Maths

Adam Boult (www.bou.lt)

July 9, 2020

Contents

Preface

This is a live document, and is full of gaps, mistakes, typos etc.

Part I

Logic

Chapter 1

Propositional logic

1.1 Introduction

1.1.1 True and False

We start off with two statements:

- True - T or $\top$
- False - F or $\bot$

1.1.2 Propositional variables

We can represent T or F using a symbol:

θ

1.2 Operators

1.2.1 Unary operators

A unary operator takes one input and returns another.

Only negation, $\neg$ is of interest.

The following statements are equivalent:

- T
- $\neg F$

1.2.2 Binary operators

A binary operator takes an additional input.

- If then - $\theta \to \gamma$
- Then if - $\theta \leftarrow \gamma$
- Iff - $\theta \leftrightarrow \gamma$
- And / Conjunction - $\theta \wedge \gamma$
- Or / Disjunction - $\theta \vee \gamma$

1.2.3 Truth tables

1.2.4 Brackets

Operators can be shown together, with brackets. For example:

$(\alpha \vee \beta) \wedge \gamma$

Is not the same as:

$\alpha \vee (\beta \wedge \gamma)$

1.2.5 Atomic formulae

Atomic formulae are those without operators taking more than one input.

Literals, and negative literals, are types of atomic formula.

A literal is a formula with no operators.

θ

These are also known as positive literals.

Negative literals are the negation of a literal.

$\neg\theta$

1.2.6 Well-formed formulae

A well-formed formula is one which can be given a truth value.

The following is not a well-formed formula:

$\theta\wedge$

1.2.7 Interpretations

An interpretation assigns meaning to propositional variables in a formula.

For example an interpretation of the formula $\theta \lor \gamma$ assigns values to each of θ and γ.

1.2.8 Satifisable

A formula is satisfisable if there is some interpretation where it is true.

For example θ is satisfisable but $\theta \land \neg\theta$ is not.

1.2.9 Tautology

A formula is a tautology if it is true in all interpretations.

Examples of tautologies include:

- $\theta \lor \neg\theta$

1.3 Multi-valued logic

1.3.1 Multi-valued logic

We can have logic with more than two states.

1.4 Semantic consequence

1.4.1 Semantic consequence

A formula, A, semantically implies another, B, if for every interpretation of A, B is true.

We show this with:

$A \vDash B$

Formula B is satisfisable if there is some A where this is true.

For example: $A \land B \vDash A$

Formula B is a tautology if this is true for any A. We can also write this as $\vDash B$.

1.4.2 Logical equivalence

If $A \vDash B$ and $B \vDash A$ we say that A and B are logically equivalent.

This is shown as $A \Leftrightarrow B$.

1.4.3 How many unique operators are there?

An arbitrary operator takes n inputs are returns T or F.

With 0 inputs there is one posible permutation. For every additional input the number of possible permutations doubles. Therefore there are 2^n possible permutations.

For the operator with one permutation there are two operators. For every additional permutation the number of operator doubles. Therefore there are $2^{(2^n)}$ possible operations.

With 0 inputs, we need 2 different operators to cover all outputs. For 1 input we need 4 and for 2 inputs we need 16.

1.4.4 We don't need 0-ary operators

There are two unique 0-ary operators. One always returns T and the other always returns F. These are already described.

1.4.5 We need one unary operator

For the operators with 1 input we have:

- one which always returns T
- one which always returns F
- one which always returns the same as the input
- one which returns the opposite of the input

It is this last one, negation, shown as $\neg$ and is of most interest.

1.4.6 We can use a subset of binary operators

The full list of binary operators are included below.

Of these, the first two are 0-ary operators, and so are not needed. The next four are unary operators, and so are not needed.

The non-implications can be rewritten using negation.

1.4.7 Brackets replace the need for n-ary operators

N-ary operators contain 3 or more inputs.

N-ary operators can be defined in terms of binary operators.

As an example if we want an operator to return positive if all inputs are true, we can use:

$(\theta \land \gamma) \land \beta$

1.4.8 De Morgan's Laws

- $\neg(A \lor B) \Leftrightarrow (\neg A \land \neg B)$
- $\neg(A \land B) \Leftrightarrow (\neg A \lor \neg B)$

This expresses the duality of normal form.

Duality is the principle that binary operators have inverses, and when they are swapped with their inverse, the truth value of the statement is unaffected.

1.4.9 Normal form

This is where a formula is shown using only:

- And / Conjunction- $\land$
- Or / Disjunction - $\lor$
- Negation - $\neg$

The conjunctive normal form (CNF) is where a formula is converted to a normal form with the following layout:

$a \land b \land c \land d$

These letters can represent complex sub-formulae, in normal form.

Statements in this form are easier to evaluate, as each subformula can be evaluated separately. The statement is true only if all formulaes within are also true.

The disjunctive normal form (DNF) is similar for $\lor$.

$a \lor b \lor c \lor d$

1.4.10 Properties of the normal form

The normal binary operators are commutitive - $A \wedge B \Leftrightarrow B \wedge A$ and $A \vee B \Leftrightarrow B \vee A$

Both binary operators are associative - $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$ and $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$

Negation is complementary.

$A \wedge \neg A \Leftrightarrow F$

$A \vee \neg A \Leftrightarrow T$

Normal binary operators are absorbative.

$A \wedge (A \vee B) \Leftrightarrow A$ $A \vee (A \wedge B) \Leftrightarrow A$

Identity.

$A \wedge T \Leftrightarrow A$

$A \vee F \Leftrightarrow A$

Distributivity.

$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$

$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$

Chapter 2

Inference in propositional logic

2.1 Inference

2.1.1 Substitution

If we have a tautology, then we can substitute the formula of any propositional variable with any formula to arrive at any other tautology.

For example, we know that $\theta \vee \neg\theta$ is a tautology. This means that an arbitrary formula for θ is also a tautology.

An example is $(\gamma \wedge \alpha) \vee \neg(\gamma \wedge \alpha)$, which we know is a tautology, without having to examine each variable.

2.1.2 Syntactic consequence

Let us call the first formula A and the second B. We can then say:

$A \vdash B$

This says that: if A is true, then we can deduce that B is true using steps such as substitution.

2.1.3 Modus Ponens

Modus Ponens is a deduction rule. This allows us to use stpes other than substitution to derive new tautologies.

If A implies B, and A is true, then B is also true.

$(\theta \rightarrow \gamma) \wedge \theta \Rightarrow \gamma$

That is, if we can show that the following are true:

$\theta \rightarrow \gamma$

θ

We can infer that the following is also true:

γ

2.1.4 Theory

Results derived from substitution or induction are called theorems. Theorems often divided into:

- Theorems - important results
- Lemmas - results used for later theorems
- Corollaries - readily deduced from a theorem

We take a set of axioms, as true, and a deduction rule which enables us to derive additional formulae, or theorems. The collection of axioms and theorems is known as the theory.

2.1.5 Principle of explosion

If axioms contradict each other then it is possible to derive anything. That is:

$P \wedge \neg P \vdash Q$

We can prove this. If P and $\neg P$ are true, then the following is also true:

$P \vee Q$

We can then use $P \vee Q$ and $\neg P$ to imply Q.

This works for any proposition Q, including $\neg Q$.

As we can derive Q and $\neg Q$, our axioms are not consistent.

2.1.6 Resolution rule

Proof by resolution

If we have a string of or statements, $A \vee B \vee C$, and another which contains the completement of one element $X \vee \neg B \vee Y$, we can infer:

$A \vee C \vee X \vee Y$

If the second statement has only one formula, then we have:

$A \vee B \vee C$ and $\neg B$ implying $A \vee C$

2.1.7 Clauses and horn clauses

A clause is a disjunction of atomic formulae.

$A \vee \neg B \vee C$

This can be written in implicative form.

$(A \vee \neg B) \vee C$

$\neg(A \vee \neg B) \rightarrow C$

$(\neg A \wedge B) \rightarrow C$

A horn clause is a clause where there is at most one positive literal. This means the implicative takes the form.

$(A \wedge B \wedge C) \rightarrow X$.

2.1.8 Inference with horn clauses

If the horn clause is true, and so is the normal form part, then X is also true.

As all inference with horn clauses uses Modus Ponens, it is sound.

Inference with horn clauses is also complete.

Chapter 3

Axioms for propositional logic

3.1 Axioms for propositional logic

3.1.1 Motivation for axioms for propositional logic

We discussed in the previous section the ability to derive new tautologies from others using substitution and Modus Ponens.

We now aim to identify a group of axioms from which all tautologies can be derived.

3.1.2 The axioms

The first is known as "Simplification". In words, this is "if it is cloudy, then if it is a Tuesday it is also cloudy."

$\theta \to (\gamma \to \theta)$

The second is called "Frege".

$(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$

The third is "Transposition". Consider the statement "If there are no clouds in the sky, it is not raining." If this is true then it is also true that "If it is raining there are clouds in the sky."

$(\neg\theta \to \neg\gamma) \to (\gamma \to \theta)$

3.1.3 Independence of axioms

These axioms are independent. That is, if you take one away, you cannot derive it from the others.

These axioms are also effective. One could define all true formulae as axioms, however this is not effective.

3.1.4 Soundness of axioms

Soundness implies that all theories are true.

$T \vdash A \Rightarrow T \vDash A$

These axioms and the deduction rule are sound. We know that the axioms are tautologies, and we know that the inference rule is valid.

As the axioms are sound, the theories are consistent. That is, it is not possible for both θ and $\neg\theta$ to be theories.

3.1.5 Completeness of axioms

Completeness implies that all true formulae are theories.

$T \vDash A \Rightarrow T \vdash A$

3.1.6 Axioms and definitions

A definition is a conservative extension of the language. A definition statement, for example that a new symbol Z is always evaluated as false allows us to make additional statements, but it does not allow us to make additional statements in the original language.

An axiom allows us to generate additional statements in the original language, a definition does not.

Chapter 4

First-order logic

4.1 Zero-order logic

4.1.1 Terms and predicates

Predicates

Zero-order logic adds predicates. Like propositional variables, these have truth values. Unlike propositional variable, predicates take terms as inputs.

For example using propositional logic we can write the statement "you are 25" as θ.

With preterites we can write this as $P(you, 25)$.

A propositional variable can be considered a special case of a predicate variable, where the number of inputs is 0.

4.1.2 Relations and equality

Relations

A special type of predicates is a relation. These take two terms and can be written differently: $P(x, y) \Leftrightarrow x \oplus y$

Equality

In preterite logic we define the relation for equality.

$a = b$

It is defined by the following:

- Reflexivity : $x = x$
- Symmetry: $x = y \leftrightarrow y = x$
- Transivity: $x = y \wedge y = z \rightarrow x = z$
- Substitution for functions: $x = y \rightarrow f(x) = f(y)$
- Substitution for formulae: $x = y \wedge P(x) \rightarrow P(y)$

4.1.3 Functions and brackets

Functions (or maps)

Functions take other terms, and are themselves terms. For example if we wanted to know if someone can legally drive in a specific country, we could use:

$P(you, age(UK))$

A function may not be able to produce an output for all inputs. For examples $age(green)$ has no interpretation.

Functions can also take different numbers of inputs. Constants, such as you and UK can be shown as functions with 0 inputs. As a result we could instead write:

$P(you(), age(UK()))$

We generally denote functions with a lower case letter, so would instead write:

$P(y(), a(b()))$

Functions are also called maps.

4.1.4 Signatures

Structures

A logical structure consists of:

- Domain
- Interpretation
- Signature

Domain

The domain is the set of variables in the structure.

We include an infinite number of variables.

Interpretation

The interpretation assigns values to propositional and predicate variables.

Signature

A logical signature describes the language of the logic which is used to construct statements. This includes:

- Functions
- Relations
- Operators

The language of a signature is all possible sentences, or formulae which can be constructed from this signature.

We include an infinite number of functions, relations and all operators.

4.1.5 Completeness of zero-order logic

A theory is complete if all true formulae are included.

Note that there are three types of formulae in a theory.

- Tautologies (always true)
- Refutable formulae (always false)
- Satisfiable formulae which are not tautologies (true in some, but not all, interpretations).

4.1.6 Injective, bijective and surjective functions

Injective functions

$f(a) = f(b) \rightarrow a = b$

Surjective functions

All points in codomain have at least one matching point in domain

Mapping info, details

Bijective

Both injective and surjective

Other

Identity function

The identity function maps a term to itself.

Idempotent

An idempotent function is a function which does not change the term if the function is used more than once. An example is multiplying by 0.

Inverse functions

An inverse function of a function is one which maps back onto the original value.

$g(x)$ is an inverse function of $f(x)$ if

$g(f(x)) = x$

Properties of binary functions

Binary functions can be written as:

$f(a, b) = a \oplus b$

A function is commutative if:

$x \oplus y = y \oplus x$

A function is associative if:

$(x \oplus y) \oplus z = x \oplus (y \oplus z)$

A function $\otimes$ is left distributive over $\oplus$ if:

$x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$

Alternatively, function $\otimes$ is right distributive over $\oplus$ if:

$(x \oplus y) \otimes z = (x \otimes z) \oplus (y \oplus z)$

A function is distributive over another function if it both left and right distributive over it.

4.1.7 Binary functions

Properties of binary functions

Binary functions can be written as:

$f(a, b) = a \oplus b$

A function is commutative if:

$x \oplus y = y \oplus x$

A function is associative if:

$(x \oplus y) \oplus z = x \oplus (y \oplus z)$

A function $\otimes$ is left distributive over $\oplus$ if:

$x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$

Alternatively, function $\otimes$ is right distributive over $\oplus$ if: $(x \oplus y) \otimes z = (x \otimes z) \oplus (y \oplus z)$

A function is distributive over another function if it both left and right distributive over it.

4.2 First-order logic

4.2.1 Writing first-order logic

Existential quantifier

We introduce a shorthand for at least one term satisfies a predicate, that is:

$P(x_0) \lor P(x_1) \lor P(x_2) \lor P(x_2) \lor P(x_3)...$

The short hand is:

$\exists x P(x)$

niversal quantifier

We introduce another shorthand, this time for:

$P(x_0) \land P(x_1) \land P(x_2) \land P(x_2) \land P(x_3)...$

The shorthand is

$\forall x P(x)$

Free and bound variables

A bound variable is one which is quantified in the formula. A free variable is one which is not. Consider:

$\forall x P(x, y)$

In this, x is bound while y is free.

Free variables can be interpreted differently, while bound variables cannot.

We can also bind a specific variable to a value. For example 0 can be defined to be bound.

Ground terms

A ground term does not contain any free variables. A ground formula is one which only includes ground terms.

$\forall x$ x is a ground term.

$\forall x P(x)$ is a ground formula.

4.2.2 Inference rules for first-order logic

Existential instantiation

If P is true for a specific input, then there exists an input for P where P is true.

$P(r) \Rightarrow \exists x P(x)$

Existential generalisation

$\exists x P(x) \Rightarrow P(r)$

Where r is a new symbol.

Universal instantiation

If P is true for all values of x, then P is true for any input to P.

$\forall x P(x) \Rightarrow P(a/x)$

Where a/x represents substituting a for x within P.

Universal generalisation

If there is a derivation for $P(x)$, then there is a derivation for $\forall x P(x)$.

$\vdash P(x) \Rightarrow \vdash \forall x P(x)$

4.2.3 Duality of first-order logic

The dual of:

$\exists x \neg P(x)$

Is:

$\neg \forall x P(x)$

Chapter 5

Gdels completeness theorem and the compactness theorem

5.0.1 Introduction

5.0.2 Gödels completeness theorem

Completeness of first-order logic

We previously showed that zero-order logic was complete. What about first-order logic?

Gödels completeness theorem says that for first order logic, a theory can include all tautologies, the first category.

If the completeness theorem is true and a formula is not in the theory, then the formula is either refutable or satisfiable under some, but not all interpretations.

That is, either the theory will contain θ, $\neg\theta$, or θ will be satisfiable in some but not all interpretations, and neither will be in the in theory.

To prove this we look for a proof that every formula is either refutable or true under some structure. So for an arbitrary formula θ we want to show it is either refutable or satisfiable under some interpretation.

Part 1: Converting the form of the formula

Remove free variables, functions

Note that if this is true, all valid formulae of the form below are provable:

$\neg\theta$

This means that there is no interpretation where the following is true:

θ

Conversely if $\neg\theta$ is not in the theory, then θ must be true under some interpretation.

That is, if all valid formulae are provable, then all

Reformulating the question:

This is the most basic form of the completeness theorem. We immediately restate it in a form more convenient for our purposes:

Theorem 2. Every formula θ is either refutable or satisfiable in some structure.

"$\theta$is refutable" means by definition "$\neg\theta$ is provable".

Decidability

Given a formula, can we find out if can be derived from the axioms? We can follow a process for doing so which would inform us if the formula was or was not a theorem. Alternative, the process could carry on forever.

If the process never carries on forever the system is decidable: there is a finite process to determine whether the formula is in or out. If the process halts for true formulas, but can carry on forever for false formulas, the system is semi-decidable. If the process takes a long time, we do not know if it is looping infinitely, or approaching its halt point.

Intuitively, use of axioms can make an existing formula shorter or longer, so finding all short formulas can require going forwards and backwards an infinite number of times.

Part II

Arithmetic of natural numbers

Chapter 6

Robinson arithmetic

Chapter 7

Peano arithmetic

7.1 Addition

7.1.1 Definition

Lets add another function: addition. Defined by:

$\forall a \in \mathbb{N}(a + 0 = a)$

$\forall ab \in \mathbb{N}(a + s(b) = s(a + b))$

That is, adding zero to a number doesnt change it, and $(a+b)+1 = a+(b+1)$.

7.1.2 Example

Lets use this to solve $1 + 2$:

$1 + 2 = 1 + s(1)$

$1 + s(1) = s(1 + 1)$

$s(1 + 1) = s(1 + s(0))$

$s(1 + s(0)) = s(s(1 + 0))$

$s(s(1 + 0)) = s(s(1))$

$s(s(1)) = s(2)$

$s(2) = 3$

$1 + 2 = 3$

All addition can be done iteratively like this.

7.1.3 Commutative property of addition

Addition is commutative:

$x + y = y + x$

7.1.4 Associative property of addition

Addition is associative:

$x + (y + z) = (x + y) + z$

7.2 Multiplication

7.2.1 Multiplication of natural numbers

7.2.2 Definition

Multiplication can be defined by:

$\forall a \in \mathbb{N}(a.0 = 0)$

$\forall ab \in \mathbb{N}(a.s(b) = a.b + a)$

7.2.3 Example

Lets calculate 2.2.

$2.2 = 2.s(1)$

$2.s(1) = 2.1 + 2$

$2.1 + 2 = 2.s(0) + 2$

$2.s(0) + 2 = 2.0 + 2 + 2$

$2.0 + 2 + 2 = 2 + 2$

$2 + 2 = 4$

7.2.4 Commutative property of multiplication

Multiplication is commutative:

$xy = yx$

7.2.5 Associative property of multiplication

Multiplication is associative:

$x(yz) = (xy)z$

7.2.6 Distributive property of multiplication

Multiplication is distributive over addition:

$a(b + c) = ab + ac$

$(a + b)c = ac + bc$

Chapter 8

Orderings

8.1 Ordering

8.1.1 Inequalities

Less than or equal

Orderings define relations between elements in a set, where one element can preceed the other.

Orderings are antisymmetric. That is, the only case where the relation is satisfied in both directions is if the elements are equal.

$(a \leq b) \land (b \leq a) \rightarrow (a = b)$

Orderings are transitive. That is:

$(a \leq b) \land (b \leq c) \rightarrow (a \leq c)$

Greater than or equal

Less than and greater than

The relation $\leq$ is refered to as non-strict.

There is a similar strict relation relation, $<$:

$(a \leq b) \land \neg(b \leq a) \rightarrow (a < b)$

Chapter 9

Subtraction and division

9.1 Integers

9.1.1 Subtraction of natural numbers

We have inverse functions for addition. This is subtraction.

For function $\oplus$, its inverse is $\oplus'$, as defined below:

$a \oplus b = c$

$b = c \oplus' a$

$f(a, b) = c \to f^{-1}(c, b) = a$

Subtraction

$a + b = c \to b = c - a$

There is no natural number b that satisfies:

$3 + b = 2$

While addition and multiplication are defined across all natural numbers, subtraction is not.

Properties of subtraction

Subtraction is not commutative:

$x - y \neq y - x$

Subtraction is not associative:

$x - (y - z) \neq (x - y) - z$

9.1.2 Division

Introduction

We have inverse functions for multiplication. This is division.

These will not necessarily have solutions for natural numbers or integers.

Division of natural numbers

$a.b = c \rightarrow b = \frac{c}{a}$

Division is not commutative

Division is not commutative:
$\frac{x}{y} \neq \frac{y}{x}$

Division is not associative

$\frac{\frac{x}{y}}{z} \neq \frac{\frac{x}{y}}{z}$

Division is not left distributive

Division is not left distributive over subtraction:
$\frac{a}{b-c} \neq \frac{a}{b} - \frac{a}{c}$

Division is right distributive

Division is right distributive over subtraction:

$\frac{a-b}{c} = \frac{a}{b} - \frac{b}{c}$

Division of integers

Chapter 10

Divisors and prime numbers

10.1 Prime numbers

10.1.1 Prime numbers and composite numbers

Definition

A prime number is a number which does not have any divisors other than 1 and itself.

By convention we do not refer to 0 or 1 as prime numbers.

Identifying prime numbers

Divisors must be smaller than the number. As a result it is easy to identify early prime numbers, as we can try to divide by all preceding numbers.

Examples of prime numbers

$[2, 35, 7, 11, 13, ...]$

Composite numbers

Composite numbers are numbers that are made up through the multiplication of other numbers.

They are not prime.

10.1.2 Relatively prime numbers

10.1.3 Euler's totient function

This functions counts numbers up to n which are relatively prime

eg for 10 we have 1, 3, 7, 9.

So $\phi(10) = 4$

10.1.4 Euler's theorem

10.1.5 Fermat's little theorem

10.1.6 Pseudoprimes

10.2 Other

10.2.1 Frobenius number

Given a set of nautral numbers, the Frobenius number is the biggest number which can't be made as linear combination of the set.

Chapter 11

Modulus and remainders

11.1 Modulus and remainders

11.1.1 Remainders

Division is defined between natural numbers. However there are many cases where this division does not map to a natural number. For example:

$\frac{7}{3}$

We can divide 6 of the 7 by 3, giving 2 with 1 remaining.

Alternatively we can divide 3 of the 7 by 3, giving 1 with 4 remaining

Or we could divide 0 of the 7 by 3 giving 0 with 7 remaining.

The remainder refers to the lowest possible number - in this case 1.

11.1.2 Residue systems

Least residue system modulo n

This is the set of numbers from 0 to $n-1$.

Complete residue system

This a set of numbers none of which are congruent $\mod n$. That is, for no pair $\{a,b\}$ does $a \mod (n) = bmod(n)$

Reduced residue system

This is a complete residue system where all numbers are relatively prime to n.

11.1.3 Congruence

5 and 11 are congrument $\mod 3$

If $a \mod (n) = bmod(n)$ then a and b are congruent mod n.

Chapter 12

GCD and LCM

12.1 Divisors and multiples

12.1.1 Divisors and Greatest Common Divisors (GCD)

Divisors

The divisors d of a natural number n are the natural numbers such that $\frac{n}{d} \in \mathbb{N}$.

For example, for 6 the divisors are $1, 2, 3, 6$.

Divisors cannot be bigger than the number they are dividing.

Universal divisors

For any number $n \in \mathbb{N}^{+}$:

$\frac{n}{n} = 1$

$\frac{n}{1} = n$

Both 1 and n are divisors.

Common divisors

A common divisor is a number which is a divisor to two supplied numbers.

Greatest common divisor

The greatest common divisor of 2 numbers is as the name suggests.

So $GCD(18, 24) = 6$

12.1.2 Multiples and Lowest Common Multiples (LCM)

Multiples

The multiple of a number is it added to itself iteratively.

The multiples of 18 for example are:

$[18, 36, 54, 72, 90, ...]$

And for 24:

$[24, 48, 72, 96, 120, ...]$

Common multiples

Lowest common multiple

The lowest common multiple of 2 numbers is again as the name suggests.

So $LCM(18, 24) = 72$.

12.1.3 Coprimes

Also known as relatively prime.

Greatest common divisor is 1.

Chapter 13

The fundamental theorem of arithmetic

13.1 The Fundamental Theorem of Arithmetic

13.1.1 Euclidian division

Euclidian division is the theory for any pair of natural numbers, we can divide one by the other and have a remainder less than the divisor. Formally: $\forall a \in \mathbb{N}, \forall b \in \mathbb{N}^+, \exists q \in \mathbb{N}, \exists r \in \mathbb{N}[(a = bq + r) \land (0 \leq r < b)]$

Where $\mathbb{N}^+$ refers to natural numbers excluding 0.

That is, every natural number a is a multiple q of any other natural number b, plus another natural number r less than the other natural number b.

These are unique. For each jump in q, r falls by b. As the range of r is b there is only one solution.

$17 = 2.8 + 1$

$9 = 3.3 + 0$

13.1.2 Bezout's identity

For any two non-zero natural numbers a and b we can select natural numbers x and y such that

$ax + by = c$

The value of c is always a multiple of the greatest common denominator of a and b.

In addition, there exist x and y such that c is the greatest common denominator itself. This is the smallest positve value of c..

Let's take two numbers of the form $ax + by$:

$d = as + bt$

$n = ax + by$

Where $n > d$. And d is the smallest non-zero natural number form.

We know from Euclidian division above that for any numbers i and j there is the form $i = jq + r$.

So there are values for q and r for $n = dq + r$.

If r is always zero that means that all values of $ax + by$ are multiples of the smallest value.

$n = dq + r$ so $r = n - dq$.

$r = ax + by - (as + bt)q$

$r = a(x - sq) + b(y - tq)$

This is also of the form $ax + by$. Recall that r is the remainder for the division of d and n, and that $d = ax + by$ is the smallest positive value.

r cannot be above or equal to d due to the rules of euclidian division and so it must be 0.

As a result we know that all solutions to $ax + by$ are multiples of the smallest value.

As every possible $ax + by$ is a multiple of d, d must be a common divisor to both numbers. This is because $a.0 + b.1$ and $a.1 + b.0$ are also solutions, and d is their divisor.

So we know that the smallest positive solution is a common mutliple of both numbers.

We now need to show that that d is the largest common denominator. Consider a common denominator c.

$a = pc$

$b = qc$

And as before:

$d = ax + by$

So:

$d = pcx + qcy$

$d = c(px + qy)$

So $d \geq c$

13.1.3 Euclid's lemma

Statement

If a prime number p divides product $a.b$ then p must divide at least of one of a or b.

Proof

From Bezout's identity we know that:

$d = px + by$

Where p and b are natural numbers and d is their greatest common denominator.

Let's choose a prime number for p. There are no common divisors, other than one. As a result there are exist values for x and y such that:

$1 = px + by$

Now, we are trying to prove that if p divides $a.b$ then p must divide at least one of a and b, so let's multiply this by a.

$a = pax + aby$

We know that p divides pax, and p divides ab by definition. As a result p can divide a.

13.1.4 Fundamental Theorem of Arithmetic

Statement

Each natural number is a prime or unique product of primes.

Proof: existance of each number as a product of primes

If n is prime, no more is needed.

If n is not prime, then $n = ab$, $a, b \in \mathbb{N}$.

If a and b are prime, this is complete. Otherwise we can iterate to find:

$n = \prod_{i=1} p_i$

Proof: this product of primes is unique

Consider two different series of primes for the same number:

$s = \prod_{i=1}^{n} p_i = \prod_{i=1}^{m} q_i$

We need to show that $n = m$ and $p = q$.

We know that p_i divides s. We also know that through Euclid's lemma that if a prime number divides a non-prime number, then it must also divide one of its components. As a result p_i must divide one of q.

But as all of q are prime then p_i=q_j.

We can repeat this process to to show that $p = q$ and therefore $n = m$.

13.1.5 Existence of an infinite number of prime numbers

Existence of an infinite number of prime numbers

If there are a finite number of primes, we can call the set of primes P.

We identify a new natural number a by taking the product of existing primes and adding 1.

$a = 1 + \prod_{p \in P} p$

From the fundamental theorem of arithmetic we know all numbers are primes or the products of primes.

If a is not a prime then it can be divided by one of the existing primes to form number n:

$\frac{\prod^{n} p_i + 1}{p_j} = n$

$\frac{p_j \prod_{i \neq j}^{n} p_i + 1}{p_j} = n$

$\prod_{i \neq j}^{n} p_i + \frac{1}{p_j} = n$

As this is not a whole number, n must prime.

We can do this process for any finite number of primes, so there are an infinite number.

Chapter 14

Finite sequences of natural numbers

14.1 Sequences

14.1.1 Definition

A sequence is an ordered list of terms.

These are commonly maps from natural numbers to real (or complex) numbers.

We can use $a_i = f(i)$ to denote this.

If $f(i)$ is defined on all $i \in \mathbb{N}$ then the sequence is infinite. Otherwise it is finite.

If a sequence is defined on $n \in \mathbb{N}$ and $n \neq 0$ then the sequence must be defined on $n - 1$.

For example $a_0, a_1, a_2, ...$ is a sequence, but $a_1, a_2, ...$ is not.

14.1.2 Monotone sequence

A monotone sequence is one where each element is succeeded ordinally by the next entry.

For example:

$< 1, 2, 3, 6, 7 >$ is monotone

$< 1, 2, 3, 3, 4 >$ is not monotone

An increasing sequence is one where:

$\forall m \in \mathbb{N} \forall n \in \mathbb{N}[m > n \leftrightarrow a_m \geq a_n]$

A strictly increasing sequence is one where:

$\forall m \in \mathbb{N} \forall n \in \mathbb{N}[m > n \leftrightarrow a_m > a_n]$

A decreasing sequence is one where:

$\forall m \in \mathbb{N} \forall n \in \mathbb{N}[m > n \leftrightarrow a_m \leq a_n]$

A strictly decreasing sequence is one where:

$\forall m \in \mathbb{N} \forall n \in \mathbb{N}[m > n \leftrightarrow a_m < a_n]$

All strictly decreasing sequences are decreasing, and all strictly increasing sequences are increasing.

A monotone sequence is one which is either increasing or decreasing.

14.1.3 Subsequences

A subsequence of a sequence is the original sequence with some elements of the original removed, not changing the order.

For example:

$< 1, 3, 5 >$ is a subsequence of $< 2, 1, 3, 4, 7, 5 >$

14.2 Series

14.2.1 Definition

A series is the summation of a sequence. For a series a_n there is a corresponding series:

$s_n = \sum_{i=0}^{n} a_n$

Where:

$\sum_{i=0}^{n} a_i = a_0 + a_1 + a_2 + ... + a_n$

14.2.2 Multiplication of summations

If all members of a sequence are multiplied by a constant, so is each member of the series.

We can take constants out of the series:

$s_n = \sum_{i=0}^{n} a_i$

$s_n = \sum_{i=0}^{n} cb_i$

$s_n = a\sum_{i=0}^{n} b_i$

14.2.3 Summation of constants

If all elements of a sequence are the same, then the series is a multiple of that constant.

$s_n = \sum_{i=0}^{n} a_i$

$s_n = \sum_{i=0}^{n} c$

$s_n = nc$

14.2.4 Addition of summations

Consider a sequence $a_i = b_i + c_i$.

$s_n = \sum_{i=0}^{n} a_i$

$s_n = \sum_{i=0}^{n}(b_i + c_i)$

We can then split this out.

$s_n = \sum_{i=0}^{n} b_i + \sum_{i=0}^{n} c_i$

14.2.5 Summation from a different start point

$\sum_{i=0}^{n} a_i = a_0 + \sum_{i=1}^{n} a_i$

14.2.6 Multiple summations

$\sum_{i=0}^{n} \sum_{j=0}^{m} a_i = n \sum_{j=0}^{m} a_i$

$\sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j = \sum_{i=0}^{n} a_i \sum_{j=0}^{m} b_i$

14.3 Products

14.3.1 Definition

A product is a repeated multiplication of a sequence.

$p_n = \prod_{i=0}^{n} s_i$

14.3.2 Multiplication of products

We can take constants out of the product.

$p_n = \prod_{i=0}^{n} ca_i$

$p_n = a^n \sum_{i=0}^{n} a_i$

14.3.3 Products of constants

If $a_i = c$ then the summation is then of the form:

$p_n = \prod_{i=0}^{n} c$

$p_n = c^n \prod_{i=j}^{n} 1$

$p_n = c^n$

14.3.4 Combining products

If a sequence is the product of to other sequences then the product of the sequence is equal to the product of the two individual sequences.

$p_n = \prod_{i=0}^{n} a_i$

$p_n = \prod_{i=0}^{n} b_i c_i$

$p_n = \prod_{i=0}^{n} b_i \prod_{i=0}^{n} c_i$

14.3.5 Factorials

A factorial is a a product across natural numbers. That is:

$n! := \prod_{i=0}^{n} i$

14.4 Summation of natural numbers

14.4.1 Goal

Let's prove that:

$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$

14.4.2 Proof by induction

We use the inference rules Modus Ponens, which says that if X is true, and $X \rightarrow Y$ is true, then Y is true.

14.4.3 True for $n = 0$

We know this is true for $n = 0$:

$0 = \frac{0(0+1)}{2}$

$0 = 0$

14.4.4 If it's true for n, it's true for $n + 1$

We can also prove that if it true for n, it is true for $n + 1$.

$\sum_{i=0}^{n+1} i = \frac{(n+1)(n+2)}{2}$

$(n+1) + \sum_{i=0}^{n} i = \frac{n^2+3n+2}{2}$

If it is true for n, then:

$(n+1) + \frac{n(n+1)}{2} = \frac{n^2+3n+2}{2}$

$\frac{n^2+3n+2}{2} = \frac{n^2+3n+2}{2}$

$1 = 1$

14.4.5 Result

So we know that it is true for $n = 0$, and if it is true for n, then it is true for $n + 1$. As a result it is true for all natural numbers.

14.5 Bounded sequences

14.5.1 Bounded sequences

A function $f(x)$ on set X is bounded if:

$\exists M \in \mathbb{R}[\forall x \in X f(x) \leq M]$

A bounded sequence is a special case of a bounded function where:

$X = \mathbb{N}$

That is, a sequence is bounded by M iff:

$\forall n \in \mathbb{R} |f(a_n)| \leq M$

Chapter 15

Powers, exponents and logarithms of natural numbers

15.1 Powers

15.1.1 Recap

Previously we defined addition and multiplication in terms of successive use of the sucessor function. That is, the definition of addition was:

$\forall a \in \mathbb{N}(a + 0 = a)$

$\forall ab \in \mathbb{N}(a + s(b) = s(a + b))$

And similarly for multiplication:

$\forall a \in \mathbb{N}(a.0 = 0)$

$\forall ab \in \mathbb{N}(a.s(b) = a.b + a)$

Additional functions could also be defined, following the same pattern:

$\forall a \in \mathbb{N}(a \oplus_n 0 = c)$

$\forall ab \in \mathbb{N}(a \oplus_n s(b) = (a \oplus_n b) \oplus_{n-1} a)$

15.1.2 Powers

Powers can also be defined:

$\forall a \in \mathbb{N} a^0 = 1$

$\forall ab \in \mathbb{N} a^{s(b)} = a^b.a$

15.1.3 Example

So 2^2 can be calculated like:

$2^2 = 2^{s(1)}$

$2^{s(1)} = 2.2^1$

$2.2^1 = 2.2.2^0$

$2.2.2^0 = 2.2.1$

$2.2.1 = 4$

Unlike addition and multiplication, exponention is not commutative. That is

$a^b \neq b^a$

15.1.4 Exponential rules

$a^b a^c = a^{b+c}$

$(a^b)^c = a^{bc}$

$(ab)^c = a^c b^c$

15.1.5 Binomial expansion

How can we expand

$(a+b)^n, n \in \mathbb{N}$

We know that:

$(a+b)^n = (a+b)(a+b)^{n-1}$

$(a+b)^n = a(a+b)^{n-1} + b(a+b)^{n-1}$

Each time this is done, the terms split, and each terms is multiplied by either a or b. That means at the end there are n total multiplications.

This can be shown as:

$(a+b)^n = \sum_{i=1}^{n} a^i b^{n-i} c_i$

So we want to identify c_i.

Each term can be shown as a series of n as and bs. For example:

- $aaba$
- $baaa$

For any of these, there are $n!$ ways or arranging the sequence, but this includes duplicates. If we were given n unique terms to multiply there would indeed by $n!$ different ways this could have arisen, but we can swap as and bs, as they were only generated once. So let's count duplicates.

There are duplicates in the as. If there are i as, then there are $i!$ ways of rearranging this. Similarly, if there are $n-i$ bs, then there are $(n-i)!$ ways or arranging this.

As a result the number of actual observed instances, c_i, is:

$c_i = \frac{n!}{i!(n-i)!}$

And so:

$(a+b)^n = \sum_{i=0}^n a^i b^{n-i} \frac{n!}{i!(n-i)!}$

We can also write this last term as:

$\binom{n}{i}$

15.1.6 Difference of two squares

$(a+b)(a-b) = a^2 - ab + ab - b^2$

$(a+b)(a-b) = a^2 - b^2$

15.2 Logarithms

15.2.1 Definition

If:

$c = a^b$

Then

$log_a c = b$

Product rule:

$a = c^{log_c a}$

$b = c^{log_c b}$

So:

$ab = c^{log_c ab}$

But also:

$ab = c^{log_c a} c^{log_c b}$

$ab = c^{log_c a + log_c b}$

So:

$log_c a + log_c b = log_c ab$

15.2.2 Power rule

$a = b^{log_b a}$

So:

$a^c = b^{log_b a^c}$

And separately:

$a^c = (b^{log_b a})^c$

$a^c = (b^{c log_b a})$

So:

$c log_b a = log_b a^c$

Chapter 16

Gdel numbering

16.1 Introduction

16.1.1 Gdel numbering

Gdel numbering assigns a unique number to each formula.

To contruct this we first assign a natural number to each symbol.

This gives us a sequence:

$\{x_1, x_2, x_3, ..., x_n\}$

We can assign a unique number to this by using the first n prime numbers.

$2^{x_1}3^{x_2}5^{x_3}...$

This number can then be prime factored to recover the sequence, and therefore the formula.

Chapter 17

The Gdel incompleteness theorems

17.0.1 Introduction

Part III

General Set Theory (GST)

Chapter 18

Axiom schema of specification

18.1 Defining sets

18.1.1 Axiom schema of specification

The axiom schema of unrestricted comprehension

We want to formalise the relationship between the preterite and the set. An obvious way of doing this is to add an axiom for each preterite in our structure that:

$\forall x \exists s[P(x) \leftrightarrow (x \in s)]$

This is known as "unrestricted comprehension" and there are problems with this approach.

Consider a predicate for all terms which are not members of themselves. That is:

$\neg(x \in x)$

This implies the following is true:

$\forall x \exists s[\neg(x \in x) \leftrightarrow (x \in s)]$

As this is true for all x, it is true for $x = s$. So:

$\exists s[\neg(s \in s) \leftrightarrow (s \in s)]$

This statement is false. As we have inferred a false formula, the axiom of unrestricted comprehension does not work. This result is known as Russel's

Paradox.

This is an axiom schema rather than an axiom. That is, there is a new axiom for each preterite.

Axiom schema of specification

To resolve Russel's paradox, we amend the axiom schema to:

$\forall x \forall a \exists s[(P(x) \land x \in a) \leftrightarrow (x \in s)]$

That is, for every set a, we can define a subset s for each predicate.

This resolves Russel's Paradox. Let's take the same steps on the above formula as in unrestricted comprehension;

$\forall x \forall a \exists s[(\neg(x \in x) \land x \in a) \leftrightarrow (x \in s)]$

$\exists s[(\neg(s \in s) \land s \in s) \leftrightarrow (s \in s)]$

So long as the subsets s are not members of themselves, this holds.

18.1.2 Implications of axiom schema of specification

All finite subsets exist

Finite subsets. Don't know about infinite subsets

If we can define a subset, by the axiom of specification it exists.

For example if set $\{a, b, c\}$ exists, we can define a preterite to select any subset of this.

For example we can use define a $P(x)$ as $x = a \lor x = b$ to extract the subset $\{a, b\}$.

If a subset is infinitely large,

Intersections of finite sets exist

Can prove exists from this axiom

If any set exists, the empty set exists

$\forall x \forall a \exists s[(P(x) \land x \in a) \leftrightarrow (x \in s)]$

18.1.3 Set-builder notation

Notation

We can use short-hand to describe sets.

$\{x \in S : P(x)\}$

This defines a set by a restriction. For example we will later be able to define natural numbers above 5 as:

$\{x \in \mathbb{N} : x > 5\}$

Class builder notation

Emuneration can be done through set builder notation too

Can define sets formally! defintion doesn't just affect sets

$\forall x(x \in C \leftrightarrow P(x))$

NB: We're not saying C exists

Can then use examples of equiv class

$\forall x(x \in C \leftrightarrow x = x)$

18.1.4 Empty set

We can use this to define the empty set - the set with no members.

$\varnothing = \{\}$

Using the above definition this is the same as writing:

$\forall x \neg(x \in \varnothing)$

18.1.5 Defining sets by enumeration

We can describe a set by the elements it contains.

$s = \{a, b, c\}$

This is a shorthand way of writing:

$\forall x(x \in s \leftrightarrow (x = a \lor x = b \lor x = c))$

18.1.6 Finite and infinite sets

Finite sets

A set is finite is there is a proper subset without a bijection.

Proper subset: $A \subset B$

For example for set $\{a, b, c\}$ There is no subset with a bijection.

Infinite sets

For the natural numbers, all natural numbers except 0 is a proper subset, and there is a bijection.

18.1.7 Cardinality

Cardinality of finite sets

The cardinality of a set s is shown as $|s|$. It is the number of elements in the set. We define it formally below.

Injectives, surjectives and bijectives

Consider 2 sets. If there is an injective from a to b then for every element in a there is a unique element in b.

If this injective exists then we say $|a| \leq |b|$.

Similarly, if there is a surjective, that is for every element in b there is a unique element in a, then $|a| \geq |b|$.

Therefore, if there is a bijection, $|a| = |b|$, and if there is only an injective or a surjective then $|a| < |b|$ or $|a| > |b|$ respectively.

Cardinality as a function

Every set has a cardinality. As a result cardinality cannot be a well-defined function, for the same reason there is no set of all sets.

Cardinality functions can be defined on subsets.

18.2 Introduction to sets

18.2.1 Membership relation

Say we have a preterite $P(x)$ which is true for some values of x. Sets allow us to explore the properties of these values.

We may want to talk about a collection of terms for which $P(x)$ is true, which we call a set.

To do this we need to introduce new axioms, however first we can add (conservative) definitions to help us do this.

We introduce a new relation: membership. If element x is in set s then the following relation is true, otherwise it is false:

$x \in s$

Sets are also terms. In first-order logic they will be included in quantifiers. Indeed, in set theory, we aim to treat everything as sets.

If a term is not a member of another term, we can write this using the non-member relation as follows:

$\forall x \forall s[\neg(x \in s) \leftrightarrow x \notin s]$

Chapter 19

Set algebra

19.1 Set algebra

19.1.1 Set union and intersection

We discuss functions. Just because we can write a function of sets which exist, does not mean the results of the functions exist. For that we need axioms discussed later.

Union function

We define a function on two sets, $a \vee b$, such that the result contains all elements from either sets.

$\forall a \forall x \forall y [a \in (x \vee y) \leftrightarrow (a \in x \vee a \in y)]$

This is commutative: $a \vee b = b \vee a$

This is associative: $(a \vee b) \vee c = a \vee (b \vee c)$

Intersection function

We define a function, $a \wedge b$, on two sets, such that the result contains all elements which are in both.

$\forall a \forall x \forall y [a \in (x \wedge y) \leftrightarrow (a \in x \wedge a \in y)]$

This is commutative: $a \wedge b = b \wedge a$

This is associative: $(a \wedge b) \wedge c = a \vee (b \wedge c)$

Distribution of union and intersection

Union is distributive over intersection: $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$

Intersection is distributive over union: $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$

19.1.2 Complements and disjoint sets

Disjoint sets

Sets are disjoint is there is no overlap in their elements. Two sets are s_i and s_j are mutually exclusive if:

$s_i \wedge s_j = \emptyset$

A collection of events s are all mutually exclusive if all pairs are mutually exclusive. That is:

$\forall s_i \in s \forall s_j \in s[s_i \wedge s_j \neq \emptyset \rightarrow s_i = s_j]$

Complement function

x^C is the completement. It is defined such that:

$\forall x[x \wedge x^C = \varnothing]$

For a set b, the complement with respect to a is all elements in a which are not in b.

$\forall x \in a \forall y \in b[x \in (a \setminus b) \wedge y \in (a \setminus b)]$

$b \wedge (a \setminus b) = \varnothing$

That is, b and $a \setminus b$ are disjoint.

Existence of the complement

For two sets a and b we can write $(a \setminus b)$. This is the set of elements of a which are not in b.

Consider the axiom of specification:

$\forall x \forall a \exists s[(P(x) \wedge x \in a) \leftrightarrow (x \in s)]$

We can also write

$\forall x \forall a \forall b \exists s[(x \notin b \wedge x \in a) \leftrightarrow (x \in s)]$

Which provides the complement, s.

19.1.3 Boolean algebra

Boolean algebra in propositional logic

We previously discussed properties of normal form, and the results from these properties.

If another structure shares these properties then they will also share the results.

Sets satisfy the definitions of a boolean algebra

If a mathematical structure has the following properties, it shares the results from normal form, and is a boolean algebra.

- Both binary operators are commutitive - $A \wedge B = B \wedge A$ and $A \vee B = B \vee A$
- Both binary operators are associative - $(A \wedge B) \wedge C = A \wedge (B \wedge C)$ and $(A \vee B) \vee C = A \vee (B \vee C)$
- Completements - $A \wedge \neg A = \emptyset$ and $A \vee \neg A = U$
- Absorption - $A \wedge (A \vee B) = A$ and $A \vee (A \wedge B) = A$
- Identity - $A \wedge U = A$ and $A \vee \emptyset = A$
- Distributivity - $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$ and $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$

These hold for sets, and so boolean algebra holds for sets.

19.1.4 Algebra on a set

Standard algebra

An algebra, Σ, on set s is a set of subsets of s such that:

- Closed under intersection: If a and b are in Σ then $a \wedge b$ must also be in Σ
- $\forall ab[(a \in \Sigma \wedge b \in \Sigma) \rightarrow (a \wedge b \in \Sigma)]$
- Closed under union: If a and b are in Σ then $a \vee b$ must also be in Σ.
- $\forall ab[(a \in \Sigma \wedge b \in \Sigma) \rightarrow (a \vee b \in \Sigma)]$

If both of these are true, then the following is also true:

- Closed under complement: If a is in Σ then $s \backslash a$ must also be in Σ

We also require that the null set (and therefore the original set, null's complement) is part of the algebra.

Chapter 20

The axiom of extensionality

20.1 Introduction

20.1.1 Axiom of extensionality

If two sets contain the same elements, they are equal.

$\forall x \forall y [\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y]$

This is an axiom, not a definition, because equality was defined as part of first-order logic.

Note that this is not bidirectional. $x = y$ does not imply that x and y contain the same elements. This is appropriate as $\frac{1}{2} = \frac{2}{4}$ for example, even though they are written differently as sets.

Reflexivity of equality

The reflexive property is:

$\forall x (x = x)$

We can replace the instance of y with x:

$\forall x [\forall z (z \in x \leftrightarrow z \in x) \rightarrow x = x]$

We can show that the following is true:

$\forall z (z \in x \leftrightarrow z \in x)$

Therefore:

$\forall x [T \rightarrow x = x]$

$x = x$

Symmetry of equality

The symmetry property is:

$\forall x \forall y[(x = y) \leftrightarrow (y = x)]$

We know that the following are true:

$\forall x \forall y[\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y]$

$\forall x \forall y[\forall z(z \in x \leftrightarrow z \in y) \rightarrow y = x]$

So:

$\forall x \forall y[\forall z(z \in x \leftrightarrow z \in y) \rightarrow (x = y \wedge y = x)]$

Transitivity of equality

The transitive property is:

$\forall x \forall y \forall z[(x = y \wedge y = z) \rightarrow x = z]$

Substitution for functions

The substitutive property for functions is:

$\forall x \forall y[(x = y) \rightarrow (f(x) = f(y))]$

Substitution for formulae

The substitutive property for formulae is:

$\forall x \forall y[((x = y) \wedge P(x)) \rightarrow P(y)]$

Doesnt this require iterating over predicates? Is this possible in first order logic??

Result 1: The empty set is unique

We can now show the empty set is unique.

Result 2: Every element of a set exists

If an element did not exist, the set containing it would be equal to a set which did not contain that element.

Result 3: Sets are unique

20.1.2 Equivalence classes

We have already ready defined the relationship equality, between terms.

$a = b$.

Sometimes we may wish to talk about a collection of terms which are all equal to each other. This is an equivalence class.

Though we have not yet defined them, integers are example of this. For example -1 can be written as $0 - 1$, $1 - 2$ and so on.

$\forall y \forall x x = y \rightarrow x \in z$

For all sets, we can call the class of all sets equal to the set an equivalence class.

This does not necessarily exist.

Chapter 21

Axiom of adjunction

21.0.1 Introduction

Chapter 22

Algebra of cardinality

22.1 Other

22.1.1 Cardinality

Cardinality of cartesian product

What about the cardinality of Cartesian products? So if we have sets:

$\{1, 2, 3\}$

$\{a, b\}$

We can have the Cartesian product set:

$\{(1, a), (2, a), (3, a), (1, b), (2, b), (3, b)\}$

We can see that:

$|A.B| = |A|.|B|$

Cardinality of union and intersection

$|A \vee B| = |A| + |B| - |A \wedge B|$

Cardinality of powerset

$|P(s)| = 2^{|s|}$

Cardinality of complement

$|a \setminus b| = |a| - |a \wedge b|$

Cardinality of even/odd natural numbers

What about the cardinality of even numbers? Well, we can define a bijective function between each:

$f(n) = 2n$

Similarly for odd numbers:

$f(n) = 2n + 1$

So these both have cardinality $\aleph_0$.

Part IV

ZermeloFraenkel set theory (ZF)

Chapter 23

Second-order logic

Chapter 24

ZermeloFraenkel set theory

24.1 Natural numbers

24.1.1 Axiom of infinity

The axiom of infinity states that:

$\exists I(\varnothing \in I \wedge \forall x \in I((x \vee \{x\}) \in I))$

There exists a set, called the infinite set. This contains the empty set, and for all elements in I the set also contains the successor to it.

Sequential function

Let's define the sequential function:

$s(n) := \{n \vee \{n\}\}$

We can now rewrite the axiom of infinity as:

$\exists \mathbb{N}(\varnothing \in \mathbb{N} \wedge \forall x \in \mathbb{N}(s(x) \in \mathbb{N}))$

Zero

This set contains the null set: $\varnothing \in \mathbb{N}$.

Zero is defined as the empty set.

$0 := \{\}$

Natural numbers

For all elements in the infinite set, there also exists another element in the infinite set: $\forall x \in \mathbb{N}((x \vee \{x\}) \in \mathbb{N})$.

We then define all sequential numbers as the set of all preceding numbers. So:

$1 := \{0\} = \{\{\}\}$

$2 := \{0, 1\} = \{\{\}, \{\{\}\}\}$

$3 := \{0, 1, 2\} = \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}\}$

Existence of natural numbers

Does each natural number exist? We know the infinite set exists, and we also know the axiom schema of specification:

Point is: For each set, all finite subsets exist. PROVE ELSEWHERE

From infinite set to natural set

We dont know I is limited to natural numbers. Could contain urelements etc.

More

Infinite set axiom written using N. should be I

I could be superset of N, for example set of all natural numbers, and also the set containing the set containing 2.

Can extract N using axiom of specification

We need a way to define the set of natural numbers:

$\forall n(n \in \mathbb{N} \leftrightarrow ([n = \emptyset \vee \exists k(n = k \vee \{k\})] \wedge))$

If we can can define N, we can show it exists from specicication

$\forall x \exists s[P(x) \leftrightarrow (x \in s)]$

$\forall n \exists s[n \in N \leftrightarrow (n \in s)]$

24.1.2 Cardinality of the natural numbers

Consider the infinite set, that is the set of all natural numbers which is defined in ZFC. Clearly there isnt a natural number cardinality of this we instead write $\aleph_0$.

We call sets with this cardinality, countably infinite.

So:

$|\mathbb{N}| = \aleph_0$

Cardinality of natural numbers

We define:

$|\emptyset| = 0$

That, the empty set has a cardinality of 0.

As we define 0 as the empty set, $|0| = 0$.

What is 1? using the definition above we know $|1| > |0|$, so let's say $|1| = 1$, and more generally:

$\forall n \in \mathbb{N} |n| = n$

24.1.3 Ordering

Ordering of the natural numbers

For natural numbers we can say that number n preceeds number $s(n)$. That is:

$n \leq s(n)$

Similarly:

$s(n) \leq s(s(n))$

From the transitive property we know that:

$n \leq s(s(n))$

We can continue this to get:

$n \leq s(s(...s(n)..))$

What can we say about an arbitary comparison?

$a \leq b$

We know that either:

- $a = b$
- $b = s(s(...s(a)...))$
- $a = s(s(...s(b)...))$

In the first case the relation holds.

In the second case the relation holds.

In the third case the relation does not hold, but antisymmetry holds.

As this is is then defined on any pair, the order on natural numbers is total.

As there is a minimum, 0, the relation is also well-ordered.

However if this does not hold then the following instead holds:

24.2 Subsets and powersets

24.2.1 Subset relation

Subset

If all terms which are members of term x are also members of term y, then x is a subset of y.

$\forall x \forall y[(\forall z(z \in x \to z \in y)) \leftrightarrow (x \subseteq y)]$

Proper subset

If two sets are equal, then each is a subset of the other. A proper subset is one which is a subset, and not equal to the other set.

$\forall x \forall y[((\forall z(z \in x \to z \in y))) \land (x \neq y) \leftrightarrow (x \subset y)]$

24.2.2 Powerset function

The power set of s, $P(s)$, contains all subsets of s.

$\forall x x \subseteq s \leftrightarrow x \in P(x)$

Do all subsets exist?? show elsewhere.

24.2.3 Cantors theorem

The cardinality of the powerset is strictly greater than the cardinality of the underlying set.

That is, $|P(s)| < |s|$.

This applies to finite sets and infinite sets. In particular, this means that the powerset of the natural numbers is bigger than the natural numbers.

Proof

If one set is at least as big as another, then then is a surjection from that set to the other.

That is, if we can prove that there is no surjection from a set to its powerset, then we have proved the theorem.

We consider $f(s)$. If there is a surjection, then for every subset of s there should be a mapping from s to that subset.

We take set s and have the powerset of this, $P(s)$.

Consider the set:

$A = \{x \in s | x \notin f(x)\}$

That is, the set of all elements of s which do not map to the surjection.

24.3 Tuples

24.3.1 Tuples

We can get a list of sets in an order. A 2-tuple is an ordered pair:

(a, b)

We can write an ordered pair of a and b as:

$\{\{a\}, \{a, b\}\}$

Ordred pair definition, and tuple

$(a, b) = (c, d) \leftrightarrow (a = c \land b = d)$

This is the characteristic property.

24.3.2 Axiom of pairing

For any pair of sets, x and y there is another set z which containing only x and y.

$\forall x \forall y \exists z \forall a [a \in z \leftrightarrow a = x \lor a = y]$

For each set, there exists a set containing only that set

Take the axiom, but replace all instance of y with x.

$\forall x \exists z \forall a [a \in z \leftrightarrow a = x \lor a = x]$

$\forall x \exists z \forall a[a \in z \leftrightarrow a = x]$

For any finite number of sets, there is a set containing only those sets

For any finite number of sets, there is a set containing the intersection of those sets

24.3.3 Cartesian product

The cartesian product takes two sets, and creates a set containing all ordered pairs of a and b.

$a \times b$

24.3.4 Direct sums

24.4 Functions

24.4.1 Constructing functions

Use of ordered pairs

We can define this as a set of ordered pairs.

$\{\{a\}, \{a, b\}\}$

24.4.2 Domains and ranges

Domain

All values on which the function can be called

$\forall x(f(x) = y) \rightarrow P(y))$

Image

$\forall x((\exists y f(x) = y) \rightarrow P(y))$

Outputs of a function.

AKA: Range

The image of x is $f(x)$.

Preimage

The preimage of y is all x where $f(x) = y$.

Codomain

Sometimes the image is a subset of another set. For example a function may map onto natural numbers above 0. Natural numbers above 0 would be the image, and the natural numbers would be the codomain.

Example

$f(n) = s(n)$

Domain is: $\mathbb{N}$

Codomain is also: $\mathbb{N}$

Image is $\mathbb{N} \land n \neq 0$

Describing functions

If function f maps from set X to set Y we can write this as:

$f : X \to Y$

24.4.3 Axiom of regularity

The axiom of regularity states that:

$\forall x[x \neq \varnothing \to \exists y \in x(y \land x) = \varnothing]$

That is, for all non-empty sets, there is an element of the set which is disjoint from the set itself.

This means that no set can be a member of itself.

Chapter 25

Axiom of union

25.1 Axiom of union

25.1.1 Axiom of union

Motivation

While we have described various sets, we have not said that they exist. That is, if A and B both exist, then currently we cannot ensure $A \wedge B$ exists, just that it can be described.

The axiom of union enables us to ensure all sets from unions and intersections exist.

Axiom of union

$\forall a \exists b \forall c [c \in b \leftrightarrow \exists d (c \in d \wedge d \in a)]$

That is, for every set a, there exists a set b where all the elements in b are the elements of the elements in a.

Here, b is the union of the sets in a.

Part V

Elementary number theory

Chapter 26

The integers

26.1 The integers

26.1.1 Integers

Defining integers

To extend the number line to negative numbers, we define:

$\forall ab \in \mathbb{N} \exists c(a + c = b)$

For any pair of numbers there exists a terms which can be added to one to get the other.

For $1 + x = 3$ this is another natural number, however for $3 + x = 1$ there is no such number.

Integers are defined as the solutions for any pair of natural numbers.

There are an infinite number of ways to write any integer. -1 can be written as $0 - 1$, $1 - 2$ etc.

The class of these terms form an equivalence class.

Integers as ordered pairs

Integers can be defined as an ordered pair of natural numbers, where the integer is valued at: $a - b$.

For example -1 could be shown as:

$-1 = \{\{0\}, \{0, 1\}\}$

$-1 = \{\{5\}, \{5, 6\}\}$

$(a, b) = a - b$

Converting natural numbers to integers

Natural numbers can be shown as integers by using:

$(n, 0)$

Natural numbers can be converted to integers:

$\{\{a\}, \{a, 0\}\}$

Cardinality of integers

26.1.2 Ordering of the integers

Ordering integers

Integers are an ordered pair of naturals.

$\{\{x\}, \{x, y\}\}$

For example -4 can be:

$\{\{4\}, \{4, 8\}\}$

$\{\{0\}, \{0, 8\}\}$

We extend the ordering to say:

$\{\{x\}, \{x, y\}\} \leq \{\{s(x)\}, \{s(x), y\}\}$

$\{\{x\}, \{x, s(y)\}\} \leq \{\{x\}, \{x, y\}\}$

So can we define this on an arbitrary pair:

$\{\{a\}, \{a, b\}\} \leq \{\{c\}, \{c, d\}\}$

We know that:

$\{\{a\}, \{a, b\}\} = \{\{s(a)\}, \{s(a), s(b)\}\}$

And either of:

$\{\{a\}, \{a, b\}\} = \{\{0\}, \{0, A\}\}$

$\{\{a\}, \{a, b\}\} = \{\{B\}, \{B, 0\}\}$

$\{\{a\}, \{a, b\}\} = \{\{0\}, \{0, 0\}\}$

As the latter is a case of either of the other 2, we consider only the first 2.

So we can define:

$\{\{a\},\{a,b\}\} \leq \{\{c\},\{c,d\}\}$

As any of:

$1 : \{\{0\},\{0,A\}\} \leq \{\{0\},\{0,C\}\}$

$2 : \{\{0\},\{0,A\}\} \leq \{\{D\},\{D,0\}\}$

$3 : \{\{B\},\{B,0\}\} \leq \{\{0\},\{0,C\}\}$

$4 : \{\{B\},\{B,0\}\} \leq \{\{D\},\{D,0\}\}$

Case 1:

$\{\{0\},\{0,A\}\} \leq \{\{0\},\{0,C\}\}$

Trivial, depends on relative size of A and C.

Case 2:

$\{\{0\},\{0,A\}\} \leq \{\{D\},\{D,0\}\}$

We can see that:

$\{\{D\},\{D,A\}\} \leq \{\{D\},\{D,0\}\}$

And therefore this holds.

Case 3:

$\{\{B\},\{B,0\}\} \leq \{\{0\},\{0,C\}\}$

We can see that:

$\{\{B\},\{B,0\}\} \leq \{\{B\},\{B,C\}\}$

And therefore this does not hold.

Case 4:

$\{\{B\},\{B,0\}\} \leq \{\{D\},\{D,0\}\}$

Trivial, like case 1.

26.1.3 Functions of integers

Addition

Then we can define addition as:

$(a,b)+(c,d)=(a+c,b+d)$

Integer addition can then be defined:

$a+b=\{\{a_1\},\{a_1,a_2\}\}+\{\{b_1\},\{b_1,b_2\}\}$

$a + b = \{\{a_1 + b_1\}, \{a_1 + b_1, a_2 + b_2\}\}$

Or:

$a + b = c$

$c_1 = a_1 + b_1$

$c_2 = a_2 + b_2$

Multiplication

Similarly, multiplication can be defined as:

$(a, b).(c, d) = (ac + bd, ad + bc)$

$ab = c$

$c_1 = a_1b_1 + a_2b_2$

$c_2 = a_2b_1 + a_1b_2$

Subtraction

$a - b = c$

$c_1 = a_1 + b_2$

$c_2 = a_2 + b_1$

26.1.4 Cardinality of the integers

Cardinality of integers

Chapter 27

The rational numbers

27.1 Rational numbers

27.1.1 Rational numbers

Defining rational numbers

We previously defined integers in terms of natural numbers. Similarly we can define rational numbers in terms of integers.

$\forall ab \in \mathbb{I}(\neg(b = 0) \to \exists c(b.c = a))$

A rational is an ordered pair of integers.

$\{\{a\}, \{a, b\}\}$

So that:

$\{\{a\}, \{a, b\}\} = \frac{a}{b}$

Converting integers to rational numbers

Integers can be shown as rational numbers using:

$(i, 1)$

Integers can then be turned into rational numbers:

$\mathbb{Q} = \frac{a}{1}$

$a = \frac{a_1}{a_2}$

$b = \frac{b_1}{b_2}$

$c = \frac{c_1}{c_2}$

Equivalence classes of rationals

There are an infinite number of ways to write any rational number, as with integers. $\frac{1}{2}$ can be written as $\frac{1}{2}$, $\frac{-2}{-4}$ etc.

The class of these terms form an equivalence class.

We can show these are equal:

$\frac{a}{b} = \{\{a\}, \{a, b\}\}$

$\frac{ca}{cb} = \{\{a\}, \{a, b\}\}$

$\frac{ca}{cb} = \{\{ca\}, \{ca, cb\}\}$

$\{\{a\}, \{a, b\}\} = \{\{ca\}, \{ca, cb\}\}$

27.1.2 Ordering of rationals

27.1.3 Functions of rational numbers

Rational addition

Then we can define addition as:

$(a, b) + (c, d) = (a.d + b.c, b.d)$

$a + b = c$

$c_1 = a_1b_2 + a_2b_1$

$c_1 = a_2b_2$

Rational subtraction

$a - b = c$

$c_1 = a_1b_2 - a_2b_1$

$c_1 = a_2b_2$

Rational multiplication

Similarly, multiplication can be defined as:

$(a, b).(c, d) = (a.c, b.d)$

$ab = c$

$c_1 = a_1 b_1$

$c_1 = a_2 b_2$

Rational division

$\frac{a}{b} = c$

$c_1 = a_1 b_2$

$c_1 = a_2 b_1$

27.1.4 Cardinality of the rationals

Cardinality of rational numbers

We can see rational numbers as cartesian products of integers. That is:

$\mathbb{Q} = Z.Z$

We can order the rational numbers like so:
$\{\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{1}{3}, \frac{2}{2}\frac{3}{1}...\}$
These can be mapped from natural numbers, so there is a bijunctive function.

So:

$|\mathbb{Q}| = |\mathbb{Z}.\mathbb{Z}| = |\mathbb{N}| = \aleph_0$

As: $|\mathbb{Z}.\mathbb{Z}| = |\mathbb{Z}|^2$

$|\mathbb{N}|^n = \mathbb{N}$

27.1.5 Fraction rules

Addition

$\frac{A}{B} + \frac{C}{D} = \frac{AD + BC}{BD}$

Multiplication

$\frac{A}{B}\frac{C}{D} = \frac{AC}{BD}$

b Scaler addition

$C + \frac{A}{B} = \frac{BC + A}{B}$

Scaler multiplication

$C\frac{A}{B} = \frac{AC}{B}$

Other

$\frac{A+B}{C} = \frac{A}{C} + \frac{B}{C}$
$\frac{A}{B} = \frac{AC}{BC}$

27.1.6 Partial fraction decomposition

We have: $\frac{1}{A.B}$
We want this in the form of:
$\frac{a}{A} + \frac{b}{B}$
First, lets define M as the mean of these two numbers, and define $\delta = M - B$.
Then:
$\frac{1}{AB} = \frac{1}{(M+\delta)(M-\delta)} = \frac{a}{M+\delta} + \frac{b}{M-\delta}$
We can rearrange the latter two to find:

$1 = a(M - \delta) + b(M + \delta)$

Now we need to find values of a and b to choose.

Let's examine a.
$a = \frac{1 - b(M+\delta)}{M - \delta}$

$a = -\frac{bM + b\delta - 1}{M - \delta}$

$a = -\frac{bM + b\delta - 1}{M - \delta}$

For this to divide neatly we need both the numerator to be a constant multiplier of the denominator. This means the ratio the multiplier for the left hand side of the denominator is equal to the right:

$\frac{bM}{M} = \frac{b\delta - 1}{-\delta}$

$b = \frac{b\delta - 1}{-\delta}$

$b = \frac{1}{2\delta}$

We can do the same for a.

$a = -\frac{1}{2\delta}$

We can plug these back into our original formula:

$$\frac{1}{(M+\delta)(M-\delta)} = \frac{-\frac{1}{2\delta}}{M+\delta} + \frac{\frac{1}{2\delta}}{M-\delta}$$

$$\frac{1}{(M+\delta)(M-\delta)} = \frac{1}{2\delta}[\frac{1}{M-\delta} - \frac{1}{M+\delta}]$$

27.1.7 Density of the rationals

Rationals are dense in rationals

For any pair of rationals, there is another rational between them:

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

$b = \frac{m}{n}$

Where $b > a$.

We define a new rational:

$c = \frac{a+b}{2}$

$c = \frac{pn + qm}{2qn}$

This is a rational number.

We can write:

$a = \frac{2pn}{2qn}$

$b = \frac{2qm}{2qn}$

As $b > a$ we know $2qm > 2pn$

So: $a < c < b$

Chapter 28

Algebraic numbers

Chapter 29

Complex numbers

29.1 Introducing complex numbers

29.1.1 Defining complex numbers

Define as an ordered pair of reals

We have a complete set of real numbers. Do we need any more?

For the real numbers, we showed there were functions on the rational numbers which did not have rational solutions. We can similarly show that there are functions on real numbers which do not have real solutions.

Consider:

$f(x) = \sqrt{x}$

This has no real solution for $x < 0$.

We define:

$i := \sqrt{-1}$

i and $-i$ can be used interchangeably.

$(-i)^2 = (-1)^2 i^2 = i^2 = -1$

Complex numbers can be shown more generally as:

$a + bi$

We define the complex conjugate of

$x = a + bi$

As

$\bar{x} = a - bi$

Note that

$x\bar{x} = (a + bi)(a - bi) = a^2 - b^2$

We can take exponents of imaginary numbers

$c^{i\theta} = a + bi$

We know the opposite is true.

$c^{-i\theta} = a - bi$

So

$c^{i\theta}c^{-i\theta} = (a + bi)(a - bi)$

$1 = a^2 - b^2$

The case where $c = e$ is of particular note. We explore this later.

29.1.2 Real numbers aren't closed

Define as an ordered pair of reals

We have a complete set of real numbers. Do we need any more?

For the real numbers, we showed there were functions on the rational numbers which did not have rational solutions. We can similarly show that there are functions on real numbers which do not have real solutions.

Consider:

$f(x) = \sqrt{x}$

This has no real solution for $x < 0$.

We define:

$i := \sqrt{-1}$

i and $-i$ can be used interchangeably.

$(-i)^2 = (-1)^2 i^2 = i^2 = -1$

Complex numbers can be shown more generally as:

$a + bi$

We define the complex conjugate of

$x = a + bi$

As

$\bar{x} = a - bi$

Note that

$x\bar{x} = (a + bi)(a - bi) = a^2 - b^2$

We can take exponents of imaginary numbers

$c^{i\theta} = a + bi$

We know the opposite is true.

$c^{-i\theta} = a - bi$

So

$c^{i\theta}c^{-i\theta} = (a + bi)(a - bi)$

$1 = a^2 - b^2$

The case where $c = e$ is of particular note. We explore this later.

29.2 Operators on complex numbers

29.2.1 Arithmetic on complex numbers

For each of these we have:

$x = a + bi$

$y = c + di$

Addition is defined as:

$x + y = a + bi + c + di$

$x + y = (a + c) + (b + d)i$

Subtraction is defined as:

$x - y = a + bi - c - di$

$x - y = (a - c) + (b - d)i$

Multiplication is defined as:

$xy = (a + bi)(c + di)$

$xy = ac - bd + adi + bci$

$xy = (ac - bd) + (ad + bc)i$

Division is defined as:

$\frac{x}{y} = \frac{a + bi}{c + di}$

$\frac{x}{y} = \frac{(a+bi)(c-di)}{(c+di)(c-di)}$

$\frac{x}{y} = \frac{(ac+bd)+(bc-ad)i}{c^2+d^2}$

29.2.2 Complex conjugate

We have $z = a + bi$.

The complex conjugate is:

$\bar{z} = a - bi$

29.2.3 Absolute value

$|z| = \sqrt{z\bar{z}}$

$|z| = \sqrt{(a+bi)(a-bi)}$

$|z| = \sqrt{a^2+b^2}$

29.3 Results

29.3.1 Roots of unity

29.3.2 Complex logarithms

29.3.3 Disks

A disk is the area contained by a circle.

An open disk at (a, b) of radius r is:

$\{(x, y) \in \mathbb{R}^2 : (x-a)^2 + (y-b)^2 < r^2\}$

For a closed disk it is:

$\{(x, y) \in \mathbb{R}^2 : (x-a)^2 + (y-b)^2 \leq r^2\}$

29.3.4 Disks

We defined an open disk at (a, b) of radius r as:

$\{(x, y) \in \mathbb{R}^2 : (x-a)^2 + (y-b)^2 < r^2\}$

For a closed disk it is:

$\{(x, y) \in \mathbb{R}^2 : (x-a)^2 + (y-b)^2 \leq r^2\}$

29.3.5 Annulus

An annulus is a disk, which excludes a smaller disk inside the disk

29.3.6 Punctured disk

If the interior disk is just a point, it is a punctured disk.

Chapter 30

Infinite sequences and limits

30.1 More on sequences

30.1.1 Limit of a sequence

A sequence converges to a limit if

Can converge to a number (1/x)

Can converge to +/- infinity (x)

Otherwise, does not converge (1,-1,1,-1)

Superior and inferior limits

A bounded increasing sequence converges to least upper bound

Identifying the limit of a sequence

Direct comparison test

Root test

30.2 Divergent series

30.2.1 Partial sum

Take a series. We can define the partial sum as:

$s_k = \sum_{i=1}^k a_i$

30.2.2 Cesro sum

The Cesro sum is the limit of the average of the first n partial sums.

That is:

$\lim_{n\rightarrow \infty } \frac{1}{n}\sum_{k=1}^n s_k$

Consider the sequence $\{1, -1, 1, -1, ...\}$

The partial sum is:

$s_k = \sum_{i=1}^k a_i$

$s_k = k \mod (2)$

The Cesro sum is: $\lim_{n\rightarrow \infty } \frac{1}{n}\sum_{k=1}^n s_k$

$\lim_{n\rightarrow \infty } \frac{1}{n}\sum_{k=1}^n k \mod (2)$

$\frac{1}{2}$

30.2.3 Abel summation

Chapter 31

Transcendental and real numbers

31.1 Constructing the real numbers

31.1.1 Cauchy sequences

Cauchy sequence

A cauchy sequence is a sequence such that for an any arbitrarily small number ϵ, there is a point in the sequence where all possible pairs after this are even closer together.

$(\forall\epsilon > 0)(\exists N \in \mathbb{N} : \forall m, n \in \mathbb{N} > N)(|a_m - a_n| < \epsilon)$

This last term gives a distance between two entries. In addition to the number line, this could be used on vectors, where distances are defined.

As a example, $\frac{1}{n}$ is a cauchy sequence, $\sum_i \frac{1}{n}$ is not.

Completeness

Cauchy sequences can be defined on some given set. For example given all the numbers between 0 and 1 there are any number of different cauchy sequences converging at some point.

If it is possible to define a cauchy sequence on a set where the limit is not in the set, then the set is incomplete.

For example, the numbers between 0 and 1 but not including 0 and 1 are not complete. It is possible to define sequences which converge to these missing points.

More abstractly, you could have all vectors where $x^2+y^2 < 1$. This is incomplete (or open) as sequences on these vectors can converge to limits not in the set.

Cauchy sequences are important when considering real numbers. We could define a sequence converging on $\sqrt{2}$, but as this number is not in the set, it is incomplete.

31.1.2 Incompleteness of the rational numbers

The square root of 2 is not a rational number

Let's prove there are numbers which are not rational. Consider $\sqrt{2}$ and let's show that it being rational leads to a contradiction.

$\sqrt{2} = \frac{x}{y}$

$2 = \frac{x^2}{y^2}$

$2y^2 = x^2$

So we know that x^2 is even, and can be shown as $x = 2n$.

$2y^2 = (2n)^2$

$y^2 = 2n^2$

So y is even. But if both x and y are even, then the fraction was not reduced.

This presents a contraction so the original statement must have been false.

So we know there isn't a rational solution to $\sqrt{2}$.

31.1.3 Density of rationals and reals

Rationals are dense in reals

Reals are dense in reals

Reals are dense in rationals

31.1.4 σ-algebra

Review of algebra on a set

An algebra, Σ, on set s is a set of subsets of s such that:

- Closed under intersection: If a and b are in Σ then $a \wedge b$ must also be in Σ
- $\forall ab[(a \in \Sigma \wedge b \in \Sigma) \rightarrow (a \wedge b \in \Sigma)]$
- Closed under union: If a and b are in Σ then $a \vee b$ must also be in Σ.
- $\forall ab[(a \in \Sigma \wedge b \in \Sigma) \rightarrow (a \vee b \in \Sigma)]$

If both of these are true, then the following is also true:

- Closed under complement: If a is in Σ then $s \backslash a$ must also be in Σ

We also require that the null set (and therefore the original set, null's complement) is part of the algebra.

σ-algebra

A σ-algebra is an algebra with an additional condition:

All countable unions of sets in A are also in A.

This adds a constraint. Consider the real numbers with an algebra of all finite sets.

This contains all finite subsets, and their complements. It does not contain $\mathbb{N}$.

However a σ-algebra requires all countable unions to be including, and the natural numbers are a countable union.

The power set is a σ-algebra. All other σ-algebras are subsets of the power set.

Part VI

Elementary algebra

Chapter 32

Solving single-variable polynomials

32.1 Single-variable polynomials

32.1.1 Introduction

A single-variable polynomial is an equation of the form:

$\sum_{i=0}^n a_i x^i = 0$

For example:

- $x = 1$
- $x^2 = 4$
- $x^2 - 3x + 2 = 0$

32.1.2 Degrees

The degree of a polynomial is the highest-order term.

For example $x^3 + x = 0$ has degree 3.

32.1.3 Roots of single-variable polynomials

A solution to a polynomial is a root.

For example 1 and 2 are roots of $x^2 - 3x + 2 = 0$

32.2 Solving quadratic polynomials

32.2.1 Quadratic polynomials

Quadratic polynomials are of the form $ax^2 + bx + c = 0$.

32.2.2 Solving quadratic polynomials

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

32.2.3 Proof

We can get the two solutions to a quadratic equation from the following manipulation.

$$ax^2 + bx + c = 0$$

$$a[x^2 + \frac{b}{a}x] = -c$$

$$a[(x + \frac{b}{2a})^2 - \frac{b^2}{4a^2}] = -c$$

$$a[(x + \frac{b}{2a})^2] = \frac{b^2}{4a} - c$$

$$(x + \frac{b}{2a})^2 = \frac{b^2 - 4ac}{4a^2}$$

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

32.3 Solving cubic polynomials

32.3.1 Cubic polynomials

Cubic polynomials are of the form $ax^3 + bx^2 + cx + d = 0$.

32.3.2 Solving specific cases

We start by solving when $b = 0$, that is:

$aX^3 + bx + c = 0$

32.3.3 Solving the general case

Chapter 33

Solving multi-variable polynomials

33.1 Multi-variable polynomials

33.2 Elliptic curves

33.2.1 Elliptic curves

Of the form $y^2 = x^3 + ax + b$.

Chapter 34

Generating functions

34.1 Generating functions

34.1.1 Generating functions

Definition

A series can be described as:

$\sum_{i=0}^{\infty} s_i x^i$

If we know the function equal to this series, we can identify the ith number.

34.1.2 Fibonacci sequence

The generating function

Let's use a generating function to create a function for the Fibonacci sequence's cth digit. $F(c) = \sum_{i=c} x^i s_i$

Let's look at it for other starts:

$F(c+k) = \sum_{i=c} x^{i+k} s_{i+k}$

$F(c+k) = \sum_{i=c+k} x^i s_i$

$F(c+1) = \sum_{i=c} x^{i+1} s_{i+1}$

$F(c+2) = \sum_{i=c} x^{i+2} s_{i+2}$

This means

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^i s_i x^2 + \sum_{i=c} x^{i+1} s_{i+1} x$

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^{i+2}s_i + \sum_{i=c} x^{i+2}s_{i+1}$

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^{i+2}(s_i + s_{i+1})$

Using the definiton of the Fibonacci sequence

From the definition of the fibonacci sequence, $s_i + s_{i+1} = s_{i+2}$.

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^{i+2}(s_{i+2})$

$F(c)x^2 + F(c+1)x = F(c+2)$

Reducing the functions

Next, we expand out $F(c+1)$ and $F(c+2)$.

$F(c) - F(c+k) = \sum_{i=c} x^i s_i - \sum_{i=c+k} x^i s_i$

$F(c) - F(c+k) = \sum_{i=c}^{c+k} x^i s_i$

$F(c+k) = F(c) - \sum_{i=c}^{c+k} x^i s_i$

So:

$F(c+1) = F(c) - \sum_{i=c}^{c+1} x^i s_i$

$F(c+1) = F(c) - x^c s_c$

$F(c+2) = F(c) - \sum_{i=c}^{c+2} x^i s_i$

$F(c+2) = F(c) - x^{c+1}s_{c+1} - x^c s_c$

Let's take our previous equation

$F(c)x^2 + F(c+1)x = F(c+2)$

$F(c)x^2 + [F(c) - x^c s_c]x = F(c) - x^{c+1}s_{c+1} - x^c s_c$

$F(c)x^2 + F(c)x - x^{c+1}s_c = F(c) - x^{c+1}s_{c+1} - x^c s_c$

$F(c)[x^2 + x - 1] = x^{c+1}s_c - x^{c+1}s_{c+1} - x^c s_c$

$$F(c) = \frac{x^c s_c + x^{c+1}s_{c+1} - x^{c+1}s_c}{1 - x - x^2}$$

Using the first element in the sequence

For the start of the sequence, $c = 0$, $s_0 = s_1 = 1$.

$$F(0) = \frac{x^0 1 + x - x}{1 - x - x^2}$$

$F(0) = \frac{1}{1-x-x^2}$

Let's factorise this:

$F(0) = \frac{-1}{(x+\frac{1}{2}+\frac{\sqrt{5}}{2})(x+\frac{1}{2}-\frac{\sqrt{5}}{2})}$

We can then use partial fraction decomposition

$\frac{1}{(M+\delta)(M-\delta)} = \frac{1}{2\delta}[\frac{1}{M-\delta} - \frac{1}{M+\delta}]$

To show that

$F(0) = \frac{-1}{\sqrt{5}}[\frac{1}{x+\frac{1}{2}-\frac{\sqrt{5}}{2}} - \frac{1}{x+\frac{1}{2}+\frac{\sqrt{5}}{2}}]$

$F(0) = \frac{-1}{\sqrt{5}}[\frac{\frac{1}{2}+\frac{\sqrt{5}}{2}}{(x+\frac{1}{2}-\frac{\sqrt{5}}{2})(\frac{1}{2}+\frac{\sqrt{5}}{2})} - \frac{\frac{1}{2}-\frac{\sqrt{5}}{2}}{(x+\frac{1}{2}+\frac{\sqrt{5}}{2})(\frac{1}{2}-\frac{\sqrt{5}}{2})}]$

$F(0) = \frac{-1}{\sqrt{5}}[\frac{\frac{1}{2}+\frac{\sqrt{5}}{2}}{x(\frac{1}{2}+\frac{\sqrt{5}}{2})-1} - \frac{\frac{1}{2}-\frac{\sqrt{5}}{2}}{x(\frac{1}{2}-\frac{\sqrt{5}}{2})-1}]$

$F(0) = \frac{1}{\sqrt{5}}[(\frac{1}{2}+\frac{\sqrt{5}}{2})\frac{1}{1-x(\frac{1}{2}+\frac{\sqrt{5}}{2})} - (\frac{1}{2}-\frac{\sqrt{5}}{2})\frac{1}{1-x(\frac{1}{2}-\frac{\sqrt{5}}{2})}]$

Finishing off

As we know

$\frac{1}{1-x} = \sum_{i=0} x^i$

So

$F(0) = \frac{1}{\sqrt{5}}[(\frac{1}{2}+\frac{\sqrt{5}}{2})\sum_{i=0} x^i(\frac{1}{2}+\frac{\sqrt{5}}{2})^i - (\frac{1}{2}-\frac{\sqrt{5}}{2})\sum_{i=0} x^i(\frac{1}{2}-\frac{\sqrt{5}}{2})^i]$

$F(0) = \frac{1}{\sqrt{5}}[\sum_{i=0} x^i(\frac{1}{2}+\frac{\sqrt{5}}{2})^{i+1} - \sum_{i=0} x^i(\frac{1}{2}-\frac{\sqrt{5}}{2})^{i+1}]$

$F(0) = \frac{1}{\sqrt{5}}\sum_{i=0} x^i[(\frac{1}{2}+\frac{\sqrt{5}}{2})^{i+1} - (\frac{1}{2}-\frac{\sqrt{5}}{2})^{i+1}]$

So the nth number in the sequence (treating $n=1$ as the first number) is:

$$\frac{1}{\sqrt{5}}[(\frac{1}{2}+\frac{\sqrt{5}}{2})^n-(\frac{1}{2}-\frac{\sqrt{5}}{2})^n]$$

Chapter 35

Diophantine equations

Part VII

Systems of linear equations

Chapter 36

Solving systems of linear equations

36.1 Introduction

36.1.1 Introduction

$m_{11}x + m_{12}y + m_{13}z = v_1$

$m_{21}x + m_{22}y + m_{23}z = v_2$

$m_{31}x + m_{32}y + m_{33}z = v_3$

36.1.2 Matrix and vector notation

We can write the above as:

$\mathbf{M}x = \mathbf{v}$

What are the properties of **M** and **v**?

They are linear and addition and scalar multiplication.

36.2 Rank

36.2.1 Matrix rank

Rank function

The rank of a matrix is the dimension of the span of its component columns.

$rank(M) = span(m_1, m_2, ..., m_n)$

Column and row span

The span of the rows is the same as the span of the columns.

36.2.2 Types of matrices

Empty matrix

A matrix where every element is 0. There is one for each dimension of matrix.

$$A = \begin{bmatrix} 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 0 \end{bmatrix}$$

36.2.3 Triangular matrix

A matrix where $a_{ij} = 0$ where $i < j$ is upper triangular.

A matrix where $a_{ij} = 0$ where $i > j$ is lower triangular.

A matrix which is either upper or lower triangular is a triangular matrix.

36.2.4 Symmetric matrices

All symmetric matrices are square.

The identity matrix is an example.

A matrix where $a_{ij} = a_{ji}$ is symmetric.

36.2.5 Diagonal matrix

A matrix where $a_ij = 0$ where $i \neq j$ is diagonal.

All diagonal matrices are symmetric.

The identity matrix is an example.

36.3 Inversion

36.3.1 Inverse matrices

An invertible matrix implies that if the matrix is multiplied by another matrix, the original matrix can be recovered.

That is, if we have matrix A, there exists matrix A^{-1} such that $AA^{-1} = I$.

Consider a linear map on a vector space.

$Ax = y$

If A is invertible we can have:

$A^{-1}Ax = A^{-1}y$

$x = A^{-1}y$

If we set $y = \mathbf{0}$ then:

$x = \mathbf{0}$

So if there is a non-zero vector x such that:

$Ax = \mathbf{0}$ then A is not invertible.

36.3.2 Left and right inverses

That is, for all matrices A, the left and right inverses of B, B_L^{-1} and B_R^{-1}, are defined such that:

$A(BB_R^{-1}) = A$

$A(B_L^{-1}B) = A$

Left and right inversions are equal

Note that if the left inverse exists then:

$B_L^{-1}B = I$

And if the right inverse exists:

$BB_R^{-1} = I$

Lets take the first:

$B_L^{-1}B = I$

$B_L^{-1}BB_L^{-1} = B_L^{-1}$

$B_L^{-1}BB_L^{-1} - B_L^{-1} = 0$

$B_L^{-1}(BB_L^{-1} - I) = 0$

36.3.3 Inversion of products

$(AB)(AB)^{-1} = I$

$A^{-1}AB(AB)^{-1} = A^{-1}$

$B^{-1}B(AB)^{-1} = B^{-1}A^{-1}$

$(AB)^{-1} = B^{-1}A^{-1}$

36.3.4 Inversion of a diagonal matrix

$DD^{-1} = I$

$D_{ii}D_{ii}^{-1} = 1$

$D_{ii}^{-1} = \frac{1}{D_{ii}}$

36.3.5 Degenerate (singular) matrices

36.3.6 Elementary row operations

Some operations to a matrix can be reversed to arrive at the original matrix. Trivially, multiplying by the identity matrix is reversible.

Similarly, some operations are not reversible. Such as multiplying by the empty matrix.

All matrix operations which can be reversed are combinations of 3 elementary row operations. These are: Swapping rows

$$T_{12} = \begin{bmatrix} 0 & 1 & ... & 0 \\ 1 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

Multiplying rows by a vector

$$D_2(m) = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & m & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

Adding rows to other rows

$$L_{12}(m) = \begin{bmatrix} 1 & 0 & ... & 0 \\ m & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

36.3.7 Gaussian elimination

Simultaneous equations

Matricies can be used to solve simultaneous equations. Condsider the following set of equations.

- $)2x + y - z = 8$
- $-3x - y + 2z = -11$
- $-2x + y + 2z = -3$

We can write this in matrix form.

$Ax = y$

$$A = \begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 2 \\ -2 & 1 & 2 \end{bmatrix}$$

$$x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$y = \begin{bmatrix} 8 \\ -11 \\ -3 \end{bmatrix}$$

Augmented matrix

Consider a form for summarising these equations. This is the augmented matrix.

$$(A|y) = \begin{bmatrix} 2 & 1 & -1 & | & -8 \\ -3 & -1 & 2 & | & -11 \\ -2 & 1 & 2 & | & -3 \end{bmatrix}$$

We can take this and recovery our original A and y.

However we can also do things to this augmented matrix which preserve solutions to the set of equations. These are:

Undertaking combinations of these can make it easier to solve the equation. In particular, if we can arrive at the form:

$$(A|y) = \begin{bmatrix} 1 & 0 & 0 & | & a \\ 0 & 1 & 0 & | & b \\ 0 & 0 & 1 & | & c \end{bmatrix}$$

The solutions for x, y, z are a, b, c.

Echeleon / triangular form

We first aim for:

$$(A|y) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & | & a \\ 0 & a_{22} & a_{23} & | & b \\ 0 & 0 & a_{33} & | & c \end{bmatrix}$$

If this cannot be reached there is no single solution. There may be infinite or no solutions.

Solving

Once we have the triangular form, we can easily solve.

$$(A|y) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & | & a \\ 0 & a_{22} & a_{23} & | & b \\ 0 & 0 & a_{33} & | & c \end{bmatrix}$$

$$(A|y) = \begin{bmatrix} 1 & 0 & 0 & | & a \\ 0 & 1 & 0 & | & b \\ 0 & 0 & 1 & | & c \end{bmatrix}$$

This process is back substitution (or forward substitution if the matrix is triangular the other way).

Matrix inversion

We can think of the inverse of a matrix as one which which takes a series of reverible operations and does these to a matrix then arriving at the identity matrix.

That is, only the three elementary row operations, and combinations of them, can transform a matrix in a way in which it can be reversed. As such All reversible matricies are combinations of the identity matrix and a series of elementary row operations. The inverse matrix is then those series of row operations, in reverse.

We can find identify an inversion by undertaking gaussian elimination. Each step done on the matrix is done to the identify matrix, reversing the process. The end result is the inverted matrix.

Instead of:

$$(A|y) = \begin{bmatrix} 2 & 1 & -1 & | & -8 \\ -3 & -1 & 2 & | & -11 \\ -2 & 1 & 2 & | & -3 \end{bmatrix}$$

Take:

$$(A|I) = \begin{bmatrix} 2 & 1 & -1 & | & 1 & 0 & 0 \\ -3 & -1 & 2 & | & 0 & 1 & 0 \\ -2 & 1 & 2 & | & 0 & 0 & 1 \end{bmatrix}$$

When we solve this we get:

$$(I|A^{-1}) = \begin{bmatrix} 1 & 0 & 0 & | & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & | & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & | & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

Chapter 37

Eigenvalues, Eigenvectors, decomposition and operations

37.1 Eigenvalues and eigenvectors

37.1.1 Eigenvalues and eigenvectors

Which vectors remain unchanged in direction after a transformation?

That is, for a matrix A, what vectors v are equal to scalar multiplication by λ following the operation of the matrix.

$Av = \lambda v$

37.1.2 Spectrum

The spectrum of a matrix is the set of its eigenvalues.

37.1.3 Eigenvectors as a basis

If eigen vectors space space, we can write

$v = \sum_i \alpha_i |\lambda_i\rangle$

Under what circumstances do they span the entirity?

37.1.4 Calculating eigenvalues and eigenvectors using the characteristic polynomial

The characteristic polynomial of a matrix is a polynomial whose roots are the eigenvalues of the matrix.

We know from the definition of eigenvalues and eigenvectors that:

$Av = \lambda v$

Note that

$Av - \lambda v = 0$

$Av - \lambda I v = 0$

$(A - \lambda I)v = 0$

Trivially we see that $v = 0$ is a solution.

Otherwise matrix $A - \lambda I$ must be non-invertible. That is:

$Det(A - \lambda I) = 0$

37.1.5 Calculating eigenvalues

For example

$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$A - \lambda I = \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$

$Det(A - \lambda I) = (2-\lambda)(2-\lambda) - 1$

When this is 0.

$(2-\lambda)(2-\lambda) - 1 = 0$

$\lambda = 1, 3$

37.1.6 Calculating eigenvectors

You can plug this into the original problem.

For example

$Av = 3v$

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

As vectors can be defined at any point on the line, we normalise $x_1 = 1$.

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3x_2 \end{bmatrix}$

Here $x_2 = 1$ and so the eigenvector corresponding to eigenvalue 3 is:

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

37.1.7 Traces

The trace of a matrix is the sum of its diagonal components.

$Tr(M) = \sum_i^n m_{ii}$

The trace of a matrix is equal to the sum of its eigenvectors.

Traces can be shown as the sum of inner products.

$Tr(M) = \sum_i^n e_i M e^i$

37.1.8 Properties of traces

Traces commute

$Tr(AB) = Tr(BA)$

Traces of 1×1 matrices are equal to their component.

$Tr(M) = m_{11}$

37.1.9 Trace trick

If we want to manipulate the scalar:

$v^T M v$

We can use properties of the trace.

$v^T M v = Tr(v^T M v)$

$v^T M v = Tr([v^T][Mv])$

$v^T M v = Tr([Mv][v^T])$

$v^T M v = Tr(M v v^T)$

37.2 Matrix operations

37.2.1 Matrix powers

For a square matrix M we can calculate $MMMM...$, or M^n where $n \in \mathbb{N}$.

37.2.2 Powers of diagonal matrices

Generally, calculating a matrix to an integer power can be complicated. For diagonal matrices it is trivial.

For a diagonal matrix $M = D^n$, $m_{ij} = d_{ij}^n$.

37.2.3 Matrix exponentials

The exponential of a complex number is defined as:

$e^x = \sum \frac{1}{j!} x^j$

We can extend this definition to matrices.

$e^X := \sum \frac{1}{j!} X^j$

The dimension of a matrix and its exponential are the same.

37.2.4 Matrix logarithms

If we have $e^A = B$ where A and B are matrices then we can say that A is matrix logarithm of B.

That is:

$\log B = A$

The dimensions of a matrix and its logarithm are the same.

37.2.5 Matrix square roots

For a matrix M, the square root $M^{\frac{1}{2}}$ is A where $AA = M$.

This does not necessarily exist.

Square roots may not be unique.

Real matrices may have no real square root.

37.3 Matrix decomposition

37.3.1 Similar matrices

In hermitian, show all symmtric matrices are hermitian

For a diagonal matrix, eigenvalues are the diagonal entries?

Similar matrix:

$M = P^{-1}AP$

M and A have the same eigenvalues. If A diagonal, then entries are eigenvalues.

37.3.2 Defective and diagonalisable matrices

37.3.3 Diagonalisable matrices and eigendecomposition

If matrix M is diagonalisable if there exists matrix P and diagonal matrix A such that:

$M = P^{-1}AP$

Diagonalisiable matrices and powers

If these exist then we can more easily work out matrix powers.

$M^n = (P^{-1}AP)^n = P^{-1}A^nP$

A^n is easy to calculate, as each entry in the diagonal taken to the power of n.

Defective matrices

Defective matrices are those which cannot be diagonalised.

Non-singular matries can be defective or not defective, for example the identiy matrix.

Singular matrices can also be defective or not defective, for example the empty matrix.

Eigen-decomposition

Consider an eigenvector v and eigenvalue λ of matrix M.

We known that $Mv = \lambda v$.

If M is full rank then we can generalise for all eigenvectors and eigenvalues:

$MQ = Q\Lambda$

Where Q is the eigenvectors as columns, and Λ is a diagonal matrix with the corresponding eigenvalues. We can then show that:

$M = Q\Lambda Q^{-1}$

This is only possible to calculate if the matrix of eigenvectors is non-singular. Otherwise the matrix is defective.

If there are linearly dependent eigenvectors then we cannot use eigen-decomposition.

37.3.4 Using the eigen-decomposition to invert a matrix

This can be used to invert M.

We know that:

$M^{-1} = (Q\Lambda Q^{-1})^{-1}$

$M^{-1} = Q^{-1}\Lambda^{-1}Q$

We know Λ can be easily inverted by taking the reciprocal of each diagonal element. We already know both Q and its inverse from the decomposition.

If any eigenvalues are 0 then Λ cannot be inverted. These are singular matrices.

37.3.5 Spectral theorem for finite-dimensional vector spaces

37.4 Other

37.4.1 Commutation

We define a function, the commuter, between two objects a and b as:

$[a, b] = ab - ba$

For numbers, $ab - ba = 0$, however for matrices this is not generally true.

37.4.2 Commutators and eigenvectors

Consider two matrices which share an eigenvector v.

$Av = \lambda_A v$

$Bv = \lambda_B v$

Now consider:

$ABv = A\lambda_B v$

$ABv = \lambda_A \lambda_B v$

$BAv = \lambda_A \lambda_B v$

If the matrices share all the same eigenvectors, then the matrices commute, and $AB = BA$.

37.4.3 Identity matrix and the Kronecker delta

37.4.4 Matrix additon and multiplication

Matrix multiplication

$A = A^{mn}$

$B = B^{no}$

$C = C^{mo} = A.B$

$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$

Matrix multiplication depends on the order. Unlike for real numbers,

$AB \neq BA$

Matrix multiplication is not defined unless the condition above on dimensions is met.

A matrix multiplied by the identity matrix returns the original matrix.

For matrix $M = M^{mn}$

$M = MI^m = I^n M$

Matrix addition

2 matricies of the same size, that is with idental dimensions, can be added together.

If we have 2 matrices A^{mn} and B^{mn}

$C = A + B$

$c_{ij} = a_{ij} + b_{ij}$

An empty matrix with 0s of the same size as the other matrix is the identity matrix for addition.

Scalar multiplication

A matrix can be multiplied by a scalar. Every element in the matrix is multiplied by this.

$B = cA$

$b_{ij} = ca_{ij}$

The scalar 1 is the identity scalar.

37.4.5 Transposition and conjugation

Transposition

A matrix of dimensions $m * n$ can be transformed into a matrix $n * m$ by transposition.

$B = A^T$

$b_{ij} = aji$

Transpose rules

$(M^T)^T = M$

$(AB)^T = B^T A^T$

$(A + B)^T = A^T + B^T$

$(zM)^T = zM^T$

Conjugation

With conjugation we take the complex conjugate of each element.

$B = \overline{A}$

$b_{ij} = \overline{a}_{ij}$

Conjugation rules

$\overline{(\overline{A})} = A$

$\overline{(AB)} = (\overline{A})(\overline{B})$

$\overline{(A + B)} = \overline{A} + \overline{B}$

$\overline{(zM)} = \overline{z}\overline{M}$

Conjugate transposition

Like transposition, but with conjucate.

$B = A^*$

$b_{ij} = \bar{a_{ji}}$

Alternatively, and particularly in physics, the following symbol is often used instead.

$(A^*)^T = A^\dagger$

37.4.6 Matrix rank

Rank function

The rank of a matrix is the dimension of the span of its component columns.

$rank(M) = span(m_1, m_2, ..., m_n)$

Column and row span

The span of the rows is the same as the span of the columns.

37.4.7 Types of matrices

Empty matrix

A matrix where every element is 0. There is one for each dimension of matrix.

$$A = \begin{bmatrix} 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 0 \end{bmatrix}$$

37.4.8 Triangular matrix

A matrix where $a_{ij} = 0$ where $i < j$ is upper triangular.

A matrix where $a_{ij} = 0$ where $i > j$ is lower triangular.

A matrix which is either upper or lower triangular is a triangular matrix.

37.4.9 Symmetric matrices

All symmetric matrices are square.

The identity matrix is an example.

A matrix where $a_{ij} = a_{ji}$ is symmetric.

37.4.10 Diagonal matrix

A matrix where $a_i j = 0$ where $i \neq j$ is diagonal.

All diagonal matrices are symmetric.

The identity matrix is an example.

Part VIII

Univariate real analysis

Chapter 38

Ordering of infinite sets

Chapter 39

Introduction

39.0.1 Ordered sets

Totally ordered sets

A totally ordered set is one where the relation is defined on all pairs:

$\forall a \forall b (a \leq b) \vee (b \leq a)$

Note that totality implies reflexivity.

Partially ordered sets (poset)

A partially ordered set, or poset, is one where the relation is defined between each element and itself.

$\forall a (a \leq a)$

That is, every element is related to itself.

These are also called posets.

Well-ordering

A well-ordering on a set is a total order on the set where the set contains a minimum number. For example the relation $\leq$ on the natural numbers is a well-ordering because 0 is the minimum.

The relation $\leq$ on the integers however is not a well-ordering, as there is no minimum number in the set.

39.0.2 Intervals

For a totally ordered set we can define a subset as being all elements with a relationship to a number. For example:

$[a, b] = \{x : a \leq x \wedge x \leq b\}$

This denotes a closed interval. Using the definition above we can also define an open interval:

$(a, b) = \{x : a < x \wedge x < b\}$

39.0.3 Infinitum and supremum

Infinitum

Consider a subset S of a partially ordered set T.

The infinitum of S is the greatest element in T that is less than or equal to all elements in S.

For example:

$\inf[0, 1] = 0$

$\inf(0, 1) = 0$

Supremum

The supremum is the opposite: the smallest element in T which is greater than or equal to all elements in S.

$\sup[0, 1] = 1$

$\sup(0, 1) = 1$

Max and min

If the infinitum of a set S is in S, then the infinimum is the minimum of set S. Otherwise, the minimum is not defined.

$\min[0, 1] = 0$

$\min(0, 1)$ isn't defined.

Similarly:

$\max[0, 1] = 1$

$\max(0, 1)$ isn't defined.

Chapter 40

Real functions

40.1 Real functions

40.1.1 Real functions

Consider a function

$y = f(x)$

$f(x)$ is a real function if:

$\forall x \in \mathbb{R} f(x) \in \mathbb{R}$

40.1.2 Support

$fX \rightarrow R$

Support of f is $x \in X$ where $f(x) \neq 0$

40.1.3 Monotonic functions

Calculus stationary points finding and monotonic functions

40.1.4 Even and odd functions

Defining odd and even functions

An even function is one where:

$f(x) = f(-x)$

An odd function is one where:

$f(x) = -f(-x)$

Functions which are even and odd

If a function is even and odd:

$f(x) = f(-x) = -f(-x)$

$f(x) = -f(x)$

Then $f(x) = 0$.

Scaling odd and even functions

Scaling an even function provides an even function.

$h(x) = c.f(x)$

$h(-x) = c.f(-x)$

$h(-x) = c.f(x)$

$h(-x) = h(x)$

Scaling an odd function provides an odd function.

$h(x) = c.f(x)$

$-h(-x) = -c.f(-x)$

$-h(-x) = c.f(x)$

$-h(-x) = h(x)$

Adding odd and even functions

Note than 2 even functions added together makes an even function.

$h(x) = f(x) + g(x)$

$h(x) = f(-x) + g(-x)$

$h(-x) = f(x) + g(x)$

$h(x) = h(-x)$

And adding 2 odd functions together makes an odd function.

$h(x) = f(x) + g(x)$

$h(x) = -f(-x) - g(-x)$

$-h(-x) = f(x) + g(x)$

$-h(-x) = h(x)$

Multiplying odd and even functions

Multiplying 2 even functions together makes an even function.

$h(x) = f(x)g(x)$

$h(-x) = f(-x)g(-x)$

$h(-x) = f(x)g(x)$

$h(-x) = h(x)$

Multiplying 2 odd functions together makes an even function.

$h(x) = f(x)g(x)$

$h(-x) = f(-x)g(-x)$

$h(-x) = (-1).(-1.)f(x)g(x)$

$h(-x) = h(x)$

40.1.5 Concave and convex functions

Convex functions

A convex function is one where:

$\forall x_1, x_2 \in \mathbb{R} \forall t \in [0,1][f(tx_1 + (1-t)x_2 \leq tf(x_1) + (1-t)f(x_2)]$

That is, for any two points of a function, a line between the two points is above the curve.

A function is strictly convex if the line between two points is strictly above the curve:

$\forall x_1, x_2 \in \mathbb{R} \forall t \in (0,1)[f(tx_1 + (1-t)x_2 < tf(x_1) + (1-t)f(x_2)]$

An example is $y = x^2$.

Concave functions

A concave function is an upside down convex function. The line between two points is below the curve.

$\forall x_1, x_2 \in \mathbb{R} \forall t \in [0,1][f(tx_1 + (1-t)x_2 \geq tf(x_1) + (1-t)f(x_2)]$

A function is strictly concave if the line between two points is strictly below the curve:

$\forall x_1, x_2 \in \mathbb{R} \forall t \in (0,1)[f(tx_1 + (1-t)x_2 > tf(x_1) + (1-t)f(x_2)]$

An example is $y = -x^2$.

Affine functions

If a function is both concave and convex, then the line between two points must be the function itself. This means the function is an affine function.

$y = cx$

40.1.6 Subadditive and superadditive functions

40.2 Limits

40.2.1 Limits of real functions

Limit operator

For a function $f(x)$,

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

We can say that L is the limit if:

$\forall \epsilon > 0 \exists \delta > 0 \forall x[0 < |x-p| < \delta \to |f(x) - L| < \epsilon]$

40.2.2 Limit superior and limit inferior

If a sequence does not converge, but stays between two points, then lim sup is upper bound, lim inf is lower bound.

40.2.3 Big O and little-o notation

Big O notation

In big O notation we are interested in t he size of a function as it getes larger. We ignore constant multiples.

$cx \in O(x)$

And addition of constants.

$cx + b \in O(x)$

If there are two terms and one is larger, we keep the largest.

$x + x^2 \in O(x^2)$

More generally we write:

$f(x) \in O(g(x))$

Little-o notation

40.3 Continuous functions

40.3.1 Continous functions

A function is continuous if:

$\lim_{x\to c} f(x) = f(c)$

For example a function $\frac{1}{x}$ is not continuous as the limit towards 0 is negative infinity. A function like $y = x$ is continous.

More strictly, for any $\epsilon > 0$ there exists

$\delta > 0$

$c - \delta < x < c + \delta$

Such that

$f(c) - \epsilon < f(x) < f(c) + \epsilon$

This means that our function is continuous at our limit c, if for any tiny range around $f(c)$, that is $f(c) - \epsilon$ and $f(c) + \epsilon$, there is a range around c, that is $c - \delta$ and $c + \delta$ such that all the value of $f(x)$ at all of these points is within the other range.

Limits

Why can't we use rationals for analysis?

If discontinous at not rational number, it can still be continous for all rationals.

Eg $f(x) = -1$ unless $x^2 > 2$, where $f(x) = 1$.

Continous for all rationals, because rationals dense in reals.

But can't be differentiated.

40.3.2 Reals or rationals for analysis

Why can't we use rationals for analysis?

If discontinous at not rational number, it can still be continous for all rationals.

eg $f(x) = -1$ unless $x^2 > 2$, where $f(x) = 1$.

Continous for all rationals, because rationals dense in reals

But can't be differentiated

40.3.3 Boundedness theorem

If $f(x)$ is closed and continuous in $[a, b]$ then $f(x)$ is bounded by m and M. That is:

$\exists m \in \mathbb{R} \exists M \in \mathbb{R} \forall x \in [a, b](m < f(x) < M$

40.3.4 Intermediate value theorem

Take a real function $f(x)$ on closed interval $[a, b]$, continuous on $[a, b,]$.

IVT says that for all numbers u between $f(a)$ and $f(b)$, there is a corresponding value c in $[a, b]$ such that $f(c) = u$.

That is:

$\forall u \in [min(f(a), f(b)), max(f(a), f(b))] \exists c \in [a, b](f(c) = u)$

40.3.5 Extreme value theorem

We can expand the boundedness theorem such that m and M are functions of $f(x)$ in the bound $[a, b]$. That is:

$\exists m \in \mathbb{R} \exists M \in \mathbb{R} \forall x \in [a, b](m < f(x) < M)$

Chapter 41

Univariate differentiation

41.1 Partial differentiation

41.1.1 The partial differential operator

Differential

When we change the value of an input to a function, we also change the output. We can examine these changes.

Consider the value of a function $f(x)$ at points x_1 and x_2.

$y_1 = f(x_1)$

$y_2 = f(x_2)$

$y_2 - y_1 = f(x_2) - f(x_1)$

$\frac{y_2 - y_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$

Let's define x_2 in terms of its distance from x_1:

$x_2 = x_1 + \epsilon$

$\frac{y_2 - y_1}{\epsilon} = \frac{f(x_1 + \epsilon) - f(x_1)}{\epsilon}$

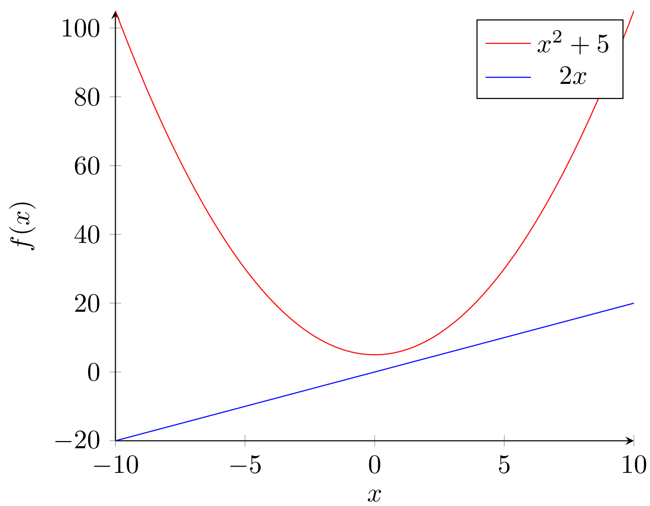
We define the differential of a function as:

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x + \epsilon) - f(x)}{\epsilon}$

If this is defined, then we say the function is differentiable at that point.

Differential operator

Graph test



41.1.2 Differentiating constants, the identity function, and linear functions

Constants

$f(x) = c$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon) - f(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{c - c}{\epsilon} = 0$

x

$f(x) = x$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon) - f(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{x + \epsilon - x}{\epsilon} = 1$

Addition

$f(x) = g(x) + h(g)$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon) + h(x+\epsilon) - g(x) - h(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon) - g(x)}{\epsilon} + \lim_{\epsilon \to 0^+} \frac{h(x+\epsilon) - h(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \frac{\delta g}{\delta x} + \frac{\delta h}{\delta x}$

41.1.3 Partial differentiation is a linear operator

Intro

41.1.4 The chain rule, the product rule and the quotient rule

Chain rule

$f(x) = f(g(x))$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(g(x+\epsilon)) - f(g(x))}{\epsilon}$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon) - g(x)}{g(x+\epsilon) - g(x)} \frac{f(g(x+\epsilon)) - f(g(x))}{\epsilon}$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon) - g(x)}{\epsilon} \frac{f(g(x+\epsilon)) - f(g(x))}{g(x+\epsilon) - g(x)}$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} [\frac{g(x+\epsilon) - g(x)}{\epsilon}] \lim_{\epsilon \to 0^+} [\frac{f(g(x+\epsilon)) - f(g(x))}{g(x+\epsilon) - g(x)}]$

$\frac{\delta f}{\delta x} = \frac{\delta g}{\delta x} \frac{\delta f}{\delta g}$

Product rule

$y = f(x)g(x)$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon) + f(x)g(x+\epsilon) - f(x)g(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon)}{\epsilon} + \lim_{\epsilon \to 0^+} \frac{f(x)g(x+\epsilon) - f(x)g(x)}{\epsilon}$

$$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} g(x+\epsilon)\frac{f(x+\epsilon)-f(x)}{\epsilon} + \lim_{\epsilon \to 0^+} f(x)\frac{g(x+\epsilon)-g(x)}{\epsilon}$$

$$\frac{\delta y}{\delta x} = g(x)\frac{\delta f}{\delta x} + f(x)\frac{\delta g}{\delta x}$$

Quotient rule

$$y = \frac{f(x)}{g(x)}$$

$$\frac{\delta}{\delta x} y = \frac{\delta}{\delta x}\frac{f(x)}{g(x)}$$

$$\frac{\delta}{\delta x} y = \frac{\delta}{\delta x} f(x)\frac{1}{g(x)}$$

$$\frac{\delta}{\delta x} y = \frac{\delta f}{\delta x}\frac{1}{g(x)} - \frac{\delta g}{\delta x}\frac{f(x)}{g(x)^2}$$

$$\frac{\delta}{\delta x} y = \frac{\frac{\delta f}{\delta x} g(x) - \frac{\delta g}{\delta x} f(x)}{g(x)^2}$$

41.1.5 Differentiating natural number power functions

Other

$$\frac{\delta}{\delta x} x^n = \lim_{\delta \to 0} \frac{(x+\delta)^n - x^n}{\delta}$$

$$\frac{\delta}{\delta x} x^n = \lim_{\delta \to 0} \frac{(\sum_{i=0}^n x^i \delta^{n-i}\frac{n!}{i!(n-i)!}) - x^n}{\delta}$$

$$\frac{\delta}{\delta x} x^n = \lim_{\delta \to 0} \sum_{i=0}^{n-1} x^i \delta^{n-i-1}\frac{n!}{i!(n-i)!}$$

$$\frac{\delta}{\delta x} x^n = \lim_{\delta \to 0} x^{n-1}\frac{n!}{(n-1)!(n-n+1)!} + \sum_{i=0}^{n-2} x^i \delta^{n-i-1}\frac{n!}{i!(n-i)!}$$

$$\frac{\delta}{\delta x} x^n = nx^{n-1}$$

41.1.6 L'Hpital's rule

L'Hpital's rule

If there are two functions which are both tend to 0 at a limit, calculating the limit of their divisor is hard. We can use L'Hopital's rule.

We want to calculate:

$\lim_{x\to c} \frac{f(x)}{g(x)}$

This is:

$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{\frac{f(x)-0}{\delta}}{\frac{g(x)-0}{\delta}}$

If:

$\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$

Then

$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{\frac{f(x)-f(c)}{\delta}}{\frac{g(x)-f(c)}{\delta}}$

$\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{f'(x)}{g'(x)}$

41.1.7 Rolle's theorem

Rolles theorem

Take a real function $f(x)$ on closed interval $[a, b]$, differentiable on $(a, b,)$, and $f(a) = f(b)$.

Rolles theorem states that:

$\exists c \in (a, b)(f(c) = 0)$

Generalised Rolles theorem states that:

Generalised Rolles theorem implies Rolles theorem, so we only need to prove the generalised theorem.

41.1.8 Mean value theorem

Mean value theorem

Take a real function $f(x)$ on closed interval $[a, b]$, differentiable on $(a, b,)$.

The mean value theorem states that:

$\exists c \in (a, b)(f(c) = \frac{f(b) - f(a)}{b - a})$

41.1.9 Elasticity

Introduction

We have $f(x)$

$Ef(x) = \frac{x}{f(x)} \frac{\delta f(x)}{\delta x}$

This is the same as:

$Ef(x) = \frac{\delta \ln f(x)}{\delta \ln x}$

41.1.10 Smooth functions

41.1.11 Analytic function

Introduction

41.2 Higher-order differentials

41.2.1 Differentiable functions

Introduction

A differentiable function is one where the differential is defined at all points on the real line.

All differentiable functions are continuous. Not all continuous functions are differentiable.

Differentiability class

We can describe a function with its differentiability class. If a function can be differentiated n times and these differentials are all continous, then the function is class C^n.

Smooth functions

If a function can be differentiated infinitely many times to produce continous functions, it is C^∞, or smooth.

41.2.2 Critial points

Critical points

Where partial derivative are 0.

Chapter 42

e

42.1 Exponentials

42.1.1 Defining e as a binomial

Lemma

$f(n,i) = \frac{n!}{n^i(n-i)!}$

$f(n,i) = \frac{(n-i)!\prod_{j=n-i+1}^{n} j}{n^i(n-i)!}$

$f(n,i) = \frac{\prod_{j=n-i+1}^{n} j}{n^i}$

$f(n,i) = \frac{\prod_{j=1}^{i}(j+n-i)}{n^i}$

$f(n,i) = \prod_{j=1}^{i} \frac{j+n-i}{n}$

$f(n,i) = \prod_{j=1}^{i}(\frac{n}{n} + \frac{j-i}{n})$

$f(n,i) = \prod_{j=1}^{i}(1 + \frac{j-i}{n})$

$\lim_{n\to\infty} f(n,i) = \lim_{n\to\infty} \prod_{j=1}^{i}(1 + \frac{j-i}{n})$

$\lim_{n\to\infty} f(n,i) = \prod_{j=1}^{i} 1$

$\lim_{n\to\infty} f(n,i) = 1$

Defining e

We know that:

$(a+b)^n = \sum_{i=0}^{n} a^i b^{n-i} \frac{n!}{i!(n-i)!}$

Let's set $b = 1$

$(a+1)^n = \sum_{i=0}^{n} a^i \frac{n!}{i!(n-i)!}$

Let's set $a = \frac{1}{n}$

$(1+\frac{1}{n})^n = \sum_{i=0}^{n} \frac{1}{n^i} \frac{n!}{i!(n-i)!}$

$(1+\frac{1}{n})^n = \sum_{i=0}^{n} \frac{1}{i!} \frac{n!}{n^i(n-i)!}$

$\lim_{n\to\infty}(1+\frac{1}{n})^n = \lim_{n\to\infty} \sum_{i=0}^{n} \frac{1}{i!} \frac{n!}{n^i(n-i)!}$

From the lemma above:

$\lim_{n\to\infty}(1+\frac{1}{n})^n = \sum_{i=0}^{\infty} \frac{1}{i!}$

$e = \sum_{i=0}^{\infty} \frac{1}{i!}$

Defining e^x

$e = \lim_{n\to\infty}(1+\frac{1}{n})^n$

$e^x = \lim_{n\to\infty}(1+\frac{1}{n})^{nx}$

$e^x = \lim_{n\to\infty} \sum_{i=0}^{nx} \frac{1}{n^i} \frac{(nx)!}{i!(nx-i)!}$

$e^x = \lim_{n\to\infty} \sum_{i=0}^{nx} \frac{x^i}{i!} \frac{(nx)!}{(nx)^i(nx-i)!}$

From the lemma:

$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$

42.1.2 Differentiating e^x

Intro

We have $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$

$\frac{\delta}{\delta x} e^x = \frac{\delta}{\delta x} \sum_{i=0}^{\infty} \frac{x^i}{i!}$

$\frac{\delta}{\delta x} e^x = \sum_{i=0}^{\infty} \frac{\delta}{\delta x} \frac{x^i}{i!}$

$\frac{\delta}{\delta x} e^x = \sum_{i=1}^{\infty} \frac{\delta}{\delta x} \frac{x^i}{i!}$

$\frac{\delta}{\delta x} e^x = \sum_{i=1}^{\infty} \frac{x^{i-1}}{(i-1)!}$

$\frac{\delta}{\delta x} e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$

$\frac{\delta}{\delta x} e^x = e^x$

42.1.3 Differentiating exponents, logarithms and power functions

Differentiating the natural logarithm

$\frac{\delta}{\delta x} \ln(x) = \lim_{\delta \to 0} \frac{\ln(x+\delta) - ln(x)}{\delta}$

$\frac{\delta}{\delta x} \ln(x) = \lim_{\delta \to 0} \frac{\ln \frac{x+\delta}{x}}{\delta}$

$\frac{\delta}{\delta x} \ln(x) = \lim_{\delta \to 0} \frac{\ln(1 + \frac{\delta}{x})}{\delta}$

$\frac{\delta}{\delta x} \ln(x) = \frac{1}{x} \lim_{\delta \to 0} \frac{x}{\delta} \ln(1 + \frac{\delta}{x})$

$\frac{\delta}{\delta x} \ln(x) = \frac{1}{x} \ln(\lim_{\delta \to 0} (1 + \frac{\delta}{x})^{\frac{x}{\delta}})$

$\frac{\delta}{\delta x} \ln(x) = \frac{1}{x} \ln(e)$

$\frac{\delta}{\delta x} \ln(x) = \frac{1}{x}$

Differentiating logarithms of other bases

$$log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$

$$log_a(x) = \frac{\ln(x)}{\ln(a)}$$

$$\frac{\delta}{\delta x}\log_a(x) = \frac{\delta}{\delta x}\frac{\ln(x)}{ln(a)}$$

$$\frac{\delta}{\delta x}\log_a(x) = \frac{1}{xln(a)}$$

Exponents

$$y = a^x$$

$$\ln(y) = x\ln(a)$$

$$\frac{\delta}{\delta x}\ln(y) = \frac{\delta}{\delta x}x\ln(a)$$

$$\frac{\delta}{\delta x}\ln(y) = \ln(a)$$

$$\frac{1}{y}\frac{\delta}{\delta x}y = \ln(a)$$

$$\frac{\delta}{\delta x}a^x = a^x\ln(a)$$

Power functions

$$y = x^n$$

$$\frac{\delta}{\delta x}y = \frac{\delta}{\delta x}x^n$$

$$\frac{\delta}{\delta x}y = \frac{\delta}{\delta x}e^{n\ln(x)}$$

$$\frac{\delta}{\delta x}y = \frac{n}{x}e^{n\ln(x)}$$

$$\frac{\delta}{\delta x}y = nx^{n-1}$$

Chapter 43

Univariate integration

43.1 The Riemann integral

43.1.1 Riemann sums

Given a function $f(x)$ and an interval $[a,b]$, we can divide $[a,b]$ into n sections and calculate:

$\sum_{j=0}^{n(b-a)} f(a+\frac{j}{n})$

This is the Riemann sum.

43.1.2 Riemann integral

We take the limit of the Riemann sum as $n \to \infty$

$\int_a^b f(x)dx := \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a+\frac{j}{n})$

43.1.3 Linearity

$\int_a^b f(x)+g(x)dx = \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a+\frac{j}{n})+g(a+\frac{j}{n})$

$\int_a^b f(x)+g(x)dx = \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a+\frac{j}{n})+\lim_{n\to\infty} \sum_{j=0}^{n(b-a)} g(a+\frac{j}{n})$

$\int_a^b f(x)+g(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$

43.1.4 Continuation

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \lim_{n\to\infty} \sum_{j=0}^{n(c-b)} f(b + \frac{j}{n})$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \sum_{j=0}^{n(c-b)} f(b + \frac{j}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \sum_{j=n(b-a)}^{n(c-b)+n(b-a)} f(b + \frac{j-n(b-a)}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \sum_{j=n(b-a)}^{n(c-a)} f(a + \frac{j}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(c-a)} f(a + \frac{j}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$

43.2 Definite and indefinite integrals

43.2.1 Definite integrals

Definite integrals are between two points.

$\int_0^1 f(x)dx$

43.2.2 Indefinite integrals

Indefinite integrals are not. Eg +c at end. The antiderivative.

$\int f(x)dx$

43.2.3 Unsigned definite integral

$\int_{[0,1]} f(x)dx$

43.3 Anti-derivatives

43.3.1 Anti-derivative

Taking the derivative of a function provides another function. The anti-derivative of a function is a function which, when differentiated, provides the original function.

As this function can include any additive constant, there are an infinite number of anti-derivatives for any function.

43.4 Integration by parts

43.4.1 Integration by parts

We have:

$\frac{\delta y}{\delta x} = f(x)g(x)$

We want that in terms of y.

We know from the product rule of differentiation:

$y = a(x)b(x)$

Means that:

$\frac{\delta y}{\delta x} = a'(x)b(x) + a(x)b'(x)$

So let's relabel $f(x)$ as $h'(x)$

δ

$\frac{\delta y}{\delta x} = h'(x)g(x)$

$\frac{\delta y}{\delta x} + h(x)g'(x) = h'(x)g(x) + h(x)g'(x)$

$y + \int h(x)g'(x) = \int h'(x)g(x) + h(x)g'(x)$

$y + \int h(x)g'(x) = h(x)g(x)$

$y = h(x)g(x) - \int h(x)g'(x)$

For example:

$\frac{\delta y}{\delta x} = x.\cos(x)$

$f(x) = \cos(x)$

$g(x) = x$

$h(x) = \sin(x)$

$g'(x) = 1$

So:

$y = x \int \cos(x)dx - \int \sin(x)dx$

$y = x\sin(x) - \cos(x) + c$

43.5 The fundamental theorem of calculus

43.5.1 Mean value theorem for integration

Take function $f(x)$. From the extreme value theorem we know that:

$\exists m \in \mathbb{R} \exists M \in \mathbb{R} \forall x \in [a, b](m < f(x) < M)$

43.5.2 Fundamental theorem of calculus

From continuation we know that:

$\int_a^{x_1} f(x)dx + \int_{x_1}^{x_1+\delta x} f(x)dx = \int_a^{x_1+\delta x} f(x)dx$

$\int_x^{x_1+\delta x} f(x)dx = \int_a^{x_1+\delta} f(x)dx - \int_a^{x_1} f(x)dx$

Indefinite integrals

43.6 Other

43.6.1 Trigonometric substitution

For later? Haven't defined trigonometry yet.

43.6.2 Getting functions from derivatives

$f(c) = f(a) + \int_a^c \frac{\delta}{\delta x} f(x)dx$

Chapter 44

The sine and cosine functions

44.1 Sine and cosine

44.1.1 Defing sine and cosine using Euler's formula

Euler's formula

Previously we showed that:

$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$

Consider:

$e^{i\theta}$

$e^{i\theta} = \sum_{j=0}^{\infty} \frac{(i\theta)^j}{j!}$

$e^{i\theta} = [\sum_{j=0}^{\infty} \frac{(\theta)^{4j}}{(4j)!} - \sum_{j=0}^{\infty} \frac{(\theta)^{4j+2}}{(4j+2)!}] + i[\sum_{j=0}^{\infty} \frac{(\theta)^{4j+1}}{(4j+1)!} - \sum_{j=0}^{\infty} \frac{(\theta)^{4j+3}}{(4j+3)!}]$

We then use this to define sin and cos functions.

$\cos(\theta) := \sum_{j=0}^{\infty} \frac{(\theta)^{4j}}{(4j)!} - \sum_{j=0}^{\infty} \frac{(\theta)^{4j+2}}{(4j+2)!}$

$\sin(\theta) := \sum_{j=0}^{\infty} \frac{(\theta)^{4j+1}}{(4j+1)!} - \sum_{j=0}^{\infty} \frac{(\theta)^{4j+3}}{(4j+3)!}$

So:

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

Alternative formulae for sine and cosine

We know

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$e^{-i\theta} = \cos(\theta) - i\sin(\theta)$

So

$e^{i\theta} + e^{-i\theta} = cos(\theta) + i\sin(\theta) + \cos(\theta) - i\sin(\theta)$

$\cos(\theta) = \dfrac{e^{i\theta} + e^{-i\theta}}{2}$

And

$e^{i\theta} - e^{-i\theta} = cos(\theta) + i\sin(\theta) - \cos(\theta) + i\sin(\theta)$

$\sin(\theta) = \dfrac{e^{i\theta} - e^{-i\theta}}{2i}$

Sine and cosine are odd and even functions

Sine is an odd function.

$\sin(-\theta) = -\sin(\theta)$

Cosine is an even function.

$\cos(-\theta) = \cos(\theta)$

44.1.2 De Moive's formula

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

Let $\theta = nx$:

$e^{inx} = \cos(nx) + i\sin(nx)$

$(e^{ix})^n = \cos(nx) + i\sin(nx)$

$(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$

44.1.3 Expanding sine and cosine

Expansion

$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$

$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$

44.1.4 Addition of sine and cosine

Adding waves with same frequency

We know that:

$a\sin(bx + c) = a\sin(bx)\cos(c) + a\sin(c)\cos(bx)$

So:

$a\sin(bx+c)+d\sin(bx+e) = a\sin(bx)\cos(c)+a\sin(c)\cos(bx)+d\sin(bx)\cos(e)+d\sin(e)\cos(bx)$

We know that:

$\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$

So:

$a\sin(bx + c) + d\sin(bx + f) = a\frac{e^{i(bx+c)} - e^{-i(bx+c)}}{2i} + d\frac{e^{i(bx+f)} - e^{-i(bx+f)}}{2i}$

$a\sin(bx + c) + d\sin(bx + f) = \frac{a(e^{i(bx+c)} - e^{-i(bx+c)}) + d(e^{i(bx+f)} - e^{-i(bx+f)})}{2i}$

$a\sin(bx + c) + d\sin(bx + f) = \frac{a(e^{ibx}e^{ic} - e^{-ibx}e^{-ic}) + d(e^{ibx}e^{if} - e^{-ibx}e^{-if})}{2i}$

$a\sin(bx + c) + d\sin(bx + f) = \frac{(e^{ibx}(ae^{ic} + de^{if}) - e^{-ibx}(ae^{-c} + d^{-if})}{2i}$

$a_i\sin(b_ix + c_i) + a_j\sin(b_jx + c_j) = a_i\sin(b_ix + c_i) + a_j\sin(b_ix + c_j)$

$a_i\sin(b_ix + c_i) + a_j\sin(b_jx + c_j) = a_i\sin(b_ix)\cos(c_i) + a_i\sin(c_i)\cos(b_ix) + a_j\sin(b_ix)\cos(c_j) + a_j\sin(c_j)\cos(b_ix)$

44.1.5 Calculus of sine and cosine

Unity

Note that with imaginary numbers we can reverse all is. So:

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$e^{-i\theta} = \cos(\theta) - i\sin(\theta)$

$e^{i\theta}e^{-i\theta} = (\cos(\theta) + i\sin(\theta))(\cos(\theta) - i\sin(\theta))$

$e^{i\theta}e^{-i\theta} = \cos(\theta)^2 + \sin(\theta)^2$

$e^{i\theta}e^{-i\theta} = e^{i\theta - i\theta} = e^0 = 1$

So:

$\cos(\theta)^2 + \sin(\theta)^2 = 1$

Note that if $\cos(\theta)^2 = 0$, then $\sin(\theta)^2 = \pm 1$

That is, if the real part of $e^{i\theta}$ is 0, the imaginary part is ± 1. And visa versa.

Similarly if the derivative of the real part of $e^{i\theta}$ is 0, the imaginary part is ± 1. And visa versa.

Sine and cosine are linked by their derivatives

Note that these functions are linked in their derivatives.

$\frac{\delta}{\delta\theta}\cos(\theta) = \sum_{j=0}^{\infty}\frac{(\theta)^{(4j+3)}}{(4j+3)!} - \sum_{j=0}^{\infty}\frac{(\theta)^{4j+1}}{(4j+1)!}$

$\frac{\delta}{\delta\theta}\cos(\theta) = -\sin(\theta)$

Similarly:

$\frac{\delta}{\delta\theta}\sin(\theta) = cos(\theta)$

Both sine and cosine oscillate

$\frac{\delta^2}{\delta\theta^2}\sin(\theta) = -\sin(\theta)$

$\frac{\delta^2}{\delta\theta^2}\cos(\theta) = -\cos(\theta)$

So for either of:

$y = \cos(\theta)$

$y = \sin(\theta)$

We know that

$\frac{\delta^2}{\delta\theta^2}y(\theta) = -y(\theta)$

Consider $\theta = 0$.

$e^{i.0} = \cos(0) + i\sin(0)$

$1 = \cos(0) + i\sin(0)$

$\sin(0) = 0$

$\cos(0) = 1$

Similarly we know that the derivative:

$\sin'(0) = \cos(0) = 1$

$\cos'(0) = -\sin(0) = 0$

Consider $\cos(\theta)$.

As $\cos(0)$ is static at $\theta = 0$, and is positive, it will fall until $\cos(\theta) = 0$.

While this is happening, $\sin(\theta)$ is increasing. As:

$\cos(\theta)^2 + \sin(\theta)^2 = 1$

$\sin(\theta)$ will equal 1 where $\cos(\theta) = 0$.

Due to symmetry this will repeat 4 times.

Let's call the length of this period τ.

Where $\theta = \tau * 0$

- $\cos(\theta) = 1$
- $\sin(\theta) = 0$

Where $\theta = \tau * \frac{1}{4}$

- $\cos(\theta) = 0$
- $\sin(\theta) = 1$

Where $\theta = \tau * \frac{2}{4}$

- $\cos(\theta) = -1$
- $\sin(\theta) = 0$

Where $\theta = \tau * \frac{3}{4}$

- $\cos(\theta) = 0$
- $\sin(\theta) = -1$

Relationship between$\cos(\theta)$ and $\sin(\theta)$

Note that $\sin(\theta + \frac{\tau}{4}) = \cos(\theta)$

Note that $\sin(\theta) = \cos(\theta)$ at

- $\tau * \frac{1}{8}$
- $\tau * \frac{5}{8}$

And that all these answers loop. That is, add any integer multiple of τ to θ and the results hold.

$e^{i\theta} = e^{i\theta + n\tau}$

$n \in \mathbb{N}$

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$e^{i\theta} = \cos(\theta + n\tau) + i\sin(\theta + n\tau)$

$e^{i\theta} = e^{i(\theta + n\tau)}$

Calculus of trig

Relationship between cos and sine

$\sin(x + \frac{\pi}{2}) = \cos(x)$

$\cos(x + \frac{\pi}{2}) = -\sin(x)$

$\sin(x + \pi) = -\sin(x)$

$\cos(x + \pi) = -\cos(x)$

$\sin(x + \tau) = \sin(x)$

$\cos(x + \tau) = \cos(x)$

Chapter 45

The tangent function, π and τ

45.1 Tangent

45.1.1 Tan

The $\tan(\theta)$ function is defined as:

$\tan(\theta) := \frac{\sin(\theta)}{\cos(\theta)}$

Behaviour around 0

$\sin(0) = 0$

$\cos(0) = 1$

$\tan(0) := \frac{\sin(0)}{\cos(0)}$

$\tan(0) = \frac{0}{1}$

$\tan(0) = 0$

Behaviour around $\cos(\theta) = 0$

$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$

So $\tan(\theta)$ is undefined where $\cos(\theta) = 0$.

This happens where:

$\theta = \frac{\tau}{4} + \frac{1}{2}n\tau$

$\theta = \frac{1}{4}\tau(1 + 2n)$

Where $n \in \mathbb{Z}$.

Derivatives

$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$

$\frac{\delta}{\delta\theta}\tan(\theta) = \frac{\delta}{\delta\theta}\frac{\sin(\theta)}{\cos(\theta)}$

$\frac{\delta}{\delta\theta}\tan(\theta) = \frac{\cos(\theta)}{\cos(\theta)} + \frac{\sin^2(\theta)}{\cos^n(\theta)}$

$\frac{\delta}{\delta\theta}\tan(\theta) = 1 + \tan^2(\theta)$

Note this is always positive. This means:

$\lim_{\cos(\theta)\to 0^+} = -\infty$

$\lim_{\cos(\theta)\to 0^-} = \infty$

45.1.2 Inverse functions

Inverse trigonometric functions

$\sin(\arcsin(\theta)) := \theta$

$\cos(\arccos(\theta)) := \theta$

$\tan(\arctan(\theta)) := \theta$

45.1.3 Integrals

Cosine and sine

$\arccos(\theta)$, $\arcsin(\theta)$ and difficulty of inversing

In order to determine τ we need inverse functions for $\cos(\theta)$ or $\sin(\theta)$.

These are the $\arccos(\theta)$ and $\arcsin(\theta)$ functions respectively.

However this is not easily calculated. Instead we look for another function.

Calculating $\arctan(\theta)$

So we want a function to inverse this. This is the $\arctan(\theta)$ function.

If $y = \tan(\theta)$, then:

$\theta = \arctan(y)$

We know the derivative for $\tan(\theta)$ is:

$\frac{\delta}{\delta\theta}\tan(\theta) = 1 + \tan^2(\theta)$

$\frac{\delta y}{\delta\theta} = 1 + y^2$

So

$\frac{\delta\theta}{\delta y} = \frac{1}{1+y^2}$

$\frac{\delta}{\delta y}\arctan(y) = \frac{1}{1+y^2}$

So the value for $\arctan(k)$ is:

$\arctan(k) = \arctan(a) + \int_a^k \frac{\delta}{\delta y}\arctan(y)\delta y$

$\arctan(k) = \arctan(a) + \int_a^k \frac{1}{1+y^2}\delta y$

What do we know about this function? We know it can map to multiple values of θ because the underlying $\sin(\theta)$ and $\cos(\theta)$ functions also loop.

We know that one of the results for $\arctan(0)$ is 0.

45.1.4 Calculating τ

As we note above, $\sin(\theta) = \cos(\theta)$ at $\theta = \tau * \frac{1}{8}$

This is also where $\tan(\theta) = 1$.

$\arctan(k) = \arctan(a) + \int_a^k \frac{1}{1+y^2}\delta y$

We start from $a = 0$.

$\arctan(k) = \arctan(0) + \int_0^k \frac{1}{1+y^2}\delta y$

We know that one of the results for $\arctan(0)$ is 0.

$\arctan(k) = \int_0^k \frac{1}{1+y^2}\delta y$

We want $k = 1$

$\arctan(1) = \int_0^1 \frac{1}{1+y^2}\delta y$

$\frac{\tau}{8} = \int_0^1 \frac{1}{1+y^2}\delta y$

$\tau = 8\int_0^1 \frac{1}{1+y^2}\delta y$

We know that the $\cos(\theta)$ and $\sin(\theta)$ functions cycle with period τ.

Therefore $cos(n.\tau) = \cos(0)$

Chapter 46

Polar coordinates

46.1 Polar coordinates

46.1.1 Polar co-ordinates

All complex numbers can be shown in polar form

Consider a complex number

$z = a + bi$

We can write this as:

$z = r\cos(\theta) + ir\sin(\theta)$

Polar forms are not unique

Because the functions loop:

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = a(\cos(\theta + n\tau) + i\sin(\theta + n\tau))$

$ae^{i\theta} = ae^{i\theta + n\tau}$

Additionally:

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = -a(\cos(\theta) - i\sin(\theta))$

$ae^{i\theta} = -a(\cos(\theta + \frac{\pi}{2}) + i\sin(\theta + \frac{\pi}{2}))$

Real and imaginary parts of a complex number in polar form

We can extract the real and imaginary parts of this number.

$Re(z) := r\cos(\theta)$

$Im(z) := r\sin(\theta)$

Alternatively:

$Re(z) = r\frac{e^{i\theta}+e^{-i\theta}}{2}$

$Im(z) = r\frac{e^{i\theta}-e^{-i\theta}}{2i}$

46.1.2 Moving between polar and cartesian coordinates

All polar numbers can be shown as Cartesian

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = a\cos(\theta) + ia\sin(\theta)$

$z = a + bi$

$e^{i\theta} =$

$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$

46.1.3 Arithmetic of polar coordinates

Addition

$z_3 = z_1 + z_2$

$z_3 = a_1e^{i\theta_1} + a_2e^{i\theta_2}$

$z_3 = a_1[\cos(\theta_1) + i\sin(\theta_1)] + a_2[\cos(\theta_2) + i\sin(\theta_2)]$

$z_3 = [a_1\cos(\theta_1) + a_2\cos(\theta_2)] + i[a_1\sin(\theta_1) + a_2\sin(\theta_2)]$

Multiplication

$z_3 = z_1.z_2$

$z_3 = a_1e^{i\theta_1}a_2e^{i\theta_2}$

$z_3 = a_1a_2e^{i(\theta_1+\theta_2)}$

$a_3 = a_1a_2$

$\theta_3 = \theta_1 + \theta_2$

Chapter 47

Other trigonometric functions

47.1 Other

47.1.1 Other functions

Reciprocal trigonometric functions

Standard

$\csc(\theta) := \frac{1}{\sin(\theta)}$

$\sec(\theta) := \frac{1}{\cos(\theta)}$

$\cot(\theta) := \frac{1}{\tan(\theta)}$

Hyperbolic

$csch(\theta) := \frac{1}{\sinh(\theta)}$

$sech(\theta) := \frac{1}{\cosh(\theta)}$

$coth(\theta) := \frac{1}{\tanh(\theta)}$

Inverse trigonometric functions

Reciprocal standard

$\csc(arccsc(\theta)) := \theta$

$\sec(arcsec(\theta)) := \theta$

$\cot(arccot(\theta)) := \theta$

Reciprocal hyperbolic

$csch(arccsch(\theta)) := \theta$

$sech(arcsech(\theta)) := \theta$

$coth(arccoth(\theta)) := \theta$

47.2 Hyperbolic functions

47.2.1 Hyperbolic functions

Hyperbolic functions

$\sinh(\theta) := \sin(i\theta)$

$\cosh(\theta) := \cos(i\theta)$

$\tanh(\theta) := \tan(i\theta)$

Inverse trigonometric functions

$\sinh(arcsinh(\theta)) := \theta$

$\cosh(arccosh(\theta)) := \theta$

$\tanh(arctan(\theta)) := \theta$

Chapter 48

Taylor and Fourier analysis

48.1 Power series

48.1.1 Power series

of the form:

$\sum_{n=0} a_n(x-c)^n$

Smoothness of power series

Power series are all smooth. That is, they are infinitely differentiable.

48.2 Taylor series

48.2.1 Taylor series

$f(x)$ can be estimated at point c by identifying its repeated differentials at point c.

The coefficients of an infinate number of polynomials at point c allow this.

$f(x)=\sum_{i=0}^\infty a_i(x-c)^i$

$f'(x)=\sum_{i=1}^\infty a_i(x-c)^{i-1}i$

$f''(x)=\sum_{i=2}^\infty a_i(x-c)^{i-2}i(i-1)$

$f^j(x)=\sum_{i=j}^\infty a_i(x-c)^{i-j}\frac{i!}{(i-j)!}$

For $x = c$ only the first term in the series is non-zero.

$f^j(c) = \sum_{i=j}^{\infty} a_i(c-c)^{i-j}\frac{i!}{(i-j)!}$

$f^j(c) = a_i j!$

So:

$a_j = \frac{f^j(c)}{j!}$

So:

$f(x) = \sum_{i=0}^{\infty}(x-c)^i\frac{f^i(c)}{i!}$

48.2.2 Convergence

If $x = c$ then the power series will be equal to a_0.

For other values the power series may not converge.

Cauchy-Hadamard theorem

Radius of convergence:

$\frac{1}{R} = \limsup_{n\to\infty}(|a_n|^{\frac{1}{n}})$

48.2.3 Maclaurin series

A Taylor series around $c = 0$.

$f(x) = \sum_{i=0}^{\infty}(x-c)^i\frac{f^i(c)}{i!}$

$f(x) = \sum_{i=0}^{\infty}(x)^i\frac{f^i(0)}{i!}$

For example, for:

$f(x) = (1-x)^{-1}$

$f^i(0) = i!$

So, around $x = 0$:

$f(x) = \sum_{i=0}^{\infty}(x)^i$

48.2.4 Fourier transforms

Taylor series of matrices

We can also use Taylor series to evaluate functions of matrices.

Consider e^M

We can evaluate this as:

$e^M = \sum_{k=0}^{\infty} \frac{1}{k!} M^k$

48.2.5 Analytic functions

(root test, direct comparison test, rate of convergence, radius of convergence)

48.3 Fourier analysis

48.3.1 Representing wave functions

Wave function are of the form:

$\cos(ax+b)$

$\sin(ax+b)$

We can use the following identities:

- $\cos(x) = \sin(x + \frac{\tau}{8})$
- $\sin(-x) = -\sin(x)$
- $\sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a)$

So we can write any function as:

Using e

48.3.2 Harmonics

48.3.3 Fourier series

Fourier series

Motivation: we have a function we want to display as another sort of function.

More specifically, a function can be shown as a combination of sinusoidal waves.

To frame this lets imagine a sound wave, with values $f(t)$ for all time values t. We can imagine this as a summation of sinusoidal functions. That is:

$f(t) = \sum_{n=0}^{\inf} a_n cos(nw_0 t)$

We want to get another function $F(\xi)$ for all frequencies ξ.

Combinations of wave functions

We can add sinusoidal waves to get new waves.

For example

$s_N(x) = 2\sin(x+3) + \sin(-4x) + \frac{1}{2}\cos(x)$

As a summation of series

We can simplify arbitrary series using the following identities:

$\cos(x) = \sin(x + \frac{\tau}{8})$

$\sin(-x) = -\sin(x)$

So we have:

$s(x) = 2\sin(x+3) - \sin(4x) + \frac{1}{2}\sin(x + \frac{\tau}{8})$

We can put this into the following format:

$s(x) = \sum_{i=1}^{m} a_i \sin(b_i x + c_i)$

Where:

$a = [2, -1, \frac{1}{2}]$

$b = [1, 4, 1]$

$c = [3, 0, \frac{\tau}{8}]$

Ordering by b

We can move terms around to get:

$s(x) = \sum_{i=1}^{m} a_i \sin(b_i x + c_i)$

Where:

$a = [2, \frac{1}{2}, -1]$

$b = [1, 1, 4]$

$c = [3, \frac{\tau}{8}, 0]$

Adding waves with same frequency

We know that:

$\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$

So:

$\sin(b_i x + c_i) = \sin(b_i x)\cos(c_i) + \sin(c_i)\cos(b_i x)$

If 2 terms have the same value for b_i, then:

$a_i \sin(b_i x + c_i) + a_j \sin(b_j x + c_j) = a_i \sin(b_i x + c_i) + a_j \sin(b_i x + c_j)$

$a_i \sin(b_i x + c_i) + a_j \sin(b_j x + c_j) = a_i \sin(b_i x)\cos(c_i) + a_i \sin(c_i)\cos(b_i x) + a_j \sin(b_i x)\cos(c_j) + a_j \sin(c_j)\cos(b_i x)$

So we now get for:

$s(x) = \sum_{i=1}^{m} a_i \sin(b_i x + c_i)$

$a = [, -1]$

$b = [, 4]$

$c = [, 0]$

48.3.4 Fourier transforms

Fourier transform

$\hat{f}(\Xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \Xi} dx$

Inverse Fourier transform

$f(x) = \int_{-\infty}^{\infty} \hat{f}(\Xi) e^{2\pi i x \Xi} d\Xi$

Fourier inversion theorem

Chapter 49

The gamma and beta functions

49.1 Expanding functions of natural numbers

49.1.1 Gamma function

The gamma function expands the factorial function to the real (and complex) numbers

We want:

$f(1) = 1$

$f(x+1) = xf(x)$

There are an infinite number of functions which fit this. The function could fluctuate between the natural numbers.

The function we use is:

$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$

49.1.2 Beta function

The beta function expand the binomial coefficient formula to the real (and complex) numbers.

We want to expand the binomial coefficient function.

$(\frac{n}{k}) = \frac{n!}{k!(n-k)!}$

We do this as:

$B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

Chapter 50

Ordinary Differential Equations (ODEs)

50.1 Introduction

50.1.1 Order of differential equations

50.1.2 Implicit and explit differential equations

An ordinary differential equation is one with only one independent variable. For example:

$\frac{dy}{dx} = f(x)$

The order of a differential equation is the number of differentials of y included. For example one with the second derivative of y is of order 2.

Ordinary equations can can either implicit or explicit. An explicit function shows the highest order derivative as a function of other terms.

An implicit function is one which is not explicit.

A linear ODE is an explicit ODE where the derivative terms of y do not multiply together, that is, in the form:

$y^{(n)} = \sum_i a_i(x)y^{(i)} + r(x)$

First-order ODEs

We have an evolution:

$\frac{dy}{dt} = f(t, y)$

And a starting condition:

$y_0 = f(t_0)$

We now discuss various ways to solve these.

50.2 First-order Ordinary Differential Equations

50.2.1 Ordinary differential equations

An ordinary differential equation is one with only one independent variable. For example:

$\frac{dy}{dx} = f(x)$

The order of a differential equation is the number of differentials of y included. For example one with the second derivative of y is of order 2.

Ordinary equations can can either implicit or explicit. An explicit function shows the highest order derivative as a function of other terms.

An implicit function is one which is not explicit.

A linear ODE is an explicit ODE where the derivative terms of y do not multiply together, that is, in the form:

$y^{(n)} = \sum_i a_i(x) y^{(i)} + r(x)$

First-order ODEs

We have an evolution:

$\frac{dy}{dt} = f(t, y)$

And a starting condition:

$y_0 = f(t_0)$

We now discuss various ways to solve these.

50.2.2 Linear first-order Ordinary Differential Equations

Linear ODEs

For some we can write:

$\frac{dy}{dt} = f(t, y)$

$\frac{dy}{dt} = q(t) - p(t)y$

This can be solved by multiplying by an unknown function $\mu(t)$:

$\frac{dy}{dt} + p(t)y = q(t)$

$\mu(t)[\frac{dy}{dt} + p(t)y] = \mu(t)q(t)$

We can then set $\mu(t) = e^{\int p(t)dt}$. This means that $\frac{d\mu}{dt} = p(t)u(t)$

$\frac{d}{dt}[\mu(t)y] = \mu(t)q(t)$

$\mu(t)y = \int \mu(t)q(t)dt + C$

In some cases, this can then be solved.

Example

$\frac{\delta y}{\delta x} = cy$

$y = Ae^{c(y+a)}$

$\frac{\delta^2 y}{\delta x^2} = cy$

$y = Ae^{\sqrt{c}(y+a)}$

50.2.3 Separable first-order Ordinary Differential Equations

For some we can write:

$\frac{dy}{dt} = f(t, y)$

$\frac{dy}{dt} = \frac{g(t)}{h(y)}$

We can then do the following:

$h(y)\frac{dy}{dt} = g(t)$

$\int h(y)\frac{dy}{dt}dt = \int g(t)dt + C$

$\int h(y)dy = \int g(t)dt + C$

In some cases, these functions can then be integrated and solved.

50.3 Second-order Ordinary Differential Equations

50.3.1 Linear second-order Ordinary Differential Equations

These are of the form

$\frac{d^2y}{dt^2} + p(t)\frac{dy}{dt} + q(t)y = g(t)$

There are two types. Homogenous equations are where $g(t) = 0$. Otherwise they are heterogenous.

We explore the case with constants:

$a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + cy = 0$

Chapter 51

Univariate optimisation

51.1 Unconstrained optimisation

51.1.1 Introduction to unconstrained optimisation

Goals

We want to identify either the maximum or the minimum.

There exist local minima and global minima.

Optimising through limits

If we are looking to minimise a function, and the limits are ∞ or $-\infty$ then we can optimise by taking large or small values.

We can examine this for each variable.

This also applies for maximising a function.

Optimisation through stationary points

Stationary points of a function are points where marginal changes do not have an impact on the value of the function. As a result they are either local maxima or minima.

Optimisation through algorithms

If we cannot identify stationary points easily, we can instead use algorithms to identify optima.

Stationary points of strictly concave and convex functions

If a function is strictly concave it will only have one stationary point, a local, and global, maxima.

If a function is strictly convex it will only have one stationary point, a local, and global, minima.

51.1.2 Local optima

51.1.3 Optimising convex functions

51.1.4 Analytic optimisation

Convex and concave functions

Convex functions only have one minimum, and concave functions have only one maximum.

If a function is not concave or convex, it may have multiple minima

If a function is convex, then there is only one critical point the local minimum. We can identify this this by looking for critical points using first-order conditions.

Similarly, if a function is concave, then there is only one critical point the local maximum.

We can identify whether a function is concave or convex by evaluating the Hessian matrix.

Evaluating multiple local optima

We can evaluate each of the local minima or maxima, and compare the sizes.

We can identify these by taking partial derivatives of the function in question and identifying where this function is equal to zero.

$u = f(x)$

$u_{x_i} = \frac{\delta f}{\delta x_i} = 0$

We can then solve this bundle of equations to find the stationary values of x.

After identifying the vector x for these points we can then determine whether or not the points are minima or maxima by examining the second derivative at these points. If it is positive it is a local minima, and therefore not an optimal point. Points beyond these will be higher, and may be higher than any local maxima.

51.1.5 Stationary points and first-order conditions

51.1.6 Local minima, maxima and inflection points

51.1.7 Optimising convex and non-convex differentiable functions

51.1.8 Hessian matrix

We can take a function and make a matrix of its second order partial derivatives. This is the Hessian matrix, and it describes the local curvature of the function.

If the function f has n parameters, the Hessian matrix is $n \times n$, and is defined as:

$$H_{ij} = \frac{\delta^2 f}{\delta x_i \delta x_j}$$

If the function is convex, then the Hessian matrix is positive semi-definite for all points, and vice versa.

If the function is concave, then the Hessian matrix is negative semi-definite for all points, and vice versa.

We can diagnose critical points by evaluating the Hessian matrix at those points.

If it is positive definite, it is a local minimum. If it is negative definite it is a local maximum. If there are both positive and negative eigenvalues it is a saddle point.

Part IX

Multivariate real analysis

Chapter 52

Multivariate functions

52.1 Multivariate space

52.1.1 Regions

A region is a subset

Type-I regions (y-simple regions)

Type-II regions (x-simple regions)

Elementary regions

An elementary region is a region which is either a type-I region or a type-II region.

Simple regions

A simple region is a region which is both a type-I and a type-II region.

52.1.2 Curves and closed curves

In a space we can identify a curve between two points. If the input in the real numbers then this curve is unique.

For more general scalar fields this will not be the case. Two points in $\mathbb{R}^2$ could be joined by an infinite number of paths.

A curve can be defined as a function on the real numbers. The curve itself is totally ordered, and homogenous to the real number line.

We can write the curve therefore as:

$r : [a, b] \to C$

Where a and b are the start and end points of the curve, and C is the resulting curve.

Closed curves

If the start and end point of the curve are the same then the curve is closed. We can write this as:

$\oint_C f(r) ds = \int_a^b f(r(t)) |r'(t)| dt$

52.1.3 Surfaces

52.1.4 Length of a curve

We have a curve from a to b in $\mathbf{R}^n$.

$f : [a, b] \to \mathbf{R}^n$

We divide this into n segments.

The ith cut is at:

$t_i = a + \frac{i}{n}(b - a)$

So the first cut is at:

$t_0 = a$

$t_n = b$

The distance between two sequential cuts is:

$||f(t_i) - f(t_{i-1}||$

The sum of all these differences is:

$L = \sum_{i=1}^n ||f(t_i) - f(t_{i-1})||$

The limit is:

$L = \lim_{n \to \infty} \sum_{i=1}^n ||f(t_i) - f(t_{i-1})||$

Method 1

$L = \lim_{n\to\infty} \sum_{i=1}^{n} ||f(t_i) - f(t_{i-1})||$

$L = \lim_{n\to\infty} \sum_{i=1}^{n} ||\frac{f(t_i) - f(t_{i-1})}{\Delta t}||\Delta t$

$L = \lim_{n\to\infty} \sum_{i=1}^{n} ||f'(t)||\Delta t$

$L = \int_a^b ||f'(t)||dt$

Method 2

$L = \lim_{n\to\infty} \sum_{i=1}^{n} ||f(t_i) - f(t_{i-1})||$

$L = \lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{(f(t_i) - f(t_{i-1}))^* M(f(t_i) - f(t_{i-1}))}$

$L = \int_a^b \sqrt{(dt)^T M(dt)}$

Chapter 53

Multivariate differentiation of scalar fields

53.1 Partial differentiation of scalar fields

53.1.1 Scalar fields

A scalar field is a function on an underlying input which produces a real output.

Inputs are not limited to real numbers. In this section we consider functions on vector spaces.

53.1.2 Del

$\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

Where e are the basis vectors.

This on its own means nothing. It is similar to the partial differentiation function.

53.1.3 Gradient

In a scalar field we can calculate the partial derivative at any point with respect to one input.

We may wish to consider these collectively. To do that we use the gradient operator.

We previously introduced the Del operator where:

$\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

Where e are the basis vectors.

This on its own means nothing. It is similar to the partial differentiation function.

We now multiply Del by the function. This gives us:

$\nabla f = (\sum_{i=1}^n e_i \frac{\delta f}{\delta x_i})$. This gives us a vector in the underlying vector space.

This is the gradient.

53.2 Directional derivative of scalar fields

53.2.1 Directional derivative

We have a function, $f(\mathbf{x})$.

Given a vector v, we can identify by how much this scalar function changes as you move in that direction.

$\nabla_v f(x) := \lim_{\delta \to 0} \frac{f(\mathbf{x} + \delta \mathbf{v}) - f(\mathbf{x})}{\delta}$

The directional derivative is the same dimension as underlying field.

Other

Differentiation of scalar field, df, can be defined as a vector field where grad is 0. can differ with orientation, scale

53.3 Total differentiation of scalar fields

53.3.1 Total differentiation

Consider a multivariate function.

$f(x)$.

We can define:

$\Delta f(x, \Delta x) := f(x + \Delta x) - f(x)$

$\Delta f(x, \Delta x) = \sum_{i=1}^n f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)$

$\Delta f(x, \Delta x) = \sum_{i=1}^{n} \Delta x_i \frac{f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)}{\Delta x_i}$

$\frac{\Delta f}{\Delta x_k} = \sum_{i=1}^{n} \frac{\Delta x_i}{\Delta x_k} \frac{f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)}{\Delta x_i}$

$\lim_{\Delta x_k \to 0} \frac{\Delta f}{\Delta x_k} = \sum_{i=1}^{n} \lim_{\Delta x_k \to 0} \frac{\Delta x_i}{\Delta x_k} \frac{f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)}{\Delta x_i}$

$\frac{df}{dx_k} = \sum_{i=1}^{n} \frac{dx_i}{dx_k} \frac{\delta f}{\delta x_i}$

53.3.2 Total differentiation of a univariate function

For a univariate function total differentiation is the same as partial differentiation.

$\frac{df}{dx} = \frac{dx}{dx} \frac{\delta f}{\delta x}$

$\frac{df}{dx} = \frac{\delta f}{\delta x}$

Chapter 54

Multivariate integration of scalar fields

54.1 Integration of scalar fields

54.1.1 Line integral of scalar fields

54.1.2 Double integral of scalar fields

54.1.3 Surface integral of scalar fields

54.1.4 Gradient theorem

In a scalar field, the line integral of the gradient field is the difference between the value of the scalar field at the start and end points.

This generalises the fundamental theorem of calclulus.

54.1.5 Green's theorem

We have a curve C on a plane.

Inside this is region D.

We have two functions: $L(x,y)$ and $M(x,y)$ defined on the region and curve.

$\oint_C (Ldx + Mdy) = \int\int_D (\frac{\delta M}{\delta x} - \frac{\delta L}{\delta y})dxdy$

54.1.6 Differential forms

Type-I

For type-I, we can integrate over y, then integrate over x.

Type-II

For type-II, we can integrate over x, then integrate over y.

Chapter 55

Multivariate differentiation of vector fields

55.1 Partial differentiation of vector fields

55.1.1 Jacobian matrix

If we have n inputs and m functions such that:

$f_i(\mathbf{x})$

The Jacobian is a matrix where:

$J_{ij} = \frac{\delta f_i}{\delta x_j}$

55.2 Scalar potential

55.2.1 Scalar potential

Given a vector field $\mathbf{F}$ we may be able to identify a scalar field P such that:

$\mathbf{F} = -\nabla P$

55.2.2 Non-uniqueness of scalar potentials

Scalar potentials are not unique.

If P is a scalar potential of $\mathbf{F}$, then so is $P + c$, where c is a constant.

55.2.3 Conservative vector fields

Not all vector fields have scalar potentials. Those that do are conservative.

For example if a vector field is the gradient of a scalar height function, then the height is a scalar potential.

If a vector field is the rotation of water, there will not be a scalar potential.

55.3 Divergence

55.3.1 Divergence

This takes a vector field and produces a scalar field.

It is the dot product of the vector field with the del operator.

$divF = \nabla.F$

Where $\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

$divF = \sum_{i=1}^n e_i \frac{\delta F_i}{\delta x_i}$

55.3.2 Divergence as net flow

Divergence can be thought of as the net flow into a point.

For example, if we have a body of water, and a vector field as the velocity at any given point, then the divergence is 0 at all points.

This is because water is incompressible, and so there can be no net flows.

Areas which flow out are sources, while areas that flow inwards are sinks.

55.3.3 Solenoidal vector fields

If there is no divergence, then the vector field is called solenoidal.

55.3.4 The Laplace operator

Cross product of divergence with the gradient of the function.

$\Delta f = \nabla.\nabla f$

$\Delta f = \sum_{i=1}^n \frac{\delta^2 f}{\delta x_i^2}$

55.4 Curl

55.4.1 Curl

The curl of a vector field is defined as:

$curl\mathbf{F} = \nabla \times \mathbf{F}$

Where: $\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

And: $\mathbf{x} \times \mathbf{y} = |||\mathbf{x}||||\mathbf{y}|| \sin(\theta)\mathbf{n}$

The curl of a vector field is another vector field.

The curl measures the rotation about a given point. For example if a vector field is the gradient of a height map, the curl is 0 at all points, however for a rotating body of water the curl reflects the rotation at a given point.

55.4.2 Divergence of the curl

If we have a vector field $\mathbf{F}$, the divergence of its curl is 0:

$\nabla.(\nabla \times \mathbf{F}) = 0$

55.4.3 Vector potential

Given a vector field $\mathbf{F}$ we may be able to identify another vector field A such that:

$\mathbf{F} = \nabla \times \mathbf{A}$

Existence:

We know that the divergence of the curl for any vector field is 0, so this applies to A:

$\nabla.(\nabla \times \mathbf{A}) = 0$

Therefore:

$\nabla.\mathbf{F} = 0$

This means that if there is a vector potential of $\mathbf{F}$, then $\mathbf{F}$ has no divergence.

55.4.4 Non-uniqueness of vector potentials

Vector potentials are not unique.

If $\mathbf{A}$ is a vector potential of $\mathbf{F}$, then so is $\mathbf{A} + \nabla c$, where c is a scalar field and ∇c is its gradient.

55.4.5 Conservative vector fields

Not all vector fields have scalar potentials. Those that do are conservative.

For example if a vector field is the gradient of a scalar height function, then the height is a scalar potential.

If a vector field is the rotation of water, there will not be a scalar potential.

55.4.6 Hodge stars

The Hodge star operator is a generalisation of cross product. In 3d space if we have a plane, we can get a vector perpendicular and visa versa. Generally, we are in n-dimensional space and we input k vectors and get out $n - k$ vectors.

55.4.7 Hodge duals

Chapter 56

Multivariate integration of vector fields

56.1 Integration of vector fields

56.1.1 Line integral of vector fields

We may wish to integrate along a curve in a vector field.

We previously showed that we can write a curve as a function on the real line:

$r : [a, b] \rightarrow C$

The integral is therefore the sum of the function at all points, with some weighting. We write this:

$\int_C f(r)ds = \lim_{\Delta srightarrow0} \sum_{i=0}^{n} f(r(t_i))\Delta s_i$

In a vector field we use

$\int_C f(r)ds = \int_a^b f(r(t)).r'(t)dt$

56.1.2 Double integral of vector fields

56.1.3 Surface integral for vector fields

56.2 Stoke's theorem

56.2.1 The divergence theorem

56.2.2 Stoke's theorem

Chapter 57

Partial Differential Equations (PDEs)

57.1 Introduction

Chapter 58

Exterior calculus

58.1 Introduction

58.1.1 Differential forms

58.1.2 Exterior derivative

58.1.3 The fundamental theorem of external calculus

Chapter 59

Calculus of variations

59.1 sort

59.1.1 Frchet derivative

59.1.2 Gteaux derivative

59.1.3 Euler-Lagrange equations

Chapter 60

Multivariate optimisation

60.1 Unconstrained multivariate optimisation

60.1.1 Introduction

60.2 Optimisation with linear equality constraints

60.2.1 Single equality constraint

Constrained optimisation

Rather than maximise $f(x)$, we want to maximise $f(x)$ subject to $g(x) = 0$.

We write this, the Lagrangian, as:

$\mathcal{L}(x, \lambda) = f(x) - \sum_k^m \lambda_k [g_k(x) - c_k]$

We examine the stationary points for both vector x and λ. By including the latter we ensure that these points are consistent with the constraints.

Solving the Langrangian with one constraint

Our function is:

$\mathcal{L}(x, \lambda) = f(x) - \lambda[g(x) - c]$

The first-order conditions are:

$\mathcal{L}_\lambda = -[g(x) - c]$

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \lambda \frac{\delta g}{\delta x_i}$

The solution is stationary so:

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \lambda \frac{\delta g}{\delta x_i} = 0$

$\lambda \frac{\delta g}{\delta x_i} = \frac{\delta f}{\delta x_i}$

$\lambda = \frac{\frac{\delta f}{\delta x_i}}{\frac{\delta g}{\delta x_i}}$

Finally, we can use the following in practical applications.

$\frac{\frac{\delta f}{\delta x_i}}{\frac{\delta g}{\delta x_i}} = \frac{\frac{\delta f}{\delta x_j}}{\frac{\delta g}{\delta x_j}}$

60.2.2 Multiple equality constraints

Solving the Langrangian with many constraints

This time we have:

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_i} = 0$

$\mathcal{L}_{x_j} = \frac{\delta f}{\delta x_j} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_j} = 0$

$\frac{\delta f}{\delta x_i} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_i} = \frac{\delta f}{\delta x_j} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_j}$

60.3 Linear programming

60.3.1 Inequality constraints

linear programming means of the form max $c^T x$ st. $Ax <= b$ $x >= 0$ this is the canonical form

Lagrangians with inequality constraints

We can add constraints to an optimisation problem. These constraints can be equality constraints or inequality constraints. We can write constrained optimisation problem as:

Minimise $f(x)$ subject to

$g_i(x) \leq 0$ for $i = 1, ..., m$

$h_i(x) = 0$ for $i = 1, , p$

We write the Lagrangian as:

$\mathcal{L}(x, \lambda, \nu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$

If we try and solve this like a standard Lagrangian, then all of the inequality constraints will instead by equality constraints.

Affinity of the Lagrangian

The Lagrangian function is affine with respect to λ and ν.

$\mathcal{L}(x, \lambda, \nu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$

$\mathcal{L}_{\lambda_i}(x, \lambda, \nu) = g_i(x)$

$\mathcal{L}_{\nu_i}(x, \lambda, \nu) = h_i(x)$

As the partial differential is constant, the partial differential is an affine function.

60.3.2 Primal and dual problems

The primal problem

We already have this.

The dual problem

We can define the Lagrangian dual function:

$g(\lambda, \nu) = \inf_{x \in X} \mathcal{L}(x, \lambda, \nu)$

That is, we have a function which chooses the returns the value of the optimised Lagrangian, given the values of λ and ν.

This is an unconstrained function.

We can prove this function is concave (how?).

The infimum of a set of concave (and therefore also affine) functions is concave.

The supremum of a set of convex (and therefore also affine) functions is convex.

Given a function with inputs x, what values of x maximise the function?

We explore constrained and unconstrained optimisation. The former is where restrictions are placed on vector x, such as a budget constraint in economics.

The dual problem is concave

The duality gap

We refer to the optimal solution for the primary problem as p^*, and the optimal solution for the dual problem as d^*.

The duality gap is $p^* - d^*$.

60.3.3 Complementary slackness for linear optimisation

60.3.4 Farkas' lemma

We have matrix A and vector b.

Either:

- $Ax = b$; $x \geq 0$
- $A^T y \geq 0$; $b^T y < 0$

60.4 Quadratic optimisation

60.4.1 The quadratic optimisation problem

60.5 Constrainted non-linear optimisation

60.5.1 Weak duality theorem

The duality gap ($p^* - d^*$ is non-negative.

60.5.2 Lagrange multipliers

60.5.3 The dual problem for non-linear optimisation

60.5.4 The weak duality theorem

60.6 Constrained convex optimisation

60.6.1 Slater's condition

Strong duality

Strong duality is where the duality gap is 0.

Slater's condition

Slater's condition says that strong duality holds if there is an input where the inequality constraints are satisified strictly.

That is they are $g(x) < 0$, not $g(x) \leq 0$

This means that the conditions are slack.

This only applies if the problem is convex. That is, if Slater's condition holds, and the problem is convex, then strong duality holds.

60.6.2 The strong duality theorem

60.6.3 Karush-Kuhn-Tucker conditions

If our problem is non-convex, or if Slater's condition does not hold, how else can be find a solution?

A solution, p^* can satisify KKT conditions.

60.7 Sort

60.7.1 Unconstrained envelope theorem

Consider a function which takes two parameters:

$f(x, \alpha$

We want to choose x to maximise f, given α.

$V(\alpha) = \sup_{x \in X} f(x, \alpha)$

There is a subset of X where $f(x, \alpha) = V(\alpha)$.

$X^*(\alpha) = \{x \in X | f(x, \alpha) = V(\alpha)\}$

This means that $V(\alpha) = f(x^*, \alpha)$ for $x^* \in X^*$.

Lets assume that there is only one x^*.

$V(\alpha) = f(x^*, \alpha)$

What happens to the value function as we relax α?

$V_{\alpha_i}(\alpha) = f_{\alpha_i}(x^*(\alpha), \alpha)$.

$V_{\alpha_i}(\alpha) = f_x \dfrac{\delta x^*}{\delta \alpha} + f_{\alpha_i}$.

We know that $f_x = 0$ from first order conditions. So:

$V_{\alpha_i}(\alpha) = f_{\alpha_i}$.

That is, at the optimum, as the constant is relaxed, we can treat the x^* as fixed, as the first-order movement is 0.

60.7.2 Identifying upper and lower bounds of linear programming

In min/max problem, any feasibly solution is an upper/lower bound.

can we get a bound at the other side? yes, by doing linear combinations of inequalities eg maximise $30x + 100y$ subject to: $4x + 10y <= 40$ $x >= 3$

We can identify a lower bound by inputting something which works, for example $x = 3$ and $y = 0$. This gives us a lower bound of 90.

To get an upper bound we can manipulate the constraints: $40x + 100y <= 400$ $10x >= 30$ And then: $40x + 100y <= 370 + 30$ $40x + 100y <= 370 + 10x$ $30x + 100y <= 370$

So we have an upper bound of 370.

This lower bound is a result of doing linear combinations of the inequalities. For different combinations, we could have a lower lower bound.

This is the dual problem. How do we choose the linear combination of inequalities such that the resulting lower bound is minimised?

Part X

Complex analysis

Chapter 61

Complex calculus

61.0.1 Complex-valued functions

61.0.2 Defining complex valued functions

We can consider complex valued functions as a type of vector fields.

61.0.3 Line integral of the complex plane

$\int_C f(r)ds = \lim_{\Delta s rightarrow 0} \sum_{i=0}^{n} f(r(t_i))\Delta s_i$

$\int_C f(r)ds = \lim_{\Delta s rightarrow 0} \sum_{i=0}^{n} f(r(t_i))\frac{\delta r(t_i)}{\delta t}\delta r_i$

$\int_C f(z)dz = \int_a^b f(r(t_i))\frac{\delta r(t_i)}{\delta t}\delta r_i$

61.0.4 Complex continuous functions

61.0.5 Open regions

61.0.6 Analytic continuation

61.0.7 Analytic functions

61.0.8 Circle of convergence

61.0.9 Complex differentiation

61.0.10 Wirtinger derivatives

Previously we had partial differentiation on the real line. We could use the partial differention operator

We want to find a similar operator for the complex plane.

61.0.11 Line integral of the complex plane

$\int_C f(r)ds = \lim_{\Delta s rightarrow 0} \sum_{i=0}^n f(r(t_i))\Delta s_i$

$\int_C f(r)ds = \lim_{\Delta s rightarrow 0} \sum_{i=0}^n f(r(t_i))\frac{\delta r(t_i)}{\delta t}\delta r_i$

$\int_C f(z)dz = \int_a^b f(r(t_i))\frac{\delta r(t_i)}{\delta t}\delta r_i$

61.0.12 Complex integration

61.0.13 Complex smooth functions

If a function is complex differentiable, it is smooth.

61.0.14 All differentiable complex functions are smooth

61.0.15 All smooth complex functions are analytic

61.0.16 Singularities

61.0.17 Contour integration

61.0.18 Line integral

61.0.19 Cauchy's integral theorem

61.0.20 Cauchy's integral formula

61.0.21 Cauchy-Riemann equations

Consider complex number z=x+iy

A function on this gives:

$f(z) = u + iv$

Take the total differential of :

$df/dz = \frac{\delta f}{\delta z} + \frac{\delta f}{\delta x}\frac{dx}{dz} + \frac{\delta f}{\delta y}\frac{dy}{dz}$

We know that:

- $\frac{dx}{dz} = 1$
- $\frac{dy}{dz} = -i$

We can see from this that

- $\frac{du}{dx} = \frac{dv}{dy}$
- $\frac{du}{dy} = -\frac{dv}{dx}$

These are the Cauchy-Riemann equations

Chapter 62

Riemann surfaces

62.1 Simply connected Riemann surfaces

62.1.1 The Riemann sphere (elliptic)

62.1.2 The complex plane (parabolic)

62.1.3 The opendisk (hyperbolic)

62.2 Other Riemann surfaces

62.2.1 The torus

62.2.2 The hyperelliptic curve

Part XI

Analytic geometry

Chapter 63

Points, lines and affine transformations

63.1 Affine spaces

63.1.1 Lines

63.1.2 Parallel lines

Chapter 64

Euclidian transformations, lengths and angles

64.1 Linear metrics

64.1.1 Metrics

We defined a norm as:

$||v|| = v^T M v$

A metric is the distance between two vectors.

$d(u, v) = ||u - v|| = (u - v)^T M(u - v)$

Metric space

A set with a metric is a metric space.

64.1.2 Inducing a topology

Metric spaces can be used to induce a topology.

64.1.3 Translation symmetry

The distance between two vectors is:

$(v - w)^T M(v - w)$

So what operations can we do now?

As before, we can do the transformations which preserve u^TMv, such as the orthogonal group.

But we can also do other translations

$(v-w)TM(v-w)$

$v^TMv+w^TMw-v^TMw-w^TMv$

so symmetry is now $O(3,1)$ and affine translations

Translation matrix

$[[1,x][0,1]]$ moves vector by x.

64.2 Specific groups

64.2.1 The affine group

64.2.2 The Euclidian group

64.2.3 The Galilean group

64.2.4 The Poincar group

64.3 Non-linear norms

64.3.1 L_p norms (p-norms)

L^P norm

This generalises the Euclidian norm.

$||x||_p=(\sum_{i=1}^n|x|_i^p)^{1/p}$

This can defined for different values of p. Note that the absolute value of each element in the vector is used.

Note also that:

$||x||_2$

Is the Euclidian norm.

Taxicab norm

This is the L^1 norm. That is:

$||x||_1 = \sum_{i=1}^n |x|_i$

Angles

Cauchy-Schwarz

64.4 To linear forms

64.4.1 Norms

We can use norms to denote the "length" of a single vector.

$||v|| = \sqrt{\langle v, v\rangle}$

$||v|| = \sqrt{v^*Mv}$

Euclidian norm

If $M = I$ we have the Euclidian norm.

$||v|| = \sqrt{v^*v}$

If we are using the real field this is:

$||v|| = \sqrt{\sum_{i=1}^n v_i^2}$

Pythagoras' theorem

If $n = 2$ we have in the real field we have:

$||v|| = \sqrt{v_1^2 + v_2^2}$

We call the two inputs x and y, and the length z.

$z = \sqrt{x^2 + y^2}$

$z^2 = x^2 + y^2$

64.4.2 Angles

Recap: Cauchy-Schwarz inequality

This states that:

$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \dot{\langle} v, v\rangle$

Or:

$\langle v, u\rangle\langle u, v\rangle \leq \langle u, u\rangle \dot{\langle} v, v\rangle$

Introduction

$\langle v, u\rangle\langle u, v\rangle \leq \langle u, u\rangle \dot{\langle} v, v\rangle$

$\frac{\langle v, u\rangle\langle u, v\rangle}{||u||.||v||} \leq ||u||.||v||$

$\frac{||u||.||v||}{\langle v, u\rangle} \geq \frac{\langle u, v\rangle}{||u||.||v||}$

$\cos(\theta) = \frac{\langle u, v\rangle}{||u||.||v||}$

Chapter 65

Volumes, perimeters and surface areas

Chapter 66

2D polygons

66.1 Elementary geometry in 2 dimensions

66.1.1 Triangles

Area of a triangle

Circumference of a triangle

Sum of angles of a triangle

Angles in a triangle add to π.

66.1.2 Quadrilaterals

66.1.3 Oblongs

Area of an oblong

Circumference of an oblong

66.1.4 Squares

Area of a square

$A = r^2$

Circumference of a square

$C = 2r$

Angles in a square

Angles in a square sum to 2π.

66.1.5 Pentagon

Chapter 67

3D polygons

67.1 Elementary geometry in 3 dimensions

67.1.1 Pyramid

67.1.2 Cubes

Volume of a cube:

$V = r^3$

Surface area of a cube:

$A = 6r^2$

Chapter 68

Algebraic geometry and spheres

68.1 Circles

68.1.1 Defining circles

$x^2 + y^2 = r^2$

68.1.2 Area of a circle

$A = \pi r^2$

68.1.3 Circumference of a circle

$C = 2\pi r$

68.2 Spheres

68.2.1 Defining spheres

$x^2 + y^2 + z^2 = r^2$

68.2.2 Volume of a sphere

$V =$

68.2.3 Surface area of a sphere

$A =$

Part XII

Abstract algebra

Chapter 69

Group theory

69.1 Introduction

69.1.1 Abstract algebra

Abstract algebra allows us to discuss properties of types of mathematical structures.

Rather than construct a specific object, and explore its properties, we can explore the properties of an abstract structure with certain definitions. We can then apply findings from this to an any structure which meets the definition.

Examples of abstract algebra

We explore:

- Groups
- Rings
- Fields
- Vector spaces
- Inner product spaces

69.1.2 Defining groups

Magma

A magma, or groupoid, is a set with a single binary operation.

These can be defined as an ordered pair $(s, \odot)$ where s is the set, and $\odot$ is the binary operation.

If a and b are in s, then $a \odot b$ is also in s.

The following are magmas:

- Natural numbers and addition
- $n \times n$ matrices with determinants other than 0
- Natural numbers above 0 and addition
- Integers and addition
- Rational numbers and division
- $\{-1, 1\}$ and multiplication

The following are not magmas:

- Natural numbers up to 10 and addition

Semigroup

A semigroup is a magma whose binary operation is associative.

The following are semigroups:

- Natural numbers and addition
- $n \times n$ matrices with determinants other than 0
- Natural numbers above 0 and addition
- Integers and addition

The following are not semigroups:

- $\{-1, 1\}$ and multiplication
- Rational numbers and division
- Natural numbers up to 10 and addition

Monoid

A monoid is a semigroup with an identity element

The following are monoids:

- Natural numbers and addition
- $n \times n$ matrices with determinants other than 0
- Integers and addition

- $\{-1, 1\}$ and multiplication

The following are not monoids:

- Natural numbers above 0 and addition
- Rational numbers and division
- Natural numbers up to 10 and addition

Group

A group is a monoid where there is an inverse operation for the binary operation.

The following are groups:

- Integers and addition
- $n \times n$ matrices with determinants other than 0
- $\{-1, 1\}$ and multiplication

The following are not groups:

- Natural numbers above 0 and addition
- Rational numbers and division
- Natural numbers and addition
- Natural numbers up to 10 and addition

69.1.3 Subgroups

A subgroup of a group is a subset of a group, which also forms a group with the same element.

For example all even numbers are a subgroup of the addition group of integers.

69.1.4 Abelian groups

A commutative group, that is where $a \odot b = b \odot a$.

The following are abelian groups:

- Integers and addition
- $\{-1, 1\}$ and multiplication

The following are not abelian groups:

- Natural numbers above 0 and addition

- Rational numbers and division
- Natural numbers and addition
- Natural numbers up to 10 and addition
- $n \times n$ matrices with determinants other than 0

69.1.5 Group order

For finite groups, each element e has:

$e^n = I$

For some $n \in \mathbb{N}$

Where I is the identity element.

The order of the group is the smallest value of n such that that holds for all elements.

For example in the multiplicative group $G = \{-1, 1\}$ the order is 2.

Or:

$|G| = 2$

Additionally

$|-1| = 2$

$|1| = 1$

69.2 Creating groups

69.2.1 Permutations and the symmetric group

A permutation is defined as a bijection from a set to itself.

For a set of size n, the number of permutations is $n!$. This is because there are n possibilities for the first item, $n-1$ for the second and so on.

The symmetric group

The set of all permutations forms a group, the symmetric group. This forms a group because:

- There is an identity element
- Each combination of permutations is also in the group.

- Each permutation has an inverse in the group.

Permutation groups

A subgroup of the symmetric group is called a permutation group.

69.2.2 Morphism

Morphisms are functions which preserve the relationships between members of a set, and specified functions.

That is, if:

$a \odot b = c$

Then $f(x)$ is morphism if:

$f(a) \odot f(b) = f(a \odot b)$

Here we discuss morphisms in the context of groups, but we can define morphisms for sets with more than one function, for example with addition and multiplication.

Morphisms are also known as homomorphisms.

The following are morphisms of the additive group of integers.

Where we refer to c, $c \neq 0 \in \mathbb{I}$.

- $f(x) = 0$
- $f(x) = x$
- $f(x) = cx$
- Converting natural numbers to integers

The following are not morphisms

- $f(x) = x + 1$

Isomorphism

An isomorphism is a morphism which has an inverse.

This means the function is bijective.

The following are isomorphisms:

- $f(x) = x$
- $f(x) = cx$

- Converting natural numbers to integers

The following are not isomorphisms

- $f(x) = 0$
- $f(x) = x + 1$

Endomorphism

An endomorphism is one where the domain and codomain are the same.

The following are endomorphisms:

- $f(x) = 0$
- $f(x) = x$
- $f(x) = cx$

The following are not endomorphisms

- Converting natural numbers to integers
- $f(x) = x + 1$

Automorphism

An endomorphism which is also an isomorphism

The following are automorphisms:

- $f(x) = x$
- $f(x) = cx$

The following are not automorphisms

- $f(x) = 0$
- $f(x) = x + 1$
- Converting natural numbers to integers

Monomorphism

A morphism which is injective. That is:

$f(a) = f(b) \to a = b$

The following are monomorphisms:

- $f(x) = x$

- $f(x) = cx$
- Converting natural numbers to integers

The following are not monomorphisms:

- $f(x) = 0$
- $f(x) = x + 1$

69.2.3 Generating sets

We can define a group through a generating set and an operation.

And define the group as $G =< S >$

69.2.4 Finite groups

Consider the set of natural numbers and addition modulo 4. This forms a group containing:

$\{0, 1, 2, 3\}$

This can be written as Z_4 or more generally as Z_n, or Z/nZ.

69.3 Group operations

69.3.1 The group commutator

The group commutator is:

$[a, b] = a^{-1}b^{-1}ab$

If the group is abelian then $[a, b] = 0$. The group commutator is a measure of how non-abelian the group is.

This has the following properties:

- Alternativity: $[A, A] = I$

69.3.2 The direct product of groups

If we have two groups G and H we can form new group $G \times H$.

For every $g \in G$ and $h \in H$ there is $(g, h) \in G \times H$.

The binary operation we have is:

$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1, h_2)$

69.4 Specifc groups

69.4.1 The trivial group

The trivial group is the group with just the identity member I.

69.4.2 The infinite cyclic group (Z)

The additive group of integers

Generating cyclic groups

We can generate a group with a single element, it is a cyclic group.

For example, we can define a group $G =< 1 >$ which gives us the additive group of integers.

Infinite cyclic groups are isomorphic to the additive group of integers

More generally, any infinite cyclic group is isomorphic to the additive group of integers.

Consider the multiplicative group of $< i >$.

This contains $\{1, -1, i, -i\}$.

This is also automorphic to the natural number and modulo addition group above.

We can define finite cyclic groups of size n using the generating element $z^{\frac{1}{n}}$. This is isomorphic to the general cyclic group C_n, and to Z/nZ.

Abelian cyclic groups

Cyclic groups are abelian.

69.4.3 The finite cyclic groups (C_n or Z_n)

69.4.4 The circle group T

The circle group, T, includes all complex numbers of magnitude 1.

69.5 Normal subgroups

69.5.1 Cosets and normal subgroups

A coset is defined between a group and a subgroup of the group.

For a group G, and its subgroup H:

- The left coset is $\{gH\}$
- The right coset is $\{Hg\}$

For $\forall g \in G$.

For abelian groups, the left and right cosets are the same.

The left and right cosets can also be the same, even if the group G is not abelian.

Normal subgroups

If the left and right cosets are the same then H is a normal subgroup.

Cosets divide a group.

Consider two left cosets, aH and bH, with a common element.

This means that $ah_i = bh_j$.

We can use this to get:

$a = bh_j h_i^{-1}$

$b = ah_i h_j^{-1}$

We know that:

$ah \in aH$

$bh \in bH$

So:

$bh_j h_i^{-1} h \in aH$

$ah_ih_j^{-1}h \in bH$

And so:

$bH \subset aH$

$aH \subset bH$

Therefore:

$aH = bH$

Example 1

Consider the group $\{-1, 1\}, \times$

For the subgroup $\{1\}, \times$, the left coset is $\{gH\} = \{1, -1\}$.

The right coset is the same.

Example 2

Consider the group of integers and addition: $(Z, +)$

For subgroup $(mZ, +)$, the left and right cosets are the same because the group is abelian.

The coset of the subgroup is the subgroup multiplied by each element in G.

This is mZ, $mZ + 1$, $mZ + 2$ and so on.

Once we reach $mZ + m$ this has looped, and is already a coset, so we only need the sets upto $mZ + m - 1$.

69.5.2 Quotient groups

We have a group G and a normal subgroup N.

We define a quotient group from this as G/N. This is the set of cosets from N.

69.5.3 Group extension

This defines a group G from a normal subgroup N and a quotient group Q.

69.6 Theorems

69.6.1 Cayley's theorem

Cayley's theorem states that every group G is isomorphic to a subgroup of the symmetric group acting on G.

Multiplication by a member of G is a bijective function, as for each g there is also a g^{-1}.

This means that multiplication of each member of G is a permutation, and so is a subset of the symmetric group on G.

69.6.2 Lagrange's theorem

Lagrange's theorem states that for any finite group G, the order of every subgroup is a divisor of the order of G.

Consider subset H. We know that all cosets are disjoint, and that the union of all cosets is G.

As cosets are the same size, we know that:

$|G| = m|H|$, where m is the number of cosets.

This means that if a group has order 10, a subgroup must have order 1, 2 5 or 10.

69.7 Group action

69.7.1 Group action

We have a group G and a set S.

We have a function $g.s$ which maps onto S such that:

- $I.s = s$
- $(gh).s = g(h.s)$

Chapter 70

Rings

70.1 Introduction

70.1.1 Rings

Consider an abelian group $(S, +)$.

A ring takes this and adds a multiplicative function which satisfies the distributive property.

Groups have an identity element for their function. Rings must have identity elements for both their functions.

The multiplicative function does not have to be a bijection. For example the set of integers, addition and multiplication form a ring.

70.1.2 Rngs

A rng is a ring without the multiplicative identity (hense no 'i').

70.2 Commutation

70.2.1 Commutative rings

The multiplication operation commutes.

70.2.2 Commutator

$[a, b] = ab - ba$

70.2.3 The Jacobi identity

70.3 Examples of rings

70.3.1 Zero (trivial) ring

The trivial ring is a ring with just one element.

0 with addition and multiplication work.

70.3.2 Integer rings

The integers with addition and multiplication form a ring.

70.3.3 Integer mod n rings

The integers mod n with addition and multiplication form a group.

Examples

The integers $\{1, 2, 3\}$ form a ring.

70.4 Properties of rings

70.4.1 Characteristic of a ring

The characteristic of a ring is the number of times the multiplicative identity must be added to get the additive identity.

If this never happens, the characteristic is 0.

Example

The integer mod 2 ring, the characteristic is 2.

70.5 Division

70.5.1 Division rings

A divison ring is a ring where every non-zero element has a multiplicative inverse.

Example

The rational numbers are a divison ring.

Relationship between divison rings and fields

Fields (not yet introduced) are different from division rings only in that multiplication for a field must be commutative.

70.5.2 Units

A unit is an element of a ring which has a multiplicative inverse.

Examples

The ring of integers with addition and multiplication, only -1 and 1 are units, as both have multiplicative inverses in the ring.

70.6 Subrings

70.6.1 Subrings

A subring is a subset of the ring, where the addition and multiplication operations on the subring result in elements also in the subring.

Example

The even numbers are a subring of the integers.

70.6.2 Ideals

An ideal is a subring where the multiplication of any element of the ideal with any element of the ring is also in the ideal.

Examples

Even numbers are an ideal of the integers.

Odd numbers are not an ideal. For example 1 is in the ideal, but multiplied by 2 gives 2, which is not in the ideal.

Chapter 71

Fields

71.1 Fields

71.1.1 Fields

A field is a ring where the multiplication function has an inverse.

The integers, addition and multiplication form a ring, but not a group.

The rational numbers (except 0), addition and multiplication form a field (and a ring).

The real numbers and complex numbers also form fields.

71.1.2 Finite (Galois) fields

Finite number of elements.

Integers mod p field

Characteristic of a field

71.2 Algebra on a field

71.2.1 Bilinear maps

A bilinear map (or function) is a map from two inputs to an output which preserves addition and scalar multiplication. This is in contrast to a linear map, which only has one input.

In addition, the function is linear in both arguments.

That is if function f is bilinear then:

$X = aM + bN$

$Y = cO + dP$

$f(X,Y) = f(aM + bN, cO + dP)$

$f(X,Y) = f(aM, cO + dP) + f(bN, cO + dP)$

$f(X,Y) = f(aM, cO) + f(aM, dP) + f(bN, cO) + f(bN, dP)$

$f(X,Y) = acf(M,O) + adf(M,P) + bcf(N,O) + bdf(N,P)$

Note that:

$f(X,Y) = f(X + 0, Y)$

$f(X,Y) = f(X,Y) + f(0,Y)$

$(0,Y) = 0$

That is, if any input is 0 in an additative sense, the value of the map must be zero.

71.2.2 Algebra on a field

Part XIII

Abstract linear algebra

Chapter 72

Vector spaces

72.1 Vector spaces

72.1.1 Vector spaces

A vector space is a group with additional structure.

The operation for each element is shown as addition. So we can say:

$\forall u, v \in V[u + v \in V]$

To this we add scalars, from a field F. We write this as multiplication.

$\forall f \in F \forall v \in V[fv \in V]$

Subspace

A subspace is a subset of V which still acts as a vector space. In practice, this means fewer dimensions.

72.1.2 Span

Span function

We can take a subset S of V. We can then make linear combinations of these elements.

This is called the linear span - $span(S)$.

72.1.3 Linear dependence

A collection of vectors in a vector space are linearly dependent if there exist values for α (other than all being 0) such that:

$\sum_i \alpha_i v_i = 0$.

If no such values for α exist we say the vectors are linearly independent.

72.1.4 Basis vectors

Basis

We can write vectors as combinations of other vectors.

$v = \sum_i \alpha_i v_i$

A subset which spans the vector space, and which is also linearly independent, is a basis of the vector space.

For an arbitrary vector of size n, we cannot use less than n elementary vectors. We could use more, but these would be redundant.

If we use n elementary vectors, there is a unique solution of weights of elementary vectors.

If we use more than n elementary vectors, there will be linear dependence, and so there will not be a unique solution.

72.1.5 Dimension function

For a basis S, the the dimension of the vector space is $|S|$.

$\dim(V) = |S|$

$S \subset V$

Finite and infinite vector spaces

If $\dim(V)$ is finite, then we say the vector space is finite.

Otherwise, we say the vector space is infinite.

72.2 Points, lines and planes

72.2.1 Points, lines and planes

$(1,0)$ is point, $(x,2x+1)$ is a line $(1,x,y)$ is a plane

72.2.2 Parallel lines and planes

Parallel lines

If we have two lines:

Parallel planes

Chapter 73

Linear endomorphisms

73.1 Endomorphisms of vector spaces

73.1.1 Endomorphisms

An endomorphism maps a vector space onto itself.

$end(V) = \hom(V, V)$

73.1.2 Endomorphisms form a vector space

An endomorphism maps a vector space onto itself.

$end(V) = \hom(V, V)$

Need to show that endomorphism is a vector space

Essentially

$v \in V$

$f \in F$

$av = f$

$bv = g$

$(a \oplus b)v = f + g$

$(a \oplus b)v = av + bv$

so there is some operation we can do on two members of endo linear in addition. That is, if we have two dual "things", we can define the addition of functions as the operation which results int he outputs being added.

what about linear in scalar? same approach.

Well we define

$c \odot a) = cav$

There is a unique endomorphism which results in two other endomorphisms being added together. define this as addition

73.1.3 Dimension of endomorphisms

$\dim(end(V)) = (\dim V)^2$

73.1.4 Basis of endomorphisms

73.1.5 Projections

A projection is a linear map which if applied again returns the original result.

A projection can drop a dimension for example.

73.1.6 Kernels and images

The kernel of a linear operator is the set of vectors such that:

$Mv = 0$

The kernel is also called the nullspace.

This can be shown as $\ker(M)$

The image of a linear operator is the set of vectors w such that:

$Mv = w$.

This can be shown as $\Im(M)$

We also know that:

$span(M) = \ker(M) + \Im(M)$

73.2 Representing endomorphisms with matrices

73.2.1 Matrix representation

Representing linear maps as matrices

We previously discsussed morphisms on vector spaces. We can write these as matrices.

Matrices represents transformations of vector spaces

Representing vectors as matrices

We can represent vectors as row or column matrices.

$v = \begin{bmatrix} a_1 & a_2 & ... & a_n \end{bmatrix}$

$v = \begin{bmatrix} a_1 \\ a_2 \\ ... \\ a_m \end{bmatrix}$

73.3 Automorphisms of vector spaces

73.3.1 Basis of an endomorphism

73.3.2 Changing the basis

For any two bases, there is a unique linear mapping from of the element vectors to the other.

73.4 The linear groups

73.4.1 General linear groups $GL(n, F)$

The general linear group, $GL(n, F)$, contains all $n \odot n$ invertible matrices M over field F.

The binary operation is multiplication.

73.4.2 Endomorphisms as group actions

We can view each member of the group g as a homomorphim on s.

Where s is a vector space V, the representation on each group member is an invertible square matrix.

If the set we use is the vector space V, then we can represent each group element with a square matrix acting on V.

Faithful means $a \neq b$ holds for repesentation too.

Representation theory. groups defined by $ab = c$. if we can match each eleemnt to amatrix where this holds we have represented the matrix.

73.4.3 Representing finite groups

Finite groups can all be represented with square matrices.

73.4.4 Representing compact groups

Chapter 74

Linear forms

74.1 Linear forms

74.1.1 Linear forms

A linear form is a linear map from a vector space to a scalar from the vector space's underlying field.

$\hom(V, F)$

Matrix operators

Linear forms can be represented as matrix operators.

$v^T M = f$

Where M has only one column.

Stuff

$f(M) = f(v)$

We introduce e_i, the element vector. This is 0 for all entries except for i where it is 1. Any vector can be shown as a sum of these vectors multiplied by a scalar.

$f(M) = f(\sum_{i=1}^m a_i e_i)$

$f(M) = \sum_{i=1}^m f(a_i e_i)$

$f(M) = \sum_{i=1}^m a_i f(e_i)$

$f(M) = \sum_{i=1}^m a_i f(e_i)$

Orthonormal basis

$f(M) = \sum_{i=1}^{m} a_i$

74.1.2 Dual space

The dual space V^* of vector space V is the set of all linear forms, $\hom(V, F)$.

The dual space is itself a vector space

$v \in V$

$f \in F$

$av = f$

$bv = g$

$(a \oplus b)v = f + g$

$(a \oplus b)v = av + bv$

So there is some operation we can do on two members of dual space

Linear in addition. That is, if we have two dual "things", we can define the addition of functions as the operation which results int he outputs being added.

what about linear in scalar? same approach.

Well we define

$(c \odot a) = cav$

The dual space has the same dimension as the underlying vector space

74.1.3 The dual space forms a vector space

The dual space forms a vector space. We can define addition and scalar multiplication on members of the dual space.

The dimension of the dual space is the same as the underlying space.

We have defined the dual space. A vector in dual space will have also have components and a basis.

$\mathbf{w} = \sum_i w_i f^j$

So how we describe the components will depend on the choice of basis.

We choose the dual basis, the basis for V^* as:

$\mathbf{e}_i\mathbf{f}^j = \delta_i^j$

If the basis changes, so does the dual basis.

We write the dual basis as e^j

74.2 Bilinear forms

74.2.1 Bilinear forms

A bilinear form takes two vectors and produces a scalar from the underyling field.

This is in contrast to a linear form, which only has one input.

In addition, the function is linear in both arguments.

$\phi(au + x, bv + y) = \phi(au, bv) + \phi(au, y) + \phi(x, bv) + \phi(x, y)$

$\phi(au + x, bv + y) = ab\phi(u, v) + a\phi(u, y) + b\phi(x, v) + \phi(x, y)$

Representing bilinear forms

They can be represented as:

$\phi(u, v) = v^T Mu$

$f(M) = f([v_1, v_2])$

We introduce e_i, the element vector. This is 0 for all entries except for i where it is 1. Any vector can be shown as a sum of these vectors multiplied by a scalar.

$f(M) = f([\sum_{i=1}^m a_{1i}e_i, \sum_{i=1}^m a_{2i}e_i])$

$f(M) = \sum_{k=1}^m f([a_{1k}e_k, \sum_{i=1}^m a_{2i}e_i])$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m f([a_{1k}e_k, a_{2i}e_i])$

Because this in linear in scalars:

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}a_{2i}f([e_k, e_i])$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}a_{2i}e_k^T Me_i$

Orthonormal basis and $M = I$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}a_{2i}e_k^T e_i$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}a_{2i}\delta_i^k$

$f(M) = \sum_{i=1}^m a_{1i}a_{2i}$

74.2.2 The dot product

$v^T Mu = f$

If the operator is I then we have the dot product.

$v^T u$

74.2.3 Orthogonal vectors

Given a metric M, two vectors v and u are orthogonal if:

$v^T Mu = 0$

For example if we have the metric $M = I$, then two vectors are orthogonal if:

$v^T u = 0$

74.2.4 Metric-preserving transformations and isometry groups

If we have a bilinear form we can write the form as:

$u^T Mv$

After a transformation P to the vectors it is:

$(Pu)^T M(Pv)$

$u^T P^T MPv$

So the value of the metric will be unaffected if:

$u^T P^T MPv = u^T Mv$

$P^T MP = M$

Equivalent metrics

Different metrics can produce the same group. For example multiplying the metric by a constant.

$P^T MP = M$

74.2.5 Orthogonal groups $O(n, F)$

Recap: Metric-preserving transformations

The bilinear form is:

$u^T M v$

The transformations which preserve this are:

$P^T M P = M$

The orthogonal group

If the metic is $M = I$ then the condition is:

$P^T P = I$

$P^T = P^{-1}$

These form the orthogonal group.

We use O instead of P:

$O^T = O^{-1}$

Rotations and reflections

The orthogonal group is the rotations and reflections.

Parameters of the orthogonal group

The orthogonal group depends on the dimension of the vector space, and the underlying field. So we can have:

- $O(n, R)$; and
- $O(n, C)$.

We generally refer only to the reals

$O(n)$ means $O(n, R)$.

The generally refer to the reals only.

74.2.6 Indefinite (pseduo) and split orthognal groups $O(n, m, F)$

Recap: Metric-preserving transformations

The bilinear form is:

$u^T M v$

The transformations which preserve this are:

$P^TMP = M$

The metric

If the metric is:

$$M = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then we have the indefinite orthogonal group $O(3,1)$

The split orthogonal group

Where $n = m$ we have the split orthogonal group.

$O(n,n,F)$

Signatures

74.2.7 The Lorentz group

The Lorentz group is the $O(1,3)$ group.

Symmetries of the Lorentz group

We can do the usual 3 rotations, however there are additional 3 symmetries, making the Lorzentz group 6-dimensional.

These are the Lorentz boosts.

A symmetry has:

$t'^2 - x'2 - y'^2 - z'^2 = t^2 - x^2 - y^2 - z^2$

We consider the case where we just boost on x, so $y = y'$ and $z = z'$.

$t'^2 - x'2 = t^2 - x^2$

Or with c:

$c^2t'^2 - x'2 = ct^2 - x^2$

New

$s^2 = t^2 - x^2 - y^z - z^2$

$s'^2 = t'^2 - x'^2 - y'^2 - z'^2$

$ds^2 = s'^2 - s^2$

$ds^2 = (t'^2 - x'^2 - y'^2 - z'^2) - (t^2 - x^2 - y^z - z^2)$

$ds^2 = (t'^2 - t^2) - (x'^2 - x^2) - (y'^2 - y^2) - (z'^2 - z^2)$

$ds^2 = dt^2 - dx^2 - dy^2 - dz^2$

boost: $s^2 = c^2t^2 - x^2 - y^z - z^2$

we want new t and x where distance is same $c^2t'^2 - x'^2 - y^z - z^2 = c^2t^2 - x^2 - y^z - z^2$ $c^2t'^2 - x'^2 = c^2t^2 - x^2$

We know that both transformations are linear [WHY??], therefore $x' = Ax + Bt$ $t' = Cx + Dt$

we transform to $x' = 0$. so $Ax + Bt = 0$
We define $v = \frac{x}{t}$

So: $x = vt$

We can plug these in: $Avt + Bt = 0$ $Av + B = 0$ $\frac{A}{B} = -v$

74.3 Sesquilinear forms

74.3.1 Sesquilinear forms

Bilinear form recap

A bilinear form takes two vectors and produces a scalar from the underyling field.

The function is linear in addition in both arguments.

$\phi(au + x, bv + y) = \phi(au, bv) + \phi(au, y) + \phi(x, bv) + \phi(x, y)$

The function is also linear in multiplication in both arguments.

$\phi(au + x, bv + y) = ab\phi(u, v) + a\phi(u, y) + b\phi(x, v) + \phi(x, y)$

They can be represented as:

$\phi(u, v) = v^T Mu$

Sesquilinear forms

Like bilinear forms, sesquilinear are linear in addition:

$\phi(au + x, bv + y) = \phi(au, bv) + \phi(au, y) + \phi(x, bv) + \phi(x, y)$

Sesqulinear forms however are only multiplictively linear in the second argument.

$\phi(au + x, bv + y) = b\phi(au, v) + \phi(au, y) + b\phi(x, v) + \phi(x, y)$

In the first argument they are "twisted"

$\phi(au + x, bv + y) = \bar{a}b\phi(u, v) + \bar{a}\phi(u, y) + b\phi(x, v) + \phi(x, y)$

The real field

For the real field, $\bar{b} = b$ and so the sesqulinear form is the same as the bilinear form.

Representing sesquilinear forms

We can show the sesquilinear form as v^*Mu

Stuff

$f(M) = f([v_1, v_2])$

We introduce e_i, the element vector. This is 0 for all entries except for i where it is 1. Any vector can be shown as a sum of these vectors multiplied by a scalar.

$f(M) = f([\sum_{i=1}^m a_{1i}e_i, \sum_{i=1}^m a_{2i}e_i])$

$f(M) = \sum_{k=1}^m f([a_{1k}e_k, \sum_{i=1}^m a_{2i}e_i])$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m f([a_{1k}e_k, a_{2i}e_i])$

Because this in linear in scalars:

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}^* a_{2i} f([e_k, e_i])$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}^* a_{2i} e_k^* M e_i$

Orthonormal basis and $M = I$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}^* a_{2i} e_k^* M e_i$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}^* a_{2i} e_k^* e_i$

$f(M) = \sum_{k=1}^m \sum_{i=1}^m a_{1k}^* a_{2i} \delta_i^k$

$f(M) = \sum_{i=1}^{m} a_{1i}^* a_{2i}$

74.3.2 Unitary groups $U(n, F)$

Metric preserving transformations for sesquilinear forms

For bilinear forms, the transformations which preserved metrics were:

$P^T = P^{-1}$

For sesquilinear they are different:

$u^* M v$

$(Pu)^* M (Pv)$

$u^* P^* M P v$

So we want the matrices where:

$P^* M P = M$

The unitary group

The unitary group is where $M = I$

$P^* P = I$

$P^* = P^{-1}$

We refer to these using U instead of P.

$U^* = U^{-1}$

Parameters of the unitary group

The unitary group depends on the dimension of the vector space, and the underlying field. So we can have:

- $U(n, R)$; and
- $U(n, C)$.

We generally refer only to the complex

For the $U(n, R)$ we have:

$U^* = U^{-1}$

$U^T = U^{-1}$

This is the condition for the orthogonal group, and so we would instead write $O(n)$.

As a result, $U(n)$ refers to $U(n, C)$.

$U(1)$: The circle group

74.4 Inner products

74.4.1 Symmetric matrices

74.4.2 Hermitian (self-adjoint) matrices

A matrix where $M = M^*$

For matrices over the real numbers, these are the same as symmetric matrices.

Sesqulinear forms on Hermitian matrices

$\phi(u, v) = u^* M v$

$(u^* M v)^* = v^* M^* u = v^* M u$

$\phi(u, v) = \overline{\phi(v, u)}$

The forms on the same vector are always real

$(v^* M v)^* = v^* M^* v = v^* M v$

So we have:

$(v^* M v)^* = v^* M v$

Which is only satisfied for reals.

If A and B are Hermitian

If A and B are Hermitian, AB is Hermitian if and only if AB commutes.

$(AB)^* = B^* A^* = BA$

If it commutes then

$(AB)^* = AB$

Real eigenvalues

Hermitian matrices have real eigenvalues.

$Hv = \lambda v$

$v^* H v = \lambda v^* v$

$v^* H v = \lambda$

Skew-Hermitian matrices

These are also known as anti-Hermitian matrices.

$M^* = -M$

If eigenvalues are different, eigenvectors are orthogonal

74.4.3 Pauli matrices

Pauli matrices are 2×2 matrices which are unitary and hermitian.

That is, $P^* = P^{-1}$.

And $P^* = P$.

The Pauli matrices

The matrices are:

$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$

$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

The identity matrix is often considered alongside these as:

$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Pauli matrices are their own inverse

$\sigma_i^2 = \sigma_i \sigma_i$

$\sigma_i^2 = \sigma_i \sigma_i^*$

$\sigma_i^2 = \sigma_i \sigma_i^{-1}$

$\sigma_i^2 = I$

Determinants and trace of Pauli matrices

$\det \sigma_i = -1$

$Tr(\sigma_i) = 0$

As the sum of eigenvalues is the trace, and the product is the determinant, the eigenvalues are 1 and -1.

74.4.4 Positive-definite matrices

The matrix M is positive definite if for all non-zero vectors the scalar is positive.

$v^T M v$

We know that the outcome is a scalar, so:

$v^T M v = (v^T M v)^T$

$v^T M v = v^T M^T v$

$v^T (M - M^T) v = 0$

74.4.5 Inner products

An inner product is a sesquilinear form with a positive-definite Hermitian matrix.

$\langle u, v \rangle = u^* H v$

If we are using the real field this is the same as:

$\langle u, v \rangle = u^T H v$

Where H is now a symmetric real matrix.

Same

$\langle v, v\rangle = v^* H v$

Always positive and real.

Properties

$\langle u, v\rangle\langle v, u\rangle = |\langle u, v\rangle|^2$

74.4.6 Cauchy-Schwarz inequality

This states that:

$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \dot{\langle} v, v\rangle$

Consider the vectors u and v. We construct a third vector $u - \lambda v$. We know the length of any vector is non-negative. $0 \leq \langle u - \lambda v, u - \lambda v\rangle$

$0 \leq \langle u, u\rangle + \langle u, -\lambda v\rangle + \langle -\lambda v, u\rangle + \langle -\lambda v, -\lambda v\rangle$

$0 \leq \langle u, u\rangle - \bar{\lambda}\langle u, v\rangle - \lambda\langle v, u\rangle + \lambda\bar{\lambda}\langle v, v\rangle$

We now look for a value of λ to simplify this equation.

$\lambda = \frac{\langle u, v\rangle}{\langle v, v\rangle}$

$0 \leq \langle u, u\rangle - \frac{\langle v, u\rangle\langle u, v\rangle}{\langle v, v\rangle} - \frac{\langle u, v\rangle\langle v, u\rangle}{\langle v, v\rangle} + \frac{\langle u, v\rangle}{\langle v, v\rangle}\frac{\langle v, u\rangle}{\langle v, v\rangle}\langle v, v\rangle$

$0 \leq \langle u, u\rangle - \frac{|\langle u, v\rangle|^2}{\langle v, v\rangle}$

$|\langle u, v\rangle|^2 \geq \langle u, u\rangle\langle v, v\rangle$

74.4.7 The orthogonal projection

in inner product space, orthogonal projection

$p_u v = \frac{< u, v >}{v, v} v$

we then know that $o = v - p_u v$ is orthogonal to u.

74.4.8 Orthogonal set / Orthogonal basis

if a set of vectors are all orthogonal, they form an orthogonal set if the set spans the vector space, it is an orthogonal basis.

74.4.9 The Gram-Schmidt process

can we form an orthogonal basis from a non-orthogonal basis? yes, using gram schmidt

we have x_1, x_2 x_3 etc we want to make v_1, v_2 etc orthognal

$v_1 = x_1$ $v_2 = x_2 - p_{x_2}v_1$ $v_3 = x_3 - p_{x_3}v_1 - p_{x_3}v_2v$

74.5 Multilinear forms and determinants

74.5.1 Multilinear forms

74.5.2 Determinants

From invertible matrix section in endo

A matrix can only be inverted if it can be created from a combination of elementary row operations.

How can we identify if a matrix is invertible? We want to create a scalar from the matrix which tells us if this possible. We can this scalar the determinant.

For a matrix A we label the determinant $|A|$, or $\det A$

We propose $|A| = 0$ when the matrix is not invertible.

So how can we identify the function we need to undertake on the matrix?

New 1

We know that linear dependence results in determinants of 0.

We can model this as a function on the columns of the matrix.

$\det M = \det([M_1, ..., M_n)$

If there is linear depednence, for example if two columns are the same then:

$\det([M_1, ..., M_i, ..., M_i, ..., M_n]) = 0$

Similarly, if there is a column of 0 then the determinant is 0.

$\det([M_1, ..., 0, ..., M_n]) = 0$

New 2

Show linear in addition

How can we identify the determinant of less simple matrices? We can use the multilinear form.

$\sum c_i \mathbf{M}_i = \mathbf{0}$

Where $\mathbf{c} \neq \mathbf{0}$

Or:

$M\mathbf{c} = \mathbf{0}$

Rule 1: Columns of matrices can be the input to a multilinear form

A matrix can be shown in terms of its columns. $A = [v_1, ..., v_n]$

$\det A = \det[v_1, ..., v_n]$

$\det A = \sum_{k_1=1}^m ... \sum_{k_n=1} \prod_{i=1}^m a_{ik_i} \det([e_{k_1}, ..., e_{k_n}])$

Multiplying a matrix by a constant multiplies the determinant by the same amount

If a whole row or columns is 0 then:

$\det A = \det[v_1, ..., v_i, ..., v_n]$

$\det A' = \det[v_1, ..., cv_i, ..., v_n]$

$\det A = \det[v_1, ..., v_i, ..., v_n]$

$\det A' = \det[v_1, ..., cv_i, ..., v_n]$

$\det A' = c \det[v_1, ..., v_i, ..., v_n]$

$\det A' = c \det A$

As a result, multiplying a column by 0 makes the determinant 0.

A matrix with a column of 0 therefore has determinant 0

Rule 2: A matrix with equal columns has a determinant of 0.

$A = [a_1, ..., a_i, ..., a_i, ..., a_n]$

$D(A) = D([a_1, ..., a_i, ..., a_i, ..., a_n])$

We know from Result 3 that swapping columns reverses the sign. Reversing columns results in the same matrix, so the determinant must be unchanged.

$D(A) = -D(A)$

$D(A) = 0$

Linear dependence

If a column is a linear combination of other columns, then the matrix cannot be inverted.

$A = [a_1, ..., \sum_{j\neq i}^n c_j a_j, ..., a_n]$

$\det A = \det([v_1, ..., \sum_{j\neq i}^n c_j v_j, ..., v_n])$

$\det A = \sum_{j\neq i}^n c_j \det([v_1, ..., v_j, ..., v_n])$

$\det A = \sum_{j\neq i}^n c_j \det([v_1, ..., v_j, , ..., v_j, ..., v_n])$

As there is a repeating vector:

$\det A = 0$

Swapping columns multiplies the determinant by -1

$A = [v_1, ..., v_i + v_j, ..., v_i + v_j, ..., v_n]$

We know.

$\det A = 0$

$\det A = \det([a_1, ..., a_i, ..., a_i, ..., a_n]) + \det([a_1, ..., a_i, ..., a_j, ..., a_n]) + \det([a_1, ..., a_j, ..., a_i, ..., a_n]) + \det([a_1, ..., a_j, ..., a_j, ..., a_n])$

So:

$\det([a_1, ..., a_i, ..., a_i, ..., a_n]) + \det([a_1, ..., a_i, ..., a_j, ..., a_n]) + \det([a_1, ..., a_j, ..., a_i, ..., a_n]) + \det([a_1, ..., a_j, ..., a_j, ..., a_n]) = 0$

As 2 of these have equal columns these are equal to 0.

$\det([a_1, ..., a_i, ..., a_j, ..., a_n]) + \det([a_1, ..., a_j, ..., a_i, ..., a_n]) = 0$

$\det([a_1, ..., a_i, ..., a_j, ..., a_n]) = -\det([a_1, ..., a_j, ..., a_i, ..., a_n])$

Calculating the determinant

We have

$\det A = \sum_{k_1=1}^m ... \sum_{k_n=1} \prod_{i=1}^m a_{ik_i} \det([e_{k_1}, ..., e_{k_n}])$

So what is the value of the determinant here?

We know that the determinant of the identity matrix is 1.

We know that the determinant of a matrix with identical columns is 0.

We know that swapping columns multiplies the determinant by -1.

Therefore the determinants where the values of k are not all unique are 0.

The determinants of the others are either -1 or 1 depending on how many swaps are required to restore to the identity matrix.

This is also shown as the Leibni formula.

$\det A = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^n a_{i,\sigma_i}$

74.5.3 Properties of determinants

Identity

$\det I = 1$

Multiplication

$\det(AB) = \det A \det B$

Inverse

$\det(M^{-1}) = \dfrac{1}{\det M}$

We know this because:

$\det(MM^{-1}) = \det I = 1$

$\det M \det M^{-1} = 1$

$\det(M^{-1}) = \dfrac{1}{\det M}$

Complex cojugate

$\det(M^*) = \overline{\det M}$

Transpose

$\det(M^T) = \det M$

Addition

$\det(A + B) = \det A + \det B$

Scalar multiplication

$\det cM = c^n \det M$

Determinants and eigenvalues

The determinant is equal to the product of the eigenvalues.

74.5.4 Determinants of 2x2 and 3x3 matrices

The determinant of a 2x2 matrix

$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$|M| = ad - bc$

The determinant of a 3x3 matrix

$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

$|M| = aei + bfg + cdh - ceg - dbi - afh$

74.6 Special groups

74.6.1 Special orthogonal groups $SO(n, F)$

The special orthogonal group, $SO(n, F)$, is the subgroup of the orthogonal group where $|M| = 1$.

As a result it includes only the rotation operators, not the flip operators.

$SO(3)$ is rotations in 3d space.

$SO(2)$ is rotations in 2d space.

Determinant of the orthogonal group

The orthogonal group has determinants of -1 or 1.

$O^T = O^{-1}$

$\det(O^T) = \det(O^{-1})$

$\det O = \frac{1}{\det O}$

$\det O = \pm 1$

74.6.2 Special unitary groups $SU(n, F)$

The special unitary group, $SU(n, F)$, is the subgroup of $U(n, F)$ where the determinants are 1.

That is, $|M| = 1$

The determinant of unitary matrices

The determinant of the unitary matrices is:

$\det U^* = \det U^{-1}$

$(\det U)^* = \frac{1}{\det U}$

$(\det U)^* \det U = 1$

$|| \det U|| = 1$

74.6.3 Special lnear groups $SL(n, F)$

The special linear group, $SL(n, F)$, is the subgroup of $GL(n, F)$ where the determinants are 1.

That is, $|M| = 1$

These are endomorphisms, not forms.

74.7 Sort

74.7.1 Normal matrices

$M^* M = M M^*$

All symmetrix matrices are normal

All hermetitian matrices (inc subset symmetric) are normal

Normal matrix never defective

Chapter 75

Linear maps

75.1 Homomorphisms of vector spaces

75.1.1 Linear maps

Homomorphisms between vector spaces

Homomorphisms map between algebras, preserving the underlying structure.

A homomorphism vetween vector space V and vector space W can be described as:

$\hom(V, W)$

Homomorphism between vector spaces must preserve the group-like structure of the vector space.

$f(u+v) = f(u) + f(v)$

The homomorphism must also preserve scalar multiplication.

$f(\alpha v) = \alpha f(v)$

A linear map (or function) is a map from one input to an output which preserves addition and scalar multiplication.

That is if function f is linear then:

$f(aM + bN) = af(M) + bf(N)$

Alternative names for homomorphisms

Vector spaces homomorphisms are also called linear maps or linear functions.

75.1.2 Homomorphisms form a vector space

If we can can show that scalars can act on morphisms, then we can shwn that morphisms on a vector space are themselves a vector space.

Scalars can act on morphisms, and so morphisms of vector spaces are themselves vector spaces.

Dimensions of homomorphisms

We can identify the dimensionality of this new vector space from the dimensions of the original vector spaces.

$\dim(\hom(V, W)) = \dim V \dim W$

75.1.3 The pseudo-inverse

The definition of the inverse is that:

$MM^{-1} = I$

$M^{-1}M = I$

We also have:

$MM^{-1}M = M$

$M^{-1}MM^{-1} = M^{-1}$

The inverse of a homomorphism

Generally we don't have inverses of homomorphisms as the number of dimensions are different.

We can, however, find a matrix M^+ which satisfies:

$MM^+M = M$

$M^+MM^+ = M^+$

This is the pseudo-inverse.

75.1.4 Linear and affine functions

Linear maps

Linear maps can be written as:

$v = Mu$

These go through the origin. That is, if $u = 0$ then $v = 0$.

Affine function

Affine functions are more general than linear maps. They can be written as:

$v = Mu + c$

Where c is a vector in the same space as v.

Affine functions where $c \neq 0$ are not linear maps. They are not homomorphisms which preserve the structure of the vector space.

If we multiply u by a scalar s, then v will not increase by the same proportion.

75.1.5 Singular value decomposition

The singular value decomposition of $m \times n$ matrix M is:

$M = U\Sigma V^*$

Where:

- U is a unitary matrix $(m \times m)$
- Σ is a diagonal matrix with non-negative real numbers $(m \times n)$
- V is a unitary matrix $(n \times n)$

Σ is unique. U and V are not.

Properties

$M^*M = U\Sigma^2U^*$

$(M^*M)^{-1} = V\Sigma^{-2}V^*$

Calculating the SVD

The SVD is generally calculated iteratively.

75.1.6 Identity matrix and the Kronecker delta

The Kronecker delta

The Kronecker delta is defined as:

p$\delta_{ij} = 0$ where $i \neq j$

$\delta_{ij} = 1$ where $i = j$

We can use this to define matrices. For example for the identity matrix:

$I_{ij} = \delta_{ij}$

Identity matrix

A square matrix where every element is 0 except where $i = j$. There is one for each square matrix.

$$I = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

Chapter 76

Lie algebra

76.1 Cross products

76.1.1 The cross product

$v \times u$

Cross product is a bilinear map

This is a bilinear map from two vectors in $\mathbb{R}^3$ to another vector in the same space.

$V \times V \to V$

Calculating the cross product

This is calcualted by:

$u \times v = ||u||||v|| \sin(\theta) n$

The resulting vector is perpendicular to both input vectors.

76.2 Lie groups

76.2.1 Lie groups

76.3 Lie algebra

76.3.1 Lie algebra

Lie groups have symmetries. We can consider only the infintesimal symmetries.

For example the unit circle has many symmetries, but we can consider only those which rotate infintesimally.

Example

Take a continous group, such as $U(1)$. Its Lie algebra is all matrices such that their exponential is in the Lie group.

$\mathfrak{u}(1) = \{X \in \mathbb{C}^{1\times 1} | e^{tX} \in U(1) \forall t \in \mathbb{R}\}$

This is satisfied by the matrices where $M = -M^*$. Note that this means the diagonals are all 0.

Scale of specific Lie algebra matrices doesn't matter

Because of t.

Commutation of Lie group algebra

Consider two members of the Lie algebra: A and B. The commutator is:

A.

The corresponding Lie group member is:

$e^{t(A+B)} = e^{tA}e^{tB}$

While the Lie group multiplication may not commute, the corresponding addition of the Lie algebra does.

76.3.2 The Lie bracket

We can define the Lie bracket from the ring commutator.

We use the Lie bracket, rather than multiplication, as the operator over a field homomorphism.

$[A, B]$

This generates another element in the algebra.

This satisifies:

- Bilinearity: $[xA + yB, C] = x[A, C] + y[B, C]$
- Alternativity: $[A, A] = 0$
- Jacobi identity: $[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$
- Anticommutivity: $[A, B] = -[B, A]$

One option for the Lie bracket is the ring commutor. So that:

$[A, B] = AB - BA$

76.3.3 Commutation of Lie groups

We can measure commutation of Lie groups using:

$ABA^{-1}B^{-1}$

If the group commutes then:

$ABA^{-1}B^{-1} = BA^{-1}B^{-1} = I$

Commutation of Lie algebra: COMPLETE THIS

This corresponds to $[A, B] = AB - BA$ in the underlying lie algebra, if we expand.

$A = e^{ta}$

$B = e^{tb}$

$ABA^{-1}B^{-1} = e$

76.4 Lie algebra of specific Lie groups

76.4.1 Lie algebra of $O(n)$

$O(n)$ forms a Lie group

Lie algebra of $O(n)$

The Lie algebra of $O(n)$ is defined as:

$\mathfrak{o}(n) = \{X \in \mathbb{R}^{n\times n} | e^{tX} \in O(n) \forall t \in \mathbb{R}\}$

This is satisfied by the skew-symmetric matrices where $M = -M^T$. Note that this means the diagonals are all 0.

76.4.2 Lie algebra of $U(n)$

$U(n)$ forms a Lie group

Lie algebra of $U(n)$

The Lie algebra of $U(n)$ is defined as:

$\mathfrak{u}(n) = \{X \in \mathbb{C}^{n\times n} | e^{tX} \in U(n) \forall t \in \mathbb{R}\}$

This is satisfied by the skew-Hermitian matrices where $M = -M^*$. Note that this means the diagonals are all 0 or pure imaginary.

76.4.3 Lie algebra of $SO(n)$

$SO(n)$ forms a Lie group

Lie algebra of $SO(n)$

The Lie algebra of $SO(n)$ is defined as:

$\mathfrak{so}(n) = \{X \in \mathbb{R}^{n\times n} | e^{tX} \in SO(n) \forall t \in \mathbb{R}\}$

This is satisfied by the skew-symmetric matrices where $M = -M^T$. Note that this means the diagonals are all 0.

76.4.4 Lie algebra of $SU(n)$

$SU(n)$ forms a Lie group

Lie algebra of $SU(n)$

The Lie algebra of $SU(n)$ is defined as:

$\mathfrak{su}(n) = \{X \in \mathbb{C}^{n\times n} | e^{tX} \in SU(n) \forall t \in \mathbb{R}\}$

This is satisfied by the skew-Hermitian matrices where $M = -M^*$ and the trace is 0. Note that this means the diagonals are all 0 or pure imaginary.

76.5 Hypercomplex numbers

76.5.1 Hypercomplex numbers

76.5.2 Quaternions

76.5.3 Clifford algebra

76.6 Sort

76.6.1 Projective line in the field

Chapter 77

Exterior algebra

77.1 Exterior algebra

77.1.1 The exterior (wedge) product

The exterior product of two vectors is:

$u \wedge v$

77.1.2 The exterior product is anticommutative

This is anticommutative (alternating).

$u \wedge v = -v \wedge u$

This implies that:

$u \wedge u = 0$

77.1.3 The exterior product is distributive

$(a + b) \wedge (c + d) = (a \wedge c) + (a \wedge d) + (b \wedge c) + (b \wedge d)$

77.1.4 Expanding the exterior product of two vectors

We can look at the exterior product in component-basis terms.

Consider 2-dimenional vector space with the following vectors:

$u = ae_1 + be_2$

$v = ce_1 + de_2$

The exterior product is:

$u \wedge v = (ae_1 + be_2) \wedge (ce_1 + de_2)$

$u \wedge v = (ae_1 \wedge ce_1) + (ae_1 \wedge de_2) + (be_2 \wedge ce_1) + (be_2 \wedge de_2)$

$u \wedge v = ac(e_1 \wedge e_1) + ad(e_1 \wedge e_2) + bc(e_2 \wedge e_1) + bd(e_2 \wedge e_2)$

$u \wedge v = ad(e_1 \wedge e_2) - bc(e_1 \wedge e_2)$

$u \wedge v = (ad - bc)(e_1 \wedge e_2)$

77.1.5 Exterior (Grassman) algebra

The exterior algebra is the algebra generated by the wedge product.

The term $u \wedge v$ can be interpreted as the area covered by the parallelogram generated by u and v.

As $a\mathbf{u} \wedge b\mathbf{v} = ab\mathbf{u} \wedge \mathbf{v}$, we can see that scaling the length of one of the vectors by a scalar, we also increase the exterior product by the same scalar.

77.1.6 Orientation

We can describe the exterior product of two vectors as $\mathbf{u} \wedge \mathbf{v}$ or $\mathbf{v} \wedge \mathbf{u}$.

77.1.7 Bivectors

77.1.8 Trivectors

Chapter 78

Tensors

78.1 Element-wise notation

78.1.1 Einstein summation convention

A vector can be written as a sum of its components.

$v = \sum_i e_i v^i$

The Einstein summation convention is to remove the $\sum_i$ symbols where they are implicit.

For example we would instead write the vector as:

$v = e_i v^i$

Adding vectors

$v + w = (\sum_i e_i v^i) + (\sum_i f_i w^i)$

$v + w = \sum_i (e_i v^i + f_i w^i)$

$v + w = e_i v^i + f_i w^i$

If the bases are the same then:

$v + w = e_i (v^i + w^i)$

Scalar multiplication

$cv = c \sum_i e_i v^i$

$cv = \sum_i ce_i v^i$

$cv = ce_i v^i$

Matrix multiplication

$AB_{ik} = \sum_j A_{ij} B_{jk}$

$AB_{ik} = A_{ij} B_{jk}$

Inner products

$\langle v, w \rangle = \langle \sum_i e_i v^i, \sum_j f_j w^j \rangle$

$\langle v, w \rangle = \sum_i v^i \langle e_i, \sum_j f_i w^j \rangle$

$\langle v, w \rangle = \sum_i \sum_j v^i \overline{w^j} \langle e_i, f_j \rangle$

If the two bases are the same then:

$\langle v, w \rangle = \sum_i \sum_j v^i \overline{w^j} \langle e_i, e_j \rangle$

We can define the metric as:

$g_{ij} := \langle e_i, e_j \rangle$

$\langle v, w \rangle = v^i \overline{w^j} g_{ij}$

78.1.2 Covariant and contravariant bases

In element form we write a vector as:

$v = e_i v^i$

The indices are raised and lowered to reflect whether the value is covariant or contravariant.

v^i is contravariant. If the basis moves one way, it moves the other.

e_i is covariant. If the basis moves, it moves with it.

78.2 Tensor product

78.2.1 Tensor product

We have spaces V and W over field F. If we have a linear operation which takes a vector from each space and returns a scalar from the underlying field, it is an element of the tensor product of the two spaces.

For example if we have two vectors:

$v = e_i v^i$

$w = e_j w^j$

A tensor product would take these and return a scalar.

There are three types of tensor products:

- Both are from the vector space
- $T_{ij} v^i w^j$
- $T_{ij} \in V \otimes W$
- Both are from the dual space
- $T^{ij} v_i w_j$
- $T_{ij} \in V^* \otimes W^*$
- One is from each space
- $T_i^j v^i w_j$
- $T_{ij} \in V \otimes W^*$

As a vector space, we can add together tensor products, and do scalar multiplication.

Basis of a tensor product

Not all elements spanned by a basis of a tensor product are themselves tensor products.

Eigenvalues and Eigenvectors of a tensor product

Homomorphisms

We can define homomorphisms in terms of tensor products.

$Hom(V) = V \otimes V^*$

T_j^i

We use the dual space for the second argument. This is because it ensures that changes to the bases do not affect the maps.

$w^j = T_i^j v^i$

78.2.2 Raising and lowering indices

We showed that the inner product between two vectors with the same basis can be written as:

$\langle v, w\rangle = \langle \sum_i e_i v^i, \sum_j f_j w^j\rangle$

$\langle v, w\rangle = v^i \overline{w^j} \langle e_i, e_j\rangle$

Defining the metric as:

$g_{ij} := \langle e_i, e_j\rangle$

$\langle v, w\rangle = v^i \overline{w^j} g_{ij}$

Metric inverse

We can use this to define the inverse of the metric.

$g^{ij} := (g_{ij})^{-1}$

We can use this to raise and lower vectors.

$v_i := v^j g_{ij}$

Raising and lowering indices of tensors

If we have tensor:

T_{ij}

We can define:

$T_i^k = T_{ij} g^{jk}$

$T^{il} = T_{ij} g^{jk} g^{kl}$

Tensor contraction

If we have:

$T_{ij} x^j$

We can contract it to:

$T_{ij} x^j = v_i$

Similarly we can have:

$T^{ij} x_j = v^i$

78.2.3 Kronecker delta

Consider matrix multiplication AI.

We have:

$AI_{ik} = A_{ij} I_{jk}$

We write this instead as:

$AI_{ik} = A_{ij} \delta_{jk}$

Where $\delta_{jk} = 0$ if $j \neq k$ and $\delta_{jk} = 1$ if $j = k$.

78.2.4 Tensors form a vector space

Recap

Tensors form a vector space

Dimension of a tensor

Basis of a tensor

78.3 Tensors

78.3.1 Tensor valence

78.3.2 Tensor inverses

For second order tensors we have:

- T^i_j
- T_{ij}
- T^{ij}

For each of these we can define an inverse:

- $T^j_i U^k_j = \delta^k_i$
- $T_{ij} U^{jk} = \delta^k_i$
- $T^{ij} U_{jk} = \delta^k_i$

Notation for inverses

If we have $T_{ij}U^{jk}=\delta_i^k$, we can instead write:

$T_{ij}T^{jk}=\delta_i^k$

78.3.3 Tensor contraction

We have a vector $v\in V$ and $w\in V^*$.

$\mathbf{v}=\sum_i v^i\mathbf{e}_i$

$\mathbf{w}=\sum_i w_i\mathbf{f}^i$

$\mathbf{wv}=[\sum_i v^i\mathbf{e}_i][\sum_i w_i\mathbf{f}^i]$

$\mathbf{wv}=\sum_i\sum_j[v^i\mathbf{e}_i][w_j\mathbf{f}^j]$

$\mathbf{wv}=\sum_i\sum_j v^iw_j\mathbf{e}_i\mathbf{f}^j$

We use the dual basis so:

$\mathbf{wv}=\sum_i\sum_j v^iw_j\mathbf{e}_i\mathbf{e}^j$

$\mathbf{wv}=\sum_i\sum_j v^iw_j\delta_i^j$

We can see that this value is unchanged when there is a change in basis.

What if these were both from V?

$\mathbf{v}=\sum_i v^i\mathbf{e}_i$

$\mathbf{w}=\sum_i w^i\mathbf{e}_i$

$\mathbf{wv}=[\sum_i v^i\mathbf{e}_i][\sum_i w^i\mathbf{e}_i]$

$\mathbf{wv}=\sum_i\sum_j v^iw^j\mathbf{e}_i\mathbf{e}_i$

This term is dependent on the basis, and so we do not contract.

So if we have v_iw^i, we can contract, because the result (calculated from the components) does not depend on the basis.

But if we have v_iw_i, the result (calcualted from the components) will change depending on the choice of basis.

We define a new object

$c=\sum_i w^iv_i$

This new term, c, does not depend on i, and so we have contracted the index.

78.3.4 Symmetric and antisymmetric tensors

Consider a tensor, e g T_{abc}.

In general, this is not symmetric, that is:

$T_{abc} \neq T_{bac}$

Symmetric part of a tensor

We can write the symmetric part of this with regard to a and b.

$T_{(ab)c} = \frac{1}{2}(T_{abc} + T_{bac})$

Clearly, $T_{(ab)c} = T_{(ba)c}$

Antisymmetric part of a tensor

We can also have an an antisymmetric part with regard to a and b.

$T_{[ab]c} = \frac{1}{2}(T_{abc} - T_{bac})$

Clearly, $T_{[ab]c} = -T_{[ba]c}$

Tensors as sums of their symmetric and antisymmetric parts

$T_{(ab)c} + T_{[ab]c} = \frac{1}{2}(T_{abc} + T_{bac}) + \frac{1}{2}(T_{abc} - T_{bac})$

$T_{(ab)c} + T_{[ab]c} = T_{abc}$

78.4 Higher-order tensors

78.4.1 Higher-order tensors

We can create higher order tensors products. For example

$V \otimes V \otimes V \otimes V^* \otimes V^*$

We write elements of these as:

$T^{i_1,\dots,i_p}_{j_1,\dots,j_q}$

We can map from matrix to matrix etc higher dimensional

Matrix has A: a_{ij}.

Tensor can have T: t_{ijk} for example

0 rank tensor: scalar

1 rank tensor: vector

2 rank tensor: matrix

page on covariance and contravariance and type (p, q)

78.5 Sort

78.5.1 Outer product

The outer product is a bilinear map

This is a bilinear map from two vectors from the same vector space to another vector space.

$V \times V \rightarrow V$

Calculating the outer product

$u \otimes v = w$

$w_{ij} = u_i v_j$

The dimensions of the tensor outer product

$\dim(V \otimes W) = \dim V \times \dim W$

Outer product on the complex numbers

Relation between the dot product and outer product

The dot product in the trace of the outer product.

78.5.2 Kronecker product

The Kronecker product takes the concept of the outer product and applies to to matrices.

We can essentially repace every element in the matrix on the left with the element multiplied by the entire matrix on the right.

Like outer products, Kronecker products are written as:

$u \otimes v = w$

78.5.3 Dot product

Dot product is a bilinear form

This is a bilinear form, a mapping from two vectors in the same vector space to the underlying field.

$V \times V \rightarrow F$

Calculating the dot product

This is calculated by multiplying each matching element, and summing the results.

$u \cdot v = \sum_{i=1}^{n} u_i v_i$

Dot product on the complex numbers

Properties don't hold. Can get zero vectors from non-zero inputs. Get complex numbers from dot product on itself.

Inner products better deal with complex number fields. However they are not bilinear maps.

78.5.4 Homomorphism as a tensor product

78.5.5 Tensors

A tensor is an element of a tensor product.

Chapter 79

Infinite-dimensional vector spaces

79.1 Real functions as infinite-dimensional vectors

79.1.1 Real functions are vectors

The real function space is a vector space because it is linear in multiplication and addition.

$g(x) = cf(x)$

$h(x) = f(x) + k(x)$

79.2 Endomorphisms of infinite-dimensional vector spaces

79.2.1 Endomorphisms on real functions

We start with our vector $f(x)$.

$h(x) = f(x)g(x)$

The equivalent of the identity matrix is where $g(x) = 1$.

These are similar to endomorphisms where all off diagonal elements are 0.

Differentiation

$h(x) = \frac{\delta}{\delta x} f(x)$

Integration

$h(x) = \int_{-\infty}^{x} f(z) dz$

79.2.2 Examples of linear operators on real functions

For a function v we can define operators Ov.

Here we consider some examples and their properties.

Real multiplication

$Rv = rf(x)$

This operator is hermitian. This is equivalent to a finite operator of the form rI.

Multiplication by underlying real number

$Xv = xf(x)$

This operator is hermitian. This is equivalent to a finite operator of the form $M_{ii} = i$ and $M_{ij} = 0$.

Differentiation

$Dv = \frac{\delta}{\delta x} f(x)$

While this operator is not hermitian, the following is:

$-iDv = \frac{\delta}{\delta x}[-if(x)]$

79.3 Eigenvalues and eigenvectors of infinite-dimensional vectors

79.3.1 Spectral theorem for infinite-dimensional vector spaces

79.4 Forms on infinite-dimensional vector spaces

79.4.1 Forms on real functions

A form takes two vectors and produces a scalar.

Integration as a form

We can use integration to get a bilinear form.

$\int f(x)g(x)dx$

If we instead want a sesquilinear form we can instead use:

$\int f(\bar{x})g(x)dx$

79.4.2 Functionals

Functionals map functions to scalars. They are the 1-forms of infinite-dimensional vector spaces.

If we have a function f, we can write functional $J[f]$.

More

We can define neighbourhoods around a function f. For example, taking y to be f with infintesimal changes. to each of the values.

The difference between the functional at both points is

$\delta J = J[y] - J[f]$

Extrema

If

$\delta J = J[y] - J[f]$

is the same sign for all y around f, then J has an extremum at f.

Functional derivatives

79.4.3 Hilbert space

A complete space with an inner product. That is, a Banach space where the norm is derived from an inner product.

79.5 Calculus of variations

79.5.1 Calculus of variations

79.5.2 Functional integration

Integrate over possible functions?

79.6 Sort

79.6.1 Banach space

A complete normed vector space

79.6.2 Wave functions

For a vector in hermitian basis, for each eigenvector we have component. wave function is function on ith component.

79.7 Other

79.7.1 Dirac delta

Kronecker delta

The function is: δ_{ij}

If $i = j$ this is 1. Otherwise it is 0.

We introduced this in linear algebra.

Dirac delta

The Dirac delta replaces the Kronecker delta for continuous functions.

That is, we want:

- $\delta(x \neq 0) = 0$
- $\delta(0) = +\infty$
- $\int_{-\infty}^{\infty} \delta(x) dx = 1$

Part XIV

Manifolds

Chapter 80

Topology of finite sets

80.1 Nearness functions

80.1.1 Topologies

80.1.2 Topologies on sets

T is a topology on set X if:

- $X \in T$
- $\varnothing \in T$
- Unions of T are in T
- Intersections of T are in T

80.1.3 Examples of topologies: The trivial topology

The trivial topology contains only the underlying set and the empty set.

80.1.4 Examples of topologies: The discrete toplogy

The discrete toplogy contains all subsets of the underlying set (is this the power set?)

80.2 Neighbourhoods

80.2.1 Neighbourhood topology

We have a set X.

For each element $x \in X$, there is a non-empty set of neighbourhoods $N \in \mathbf{N}(x)$ where $x \in N \subseteq X$ such that:

- If N is a subset of M, M is a neighbourhood.
- The intersection of two neighbourhoods of x is a neighbourhood of x.
- N is a neighbourhood for each point in some $M \subseteq N$

80.2.2 Topological distinguishability

If two points have the same neighbourhoods then they are topologically indistinguishable.

For example in the trivial topology, all points are topologically indistinguishable.

80.2.3 Open sets

U is an open set if it is a neighbourhood for all its points.

80.3 Open and closed sets

80.3.1 Limit points and closure

Limit points

A point x in the topological set X is a limit point for $S \subset X$ if every neighbourhood of x contains another point in S.

For example -1 is a limit point for the real numbers where S is $[0, 1]$ (or $(0, 1)$.

Closure

The closure of a subset of a topological space is the subset itself along with all limit points.

So the closure of $|x| < 1$ includes -1 and 1.

80.3.2 Boundries and interiors

The boundry of the subset S of a topology is the intersection with the closure of S with the closure of the complement of S.

So the boundry of both $(0, 1)$ and $[0, 1]$ are 0 and 1.

The interior of S is S without the boundry.

So the interior of $(0, 1)$ and $[0, 1]$ are both $(0, 1)$.

80.3.3 Closed sets

The complement of any open set is a closed set.

A set can be open, closed, both or neither.

80.4 Compactness

80.4.1 Covers

A space X is covered by a set of subsets of X, C, if the union of C is X.

80.4.2 Subcover

A subset of C which still covers X is a a subcover.

80.4.3 Open cover

C is an open cover if each member is an open set.

80.4.4 Universal cover

80.4.5 Bases of topologies

Subset B of topology T is a base for T if all elements of T are unions of members of B.

Second-countable space

If B is finite then the toplogy is a second-countable space.

80.5 Separation

80.5.1 Connected and separated sets

Two subsets of X in topological space T are separated if each subset is disjoint from the other's closure.

So $[-1, 0)$ and $(0, 1)$ are separated.

$[-1, 0]$ and $(0, 1)$ are not separated.

Sets which are not separated are connected.

80.6 Cartesian products

80.6.1 Box topology

80.6.2 Product topology

80.7 Creating topologies from sets

80.7.1 The trivial topology

A topology which contains just X and $\varnothing$ is the trivial topology.

80.7.2 Discrete topology

80.8 Taxonomy of spaces

80.8.1 Lindelf space

In a Lindelf space all open covers have countable subcovers.

This is weaker than compactness, which requires that every open cover has a finite subcover.

80.8.2 Kolmogorov space

In a Kolmogorov (or T_0) space, for every pair of points there is a neighbourhood containing one but not the other.

80.9 Local properties

80.9.1 Local properties

Locally, a topology may have properties which are not present globally.

80.9.2 Locally compact spaces

80.9.3 Locally connected spaces

80.10 TO INF

80.10.1 Hausdorf space

In a Hausdorf (or T_2) space, any two different points have neighbourhoods which are disjoint.

80.10.2 Compact spaces

A space X is compact if each open cover has a finite subcover.

If we can define a cover which does not have a finite subcover, then the space is not compact.

For example an infinite cover could be tend towards $(0, 1)$, eg as $\frac{1}{n}, 1 - \frac{1}{n}$

This covers $(0, 1)$, but there is no finite subcover. As a result $(0, 1)$ is not compact.

Chapter 81

Topological manifolds

81.1 Introduction

81.1.1 Manifolds, charts and atlases

A manifold is a set of points and associated charts.

A chart is a mapping from each point in a subset of the manifold to a point in a vector space.

These charts are invertible. If we are given coordinates, we can identify the point in the manifold it comes from.

For each point we have a topological neighbourhood. For each point in the neighbourhood, we can map to an element in the tangent space.

Example: The sphere

We can map a hemisphere to a subset of R^2. Given a point in R^2 we can identify a specific point on the hemisphere, and given a s specific point on the hemisphere we can identify a point in R^2.

Universal charts

If the vector space is flat and non-repeating, then a single chart can be used to map the whole manifold.

Atlases

If we have a collection of charts which covers each point needs to be covered at least once, we have an atlas. Each chart needs to be to the same dimensional vector space.

81.1.2 Transition maps

Where two charts overlap we can express the points where the charts overlap as two different coordinates.

We can express the mapping from these coordinates as a function. This is a transition map.

Overlapping charts

If two charts cover some of the same points on a manifold then we can define a function for those points where we move from one vector to another.

We can represent moving between charts as:

ab^{-1}

81.1.3 Mapping 2D manifolds to Riemann surfaces

Needs to be orientable and metricisable.

81.1.4 Connections of topological manifolds

Connected vs path-connected topological manifods.

81.2 Dimension theory

81.2.1 Refinement

81.2.2 Ply (order) of a cover

81.2.3 Small inductive dimension

81.2.4 Large inductive dimension

81.2.5 Lebesgue covering dimension

81.3 Paths

81.3.1 Paths and loops

Paths

We have the set X. We define a mapping $[0,1] \to X$

If a path exists between any two points, then the space is path-connected.

Loops

This is a path which ends on itself.

If $f(0) = f(1)$ then it is a loop.

81.3.2 Holes and genuses

Holes

Genes

The genus of a topology is the number of holes in the topology.

81.3.3 Path-connect spaces

81.4 Simply-connected 2D manifolds

81.4.1 Elliptic (Riemann sphere)

81.4.2 Parabolic (complex plane)

81.4.3 Hyperbolic (open disk)

81.5 Not simply-connected 2D manifolds

81.5.1 Torus

81.5.2 Hyper-elliptic curves

81.6 Functions between topologies

81.6.1 Functions between topologies

We can define a function from topology to another.

$f(X) = Y$

Continuous functions between topologies

If $f(X)$ is continous, then we have a continous function between topologies.

Inverse functions between topologies

If $f(X)$ is invertible then there is a inverse mapping.

81.6.2 Homotopy

81.6.3 Homeomorphisms

If there is a mapping which is invertible and continuous, it is a homeomorphism.

81.7 Fibre bundles

81.7.1 Vector bundles

A vector bundle consists of a base manifold (a base space), and a real vector space at each point in the base manifold.

Example

For example we can have a base manifold of a circle, and have a 1-dimensional vector space at each point on the circle to create an infinitely extended cylinder.

81.7.2 Bundle projection

This is a projection from any point on any of the fibres, to the underlying base manifold.

81.7.3 Trivial and twisted bundles

81.7.4 Cross-sections and zero-sections of fibre bundles

81.7.5 Trivial bundles and the torus

Trivial bundles

The torus

$S_1 \times S_1$

81.7.6 Twisted bundles and the Klein bottle

Twisted bundles

Klein bottles

$S_1 \times S_1$, but twisted

81.7.7 Mobius strips

$S_1\times$ line segment.

81.8 Other

81.8.1 Submanifolds

Submanifold: subset of manifold which is also manifold

Eg: circle inside a sphere

81.8.2 Boundries and interiors

Around every manifold of dimension n is a boundry of dimension $(n-1)$.

Homeomorphism at boundry: one coordinate always ≥ 0. reduced dimension.

Interior is rest.

81.8.3 Embeddings and immersions

Whitney embedding theroem: all manfiolds can be embedded in R^n space for some n.

81.8.4 Topological groups

We have two operations for groups: multiplication and inversion.

A group is topological if these functions are continuous. + need to just read up on this. where is this relevant? + topological space

For these functions to be continous we need a metric defined on the group.

Chapter 82

Differentiable manifolds

82.1 Introduction

82.1.1 Differentiable transition maps

Transition map recap

Given two charts with an overlap, we have a transition mapping between the two charts of the overlap, where the mapping corresponds to a position on the manifold.

Differentiable transition maps

If this mapping is differentiable, we have a differentiable manifold.

Smooth manifolds

If transition maps are smooth (C^∞) then the manifold is smooth.

82.1.2 Differentiable and smooth manifolds

82.1.3 Diffeomorphisms

82.2 Tangent space

82.2.1 Tangent space and tangent vectors

Take a topological space: can all subsets in the toplogy be mapped to n dimensional space? if so, manifold

For this we need openness: a graph for example isn't open and so isn't a manifold

We also need the same number of dimensions at each point

Isn't always the case. eg two circles conneceted by a line is not a manifold. it's 2d in circles, 1d on line (and 3d at connections)

We have a homeomorphism from each point in the toplogy to an n dimensional coordinate system

We also have homeomorphisms of transformation maps, between different points on the topology

The vector space from the homeomorphism is tangent to the manifold at that point. the set of all tangents forms a tangent space

Interior: M; boundry δM Tangent on a manifold:

The tangent space of manifold M at point p is denoted TM_p.

If we have a normal field

$v=v^ie_i$

Then we can differentiate wrt a direction x.

$$\frac{\delta}{\delta x}v=\frac{\delta}{\delta x}v^ie_i$$

$$\frac{\delta}{\delta x}v=e_i\frac{\delta v^i}{\delta x}$$

Because the basis does not change.

If the basis does change we instead have:

$$\frac{\delta}{\delta x}v=\frac{\delta}{\delta x}v^ie_i$$

$$\frac{\delta}{\delta x}v=e_i\frac{\delta v^i}{\delta x}+v^i\frac{\delta e_i}{\delta x}$$

General point. basis can vary across manifold

After this basis diff

Tangent space as vector bundle

Christoffel symbols (page)

Christoffel symbols are connections.

The torsion tensor (own page)

Torsion tensor is

$T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}$

If torsion is 0, then the connection is symmetric.

Basis of tangent space

We can use as the basis for tangent space:

$\{\frac{(\delta}{\delta x^1})_p, (\frac{\delta}{\delta x^2})_p, ...\}$

This means we can write a tangent vector as:

$u = u^i(\frac{\delta}{\delta x^i})_p$

Basis of contant space

We can use as the basis for the contangent space:

$\{dx^1, dx^2, ...\}$

Metric on the tangent space (to Riemann)

Basis of metric (to Riemann)

The metric depends on the basis too:

$g_{ij}(p) = g((\frac{\delta}{\delta x^i})_p, (\frac{\delta}{\delta x^j})_p)$

The metric on two tangent vectors is defined on the components.

$g = g_{ij}(p)u^i v^j$

82.2.2 Cotangent space and cotangent vectors

82.3 Connections

82.3.1 Transport

82.3.2 Covariant derivative

Essentially as we move across path, we are changing the basis.

We can look at how basis vector change as we translate

We can define as basis as:

$e_i = \frac{\delta x}{\delta x_i}$

How to measure transport

If we take a vector and move it around a curved surface and return it to the same point, it may not face the same way

Eg if you're on equator, move east, north, south to equator, you'll face diffrent direction

This is true on smaller movements of a curved surface

We can use this to measure curvature of a manifold without coordinates

82.3.3 New

covariant derivative. how does change in field compare to parallel transport from curent position?

$\nabla_v(X) = \lim_{t\to 0} \frac{X(p+tv)-X(p)}{t}$

We have point p. We can compare how field in tangent space varies in direction of v.

we don't define basis as each point, but rather how basis changes as you move along a curve

82.3.4 Affine connections

If we have a tangent vector at one point of the manifold, we can map it to a tangent vector at a nearby point on the manifold.

We can use chain rule. so we can have coordinate maps where there is no overlap.

Smooth connections

Affine connection

We have a vector in a tangent space

We have a curve on the manifold from the start point

As we "roll" the tangent, there is a unique vector in each new tangent, determined by transition map

These are affine transformations

Given two points, what path? what transfomration? if curuved then different paths will given different transformation.

82.3.5 Parallel transport

Move and therefore change basis, but components are the same.

82.4 Sort

82.4.1 Orientability of surfaces

Chapter 83

Riemann manifolds

83.1 Introduction

83.1.1 Metric tensors

A metric tensor assigns a bilinear form to each point on the manifold.

We can then take two vectors in the tangent space and return a scalar.

83.1.2 Riemann manifolds and pseudo-Riemann manifolds

Riemann manifolds

Metric is positive definite.

Pseudo-Riemann manifolds

The metric isn't necessarily positive definite.

83.1.3 Metric tensor field

metric tensor field assigns a metric tensor to each point. metric tensor is defined on the tangent bundle. so we have metric on each tangent bundle, but the metric can change thoughout the manifold

83.1.4 Length of paths in Riemann manifolds

We can work out the length of a path through a Riemann manifold.

The geodesic is the shortest such path.

The Riemann metric between two points is the length of the geodesic.

83.2 Connections on Riemann manifolds

83.2.1 Metric compatibility

If we have two vectors in the tangent space of a manifold with a metric tensor, we can get a scalar:

$v^i u^j g_{ij}$

Transported metric

If we transport two vectors along a connection, we have the metric at the new point.

Metric preserving connections

If the connection preserves the metric, then the connection is metric compatible.

83.2.2 Torsion tensor

83.2.3 The Levi-Civita connection

For any metric tensor there is only one connection which preserves the metric and is torsion free.

83.3 Sort

83.3.1 The circle as a topology

83.3.2 Cylinders

83.3.3 Embeddings and immersions

83.3.4 Conformal maps

83.3.5 Geodesics

How do we have straight line on a curve? eg going round equator, but not going via uk.

Take start direction and find tangent vectors. geodesic is where tangent vectors stay parallel.

83.3.6 Curvature tensor

83.3.7 Ricci curvature

Part XV

Other topology

Chapter 84

Measure space

84.1 Defining measure spaces

84.1.1 Measure space

In a metric space, the structure was defining a value for each two elements of the set.

In a measure space, the structure defines a value of subsets of the set.

A measure space includes the set X, subsets of the set, Σ, and a function μ which maps from Σ to $\mathbb{R}$.

Sigma algebra

Requirement for Σ.

84.1.2 Axioms for measures

Measures are non-negative

$\forall E \in \Sigma : \mu(E) \geq 0$

The measure for the null set is 0.

$\mu() = 0$

Disjoint sets are additive

$\mu(\vee_{k=1}^{\infty} E_k) = \sum \mu(E_k)$

Where all elements E_k are disjoint. That is, they have no elements in common.

84.2 Examples of measure spaces

84.2.1 The counting measure

$\mu(E)$

This provides the number of elements in E.

84.2.2 The probability measure

This is discussed in more detail in Statistics.

Chapter 85

Graph theory

85.1 Undirected graphs

85.1.1 Vertices and edges

A graph is a set of vertices V, a set of edges E which are subset of pairs from V.

undirected so each edge is a set

Degree of a vertex

The degree of a vertex is the number of edges connections to it.

85.1.2 Order and size of graphs

The order of a graph is the number of vertices, $|V|$.

The size of a graph is the number of edges, $|E|$.

85.1.3 Subgraphs

We can take a subset of vertices, and all edges which only depend on these vertices. This is an induced subgraph.

Induced subgraph

85.1.4 Loops, multiple edges and simple graphs

Loops

A loop is an edge where both the vertices are the same.

Multiple edges

If there are two edges with the same pair of indices, there are multiple edges.

Simple graphs

No loops or multiple edges.

85.2 Directed graphs

85.2.1 Direct acyclic graphs

85.3 Weighted graphs

85.3.1 Edge-weighted graph

An edge-weighted graphs has weights for each edge.

85.3.2 Vertex-weighted graph

A vertex-weighted graph has weights for each vertex.

85.4 Graph representation

85.4.1 Adjacency matrix

We can represent a finite graph as a square matrix. m_{ij} is the number of edges connecting vertex i to vertex j.

85.4.2 Incidence matrix

An incidence matrix has m_{ij} representing the number of connections from vertex i to edge j.

85.4.3 Degree matrix

A degree matrix is a diagonal matrix. Each diagonal contains the degree of the a vertex.

85.4.4 Laplacian matrix

The Laplacian matrix L is formed using the degree matrix D and the adjacency matrix A. $L = D - A$.

85.5 Representing manifolds

85.5.1 Nearest-neighbour graph

85.5.2 Triangular mesh

Chapter 86

Category theory

86.1 Category theory

86.1.1 Introduction