

# M1J2 Summary Notes

JMC Year 1, 2017/2018 syllabus

Fawaz Shah (original notes by Dr Marie-Amelie Lawn)

(Dr Lawn refers to propositions, theorems, corollaries and lemmas. In this document I will refer to them all as 'theorems'.)

This document contains a list of definitions and a list of theorems.

Note that the exam will probably require you to PROVE some of these theorems, so you should refer back to the original notes for the proofs.

Boxes cover content in more detail. Titles of some theorems are given in italics.

## Contents

<b>I. Abstract Linear Algebra</b>	<b>3</b>
<b>1. Definitions</b>	<b>3</b>
<b>2. Theorems</b>	<b>5</b>
2.1. Vector spaces . . . . .	5
2.2. Subspaces . . . . .	6
2.3. Spanning sets, linear independence, bases, dimension . . . . .	6
2.4. Linear maps . . . . .	8
<b>II. Group Theory</b>	<b>11</b>
<b>3. Definitions</b>	<b>11</b>
<b>4. Theorems</b>	<b>14</b>
4.1. Groups . . . . .	14

4.2. Modular arithmetic and $\mathbb{Z}_n$ . . . . .	14
4.3. Cyclic groups . . . . .	14
4.4. Symmetric groups . . . . .	15
4.5. Subgroups . . . . .	16
4.6. Cosets and Lagrange's Theorem . . . . .	16
<b>III. Analysis</b>	<b>17</b>
<b>5. Definitions</b>	<b>17</b>
<b>6. Theorems</b>	<b>21</b>
6.1. Sequences . . . . .	21
6.2. Subsequences . . . . .	22
6.3. Summability . . . . .	23
6.4. Power series . . . . .	24
6.5. Continuity . . . . .	25
6.6. Differentiable functions . . . . .	25

# Part I.

## Abstract Linear Algebra

### 1. Definitions

**Vector space** A vector space is a set  $V$  coupled with:

- a function  $+: V \times V \rightarrow V$  (addition)
- a function  $\cdot: \mathbb{R} \times V \rightarrow V$  (scalar multiplication)

(For the rest of this part, we will assume  $V$  is a vector space)

**Subspace** A subset  $U \subseteq V$  is a subspace if:

- $\mathbf{0}_V \in U$
- If  $\mathbf{x}, \mathbf{y} \in U$  then  $\mathbf{x} + \mathbf{y} \in U$  (closure under addition)
- If  $\mathbf{x} \in U$  then for all  $\lambda \in \mathbb{R}$ ,  $\lambda\mathbf{x} \in U$  (closure under scalar multiplication)

**Linear combination** A linear combination of a set of vectors  $\{\mathbf{v}_1 \dots \mathbf{v}_n\}$  is any vector  $\mathbf{x}$  of the form:

$$\mathbf{x} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n \quad (1)$$

for some real numbers  $\lambda_1 \dots \lambda_n$

**Span** The span of a set  $S \subseteq V$  is the set of all linear combinations of elements of  $S$ . We define  $\text{span}(\emptyset) = \{\mathbf{0}_V\}$ .

**Spanning set** A subset  $S \subseteq V$  is called a spanning set of  $V$  if  $\text{span}(S) = V$ .

**Linear dependence** A subset of vectors  $\{\mathbf{v}_1 \dots \mathbf{v}_n\} \subseteq V$  is linearly dependent if there exists some real numbers  $\lambda_1 \dots \lambda_n$  (which are not just all 0s) such that:

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n = \mathbf{0}_V \quad (2)$$

**Basis** A basis of a vector space is a linearly independent spanning set.

We can also think of a basis as a spanning set of minimum possible size, or a linearly independent set of maximum possible size (theorems to show this later).

**Standard basis of  $\mathbb{R}^n$**  We define the standard basis elements of any  $\mathbb{R}^n$  to be:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \dots e_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (3)$$

The standard basis of  $\mathbb{R}^n$  is therefore  $\{e_1, e_2 \dots e_n\}$ .

**Dimension** The dimension of a vector space is the size of any basis of that vector space.

**Linear map** Let  $U$  and  $V$  be vector spaces. A linear map is a function  $f : U \rightarrow V$  such that:

- for all  $\mathbf{x}, \mathbf{y} \in U$ ,  $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$
- for all  $\mathbf{x} \in U$  and  $\lambda \in \mathbb{R}$ ,  $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$

**Image** The image of a linear map  $f : U \rightarrow V$  is the set of all  $f(\mathbf{u}) \in V$  where  $\mathbf{u} \in U$ .

$$\text{image}(f) = \{f(\mathbf{u}) \mid \mathbf{u} \in U\} \quad (4)$$

**Kernel** The kernel of a linear map  $f : U \rightarrow V$  is the set of all  $\mathbf{u} \in U$  such that  $f(\mathbf{u}) = \mathbf{0}_V$ .

$$\text{kernel}(f) = \{\mathbf{u} \mid \mathbf{u} \in U, f(\mathbf{u}) = \mathbf{0}_V\} \quad (5)$$

**Isomorphism** A linear map  $f : U \rightarrow V$  is an isomorphism if it is bijective. We say  $U \simeq V$ .

**Rank** The rank of  $f$  is defined as  $\dim(\text{image}(f))$ .

**Nullity** The nullity of  $f$  is defined as  $\dim(\text{kernel}(f))$ .

$T_A$  We define a function  $T_A$  that pre-multiplies a vector by a matrix  $\mathbf{A}$ :

$$T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{v} \mapsto \mathbf{A}\mathbf{v}, \mathbf{A} \in \text{Mat}_{m \times n}(\mathbb{R}) \quad (6)$$

where  $\text{Mat}_{m \times n}(\mathbb{R})$  denotes the set of all  $m \times n$  matrices with real entries. Note that if  $\mathbf{A}$  is an  $m \times n$  matrix, then  $T_A$  transforms a vector in  $\mathbb{R}^n$  to a vector in  $\mathbb{R}^m$ .

**Matrix representing  $f$**  Following from the previous definition, if we have:

- $f : U \rightarrow V$
- $B$  is a basis of  $U$
- $C$  is a basis of  $V$
- There is an isomorphism  $f_B : \mathbb{R}^n \rightarrow U$
- There is an isomorphism  $f_C : \mathbb{R}^m \rightarrow V$

We can define a version of  $T_A$  that is equivalent to  $f$ . The matrix  $\mathbf{A}$  in this case is called the matrix representing  $f$  with respect to  $B$  and  $C$ . This is denoted by:

$$\mathbf{A} = [f]_B^C \quad (7)$$

**Change-of-basis matrix** Let  $B$  and  $C$  be two bases for  $V$ . The matrix:

$$\mathbf{A} = [\text{Id}_V]_B^C \quad (8)$$

is called the change-of-basis matrix from  $B$  to  $C$ .  $\text{Id}_V$  denotes the identity function in the vector space  $V$  (maps every vector to itself).

In this case the linear map  $T_A$  will convert a vector given with respect to the basis  $B$  into a vector with respect to the basis  $C$ .

**'Vector with respect to a basis'** If we have an  $n$ -dimensional vector space  $V$  and a basis  $B = \{\mathbf{b}_1 \dots \mathbf{b}_n\}$ , then we say any  $\mathbf{v} \in V$  is given with respect to  $B$  if:

$$\mathbf{v} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}, \quad \mathbf{v} = \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \dots + \lambda_n \mathbf{b}_n \quad (9)$$

## 2. Theorems

### 2.1. Vector spaces

*Vector space axioms*

- $(V, +)$  is an Abelian group (the identity element being  $\mathbf{0}_V$ )
- for any  $\mathbf{v} \in V$ ,  $1\mathbf{v} = \mathbf{v}$
- for any  $\mathbf{v} \in V, \lambda, \mu \in \mathbb{R}, \lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$  (commutative w.r.t. scalar multiplication)

- for any  $\mathbf{u}, \mathbf{v} \in V, \lambda \in \mathbb{R}, \lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$  (scalar multiplication distributes over addition)
- for any  $\mathbf{v} \in V, \lambda, \mu \in \mathbb{R}, (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$  (scalar multiplication distributes over scalar addition)

For any  $\mathbf{v} \in V$ :

- For any  $n \in \mathbb{Z}, n\mathbf{v} = \mathbf{v} + \mathbf{v} + \dots + \mathbf{v}$  ( $n$  times)
- $0\mathbf{v} = \mathbf{0}_V$
- $(-1)\mathbf{v}$  is the additive inverse of  $\mathbf{v}$

## 2.2. Subspaces

Every vector space  $V$  has two trivial subspaces, itself and  $\{\mathbf{0}_V\}$ .

For any subspaces  $U, W \subseteq V$ :

- $U \cap W$  is a subspace
- $U \cup W$  is NOT a subspace

Any  $U \subseteq V$  is a subspace iff every linear combination of vectors in  $U$  is again in  $U$  (i.e.  $\text{span}(U) \subseteq U$ ).

For any  $S \subseteq V$ ,  $\text{span}(S)$  is a subspace.

If  $U \subset V$  is a subspace and  $S \subset U$  then  $\text{span}(S) \subset U$ .

## 2.3. Spanning sets, linear independence, bases, dimension

Every element of a vector space  $V$  can be written as a unique linear combination of basis vectors (for any basis).

For any set  $S \subseteq V$ :

- If  $\mathbf{v}_1 = \lambda\mathbf{v}_2$  for any  $\mathbf{v}_1, \mathbf{v}_2 \in S$  then  $S$  is linearly dependent
- If  $\mathbf{0}_V \in S$  then  $S$  is linearly dependent

If a set  $S$  is linearly independent/dependent then any subset of  $S$  is also linearly independent/dependent respectively.

A vector space is finite dimensional if it contains a finite spanning set.

Every finite spanning set contains a basis.

Therefore, a vector space is finite dimensional if it has a finite basis.

If a finite dimensional vector space has a basis, then there exists a finite dimensional spanning set.

If  $S \subseteq V$  is a linearly DEPENDENT spanning set, there exists some  $\mathbf{s} \in S$  such that  $S - \{\mathbf{s}\}$  is still a spanning set.

In other words, we can keep removing elements from a spanning set until it is linearly independent. At this point the spanning set is now a basis, by definition. This gives us our alternate definition of a basis as a spanning set of minimum size.

*Steinitz exchange lemma - base case*

Let  $S \subset V$  be a spanning set, and let  $\mathbf{v} \in V$ . There always exists an  $\mathbf{s} \in S$  such that

$$(S \setminus \{\mathbf{s}\}) \cup \{\mathbf{v}\} \quad (10)$$

is still a spanning set.

*Steinitz exchange lemma - in full*

Let  $S \subset V$  be a spanning set, and let  $\mathbf{v}_1 \dots \mathbf{v}_n \in V$  be a linearly independent set. There always exists some  $\mathbf{s}_1 \dots \mathbf{s}_n \in S$  such that

$$(S \setminus \{\mathbf{s}_1 \dots \mathbf{s}_n\}) \cup \{\mathbf{v}_1 \dots \mathbf{v}_n\} \quad (11)$$

is still a spanning set.

In other words, we can substitute in any linearly independent set, and  $S$  will still be a spanning set.

Any linearly independent set is smaller than or equal to any spanning set. (Consequence of Steinitz exchange lemma).

If  $L \subset V$  linearly independent and  $\mathbf{v} \notin \text{span}(L)$  then  $L \cup \mathbf{v}$  is linearly independent.

In other words, we can keep adding elements to a linearly independent set until it is a spanning set. At this point the linearly independent set is a basis, by definition. This gives us our alternate definition of a basis as a linearly independent set of maximum size.

If  $\dim(V) = n$  then every basis of  $V$  has size  $n$ .

If  $V$  is infinite-dimensional, we can always find a linearly independent subset of  $V$  with size  $n$ , for any  $n$ .

Any linearly independent set is contained in a basis.

Any linearly independent set  $L$  where  $\#L = \dim(V)$  is a basis.

If  $V$  is finite dimensional and  $U \subseteq V$ :

- $U$  is finite dimensional
- $\dim(U) \leq \dim(V)$
- if  $\dim(U) = \dim(V)$  then  $U = V$

## 2.4. Linear maps

(For the rest of this subsection assume  $f, g$  are linear maps, and let  $f : U \rightarrow V$ )

$g \circ f$  is also a linear map.

$$f(\mathbf{0}_U) = f(\mathbf{0}_V).$$

$\text{image}(f)$  is a subspace of  $V$ .

$\text{kernel}(f)$  is a subspace of  $U$ .

If  $f$  surjective then  $\text{image}(f) = V$ .

If  $f$  injective then  $\text{kernel}(f) = \{\mathbf{0}_U\}$ .

If  $f(\mathbf{x}) = \mathbf{y}$  then  $f^{-1}(\mathbf{y}) = \{\mathbf{x} + \mathbf{w} \mid \mathbf{w} \in \text{kernel}(f)\}$ .

If  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  then  $f \equiv T_A$  for some matrix  $\mathbf{A} \in \text{Mat}_{m \times n}(\mathbb{R})$ .



Specifically  $f : \lambda_1 \mathbf{e}_1 + \dots + \lambda_n \mathbf{e}_n \mapsto \lambda_1 f(\mathbf{e}_1) + \dots + \lambda_n f(\mathbf{e}_n)$

Therefore we can set:

$$\mathbf{A} = [f(\mathbf{e}_1) \mid f(\mathbf{e}_2) \mid \dots \mid f(\mathbf{e}_n)] \quad (12)$$

so that for any  $\mathbf{v} \in U$ :

$$T_A(\mathbf{v}) = \mathbf{A} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \lambda_1 f(\mathbf{e}_1) + \dots + \lambda_n f(\mathbf{e}_n) \quad (13)$$

Let  $g : U \rightarrow V$ , let  $B = \{\mathbf{b}_1 \dots \mathbf{b}_n\}$  be a basis of  $U$ .

If  $f(\mathbf{b}_i) = g(\mathbf{b}_i)$  for all  $\mathbf{b}_i$  then  $f \equiv g$ .

There is always a linear map between a basis of  $U$  and any set of vectors in  $V$ .

If  $U \simeq V$  then  $\dim(U) = \dim(V)$

If  $\dim(V) = n$  then  $f \simeq \mathbb{R}^n$ .

Let  $B = \{\mathbf{b}_1 \dots \mathbf{b}_n\}$  be a basis of  $U$  and  $C = \{f(\mathbf{b}_1) \dots f(\mathbf{b}_n)\}$  a subset of  $V$ :

- $\text{span}(C) = \text{image}(f)$
- $C$  is a spanning set  $\Leftrightarrow f$  is surjective
- $C$  is linearly independent  $\Leftrightarrow f$  is injective
- $C$  is a basis  $\Leftrightarrow f$  is bijective (aka an isomorphism)

If  $\dim(U) = \dim(V)$  then  $f$  bijective  $\Leftrightarrow f$  surjective  $\Leftrightarrow f$  injective

*Rank-Nullity Theorem*

$$\text{rank}(f) + \text{nullity}(f) = \dim(U)$$

Any  $f : U \rightarrow V$  can be represented as  $T_A$  for some matrix  $\mathbf{A}$ .

We denote  $\mathbf{A}$  by  $[f]_B^C$ , also known as the matrix representing  $f$  with respect to  $B$  and  $C$ .

*Steps for computing  $\mathbf{A}$ :*

Let  $B = \{\mathbf{b}_1 \dots \mathbf{b}_n\}$  be a basis of  $U$

Let  $C = \{\mathbf{c}_1 \dots \mathbf{c}_m\}$  be a basis of  $V$

We have isomorphisms:

$$f_B : \mathbb{R}^n \rightarrow U, \lambda_1 \mathbf{e}_1 + \dots + \lambda_n \mathbf{e}_n \mapsto \lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n$$

$$f_C : \mathbb{R}^m \rightarrow V, \lambda_1 \mathbf{e}_1 + \dots + \lambda_m \mathbf{e}_m \mapsto \lambda_1 \mathbf{c}_1 + \dots + \lambda_m \mathbf{c}_m$$

Note that the linear map  $(f_C)^{-1} \circ f \circ f_B$  sends vectors from  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ , therefore we can define:

$$T_A \equiv (f_C)^{-1} \circ f \circ f_B \quad (14)$$

since, from earlier,  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

1. Take basis vectors of  $U$  ( $\mathbf{b}_j$ ) in some order. Compute  $f(\mathbf{b}_j)$ . We have just applied  $f_B$ , followed by  $f$ .
2. Express each  $f(\mathbf{b}_j)$  as a linear combination of basis vectors of  $V$  ( $\mathbf{c}_i$ ).
3. Applying  $(f_C)^{-1}$  sends vectors in  $V$  to their coefficients w.r.t the basis vectors  $\mathbf{c}_i$ .

The matrix  $A$  is such that the  $j^{\text{th}}$  column of  $A$  is the vector  $(f_C)^{-1} \circ f \circ f_B(\mathbf{e}_j) = (f_C)^{-1} \circ f(\mathbf{b}_j)$

### *Change-of-basis matrix*

Let  $B$  and  $C$  both be bases of  $V$ . To convert a vector  $\mathbf{v} \in V$  given with respect to a basis  $B$  to a different basis  $C$ , we can premultiply by  $A = [Id_V]_B^C$ . If we denote the two representations as  $\mathbf{v}_B$  and  $\mathbf{v}_C$ :

$$\mathbf{v}_C = [Id_V]_B^C \mathbf{v}_B \quad (15)$$

Let  $B$  and  $B'$  be two bases of  $U$ ,  $C$  and  $C'$  two bases of  $V$ , and  $f : U \rightarrow V$ . The following equation holds:

$$[f]_{B'}^{C'} = [Id_V]_C^{C'} [f]_B^C [Id_U]_{B'}^B \quad (16)$$

# Part II.

## Group Theory

### 3. Definitions

**Binary operation** A binary operation on a set  $G$  is a any function  $f : G \times G \rightarrow G$

**Associative** A binary operation  $\star$  on a set  $G$  is associative if it satisfies:

$$(a \star b) \star c = a \star (b \star c) \quad (17)$$

for all  $a, b, c \in G$ .

**Commutative** A binary operation  $\star$  on a set  $G$  is commutative if it satisfies:

$$a \star b = b \star a \quad (18)$$

for all  $a, b \in G$ .

**Left/right identity** An element  $e \in G$  is called the left identity if:

$$e \star g = g \quad (19)$$

for all  $g \in G$ . Similar statement for right identity.

**(Two sided) Identity element** An element  $e \in G$  is a two-sided identity element if it is both a left identity and a right identity.

From now on the two-sided identity element will be referred to as  $e$ .

**Left/right inverse** An element  $h \in G$  is called the left inverse of  $g \in G$  if:

$$h \star g = e \quad (20)$$

Similar statement for right inverse.

**Two sided inverse** A two sided inverse of an element  $g \in G$  is both a left inverse and a right inverse of  $g$ .

From now on the two-sided inverse of  $g$  will be referred to as  $g^{-1}$ .

**Group** A group  $(G, \star)$  is a set  $G$  equipped with a binary operation  $\star$  such that:

- $\star$  is associative
- $\star$  has an identity element  $e \in G$
- Every  $g \in G$  has an inverse  $g^{-1} \in G$

The above three suffice for the exam, however there is technically a fourth requirement:

- $G$  is closed under  $\star$ , i.e. for all  $g, h \in G, g \star h \in G$

(For the rest of this part, we will assume  $(G, \star)$  is a group)

**Order (group)** The order of a group  $(G, \star)$  is the size of  $G$ .

**Abelian group** An Abelian group is a group with a commutative binary operation  $\star$ .

**Powers of  $g$**  We can define the powers of any  $g \in G$  to be:

$$g^n = \begin{cases} g \star g \star \dots g & n > 0 \\ g^{-1} \star g^{-1} \star \dots g^{-1} & n < 0 \\ e & n = 0 \end{cases} \quad (21)$$

where in the first cases there are  $n$  copies of  $g$ , and in the second case there are  $-n$  copies of  $g^{-1}$ .

**Definition of  $[a]_n$  and  $\mathbb{Z}_n$**  For any  $a \in \mathbb{Z}$ :

$$[a]_n = \{b \in \mathbb{Z} \mid b \equiv a \pmod{n}\} \quad (22)$$

Note that  $[a]_n$  forms an equivalence class, and there are exactly  $n$  of these equivalence classes.  $\mathbb{Z}_n$  is the set of all these equivalence classes.

$$\mathbb{Z}_n = \{[a]_n \mid a \in \mathbb{Z}\} \quad (23)$$

**Definition of  $\mathbb{Z}_n^*$**   $\mathbb{Z}_n^*$  is the set of all invertible  $[a]_n$ . Note in this case the identity element is  $[1]_n$ .

$$\mathbb{Z}_n^* = \{[a]_n \mid \exists [b]_n \in \mathbb{Z}_n \text{ s.t. } [a]_n [b]_n = [1]_n\} \quad (24)$$

Note that  $[a]_n [b]_n = [1]_n \Leftrightarrow \gcd(a, n) = 1$ .

**Order (element)** The order of any  $g \in G$  is the smallest positive integer such that:

$$g^n = e \quad (25)$$

**Cyclic group + generator** A group  $(G, \star)$  is cyclic if:

$$G = \{g^n \mid n \in \mathbb{Z}\} \quad (26)$$

$g$  is called the generator of the group.

**Permutation** A permutation  $\sigma$  on  $n$  symbols is a bijection:

$$\sigma : \{1 \dots n\} \rightarrow \{1 \dots n\} \quad (27)$$

**Symmetric group** The symmetric group  $S_n$  on  $n$  symbols is the set of all permutations of  $n$  symbols.

$$S_n = \{\sigma : \{1 \dots n\} \rightarrow \{1 \dots n\}\} \quad (28)$$

Note that  $S_n$  is a set of functions. Therefore the identity element is the identity function.

**$k$ -cycle** A permutation  $\sigma \in S_n$  is a  $k$ -cycle if there exists some  $a_1 \dots a_k \in \{1 \dots n\}$  such that:

$$\sigma(a_1) = a_2, \quad \sigma(a_2) = a_3 \quad \dots \quad \sigma(a_k) = a_1 \quad (29)$$

and  $\sigma(i) = i$  for all  $i \notin \{1 \dots n\}$ .  $k$  is called the length of the cycle. The notation for a cycle is  $(a_1 \dots a_k)$ .

**Disjoint cycles** Two cycles  $(a_1 \dots a_m)$  and  $(b_1 \dots b_n)$  are disjoint if no  $a_i$  is equal to any  $b_j$ .

**Subgroup** Let  $(G, \star)$  be a group, and  $H \subseteq G$ .  $(H, \star)$  is a subgroup of  $G$  if:

- $e \in H$
- For any  $g, h \in H$ ,  $g \star h \in H$
- For any  $g \in H$ ,  $g^{-1} \in H$

**Cyclic subgroup** Let  $(G, \star)$  be a group. For any  $g \in G$ , the cyclic subgroup  $\langle g \rangle$  generated by  $g$  is defined as:

$$\langle g \rangle = (\{g^i \mid i \in \mathbb{Z}\}, \star) \quad (30)$$

Note that order of  $g$  = size of cyclic subgroup  $\langle g \rangle$ .

**Left/right cosets** Let  $(G, \star)$  be a group and  $(H, \star)$  a subgroup. For any  $g \in G$ , the left coset of  $H$  by  $g$  (denoted by  $gH$ ) is defined as:

$$gH = \{g \star h \mid h \in H\} \quad (31)$$

Similar definition for right coset of  $H$  by  $g$  (denoted by  $Hg$ ).

The set of all left cosets of  $H$  by  $g$  is denoted by  $G : H$ .

The set of all right cosets of  $H$  by  $g$  is denoted by  $H : G$ .

## 4. Theorems

### 4.1. Groups

Any identity element  $e$  is unique for that group.

Any two-sided inverse  $g^{-1}$  of an element  $g \in G$  is unique.

For any  $g, h \in G$

$$(g \star h)^{-1} = h^{-1} \star g^{-1} \quad (32)$$

The normal exponent rules apply within groups, e.g.

$$g^n \star g^m = g^{n+m} \quad (33)$$

$$(g^n)^{-1} = g^{-n} \quad (34)$$

$$(g^n)^m = g^{nm} \quad (35)$$

Some examples of groups:  $(\mathbb{R}, +)$ ,  $(\mathbb{Z}, +)$ ,  $(\mathbb{Z}^*, \times)$

### 4.2. Modular arithmetic and $\mathbb{Z}_n$

$(\mathbb{Z}_n, +)$  is an Abelian group.

$(\mathbb{Z}_n^*, \cdot)$  is an Abelian group.

### 4.3. Cyclic groups

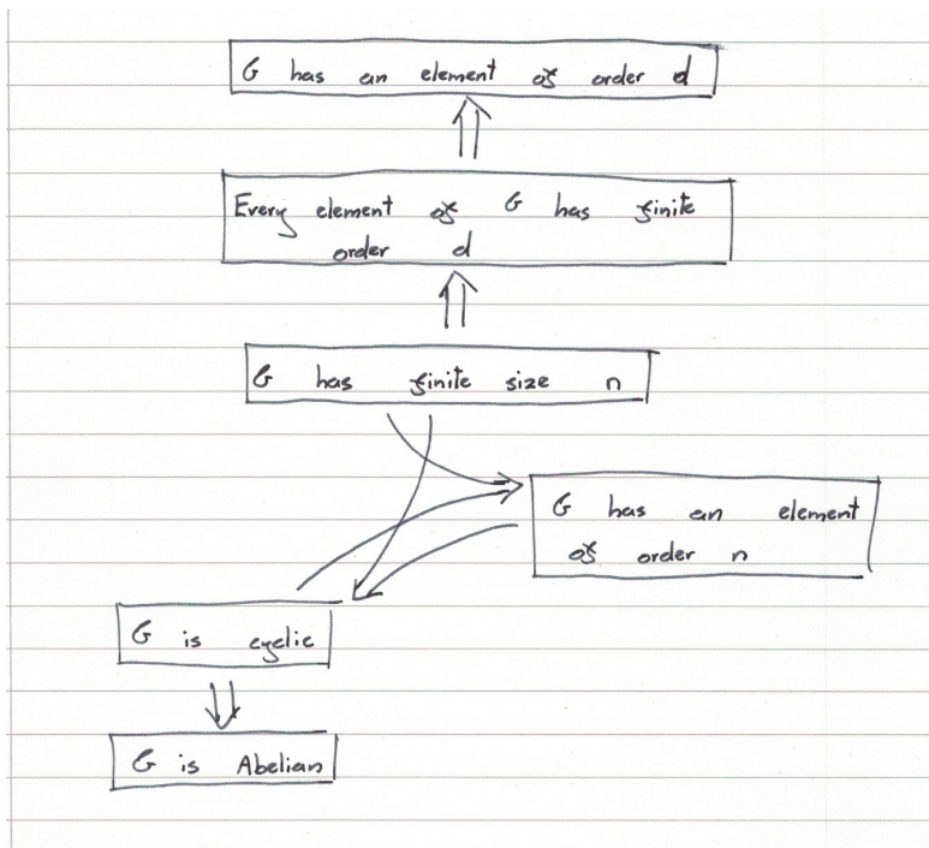
If  $(G, \star)$  is a finite group then every  $g \in G$  has finite order.

Any  $g \in G$  with order  $n$  has distinct powers  $g^0, g^1, g^2 \dots g^{n-1}$ .

All cyclic groups are Abelian.

Assume  $G$  is finite with size  $n$ .

$G$  is cyclic  $\Leftrightarrow G$  contains an element of order  $n$ .



#### 4.4. Symmetric groups

$(S_n, \circ)$  is a group.

The size of any  $S_n$  is  $n!$

The order of a  $k$ -cycle is  $k$ .

For any  $\sigma \in S_n$ :

- for any  $i \in \{0 \dots n\}$  there exists a  $d > 0$  such that  $\sigma^d(i) = i$  (i.e.  $\sigma^d \equiv Id = e$ )
- if  $d$  is the smallest integer such that  $\sigma^d(i) = i$  then the numbers  $i, \sigma^1(i), \sigma^2(i) \dots \sigma^{d-1}(i)$  are distinct
- If  $j$  is not in the set  $\{i, \sigma(i), \sigma^2(i) \dots \sigma^{d-1}(i)\}$  then neither is  $\sigma(j)$

Any permutation  $\sigma$  can be expressed as the product of disjoint  $k$ -cycles.

## 4.5. Subgroups

Any group  $(G, \star)$  has two trivial subgroups,  $(e, \star)$  and itself.

*Subgroup test*

Any  $H \subseteq G$  is a subgroup if:

- $H \neq \emptyset$
- for all  $x, y \in H, x \star y^{-1} \in H$

## 4.6. Cosets and Lagrange's Theorem

For any  $g_1, g_2 \in G$  and subgroup  $H$ :

$$g_1H = g_2H \Leftrightarrow g_1 \in g_2H \quad (36)$$

The left cosets of  $H$  form a partition of  $G$ . This means any  $g \in G$  is in exactly one left coset of  $H$ . The right cosets also form a (different) partition.

For any  $g \in G$ :

$$\#gH = \#hG = \#H \quad (37)$$

*Lagrange's Theorem*

For any subgroup  $(H, \star)$  where  $H \subseteq G$ :

$$\#G = \#H \cdot \#(G : H) \quad (38)$$

For any subgroup  $(H, \star)$ ,  $\#H$  divides  $\#G$ . (Consequence of Lagrange's Theorem).

For any  $g \in G$ , the order of  $g$  divides  $\#G$ . (Since the order of  $g$  is the size of the subgroup  $\langle g \rangle$ ).

If  $\#G = p$ , where  $p$  is prime, then  $G$  is cyclic.



# Part III.

## Analysis

### 5. Definitions

**Sequence** A sequence is simply a map  $f : \mathbb{N} \rightarrow \mathbb{R}$ , denoted by  $a_n$

**Limit (sequence)** A sequence  $a_n$  converges to a limit  $L$  if for all real numbers  $\epsilon > 0$ , there exists an  $N \in \mathbb{N}$  such that for all  $n > N$  we have  $|a_n - L| < \epsilon$ .

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \quad \text{s.t.} \quad \forall n > N \quad |a_n - L| < \epsilon \quad (39)$$

**Tends to infinity (sequence)** We say a sequence tends to infinity if for all  $R \in \mathbb{R}$ , the sequence  $a_n$  is eventually bigger than  $R$ .

$$\forall R \in \mathbb{R} \quad \exists N \in \mathbb{N} \quad \text{s.t.} \quad \forall n > N \quad a_n > R \quad (40)$$

**Shift** The shift of a sequence by say,  $k$ , is the sequence  $b_n = a_{n+k}$

**Triangle inequality** The general triangle inequality is:

$$|x - y| < |x - z| + |z - y| \quad (41)$$

Setting  $z = 0$  gives us:

$$|x - y| > |x| - |y| \quad (42)$$

Then setting  $y = -y$  gives us the familiar case:

$$|x + y| < |x| + |y| \quad (43)$$

**Bounded above** A sequence  $a_n$  is bounded above if there's a real number  $A$  such that  $a_n < A$  for all  $n$ .

**Bounded below** A sequence  $a_n$  is bounded below if there's a real number  $A$  such that  $a_n > A$  for all  $n$ .

**Bounded** A sequence  $a_n$  is bounded if there's a real number  $A$  such that  $|a_n| < A$  for all  $n$ .

**Increasing** A sequence is increasing if  $a_{n+1} \geq a_n$  for all  $n$ .

**Strictly increasing** A sequence is strictly increasing if  $a_{n+1} > a_n$  for all  $n$ .

**Decreasing** A sequence is decreasing if  $a_{n+1} \leq a_n$  for all  $n$ .

**Strictly decreasing** A sequence is strictly decreasing if  $a_{n+1} < a_n$  for all  $n$ .

**Monotonic** A sequence is monotonic if it is increasing or decreasing.

**Supremum (set)** The supremum  $A$  of a set  $S$  is the least upper bound of that set i.e. the smallest number such that  $s \leq A$  for all  $s \in S$ .

**Supremum (function)** The supremum of a function  $f$  is the sup of  $\{f(x) \mid x \in \text{dom}(f)\}$ .

**Infimum (set)** The infimum  $B$  of a set  $S$  is the greatest lower bound of that set i.e. the largest number such that  $s \geq B$  for all  $s \in S$ .

**Infimum (function)** The infimum of a function  $f$  is the inf of  $\{f(x) \mid x \in \text{dom}(f)\}$ .

**Subsequence** A subsequence of  $a_n$  is a sequence  $a_{f(n)}$ , where  $f(n)$  is a strictly increasing function.

**Cauchy sequence** A sequence is Cauchy if all the terms get arbitrarily close to one another. To put it mathematically:

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \quad \text{s.t.} \quad \forall m, n \geq N \quad |a_n - a_m| < \epsilon \quad (44)$$

**Partial sum** The  $n^{\text{th}}$  partial sum  $S_n$  of a sequence  $a_n$  is the sum of terms up to that point:

$$S_n = \sum_{i=1}^n a_n \quad (45)$$

**Summable** A sequence  $a_n$  is summable if the sequence of its partial sums converges. The limit of the sequence of partial sums will be:

$$L = \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n a_n \right) = \sum_{i=1}^{\infty} a_n \quad (46)$$

**Absolutely summable** A sequence  $a_n$  is absolutely summable if  $|a_n|$  is summable.

**Conditionally summable** A sequence is conditionally summable if it is summable but not absolutely summable.

**Power series** The power series associated with a sequence  $a_n$  is the sequence of partial sums:

$$\sum_{i=1}^n a_i x^i \quad (47)$$

**Radius of convergence** The radius of convergence  $R$  of a power series  $P(x)$  is defined as the largest  $x$  for which  $P(x)$  is convergent.

$$R = \sup\{x \in \mathbb{R} \mid P(x) \text{ convergent}\} \quad (48)$$

**Limit as  $x \rightarrow \infty$  (function)** A function  $f(x)$  tends to a limit  $L$  as  $x \rightarrow \infty$  if for all real numbers  $\epsilon > 0$ , there exists an  $R \in \mathbb{R}$  such that for all  $x \geq R$  we have  $|f(x) - L| < \epsilon$ .

$$\forall \epsilon > 0 \quad \exists R \in \mathbb{R} \quad \text{s.t.} \quad \forall x > R \quad |f(x) - L| < \epsilon \quad (49)$$

**Tends to infinity (function)** A function  $f(x)$  tends to infinity as  $x \rightarrow \infty$  if for any  $M \in \mathbb{R}$  there exists an  $R \in \mathbb{R}$  such that if  $x > R$  then  $f(x) > M$ .

$$\forall M \in \mathbb{R} \quad \exists R \in \mathbb{R} \quad \text{s.t.} \quad x > R \Rightarrow f(x) > M \quad (50)$$

**One-sided limit** A function  $f(x)$  tends to a limit  $L$  as  $x \rightarrow a^-$  if for any  $\epsilon > 0$  there exists a  $\delta > 0$  such that if  $x \in (a - \delta, a)$  then  $|f(x) - L| < \epsilon$ .

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{s.t.} \quad x \in (a - \delta, a) \Rightarrow |f(x) - L| < \epsilon \quad (51)$$

Same format for the other sided limit ( $x \rightarrow a^+$ )

(Note that  $\epsilon - \delta$  definition is only used for limits as  $x$  tends to a finite number  $a$ , not infinity)

**Limit as  $x \rightarrow a$**  A function  $f(x)$  tends to a limit  $L$  as  $x \rightarrow a$  if we have both:

$$\lim_{x \rightarrow a^-} f(x) = L \quad \text{and} \quad \lim_{x \rightarrow a^+} f(x) = L \quad (52)$$

**Limit as  $x \rightarrow a$  ( $\epsilon - \delta$  def.)** A function  $f(x)$  tends to a limit  $L$  as  $x \rightarrow a$  if:

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{s.t.} \quad |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon \quad (53)$$

**Continuous** A function  $f(x)$  is continuous at  $a$  if:

$$\lim_{x \rightarrow a} f(x) = f(a) \quad (54)$$

**Continuous ( $\epsilon$  -  $\delta$  def.)** A function  $f(x)$  is continuous at  $a$  if for all  $\epsilon > 0$  there is a  $\delta > 0$  such that if  $|x - a| < \delta$  then  $|f(x) - f(a)| < \epsilon$ .

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{s.t.} \quad |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon \quad (55)$$

**Continuous everywhere** A function  $f(x)$  is continuous everywhere if it is continuous at  $a$  for all  $a \in \text{dom}(f)$ .

**Open interval** An open interval  $I$  is a set  $I \subseteq \mathbb{R}$  of the form:

- $I = (a, b)$  for some  $a, b \in \mathbb{R}$ , or
- $I = (-\infty, b)$ , or
- $I = (a, +\infty)$ , or
- $I = \mathbb{R}$

**Discontinuity** Discontinuity is the negation of continuity. Hence a function  $f(x)$  is discontinuous at  $a$  if there exists  $\epsilon > 0$  such that for all  $\delta > 0$ ,  $|x - a| < \delta$  AND  $|f(x) - f(a)| > \epsilon$ .

$$\exists \epsilon > 0 \quad \text{s.t.} \quad \forall \delta > 0 \quad |x - a| < \delta \text{ AND } |f(x) - f(a)| > \epsilon \quad (56)$$

**Bounded (function)** A function  $f(x)$  is bounded if the set of all possible values of  $f(x)$  is bounded.

**Differentiable (ver. 1)** A function  $f(x)$  is differentiable at  $a$  if:

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \quad (57)$$

exists.

**Differentiable (ver. 2)** A function  $f(x)$  is differentiable at  $a$  if:

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h} \quad (58)$$

exists.

**Differentiable everywhere** A function  $f(x)$  is differentiable everywhere if it is differentiable at  $a$  for all  $a \in \text{dom}(f)$ .

**Global maximum** A function  $f(x)$  has a global maximum at  $a$  if  $f(a) \geq f(x)$  for all other values of  $f(x)$ .

Similar definition for global minimum.

**Local maximum** A function  $f(x)$  has a local maximum at  $a$  if  $f(a) \geq f(x)$  for all  $x$  in the set  $(a - \epsilon, a + \epsilon)$ , for some  $\epsilon$ .

Similar definition for local minimum.

**Lipschitz continuous** A function is Lipschitz continuous if:

$$|f'(x)| \leq L \Rightarrow |f(x_1) - f(x_2)| \leq L|x_1 - x_2| \quad (59)$$

## 6. Theorems

### 6.1. Sequences

Every convergent sequence has a unique limit.

Every convergent sequence is bounded.

If all terms of a convergent sequence are larger than a number  $B$ , then so is its limit.

Some properties of limits:

$$\lim_{x \rightarrow \infty} (a_n + b_n) = \lim_{x \rightarrow \infty} a_n + \lim_{x \rightarrow \infty} b_n \quad (60)$$

$$\lim_{x \rightarrow \infty} (\lambda a_n) = \lambda \lim_{x \rightarrow \infty} a_n \quad (61)$$

$$\lim_{x \rightarrow \infty} (a_n b_n) = \lim_{x \rightarrow \infty} a_n \lim_{x \rightarrow \infty} b_n \quad (62)$$

$$\lim_{x \rightarrow \infty} \left( \frac{a_n}{b_n} \right) = \frac{\lim_{x \rightarrow \infty} a_n}{\lim_{x \rightarrow \infty} b_n} \quad (63)$$

where  $\lambda$  is any real number.

If  $a_n \rightarrow \infty$  and  $b_n$  is bounded below,  $a_n + b_n \rightarrow \infty$ .

If  $a_n \rightarrow \infty$  and  $b_n$  is bounded below by a positive number,  $a_n b_n \rightarrow \infty$ .

If  $a_n$  is bounded and  $b_n \rightarrow \infty$ , then  $\frac{a_n}{b_n} \rightarrow 0$ .

If  $a_n \rightarrow \infty$ , for any real number  $\lambda$ :

- $\lambda < 0 \Rightarrow \lambda a_n \rightarrow -\infty$
- $\lambda = 0 \Rightarrow \lambda a_n \rightarrow 0$
- $\lambda > 0 \Rightarrow \lambda a_n \rightarrow \infty$

If  $a_n \rightarrow a$  and  $b_n \rightarrow b$ , and for all  $n$   $a_n < b_n$ , then  $a < b$ .

*Sandwich Theorem*

If  $a_n \leq b_n \leq c_n$  for all  $n$ , and  $a_n$  and  $c_n$  tend to the same limit  $L$ , then  $b_n \rightarrow L$ .

Every bounded monotonic sequence is convergent.

*Completeness Axiom*

Every non-empty subset of the real numbers which is bounded above has a supremum. Corresponding statement for infimum.

Useful results for sequences:

$$\lim_{n \rightarrow \infty} \lambda^n = \begin{cases} \infty & \lambda > 1 \\ 1 & \lambda = 1 \\ 0 & -1 < \lambda < 1 \end{cases} \quad (64)$$

$\lambda^n$  diverges if  $\lambda = -1$ .

If  $m > 0$  and  $\lambda > 1$  then  $\frac{\lambda^n}{n^m} \rightarrow \infty$  (exponentials beat powers).

If  $m > 0$  then  $\frac{\log(n)}{n^m} \rightarrow 0$  (powers beat logs).

## 6.2. Subsequences

If  $a_n \rightarrow L$  then any subsequence  $a_{f(n)} \rightarrow L$ .

If two subsequences of  $a_n$  converge to different limits,  $a_n$  doesn't converge to a limit.

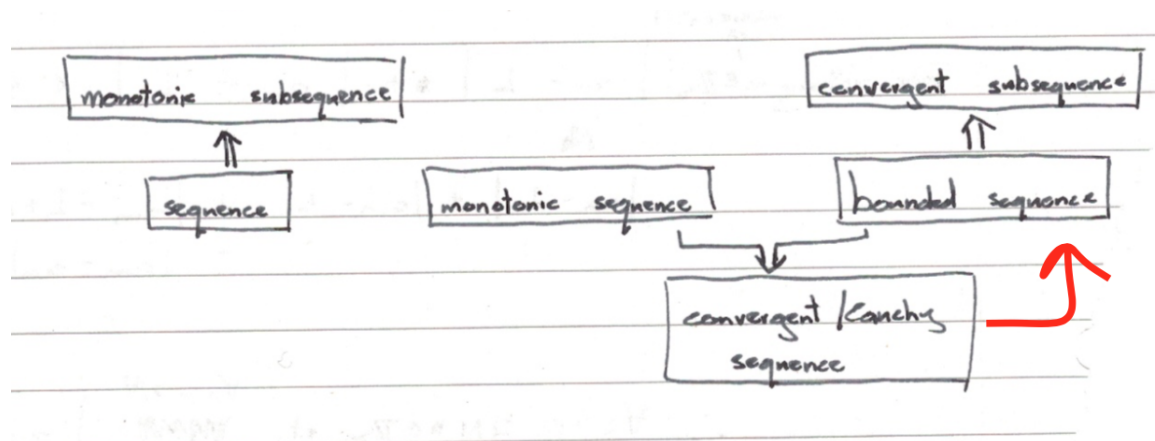
Every sequence has a monotonic subsequence.

*Bolzano-Weierstrass Theorem*

Every bounded sequence has a convergent subsequence.

Every Cauchy sequence is bounded.

Cauchy sequence  $\Leftrightarrow$  convergent sequence (for real numbers).



### 6.3. Summability

A sequence is summable iff the sequence of its partial sums converges.

If two subsequences of a sequence  $a_n$  converge to two different limits,  $a_n$  is not summable.

If  $a_n$  and  $b_n$  are summable with  $\sum_{i=0}^{\infty} a_i = a$  and  $\sum_{i=0}^{\infty} b_i = b$ :

- $a_n + b_n$  is summable with  $\sum_{i=0}^{\infty} (a_i + b_i) = a + b$ .
- $\lambda a_n$  is summable with  $\sum_{i=0}^{\infty} \lambda a_i = \lambda a$  (for any real number  $\lambda$ )

If  $b_n = a_{n+k}$  then  $a_n$  summable  $\Leftrightarrow b_n$  summable.

$a_n$  is summable  $\Rightarrow a_n \rightarrow 0$ .

Note that the converse is not necessarily true, take  $a_n = \frac{1}{n}$  as an example.

Let  $S_n$  denote the sequence of partial sums of  $a_n$  ( $S_n = \sum_{i=0}^n a_i$ ). A sequence of non-negative numbers  $a_n$  is summable iff  $S_n$  is bounded above. Similar statement for sequences of non-positive numbers.

Every absolutely summable sequence is summable.

#### *Comparison test*

If  $b_n > a_n$  for all  $n$  then  $b_n$  summable  $\Rightarrow a_n$  summable.

#### *Alternating series test*

If  $a_n$  is a decreasing sequence AND  $a_n \geq 0$  for all  $n$  AND  $a_n \rightarrow 0$  then  $(-1)^{n+1}a_n$  is summable.

#### *Ratio test for sequences*

Let  $r = \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$ :

- $r < 1 \Rightarrow a_n$  is absolutely summable
- $r > 1 \Rightarrow a_n$  is not summable
- $r = 1$  is an indeterminate case

### **6.4. Power series**

The power series associated with a sequence  $a_n$  converges iff the sequence of partial sums of  $a_n x^n$  converges (i.e. if  $\sum_{i=0}^n a_i x^i$  converges).

Let  $P(x)$  be a power series. If  $P(a)$  converges absolutely for some  $a$ , then  $P(x)$  converges absolutely for all  $x$  such that  $|x| < |a|$

Let  $R$  be the radius of convergence of  $P(x)$ . For all real numbers  $a$ :

- $|a| < R \Rightarrow P(a)$  converges absolutely
- $|a| > R \Rightarrow P(a)$  diverges

#### *Ratio test for power series*

Let  $r = \frac{a_{n+1}}{a_n}$ . Let  $P(x) = \sum_{i=0}^n a_i x^i$  (i.e. the power series associated with  $a_n$ ):

- $r \rightarrow 0 \Rightarrow R = \infty$
- $r \rightarrow L$  for some  $L \Rightarrow R = \frac{1}{L}$
- $r \rightarrow \infty \Rightarrow R = 0$

Note: if  $r = 1$  here then  $R = 1$ . This is DIFFERENT to the ratio test for sequences, where  $r = 1$  is an indeterminate case.



## 6.5. Continuity

The limit of a function at any specific point is unique.

If functions  $f$  and  $g$  are continuous at  $a$ :

- $(f + g)$  is continuous at  $a$
- $fg$  is continuous at  $a$
- $\frac{1}{f(x)}$  and  $\frac{1}{g(x)}$  are continuous at  $a$
- $g \circ f$  is continuous at  $a$

Any polynomial in  $\mathbb{R}$  is continuous

Any rational function in  $\mathbb{R}$  is continuous

*Sequential continuity*

A function  $f$  is continuous at  $a$  iff  $f(a_n) \rightarrow f(a)$  for all sequences  $a_n$  such that  $a_n \rightarrow a$ .

Any continuous function on a closed bounded interval is bounded.

*Intermediate Value Theorem*

If  $f$  continuous and  $f(a) \leq f(b)$  for some  $a, b$ , then there exists some  $c \in [a, b]$  such that  $f(a) \leq f(c) \leq f(b)$ .

*Fixed Point Theorem*

If  $f$  continuous and  $f : [a, b] \rightarrow [a, b]$ , then there exists some  $c \in [a, b]$  such that  $f(c) = c$ .

Polynomials of odd degree have at least 1 root.

$f$  differentiable  $\Rightarrow f$  continuous.

## 6.6. Differentiable functions

If functions  $f$  and  $g$  are differentiable at  $a$ :

- $(f + g)$  is differentiable at  $a$
- $fg$  is differentiable at  $a$
- $\frac{1}{f(x)}$  and  $\frac{1}{g(x)}$  are differentiable at  $a$
- $g \circ f$  is differentiable at  $a$

- $g^{-1}$  and  $f^{-1}$  are differentiable at  $a$

Let  $f$  be continuous and differentiable. If  $f$  has a local extremum at  $a$  then  $f'(a) = 0$  (except at endpoints of the interval).

Let  $f$  be continuous and differentiable. If  $f$  has a local extremum at  $c$  (say in the interval  $[a, b]$ ), there are 3 possibilities:

- $c$  is an endpoint of  $[a, b]$
- $f'(c) = 0$
- $c$  is a non-differentiable point

#### *Mean Value Theorem*

Let  $f$  be continuous on  $[a, b]$  and differentiable on  $(a, b)$ . There exists a point  $c \in (a, b)$  such that:

$$f'(c) = \frac{f(b) - f(a)}{b - a} \quad (65)$$

#### *Rolle's Theorem*

Let  $f$  be continuous and differentiable on  $(a, b)$ . If  $f(a) = f(b)$  then there exists some  $c \in (a, b)$  such that  $f'(c) = 0$ . This is a special case of the Mean Value Theorem.