$. It is also injective; let $\lambda_g(a) = \lambda_g(b)$, then $ga = gb$ and so $a = b$. This means that λ_n is bijective and thus a permutation of G. But also the set of all λ_g, ranging g to all elements of G, themselves form a group, as the group axiom holds for λ_g as it holds for G.

Define the map $\lambda \colon G \to \mathcal{P}(G)$ from G to the set of permutations of G, by $\lambda(g) = \lambda_g$. It is easy to see that λ is a homomorphism; for $g, h \in G$ we have

$$\lambda_{gh}(a) = (gh)a = g(ha) = \lambda_g(ha) = \lambda_g \lambda_h a,$$

this shows that $\lambda_{gh} = \lambda_g \lambda_h$, as it holds for all $a \in G$. Thus $\lambda \colon G \to \mathcal{P}(G)$ is a homomorphism.

Finally consider the image $\operatorname{Im} \lambda$ of G under λ. It is a subgroup of $\mathcal{P}(G)$, precisely since for $\lambda_g, \lambda_h \in \operatorname{Im} \lambda$, we have $\lambda_g \lambda_h^{-1} = \lambda_g \lambda_{h^{-1}} = \lambda_{gh^{-1}} \in \operatorname{Im} \lambda$. By definition of $\operatorname{Im} \lambda$, λ is an isomorphism from G to a subgroup $\operatorname{Im} \lambda$ of a symmetric group $\mathcal{P}(G)$. □

This theorem holds even for infinite group, with extra care for cardinality of G. Another proof of Cayley's theorem will be given using isomorphism theorems.

Theorem 2.23. Any subgroup H of a group G is the stabilizer of some group action.

4.3 Conjugacy classes

In the previous section, we let G acts on itself in order to prove Cayley's theorem.

4.4 Cauchy's theorem

5 Quotient groups

5.1 Normal subgroups

5.2 Quotient groups

5.3 The isomorphism theorem

6 Matrix groups

6.1 The general and special linear groups

6.2 The orthogonal and special orthogonal groups

6.3 Basis change

7 Permutations

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 are.

Example 2.5. Let $\alpha, \beta \in S_6$, with

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

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 2.19. 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 2.24. 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 $\alpha(k') = k$. Consider the following two possibility of $\beta(k')$.

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 2.20. 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 to any elements in $\{1, 2, \ldots, n\}$, it must eventually reappear after some finite repetitions. 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 2.21. 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 2.22. 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, in the above example we have $\alpha = (1\,5)(2\,3\,4) = (2\,3\,4)(1\,5)$. We will show that any permutation can be written in this manner, as the compositions of cycles.

Theorem 2.25. Any permutation α in the symmetric group S_n can be written as a composition of disjoint cycles.

Proof. The proof employs a similar strategy used above. For any integer $k \in \{1, \ldots, n\}$, we apply α repeatedly, and so we have the sequence $k, \alpha(k), \alpha^2(k), \ldots$; 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 therefore be distinct.

Now we construct the cycle

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

Any two cycles constructed this way are either disjoint or identical, for if $y = \alpha^d(x)$ for some integer d, then $\gamma_x = \gamma_y$, and we see 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, which is indeed the **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.

7.2 Sign of Permutations

7.3 Conjugacy in S_n and A_n

7.4 Simple Groups

Chapter 3

Vectors and Matrices

1 Complex Numbers

1.1 Complex logarithm

2 Vectors

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 $(\mu\lambda)\mathbf{u} = \mu(\lambda\mathbf{u})$.

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 3.1. Using the standard basis $e_1, \dots, e_n$ 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.*

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

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

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

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

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

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

Consider the equation

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

so $x \cdot y \le \|x\|\|y\|$; similarly, put $-x$ as x and we have $-x \cdot y \le \|x\|\|y\|$. Therefore $|x \cdot y| \le \|x\|\|y\|$. 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 $\mathbb{R}^n$,

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

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

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

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

2.3 Concepts in linear algebra

We shall begin with a concept of a basis.

Definition 3.2. The set of vectors $v_1, \dots, v_n$ of vectors in a vector space V is called a *basis* of V if, for every v in V, there exists unique scalars λ_j such that $\sum_{j=1}^n \lambda_j v_j = v$.

Note that the scalars must exists, and are unique for any v. The following theorem follows readily.

Theorem 3.3. For any vector space V, if $v_1, \dots, v_n$ and $w_1, \dots w_m$ both form the bases of V, then $n = m$.

The above theorem allows us to meaningfully assign *dimension* to non-trivial vector spaces V, which is the number of elements in any basis of V, if it is finite. This is called the dimension of V.

Definition 3.3. Let V be a vector space. We say that V is a *finite dimensional* vector space, or $\dim V$ is finite, if the basis of V is finite. Then we take $\dim V$ as the number of elements of the basis of V. If $V = \{0\}$ we put $\dim V = 0$. If the basis of V is infinite, we say that V is infinite dimensional.

The following definition captures two essential properties of a basis.

Definition 3.4. Let $S = \{v_1, \dots, v_k\}$ be a finite set of vectors in V. The set $\operatorname{span}(S)$ is the set of all linear combination of elements of S. If $\operatorname{span}(S) = V$, we say that S *spans* V, or S *generates* V.

Definition 3.5. The set $S = \{v_1, \dots, v_k\}$ is *linearly independent* if, for all scalars λ_j, μ_j,

$$\sum_j \lambda_j v_j = \sum_j \mu_j v_j \text{ implies } \lambda_j = \mu_j \text{ for all } j.$$

Otherwise we say that S is *linearly dependent.*

An equivalence definition of linear independence and be given as follows:

Definition 3.6. The set $S = \{v_1, \dots, v_k\}$ is *linearly independent* if, whenever we have

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0,$$

then $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

Theorem 3.4. The set S forms a basis of V if and only if it spans V and is linearly independent.

2.4 Suffix notation

2.5 Vector product and triple product

2.6 Solution of linear vector equations

2.7 Applications

3 Matrices

Definition 3.7. 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, the index i stands for the *row* and j for the *column* of the element.

Definition 3.8. An $n \times n$ matrix is called a *square* matrix.

3.1 Algebra of matrices

3.2 Determinant and trace

Definition 3.9. The *trace* of an $n \times n$ matrix $\mathbf{A}$, denoted $\operatorname{tr} \mathbf{A}$ is the sum of its diagonal entries, that is

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

Theorem 3.5. It is evident that, for two square matrices $\mathbf{A}$ and $\mathbf{B}$ with same dimension,

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

3.3 Matrix as linear transformation

We start with the definition of linear transformations.

Definition 3.10. 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.

3.4 Simultaneous linear equations

4 Eigenvalues and Eigenvectors

Chapter 4

Differential Equations

1 Basic Calculus

1.1 Differentiation

Definition 4.1. 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, $f'(2x) = \frac{df}{d(2x)}$.

1.2 Big O and small o notation

Definition 4.2. 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 4.3. 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. ☐

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

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

☐

2 First-order Linear Differential Equations

2.1 Equations with constant coefficients

2.2 Equations with non-constant coefficients

3 Nonlinear first-order equations

3.1 Separable equations

3.2 Exact equations

4 Higher-order Linear Differential Equations

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. ☐

1 Limit and Convergences

Let's start with sequences.

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

Definition 5.1. A *sequence* is an ordered list of number, with a natural number n corresponding to the nth term in the sequence.

Alternatively, a sequence is a function from $\mathbb{N}$ to a particular set, namely $\mathbb{R}$ or $\mathbb{C}$.

Definition 5.2. 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 5.3. 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 after some index N.

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 5.4. 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. ☐

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 numerically 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.

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. 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 $s_n \leq t_n$ for all n, then $s \leq t$.

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

1.3 Bolzano-Weierstrass theorem

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

1.4 Absolute convergence

1.5 Comparison and ratio test

1.6 Alternating series test

2 Continuity

2.1 Continuity of real and complex function

2.2 The intermediate value theorem

Theorem 5.13 (Intermediate value theorem)**.** Suppose f is a continuous function on an interval I with $a, b \in I$ satisfying $a < b$ and $f(a) < v < f(b)$. Then there exists $a < c < b$ such that $f(c) = v$

We shall give two proofs.

Proof. This used repeated dissection. Set $a_0 = a$ and $b_0 = b$. Let $d_0 = \frac{(a_0+b_0)}{2}$. If $f(d_0) \geq v$, set $a_1 = a_0$ and $b_1 = d_0$. Otherwise set $a_1 = d_0$, $b_1 = b_0$ if it is the case that $f(d_0) < v$. Now we have $b_j - a_j = (b_0 - a_0)/2^j$ and $f(a_j) \leq v \leq f(b_j)$ for each $j \in \mathbb{N}$. Continuing this, we have two sequences a_j and b_j. The sequence a_j is increasing, b_j is decreasing, and both converge to a same limit c. □

Theorem 5.14 (Boundedness theorem)**.** content...

Theorem 5.15 (Extreme value theorem)**.** content...

3 Differentiability

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

3.2 Rolle's theorem

3.3 Derivative of sums and products

3.4 Taylor's theorem

4 Power series

Definition 5.5. An infinite series of the form

$$\sum_{n=0}^{\infty} a_n (z-z_0)^n = a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \cdots ,$$

composed of multiples of powers of $(z - z_0)$, is called a *power series.* Both the variable z, the centre z_0, and the coefficients a_n might be real of complex.

It suffices to consider only when $z_0 = 0$.

There are three possibilities with convergence of a power series.

1. The series converges for all $z \in \mathbb{C}$.
2. The series converges only within a finite radius of convergence R, that is, only for $|z| < R$.
3. The series converges only for $z = 0$.

4.1 Differentiability of power series

5 Integration

Riemann integral.

5.1 Integrability of functions

monotonic functions are integrable. piecewise continuous functions are integrable.

5.2 Fundamental theorem of calculus

5.3 Improper integrals

Chapter 6

Probability

1 Basic concepts

2 Axiomatic approach

3 Discrete random variables

4 Continuous random variables

5 Inequalities and limits

5.1 Markov's and Chebyshev's inequality

5.2 Weak law of large numbers

5.3 Convexity and Jensen's inequality

5.4 AM-GM inequality

Chapter 7

Vector Calculus

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

1.1 Parameterised curves

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

3 Vector operators

3.1 Directional derivatives

4 Integration theorems

4.1 Divergence theorem

Theorem 7.1 (Divergence theorem)**.** Let $\mathbf{u}$ be a continuously differentiable vector field, defined in a volume V. Let S be the closed surface forming the boundary of V and let $\mathbf{n}$ be the unit outward normal to S. Then

$$\iiint_V \nabla \cdot \mathbf{u}\, dV = \oiint_S \mathbf{u} \cdot \mathbf{n}\, dS.$$

The theorem is also known as Gauss's theorem.

5 Laplace’s equation

5.1 Solution of Poisson’s equation

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

Part IB

Chapter 8

Analysis and Topology

1 Uniform convergence and uniform continuity

We first studies the notions of convergence.

Definition 8.1. Suppose $(f_n)_{n=1}^{\infty}$ is a sequence of real-valued functions, defined on a set S. Then $(f_n)_{n=1}^{\infty}$ *converges point-wise* to f if for each $s \in S$ and each $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $|f_n(s) - f(s)| < \epsilon$ for all $n \geq n_0$.

But point-wise convergence is not adequate. We consider instead uniform convergence, which the number n_0 works for all s.

Definition 8.2. We say that the sequence $(f_n)_{n=1}^{\infty}$ *converges uniformly* to f on S if for each $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $|f_n(s) - f(s)| < \epsilon$ for all $n \geq n_0$ and $s \in S$.

2 Metric spaces

3 Topological spaces

Let's start with the definition of topological spaces. We take the open set definition of topological spaces here.

Definition 8.3. A *topological space* $T = (X, \tau)$ consists of a non-empty set X with a family τ of subset of X such that.

(T1) $X, \emptyset \in \tau$,

(T2) the intersection of any two sets of τ is itself in τ, and

(T3) the union of any collection of sets in τ is in τ.

The family τ is called a *topology* for X, and the members of τ are called the *open set* of T. Elements of X are called *points in the space* T. We also say '$U \in \tau$' to be 'U is open in T'. Later on we shall interchangeably use X and T. For example, the 'topological space X', 'points of X', etc.

Theorem 8.1. A subset U of a topological space X is open in X if and only if for every $x \in U$ there is an open subset U_x of X such that $x \in U_x \subseteq U$.

Proof. If U is open in X, for each $x \in U$, let $U_x = U$. Then the conditions hold.

Conversely, if for every $x \in U$ there is an open subset U_x of X such that $x \in U_x \subseteq U$. We shall show that

$$U = \bigcup_{x \in U} U_x.$$

Suppose $x \in U$, then $x \in U_x \subseteq \bigcup_{x \in U} U_x$. Now consider when $x \in \bigcup_{x \in U} U_x$, then $x \in U_{x_0}$ for some x_0, and we have $x \in U_{x_0} \subseteq U$ from the hypothesis.

Since U is a union of sets open in X, it follows that U is open in X. ☐

3.1 Metric spaces as topological spaces

It is easy to construct a topological space given a metric space, since the definition of a topological space is, in fact, stemmed from metric spaces. Here we show how.

Example 8.1. Given a metric space (X, d), we can construct a topological space (X, τ_d), where τ_d is exactly the family of all d-open subsets of X.

We call such topological space created from a metric space *metrisable*. The space (X, τ_d) *underlies* the metric space (X, d) and τ_d is the topology *induced* by the metric d

Different metrics can give rise to the same topology.

3.2 Further examples

Example 8.2. Consider the family $\tau = \{\emptyset, X\}$, for any non-empty set X. It is trivial to see that τ forms a topology of X, called *indiscrete topology*.

There is also a topology constructed from all possible subset of X, or more precisely:

Example 8.3. The family τ consists of all subset of X is a topology. This is called *discrete topology.*

Direct verification of the axioms should be rather easy.

Definition 8.4. Given two topologies τ_1, τ_2 on the same set, we say that τ_1 is *coarser* than τ_2 if $\tau_1 \subseteq \tau_2$. The topology τ_2 is *finer* than τ_1.

Example 8.4. The *Sierpinski space* $\mathbb{S}$ consists of two points $\{0, 1\}$ with the topology $\{\emptyset, \{1\}, \{0, 1\}\}$. It is finer than the indiscrete topology $\{\emptyset, \{0, 1\}\}$, but courser than the discrete topology $\{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.

Let's see the final example of a topology.

Example 8.5. Let X be a non-empty set. The *co-finite* topology on X consists of the empty set and every subset U of X such that $X \setminus U$ is finite.

Remark. The co-finite topology of a finite set X is the discrete topology.

3.3 Concepts in topological spaces

We introduce the idea of closeness, closure, interior, neighbourhood, etc. borrowed from those of a metric space.

Definition 8.5. Let (X, τ) be a topological space. A subset V of X is said to be *closed* in X if $X \setminus V$ is open in X.

Theorem 8.2. Let X be a topological space. Then

(C1) $X, \emptyset$ are closed in X;

(C2) if V_1, V_2 are closed in X then $V_1 \cup V_2$ is closed in X; and

(C3) if V_i is closed in X for all $i \in I$ then $\cap_{i \in I} V_i$ is closed in X.

This is one of the possible definition of a topological space, in terms of closed sets.

4 Connectedness

5 Compactness

We shall define compactness by first introducing the idea of covers.

Definition 8.6. Suppose X is a set and $A \subseteq X$. A family $\{U_i \colon i \in I\}$ of subsets of X is called a *cover* for A if

$$A \subseteq \bigcup_{i \in I} U_i.$$

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

Chapter 9

Variational Principles

Chapter 10

Linear Algebra

1 Vector Spaces

Theorem 10.1. Let V be a vector space which is spanned by a finite set of vectors $\beta_1, \beta_2, \ldots, \beta_m$. Then any independent set of vectors in V is finite and contains no more than m elements.

1.1 Linear independence

2 Linear maps

3 Determinant

4 Eigenvalues and Eigenvectors

5 Duals

6 Bilinear Forms

7 Inner Product Spaces

Chapter 11

Groups, Rings and Modules

1 Groups

We have gone into details of groups in Part IA.

1.1 Basics concepts

1.2 Normal subgroups

1.3 Sylow subgroups and Sylow theorems

2 Rings

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.

Definition 11.1. A set R, together with two binary operations $+$ and $\cdot$, is a *ring* if it satisfies all of the following properties

(A1) $a, b \in R$ implies $a + b \in \mathbb{R}$,

(A2) $a + b = b + a$ for all $a, b \in R$,

(A3) $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$,

(A4) there exists an element, denoted 0 such that $a + 0 = a$ for all $a \in R$,

(A5) for each $a \in R$, there exists some element $b \in R$ such that $a + b = 0$,

(M1) $a, b \in R$ implies that $a \cdot b \in R$

(M2) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in R$

(D) $a \cdot (b + c) = a \cdot b + a\dot{c}$ and $(b + c) \cdot a = b \cdot a + c \cdot a$ for $a, b, c \in R$

We write $x \cdot y$ as xy for comprehensiveness. There is another formulation of the definition of a ring.

Definition 11.2. 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.

Note that both definition requires multiplication to be associative. Of course, non-associative ring exists, but we shall not confer with them now.

Furthermore, even if our prototype is $\mathbb{Z}$, there are few properties of $\mathbb{Z}$ missing from the definition. For example, we do not impose that there exists an element $1 \in R$ so that $a \cdot 1 = 1 \cdot a = a$ for every $a \in R$. Such ring is called *ring with unit.*

You might see that elements of a ring do not need to commute under multiplication. But for such special occasion, i.e. $a \cdot b = b \cdot a$ for $a, b \in R$, they are called *commutative ring.*

Lastly, it is not true in general that, if $ab = 0$ then $a = 0$ or $b = 0$. The ring with the above property is called a *domain.*

Definition 11.3. A commutative ring R is an *integral domain* if $ab = 0$ implies $a = 0$ or $b = 0$.

Definition 11.4. A ring R is called a division ring if for every $a \neq= 0$ there is an element $b \in R$ such that $ab = ba = 1$. Later we shall denote b by a^{-1}.

Definition 11.5. A ring R is said to be a field if it is a commutative division ring.

Direct verification should show that $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are fields.

2.2 Ideals

2.3 Fields

2.4 Factorisation in rings

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

3 Modules

3.1 Definition

3.2 Submodules

3.3 Equivalence of matrices

3.4 Finitely generated modules over Euclidean domains

Chapter 12

Complex Analysis

1 Analytic Functions

1.1 Complex differentiation

We first start with the basic definition of limit for complex functions.

Definition 12.1. The function f is said to have the limit A as x tends to a,

$$\lim_{x \to a} f(x) = A,$$

if and only if the following is true:

For every $\epsilon > 0$ there exists a real number $\delta > 0$ with the property that $|f(x) - A| < \epsilon$ for all values of X such that $|x - a| < \delta$ and $x \neq a$.

Note that the definition is the same to those of limit of a real function. This is possible since the absolute function admits both real and complex numbers. The well-known results concerning the limit of a sum, a product and a quotient of limits is preserved.

Note that we also have the following properties:

1. $\lim_{x \to a} \overline{f(x)} = \overline{A}$.
2. $\lim_{x \to a} \Re f(x) = \Re A$.
3. $\lim_{x \to a} \Im f(x) = \Im A$.

Definition 12.2. The function f is said to be continuous if $\lim_{x \to a} f(x) = f(a)$.

Definition 12.3. A complex-valued function f defined on an *open* subset G of $\mathbb{C}$ is differentiable at $z \in G$ if

$$\lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

exists. When the limit does exist it is denoted by $f'(z)$.

Let us consider when f is a real function, that is $f(z)$ is real for all value of z.

Theorem 12.1. A real function of a complex variable has the derivative zero, or else does not exists.

Proof. Suppose the derivative $f'(z)$ exists. Then the quotient

$$\frac{f(z+h) - f(z)}{h}$$

is real if h is real; and if $h = ia$ is purely imaginary, then

$$\frac{f(z+ia) - f(z)}{ia}$$

is imaginary. It follows that $f'(z) = 0$ for all z in the domain. Thus a real function of a complex variable must either has the derivative zero, else the derivative does not exist. ☐

1.2 Conformal mappings

2 Contour Integration and Cauchy's theorem

2.1 Maximum modulus theorem

3 Expansions and singularities

4 The residue theorem

Theorem 12.2 (Rouché's theorem)**.** Let f and g be holomorphic inside and on a contour γ and suppose that $|f(z)| > |g(z)|$ on γ^*. Then f and $f + g$ have the same number of zeros inside γ.

4.1 Open mapping theorem

Chapter 13

Complex Methods

1 Analytic Functions

2 Contour Integration and Cauchy's Theorem

3 Residue Calculus

4 Fourier and Laplace Transforms

Chapter 14

Geometry

1 Surfaces

Definition 14.1 (Continuity in metric space)**.** A function $f\colon X \to Y$ is continuous at $P \in X$ if, for any $\varepsilon > 0$, there exists $\delta > 0$ such that $d_X(Q,P) < \delta$ implies $d_Y(f(Q), f(P)) < \varepsilon$.

This can be stated as, the inverse image of $B(f(P), \varepsilon)$ under f contains $B(P, \delta)$. In other notation,

$$B(P,\delta) \subseteq f^{-1}``(B(f(P),\varepsilon)).$$

Lemma 14.1. A map $f\colon X \to Y$ of metric spaces is continuous if and only if under f, the inverse image of every open subset of Y is open in X.

Proof. The necessary condition is equivalent to that there exists an open neighbourhood for all point $P \in f^{-1}``U$, i.e.

$$(\forall P \in f^{-1}``U)(\exists \delta > 0)[B(P,\delta) \subseteq f^{-1}``U]$$

for any open set $U \subseteq Y$. By proving this for all point P, we can conclude that $f^{-1}``U$ is open.

Since $f(P) \in U$ and U is open, there exists a ball $B(f(P),\varepsilon) \subseteq U$ for some ε. And since f is continuous, there exists $\delta > 0$ such that

$$B(P,\delta) \subseteq f^{-1}``(B(f(P),\varepsilon)) \subseteq f^{-1}``U.$$

The last inclusion follows from $B(f(P),\varepsilon) \subseteq U$.

The sufficient condition is straightforward. For each $P \in X$ and $\varepsilon > 0$, consider the image $f^{-1}``(B(f(P),\varepsilon))$ which maps to some open neighbourhood of $P$ by the hypothesis. It follows that there is $\delta > 0$ so that $B(P,\delta) \subseteq f^{-1}``(B(f(P),\varepsilon))$. □

This allows us to translate the idea of continuity to more general settings using open sets.

Chapter 15

Methods

1 Self-adjoint ODEs

2 PDEs on bounded domains: separation of variables

3 Inhomogeneous ODEs: Green's functions

4 Fourier transforms

5 PDEs on unbounded domains

Chapter 16

Statistics

1 Estimation

2 Hypothesis testing

3 Linear models

Chapter 17

Markov chains

Part II

Chapter 18

Number Theory

1 Basics

Most fundamentals are covered in part IA.

2 Chinese Remainder Theorem

3 Law of quadratic reciprocity

4 Binary quadratic forms

5 Distribution of the primes

6 Continued fractions and Pell's equation

7 Primality testing

8 Factorisation

Chapter 19

Topics in Analysis

Chapter 20

Coding and Cryptography

Chapter 21

Automata and Formal Languages

1 Register machines

2 Regular languages and finite-state automata

3 Pushdown automata and context-free languages

Chapter 22

Logic and Set Theory

1 Ordinals and Cardinals

1.1 Well-orderings and order-types

2 Posets and Zorn's Lemma

3 Propositional Logic

3.1 Deduction and completeness theorem

4 Predicate Logic

4.1 Löwenheim-Skolem theorems

5 Set Theory

6 Consistency

Chapter 23

Graph Theory

1 Introduction

2 Connectivity and matchings

3 Extremal graph theory

4 Eigenvalue methods

5 Graph colouring

6 Ramsey theory

7 Probabilistic methods

Chapter 24

Galois Theory

1 Fields extensions

2 Separability

Chapter 25

Representation Theory

1 Representations of Finite Groups

1.1 Representations on vector spaces

2 Character Theory

3 Arithmetic Properties of Characters

4 Tensor Products

5 Representations of S^1 and SU_2

6 Further Worked Examples

Chapter 26

Number Fields

1 Algebraic Number Fields

2 Ideals

3 Units

4 Ideal classes

5 Dedekind's theorem on the factorisation of primes

Chapter 27

Algebraic Topology

1 The Fundamental Group

2 Covering Spaces

3 The Seifert-Van Kampen Theorem

$$\begin{tikzcd}
\pi_1(X, x_1) \arrow[r, "\beta_h"] \arrow[d, "f_*"] & \pi_1(X, x_0) \arrow[d, "f_*"] \\
\pi_1(Y, f(x_1)) \arrow[r, "\beta_{fh}"] & \pi_1(Y, f(x_0))
\end{tikzcd}$$

4 Simplicial Complexes

5 Homology

6 Homology Calculations

Chapter 28

Linear Analysis

Chapter 29

Analysis of Functions

1 Lebesgue integration theory

Chapter 30

Riemann Surfaces

φ

Chapter 31

Algebraic Geometry

Chapter 32

Differential Geometry

Chapter 33

Probability and Measure

Part III

Chapter 34

Topics in Algebra

Examples, tensor products. Ideal theory for commutative Noetherian algebras, localisations. Artinian algebras (commutative and non-commutative), Artin-Wedderburn theorem. Integral dependence. Dimension theory. Filtrations and associated graded algebras. Injective and Projective modules; Ext and Tor. Derivations and differential operators. Hochschild (co-)homology.

Chapter 35

Topics in Set Theory

1 Model theory of set theory

1.1 Models of set theory

1.2 Transitive models of set theory

1.3 Inaccessible cardinals

1.4 Absoluteness

1.5 Simple independence results

1.6 Reflection principles

2 Inner models

2.1 Definability

2.2 Ordinal definability

2.3 Constructibility

3 Forcing

3.1 Generic extensions

Chapter 36

Category Theory

1 Categories, functors and natural transformations

2 Locally small categories

2.1 Yoneda lemma

3 Adjunctions

4 Limits

5 Monads

6 Filtered colimits

7 Regular categories

8 Abelian categories

9 Monoidal categories

Chapter 37

Model Theory

Chapter 38

Modular Forms and L-Functions

Chapter 39

Algebraic Number Theory

Chapter 40

Elliptic Curves

Chapter 41

Analytic Number Theory

Index

Basis, 29
Burnside's lemma, 22

Cayley's theorem, 23
Cover, 54

Dimension, 29

Fixed point, 20, 25

Group, 9

integral domain, 60

Lagrange's theorem, 18
Linearly independent, 29

Orbit, 21

Span, 29
Stabilizer, 20
Symmetric group, 11

Trace, 31