Contents

Chapter 1

Differential and Integral Calculus

1.1 Numbers

1.1.1 Axioms for $\mathbb{R}$

The real number system consists of a set $\mathbb{R}$, and two binary operations $(a, b) \to a + b$ and $(a, b) \to a \cdot b$ such that the following axioms are satisfied:

A1) (associative law)

$$(a+b)+c = a+(b+c) \qquad \forall a, b, c \in \mathbb{R}$$

A2) (identity element)
There exists a unique element $0 \in \mathbb{R}$ such $a + 0 = a \quad \forall a \in \mathbb{R}$

A3) (inverse element)
$\forall a \in \mathbb{R}$ there exists a unique element $-a \in \mathbb{R}$ $a + (-a) = 0$

A4) (commutative law)

$$a+b = b+a \qquad \forall a, b \in \mathbb{R}$$

M1) (associative law)

$$(ab)c = a(bc) \quad \forall a, b, c \in \mathbb{R}$$

M2) (identity element)
There exists a unique element $1 \in \mathbb{R}$, $1 \neq 0$ such that $a \cdot 1 = a \quad \forall a \in \mathbb{R}$

M3) (inverse element)
$\forall a \in \mathbb{R}$ and $a \neq 0$ there exists a unique element $a^{-1} \in \mathbb{R}$ such that $aa^{-1} = 1$

M4) (commutative law)

$$ab = ba \quad \forall a, b \in \mathbb{R}$$

D) (distributive law)

$$a(b + c) = ab + ac \quad \forall a, b, c \in \mathbb{R}$$

Remark
A field is any set M together with two binary operations $(x, y) \to x + y$ and $(x, y) \to xy$ such that Axioms (A1)-(D) obtain.

Corollary 1.1.1
Let $a, b \in \mathbb{R}$,

i) then there is one and only number $x \in \mathbb{R}$ such that $a + x = b$. This number x is given by $x = b + (-a)$.

ii) then there is one and only number $x \in \mathbb{R}$ such that $ax = b$. This number x is given by $x = a^{-1}b$.

Remark
Let $a, b, c \in \mathbb{R}$.

i) $b + a = c + a \Rightarrow b = c$

ii) $ba = ca \Rightarrow b = c$ *if* $(a \neq 0)$

iii) $-(-a) = a$

iv) If $a \neq 0$, *then* $a^{-1} \neq 0$ *and* $(a^{-1})^{-1} = a$

v) $a \cdot 0 = 0 \qquad \forall a \in \mathbb{R}$

vi) We have $ab = 0$ *if and only if* $a = 0$ *or* $b = 0$ *or both.*

vii) If $ab \neq 0$, *then* $(ab)^{-1} = b^{-1}a^{-1}$

1.1.2 Order Properties of the Real Numbers

We also have a relation $<$ which satisfies the following axioms:
For all $a, b, c \in \mathbb{R}$ we have

O1) Exactly one of the relation holds:

$$a < b, \quad a > b, \quad a = b \qquad \text{(trichotomy)}$$

O2) $a < c$ and $b < c \Rightarrow a < c$ (transitivity)

O3) $a < b \Rightarrow a + c < b + c$
$a < b \quad , c > 0 \Rightarrow ac < bc$

Remark
We write

$$a > b \Leftrightarrow b < a$$
$$a \leq b \Leftrightarrow a < b \text{ or } a = b$$

Example
Let $a, b \in \mathbb{R}$.

i) $a < 0 \Rightarrow -a > 0$

ii) $a < 0, b < 0 \Rightarrow ab > 0$

iii) $a^2 > 0 \quad \forall a \in \mathbb{R}, a \neq 0$

Definition 1.1.2
For all $a \in \mathbb{R}$ we define

$$|a| := \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

the absolute value of a.

Theorem 1.1.3
For all $a, b \in \mathbb{R}$ we have

i) $|ab| = |a||b|$

ii) $|a + b| \leq |a| + |b|$

iii) $||a| - |b|| \leq |a - b|$

1.1.3 Completeness Axiom

Definition 1.1.4
Let $M \subset \mathbb{R}$.

i) $x \in \mathbb{R}$ is called an upper bound for the set M, if $x \geq m \quad \forall m \in M$.

ii) If there exists such a number x, then M is said to be bounded from above.

Definition 1.1.5 (Infimum, Supremum)
Let $M \subset \mathbb{R}, M \neq \emptyset$. $x \in \mathbb{R}$ is called a least upper bound, or a supremum $(x = \sup M)$ of M

i) if x is an upper bound for M.

ii) if $y \in \mathbb{R}$ *is any upper bound for* M*, then* $x \leq y$*.*

(The greatest lower bound $\inf M$ *is similar defined if* M *is bounded from below)*

If $\sup M = x \in M$*, we write* $x = \max M$ *and* $\inf M = y \in M$ *is denoted by* $y = \min M$*)*

Example
Let $a, b \in \mathbb{R}$*.*

i) $M = [a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$
$\Rightarrow a = \min M = \inf M, b = \max M = \sup M$

ii) $M = (a, b) = \{x \in \mathbb{R} | a < x < b\}$
$\Rightarrow a = \inf M, b = \sup M, \quad \inf M \notin M, \sup M \notin M$

iii) $M = \{x | x > 0 \text{ and } x^2 > 2\}$*. In this case there is no supremum, but* $\inf M = \sqrt{2}$*.*

Theorem 1.1.6 (Completeness axiom)
If a nonempty subset of real numbers has an upper bound, then it has a least upper bound. If a nonempty subset of real numbers has a lower bound, it has a greatest lower bound.

Example
There exists a real number $x \in \mathbb{R}$*,* $x > 0$ *such that* $x^2 = 2$*.*

1.1.4 Natural Numbers

Definition 1.1.7
A set $M \subset \mathbb{R}$ *is called an inductive set if*

i) $1 \in M$

ii) $a \in M \Rightarrow a + 1 \in M$

Definition 1.1.8
i) $\mathbb{N} = \{n \in \mathbb{R} : n \in M \text{ for every inductive set } M \subset \mathbb{R}\}$

ii) $\mathbb{Z} := \{z \in \mathbb{R} : z \in \mathbb{N} \text{ or } -z \in \mathbb{N} \text{ or } z = 0\}$

iii) $\mathbb{Q} := \{\frac{p}{q} | p \in \mathbb{Z} \text{ and } q \in \mathbb{Z}, q \neq 0\}$

Corollary 1.1.9
$\mathbb{N}$ *is an inductive set.*

Theorem 1.1.10 (Principle of mathematical induction)
Let $M \subset \mathbb{N}$ *such that*

i) $1 \in M$

ii) $a \in M \Rightarrow a+1 \in M$

Then $M = \mathbb{N}$.

Proof. $i), ii) \Rightarrow M$ is an inductive set $\Rightarrow \mathbb{N} \subset M$. Since $M \subset \mathbb{N} \Rightarrow M = \mathbb{N}$. $\square$

Example

i) $a_0 := 0, a_1 := 1$
$a_{n+1} := a_n + a_{n-1} \quad \forall n \in \mathbb{N}$

$$\Rightarrow a_n = \frac{1}{\sqrt{5}} \left\{ \left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right\} \quad \forall n \in \mathbb{N}_0$$

ii) *(binomial theorem) Let* $a, b \in \mathbb{R}, n \in \mathbb{N}$

$$(a+b)^n = \sum_{k=o}^{n} \binom{n}{k} a^k b^{n-k}$$

iii) *Let* $a, b \in \mathbb{R}, n \in \mathbb{N}$

$$a^n - b^n = (a-b) \sum_{k=1}^{n} a^{n-k} b^{k-1}$$

Theorem 1.1.11
Any nonempty subset of $\mathbb{N}$ *contains a smallest element.*

Theorem 1.1.12 (Archimedes)
If $a, b \in \mathbb{R}$ *and* $a > 0$, *then there exists* $n \in \mathbb{N}$ *such that* $na > b$.

1.1.5 Complex Numbers

Definition 1.1.13

i) *The set of all complex numbers is denoted by* $\mathbb{C}$.

$$\mathbb{C} = \{x + iy | x, y \in \mathbb{R}\} \quad (i^2 = -1)$$

Addition and multiplication of complex numbers are defined by

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2) \tag{1.1}$$
$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \tag{1.2}$$

ii) *The real part of* z: $\Re(z) = x$

iii) The imaginary part of z: $\Im(z) = y$

iv) The complex conjugate of z: $\bar{z} = x - iy$

v) The absolute value of z: $|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$

Remark

i) $\mathbb{C}$*, together with (1.1) and (1.2) is a field.*

ii) $i^2 = -1 < 0$

iii) $|\sum_{i=1}^n z_i| \leq \sum_{i=1}^n |z_i|$

Theorem 1.1.14 (Cauchy's inequality)
If $a_1 \dots a_n, b_1, \dots, b_n \in \mathbb{C}$*, then*

$$|\sum_{i=1}^n a_i \overline{b_i}|^2 \leq (\sum_{i=1}^n |a_i|^2)(\sum_{i=1}^n |b_i|^2)$$

Theorem 1.1.15 (Minkowski's inequality)
Let $a_1, a_2, \dots, a_n, b_1, \dots, b_n \in \mathbb{C}$*, then*

$$(\sum_{k=1}^n |a_k + b_k|^2)^{1/2} \leq (\sum_{k=1}^n |a_k|^2)^{1/2} + (\sum_{k=1}^n |b_k|^2)^{1/2}$$

Theorem 1.1.16 (Bernoulli's inequality)
Let $x \in \mathbb{R}, n \in \mathbb{N}$*. If* $x > -1$ *and* $x \neq 0$ *then*

$$(1 + x)^n > 1 + nx \quad (\forall n \geq 2)$$

Theorem 1.1.17 (Geometric Mean - Arithmetic Mean inequality)
Let $a_1, \dots, a_n \in \mathbb{R}$ *nonnegative real numbers, then*

$$\sqrt[n]{(a_1 a_2 \dots a_n)} \leq \frac{(a_1 + a_2 + \dots + a_n)}{n}$$

1.2 Sequences and Series

1.2.1 Sequences

Definition 1.2.1
Let M be a set and $f : \mathbb{N} \to M$. Then f is called a sequence in M. If $f(n) = a_n$, we write such a sequence as $(a_n)_{n=1}^{\infty} \subset M$.

Definition 1.2.2
A sequence $(a_n) \subset \mathbb{R}$ is said to converge if there exists $a \in \mathbb{R}$ such that

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} : \quad |a_n - a| < \epsilon \quad \forall n \geq N$$

In this case, we call a the limit for (a_n) and write $\lim_{n\to\infty} a_n = a$ or $a_n \to a$ as $n \to \infty$.

Example

i) $a_n = 2 \quad \forall n \in \mathbb{N}$ *converges and* $\lim a_n = 2$.

ii) $a_n = \frac{1}{n}$ *converges and* $\lim a_n = 0$.

Theorem 1.2.3
If $a_n < b_n$ for all $n \geq N \in \mathbb{N}$, then $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.

Theorem 1.2.4 (Sandwiching theorem)
Suppose that $a_n \leq c_n \leq b_n$ for all $n \in \mathbb{N}$. If $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = a$, then $\lim_{n\to\infty} c_n = a$.

Definition 1.2.5
A sequence (a_n) is bounded, if there exists a real number $K > 0$ such that $|a_n| \leq K \quad \forall n \in \mathbb{N}$.

Theorem 1.2.6
If the sequence (a_n) converges, then

i) the limit is unique.

ii) the sequence is bounded.

Theorem 1.2.7
Let $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$. Then

i) $\lim_{n\to\infty} \alpha a_n = \alpha a \quad (\alpha \in \mathbb{R})$

ii) $\lim_{n\to\infty} |a_n| = |a|$

iii) $\lim_{n\to\infty}(a_n + b_n) = a + b$

iv) $\lim_{n\to\infty}(a_n b_n) = ab$

v) $\lim_{n\to\infty}(\frac{a_n}{b_n}) = \frac{a}{b}$, *if* $b \neq 0$

Definition 1.2.8

i) A sequence is said to be increasing (strictly increasing), if $a_{n+1} \geq a_n \quad \forall n \in \mathbb{N}$ $(a_{n+1} > a_n \quad \forall n \in \mathbb{N})$.

ii) A sequence is said to be decreasing (strictly decreasing), if $a_{n+1} \leq a_n \quad \forall n \in \mathbb{N}$ $(a_{n+1} < a_n \quad \forall n \in \mathbb{N})$.

iii) All such sequences are said to be monotone.

Example

i) $S_n = \sum_{k=1}^{n} \frac{1}{k(k+1)}$ *is strictly increasing.*

ii) $a_n = (-1)^n$ *is not monotone.*

Theorem 1.2.9

Every bounded, monotone sequence in $\mathbb{R}$ *have a limit in* $\mathbb{R}$.

Example

i) $S_n = \sum_{k=0}^{n} \frac{1}{k!} \leq 2 + \sum_{k=1}^{n} \frac{1}{k(k+1)} \leq 3 \quad \forall n \in \mathbb{N}$

ii) $a_n = (1+\frac{1}{n})^n$ *is strictly increasing.*

iii) $b_n = (1+\frac{1}{n})^{n+1}$ *is strictly decreasing.*

iv) $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n =: e$

Theorem 1.2.10 (Nested intervals theorem)

Suppose that $I_n = \{x : a_n \leq x \leq b_n\}$ *for all* $n \in \mathbb{N}$ *is a sequence of closed intervals such that* $I_{n+1} \subset I_n$ *for each* n. *If* $\lim_{n\to\infty}(b_n - a_n) = 0$, *then there is one and only number* $x_0 \in I_n \quad \forall n$.

1.2.2 Subsequences

Definition 1.2.11

Let $(a_n)_{n=1}^{\infty}$ *be a sequence.*

i) If $r_1 < r_2 < r_3 < \ldots \in \mathbb{N}$, *then the sequence* $(a_{r_n})_{n=1}^{\infty}$ *is called a subsequence of* $(a_n)_{n=1}^{\infty}$.

ii) $a \in \mathbb{R}$ *is called point of accumulation of* (a_n) *if there exists a subsequence* (a_{r_n}) *of* (a_n) *such that* $\lim_{n\to\infty} a_{r_n} = a$

Theorem 1.2.12

Suppose $a_n \to a \quad (n \to \infty)$. *If* (a_{r_n}) *is a subsequence of* (a_n), *then* $(a_{r_n}) \to a \quad (n \to \infty)$

Theorem 1.2.13 (Bolzano-Weierstrass)
Any bounded infinite sequence of real numbers contains a convergent subsequence.

Definition 1.2.14 (Cauchy sequence)
A sequence (a_n) is called Cauchy sequence if and only if

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

Corollary 1.2.15 (Cauchy criterion)
A sequence is convergent if and only if it is a Cauchy sequence.

Remark
If (a_n) converges then $a_n - a_{n-1} \to 0 \quad (n \to \infty)$

1.2.3 Series

Definition 1.2.16
Let the sequence (a_k) be given. We define $S_n := \sum_{k=1}^{n} a_k \quad \forall n \in \mathbb{N}$. The symbol $\sum_{k=1}^{\infty} a_k$ is called an infinite series having kth term a_k and nth partial sum S_n. This series is said to converge in the case the sequence S_n of its partial sums converges.

Example

i) (geometric series) $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ if $|q| < 1$.

ii) (harmonic series) $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges.

iii) $\sum_{k=0}^{\infty} \frac{1}{k!}$ converges.

Theorem 1.2.17 (Cauchy criterion)
$\sum_{k=0}^{\infty} a_k$ converges if and only if

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N}: \left| \sum_{k=n+1}^{n+p} a_n \right| < \epsilon \quad \forall n \geq N, p \in \mathbb{N}$$

Corollary 1.2.18
Suppose $\sum_{k=0}^{\infty} a_k$ converges. Then $\lim_{n\to\infty} a_n = 0$.

1.2.4 Series of Nonnegative Terms

Theorem 1.2.19
The series $\sum_{k=0}^{\infty} a_k$ with nonnegative terms ($a_k \geq 0 \quad \forall k \in \mathbb{N}_0$) converges if and only if the partial sums $S_n = \sum_{k=0}^{n} a_k$ are bounded.

Definition 1.2.20
Let $\sum_{k=0}^{\infty} a_k$ be a given series.

i) $\sum_{k=0}^{\infty} a_k$ *is said to converge absolutely if* $\sum_{k=0}^{\infty} |a_k|$ *converges. (In that case we write* $\sum_{k=0}^{\infty} |a_k| < \infty$*).*

ii) If $\sum_{k=0}^{\infty} a_k$ *converges and* $\sum_{k=0}^{\infty} |a_k|$ *diverges, then the series* $\sum_{k=0}^{\infty} a_k$ *is said to be conditionally convergent.*

Theorem 1.2.21
If $\sum_{k=0}^{\infty} |a_k|$ *converges, then* $\sum_{k=0}^{\infty} a_k$ *converges and* $|\sum_{k=0}^{\infty} a_k| \leq \sum_{k=0}^{\infty} |a_k|$

Theorem 1.2.22 (Comparison test)
Suppose that $b_k \geq 0 \quad \forall k \in \mathbb{N}_0$*.*

i) If $|a_k| \leq b_k \quad \forall k$ *and* $\sum_{k=0}^{\infty} b_k < \infty$*, then* $\sum_{k=0}^{\infty} a_k$ *is absolutely convergent.*

ii) Assume $a_k \geq 0 \quad \forall k \in \mathbb{N}_0$*. If* $b_k \geq a_k \quad \forall k$ *and* $\sum_{k=1}^{\infty} a_k$ *diverges, then* $\sum_{k=0}^{\infty} b_k$ *diverges.*

Example
$\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} = \infty$ *because* $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$ *and* $\frac{1}{\sqrt{k}} \geq \frac{1}{k} \quad \forall k \in \mathbb{N}$

Theorem 1.2.23 (Root test)
Let $\sum_{k=0}^{\infty} a_k$ *be a series with* $\limsup \sqrt[n]{|a_n|} = \rho$*. Then*

i) if $\rho < 1$*, the series* $\sum_{k=0}^{\infty} a_k$ *converges absolutely.*

ii) if $\rho > 1$*, the series* $\sum_{k=0}^{\infty} a_k$ *diverges.*

iii) if $\rho = 1$*, the test gives no information.*

Theorem 1.2.24 (Ratio test)
Suppose that $a_k \neq 0, k = 0, 1, 2, \ldots$ *and that*
$\limsup \left| \frac{a_{n+1}}{a_n} \right| = \rho$*. Then*

i) if $\rho < 1$*, the series* $\sum_{k=0}^{\infty} a_k$ *converges absolutely.*

ii) if $\rho > 1$*, the series* $\sum_{k=0}^{\infty} a_k$ *diverges.*

iii) if $\rho = 1$*, the test gives no information.*

Theorem 1.2.25
Suppose $a_k \geq 0 \quad \forall k \in \mathbb{N}$*. The series* $\sum_{k=1}^{\infty} a_k$ *converges if and only if* $\sum_{k=0}^{\infty} 2^k a_{2^k}$ *converges.*

Corollary 1.2.26
$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}}$ *converges if* $\alpha > 1$ *and diverges if* $\alpha \leq 1$*.*

1.2.5 Alternating Series

Theorem 1.2.27 (Leibnitz series test)
Let $(a_k) \geq 0$ and monotone decreasing ,and suppose that $\lim_{k\to\infty} a_k = 0$. Then the series $\sum_{k=0}^{\infty}(-1)^{k+1}a_k$ is convergent.

Example

i) $\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}$ *is convergent.*

ii) $\sum_{k=0}^{\infty} \frac{(-1)^k}{\sqrt{k+1}}$ *is convergent.*

iii) $\sum_{k=1}^{\infty}(-1)^k(1+\frac{1}{k})^{-k}$ *diverges, because* $(1+\frac{1}{k})^{-k} \to \frac{1}{e} \neq 0$.

1.2.6 Multiplication of Series

Definition 1.2.28 (Cauchy product)
Let $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ be given series. We define $c_n := \sum_{\nu=0}^{n} a_\nu b_{n-\nu}$ and call $\sum_{n=0}^{\infty} c_n$ the product of the two series.

Theorem 1.2.29 (Cauchy product)
If $\sum_{k=0}^{\infty} |a_k| < \infty$ and $\sum_{k=0}^{\infty} |b_k| < \infty$ then $\sum_{n=0}^{\infty} c_n$ is absolutely convergent and $\sum_{n=0}^{\infty} c_n = (\sum_{k=0}^{\infty} a_k)(\sum_{k=0}^{\infty} b_k)$.

Example

i) $\sum_{k=0}^{\infty} \frac{1}{2^k} \sum_{k=0}^{\infty} \frac{1}{2^k} = 4 = \sum_{k=0}^{\infty}(k+1)2^{-k}$

ii) $\frac{1}{(1-x)^2} = \sum_{k=0}^{\infty}(k+1)x^k$

iii)

$$\left(\sum_{k=0}^{\infty} \frac{x^k}{k!}\right)\left(\sum_{k=0}^{\infty} \frac{y^k}{k!}\right) = \sum_{k=0}^{\infty}\left(\sum_{\nu=0}^{k} \frac{x^\nu}{\nu!}\frac{y^{k-\nu}}{(k-\nu)!}\right) = \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!}$$

Example
Let $a_0 = b_0 := 0, a_k = b_k := \frac{(-1)^{k-1}}{\sqrt{k}} \quad \forall k \geq 1$. Then $\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} b_k$ converges, but $|c_n| = |(-1)^n \sum_{\nu=1}^{n-1} \frac{1}{\sqrt{\nu}\sqrt{n-\nu}}| \geq 1$ hence $\sum_{k=0}^{\infty} c_n$ diverges

Theorem 1.2.30 (Abel)
If $\sum_{k=0}^{\infty} a_k = A$ and $\sum_{k=0}^{\infty} b_k = B$ are convergent series and if their Cauchy product $\sum_{n=0}^{\infty} c_n = C$ converges, then $C = AB$.

1.2.7 Rearrangements

Definition 1.2.31
Let $\phi : \mathbb{N}_0 \to \mathbb{N}_0$ *be a bijective function. We say that* $\sum_{k=0}^{\infty} a_{\phi(k)}$ *is a rearrangement of* $\sum_{k=0}^{\infty} a_k$.

Theorem 1.2.32
If $\sum_{k=0}^{\infty} a_k$ *is an absolute convergent series, then every rearrangement of* $\sum_{k=0}^{\infty} a_k$ *also converges absolutely and* $\sum_{k=0}^{\infty} a_{\phi(k)} = \sum_{k=0}^{\infty} a_k$.

Corollary 1.2.33
Let $\sum_{k=0}^{\infty} a_k$ *be convergent, but not absolutely convergent. Then there exists for all* $s \in \mathbb{R}$ *a rearrangement* $\sum_{k=0}^{\infty} a_{\phi(k)}$ *such that* $\sum_{k=0}^{\infty} a_{\phi(k)} = s$.

Example
Consider the convergent series $s = 1 - 1/2 + 1/3 - 1/4 + 1/5 \ldots (= \log 2)$. *Then* $s/2 = 1/2 - 1/4 + 1/6 - 1/8 \ldots$ *and* $\frac{3}{2}s = 1 + 1/3 - 1/2 + 1/5 + 1/7 - 1/4 \ldots \neq s$

1.2.8 Power Series

Definition 1.2.34
An infinite series of the form $\sum_{k=0}^{\infty} a_k(z - z_0)^k$ *is called power series in* $(z - z_0)$, *where* $(a_k) \subset \mathbb{C}$ *is a fixed sequence and* $z, z_0 \in \mathbb{C}$.

Theorem 1.2.35 (Cauchy Hadamard)
Let $\sum_{k=0}^{\infty} a_k(z - z_0)^k$ *be a given power series and define* $R = \frac{1}{\limsup \sqrt[k]{|a_k|}}$ *(radius of convergence). Then*

i) the series converges only for $z = z_0$; *or*

ii) the series converges for all values of z; *or*

iii) the series converges absolutely for all $|z - z_0| < R$ *and diverges if* $|z - z_0| > R$.

Example

i) $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$ *if* $|z| < 1 \quad , z \in \mathbb{C} \quad (R = 1)$

ii) $\sum_{k=0}^{\infty} \frac{z^k}{k!} =: e^z \quad \forall z \in \mathbb{C} \quad (R = \infty)$

iii) $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1} =: \sin z \quad \forall z \in \mathbb{C} \quad (R = \infty)$
$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k} =: \cos z \quad \forall z \in \mathbb{C} \quad (R = \infty)$

iv) $\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} z^{k+1} =: \log(1 + z) \quad (R = 1)$

v) $\sum_{k=0}^{\infty} \binom{\alpha}{k} z^k = (1+z)^\alpha$ *if* $|z| < 1, (\alpha \notin \mathbb{N}_0) \quad (R = 1)$

Theorem 1.2.36 (Cauchy product of power series)
Let $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ *and* $\sum_{k=0}^{\infty} b_k(z-z_0)^k$ *be absolutely convergent for all* $|z-z_0| < R$ *and* $c_n = \sum_{k=0}^{n} a_\nu b_{k-\nu}$. *Then* $\sum_{k=0}^{\infty} c_k(z-z_0)^k$ *converges absolutely* $\forall |z-z_0| < R$ *and*

$$\sum_{k=0}^{\infty} c_k(z-z_0)^k = \left(\sum_{k=0}^{\infty} a_k(z-z_0)^k\right)\left(\sum_{k=0}^{\infty} b_k(z-z_0)^k\right) \quad \forall |z-z_0| < R$$

.

Theorem 1.2.37

i) $\cos z = \frac{1}{2}(e^{iz} + e^{-iz}) \quad \sin z = \frac{1}{2i}(e^{iz} - e^{-iz})$

ii)

$$\begin{aligned} e^{z_1+z_2} &= e^{z_1}e^{z_2} \\ \sin(z_1+z_2) &= \sin z_1 \cos z_2 + \sin z_2 \cos z_1 \\ \cos(z_1+z_2) &= \cos z_1 \cos z_2 - \sin z_1 \sin z_2 \end{aligned}$$

iii)

$$e^{ix} = \cos x + i \sin x \quad (x \in \mathbb{R})$$

$$|e^{ix}| = 1, |\cos x| \leq 1, |\sin x| \leq 1 \quad (x \in \mathbb{R})$$

1.3 Continuity

Definition 1.3.1 (Limits of functions)
Let $I \subset \mathbb{R}$ *be an interval,* $f : I \to \mathbb{R}, x_0 \in \mathbb{R}$.

i) We say that the limit of $f(x)$ *as* x *approaches* x_0 *exists if there exists a number* $\alpha \in \mathbb{R}$ *such that*

$$\forall \epsilon > 0 \quad \exists \delta > 0 : |f(x) - \alpha| < \epsilon \quad \forall x \in U_\delta(x_0) \cap I \quad (x \neq x_0)$$

We write: $\lim_{x \to x_0} f(x) = \alpha$ *or* $f(x) \to \alpha \quad (x \to x_0)$ *. The condition* $x \neq x_0$ *is used since* f *may not be defined at* x_0.

ii) One sided limits:
If $\lim_{x \to x_0, x > x_0} f(x) = \alpha_1$ *and* $\lim_{x \to x_0, x < x_0} f(x) = \alpha_2$ *we call* α_1 *and* α_2 *the right and left hand limits of* $f(x)$ *at* x_0 *and denote them by* $f(x_0^+)$ *and* $f(x_0^-)$.
The ϵ, δ *definition of the limit of* $f(x)$ *as* $x \to x_0^+$ *or* $x \to x_0^-$ *remain the same except that the values of* x *are restricted to* $x > x_0$ *or* $x < x_0$.

Remark
$\lim_{x \to x_0} f(x) = \alpha$ *if and only if* $f(x_0^-) = f(x_0^+) = \alpha$.

Theorem 1.3.2
Let $f : I \to \mathbb{R}$. *Then* $\alpha = \lim_{x \to x_0} f(x)$ *if and only if* $\lim f(x_n) = \alpha$ *for all sequences* $(x_n) \subset I$ $x_n \neq x_0, \lim_{n \to \infty} x_n = x_0$

Corollary 1.3.3 (Properties)
Assume $f(x) \to \alpha, g(x) \to \beta \quad (x \to x_0)$. *Then*

i) $f(x) + g(x) \to \alpha + \beta$

ii) $cf(x) \to c\alpha \quad (c \in \mathbb{R})$

iii) $f(x)g(x) \to \alpha\beta$

iv) $\frac{f(x)}{g(x)} \to \frac{\alpha}{\beta}, \quad$ *if* $\beta \neq 0$

Theorem 1.3.4
Let $f : I \to T$ *be continuous at* x_0, $f(x_0) = y_0$, *and let* $g(x) : T \to \mathbb{R}$ *be continuous at* y_0. *Then* $g \circ f$ *is continuous at* x_0.

Remark
For every polynomial $P(x) = \sum_{k=0}^{n} a_k x^k$ *we have* $\lim_{x \to x_0} P(x) = P(x_0) \quad \forall x_0 \in \mathbb{R}$

Definition 1.3.5 (Continuity of functions)
Let $I \subset \mathbb{R}$ *be an interval,* $f : I \to \mathbb{R}$. *Then*

i) f *is continuous at* $x_0 \in I$ *if* $\lim_{x \to x_0} f(x) = f(x_0)$

ii) f *is continuous on* I *(Notation:* $f \in C(I)$*) if* f *is continuous at* x_0 *for all* $x_0 \in I$*.*

Remark
In other words:
f *is continuous at* x_0 *if and only if*

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \quad |f(x) - f(x_0)| < \epsilon \quad \forall |x - x_0| < \delta$$

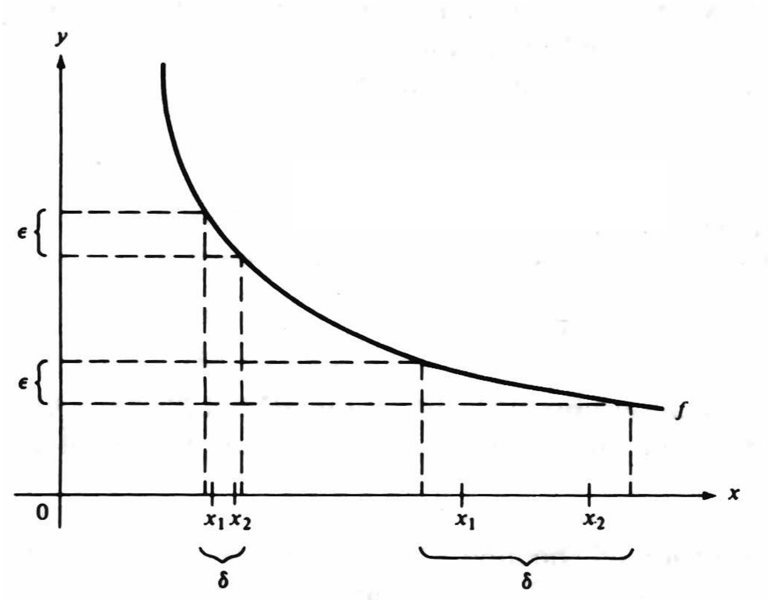

Figure 1.1.

Theorem 1.3.6
Let $f, g : I \to \mathbb{R}$ *be continuous at* $x_0 \in I$*. Then* $f+g$*,* cf*,* fg*,* $\frac{f}{g}$ *(if* $g(x_0) \neq 0$*) are continuous at* x_0*.*

Theorem 1.3.7 (Intermediate value theorem)
Let $f : [a,b] \to \mathbb{R}$ *be continuous on* $[a,b]$*. Then there is for every* $y \in \mathbb{R}$ *:* $f(a) \leq y \leq f(b)$ *at least one number* $x_0 \in [a,b]$ *such that* $f(x_0) = y$*.*

Corollary 1.3.8
Suppose that the domain of f *is an interval* I *and* f *is continuous on* I*. Then the range of* f *is an interval.*

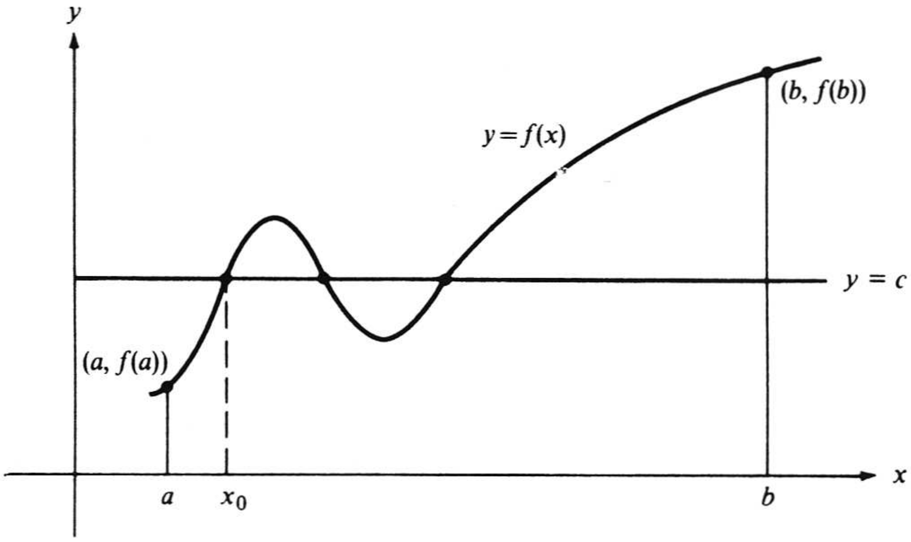

Figure 1.2. Illustrating the Intermediate-value theorem

Theorem 1.3.9 (Inverse functions)
Let $f : I \to \mathbb{R}$ be continuous on an interval I.

i) f is invertible if and only if f is strictly monotone on I.

ii) If f^{-1} exists, then f^{-1} is continuous on $f(I)$.

Theorem 1.3.10 (Extreme–value-theorem)
Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on the closed interval $[a, b]$. Then there are numbers $x_0 \in [a, b]$ and $x_1 \in [a, b]$ such that $f(x_0) \leq f(x) \leq f(x_1) \quad \forall x \in [a, b]$.

That is, f takes its maximum and minimum values on $[a, b]$. (Notation: $x_0 = \max_{x\in[a,b]} f(x), \quad x_1 = \min_{x\in[a,b]} f(x)$)

Definition 1.3.11 (Uniformly continuous)
$f : I \to \mathbb{R}$ is uniformly continuous on an interval I if

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \quad |f(x) - f(y)| < \epsilon \quad \forall |x - y| < \delta \text{ (δ depends on ϵ only)}$$

Theorem 1.3.12 (Uniform continuity theorem)
If $f : [a, b] \to \mathbb{R}$ is continuous on the closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

1.3.1 Sequences of Functions

Definition 1.3.13 (Uniform convergence)
Let (f_n) be a sequence of functions $f_n : I \to \mathbb{R} \quad n \in \mathbb{N}$ and
Let $f_n : I \to \mathbb{R}$ be a sequence of functions $(f_n) \quad n \in \mathbb{N}$ and $f : I \to \mathbb{R}$.

i) We say the sequence (f_n) converges point-wise to the function f, if $\lim_{n\to\infty} f_n(x) = f(x)$ for each fixed $x \in I$.

ii) We say (f_n) converges uniformly on the interval I to the function f, if

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} : \quad |f_n(x) - f(x)| < \epsilon \quad \forall n \geq N \text{ and } \forall x \in I$$

Uniform convergence differs from ordinary (point-wise) convergence in that the integer N does not depend on x.

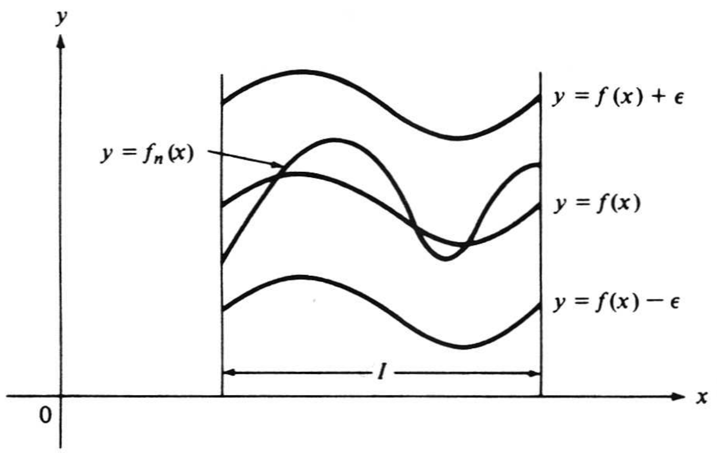

Figure 1.3. Illustrating uniform convergence

Remark

i) $f(x) = \sum_{k=0}^{\infty} f_k(x)$ converges uniformly on I if $S_n(x) = \sum_{k=0}^{n} f_k(x) \to f(x)$ uniformly on I.

ii) Cauchy criterion: (f_n) converges uniformly on an interval I if and only if

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} : |f_n(x) - f_m(x)| < \epsilon \quad \forall n, m \geq N \text{ and } \forall x \in I$$

Theorem 1.3.14 (Weierstrass M-test)
Let $f_n : I \to \mathbb{R}$ and $a_n \geq 0$ such that $|f_n(x)| \leq a_n \quad \forall n \in \mathbb{N}$ and $\forall x \in I$. If $\sum_{k=1}^{\infty} a_k < \infty$, then $\sum_{k=1}^{\infty} f_k(x)$ is uniformly convergent on I.

Example

Suppose the radius of convergence of $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ *is* R*, then* $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ *is uniformly convergent on* $[x_0 - r, x_0 + r] \quad 0 < r < R$.

Theorem 1.3.15

Let $f_n : I \to \mathbb{R}$*,* $f : I \to \mathbb{R}$ *such that* f_n *is continuous on* I $\forall n \in \mathbb{N}$ *and suppose* (f_n) *converges uniformly to* f *on* I*. Then* f *is continuous on* I*.*

Corollary 1.3.16

i) Power series $f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k$ *are continuous on* (x_0-R, x_0+R) *(R: radius of convergence).*

ii) $e^x, \sin x, \cos x, \sinh x, \cosh x$ *are continuous on* $\mathbb{R}$*.*

Theorem 1.3.17 (Abel)

Suppose $\sum_{k=0}^{\infty} a_k$ *converges. Put* $f(x) = \sum_{k=0}^{\infty} a_k x^k$ *for* $|x| < 1$*. Then*

$$\lim_{x \to 1^-} f(x) = \sum_{k=0}^{\infty} a_k$$

1.4 Differentiation

Definition 1.4.1
Let $I \subset \mathbb{R}$ be an interval, $f : I \to \mathbb{R}, x_0 \in I$.

i) f is differentiable at x_0 if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists.

If this limit exists, it is called the derivative f' of f at x_0.

ii) f is differentiable on I, if f' is defined at every point of I. f' is called the derivative of f.

Example
i) If f is constant, then $f'(x) = 0 \quad \forall x$.

ii) If $f(x) = x^n, \quad n \in \mathbb{N}$, then $f'(x) = nx^{n-1}$.

iii) $f(x) = e^x$

$$f'(x) = \lim_{x \to x_0} \frac{e^x - e^{x_0}}{x - x_0} = e^{x_0} \lim_{x \to x_0} \sum_{k=1}^{\infty} \frac{(x - x_0)^{k-1}}{k!} = e^{x_0}$$

iv) $f(x) = |x|$ is not differentiable at $x_0 = 0$ because

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = -1 \neq 1 = \lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0}$$

Theorem 1.4.2
Suppose that f has a derivative at x_0. Then there is a function η defined in an interval about x_0 such that $f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + \eta(x - x_0) \cdot (x - x_0)$. Also η is continuous at x_0 with $\eta(x_0) = 0$.

Theorem 1.4.3
If $f'(x_0)$ exists at x_0, then f is continuous at x_0.

Theorem 1.4.4 (Properties)
Let $f, g : I \to \mathbb{R}$ and assume f, g are differentiable on an interval I.
Then $f + g$, cf, $f \cdot g$, f/g $(g(x_0) \neq 0)$ are differentiable at x_0, and

i) $(f + g)'(x_0) = f'(x_0) + g'(x_0)$

ii) $(cg)'(x_0) = cf'(x_0)$

iii) $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$

iv) $(f/g)'(x_0) = \frac{1}{g^2(x_0)}(f'(x_0)g(x_0) - f(x_0)g'(x_0))$

Theorem 1.4.5 (Chain rule)
Suppose that $g : I \to J$ is differentiable at $x_0 \in I$ and $f : J \to \mathbb{R}$ is differentiable at $y_0 = g(x_0)$. Then $(f \circ g) : I \to \mathbb{R}$ has a derivative at x_0 and

$$(f \circ g)'(x_0) = f'(g(x_0))g'(x_0)$$

Theorem 1.4.6
Suppose $f : I \to \mathbb{R}$ is continuous on an interval I, strictly monotone on I and differentiable at $x_0 \in I$ such that $f'(x_0) \neq 0$. Then the inverse function f^{-1} is differentiable at $y_0 = f(x_0)$, and

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'\left(f^{-1}(y_0)\right)}$$

Theorem 1.4.7 (Differentiation of sequence)
Let $f_n : I \to \mathbb{R}$ be a sequence of functions that are differentiable on an interval I $\forall n \in \mathbb{N}$. Suppose $f_n(x)$ converges point-wise to $f(x)$ $\forall x \in I$ and $f_n'(x)$ converges uniformly on I. Then $f(x)$ is differentiable on I and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad \forall x \in I.$$

Corollary 1.4.8
Power series $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ are differentiable in $(-R, R)$ and

i) $\left(\sum_{k=0}^{\infty} a_k(x - x_0)^k\right)' = \sum_{k=1}^{\infty} ka_k(x - x_0)^{k-1}$ $\forall |x - x_0| < R$. *Both series have the same Radius of convergence.*

ii) $\left(\sum_{k=0}^{\infty} a_k(x - x_0)^k\right)^{(n)}(x_0) = n!a_n \quad (n = 1, 2, \dots)$

Definition 1.4.9 (Local extrema)
Suppose $f : I \to \mathbb{R}$, $x_0 \in I$. We say f has a local maximum (minimum) at x_0, if there exists $\epsilon > 0$ such that $f(x) \geq f(x_0)$ $(f(x) \leq f(x_0))$ $\forall x \in U_\epsilon(x_0) \cap I$.

Theorem 1.4.10 (Necessary condition for local extrema)
Let $f : [a, b] \to \mathbb{R}$ and $x_0 \in (a, b)$ be given. If f has a local extrema at x_0 and $f'(x_0)$ exists, then $f'(x_0) = 0$.

Theorem 1.4.11 (Rolle's theorem)
Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and f has a derivative at each point of (a, b). If $f(a) = f(b)$, then there exists $\xi \in (a, b)$ such that

$$f'(\xi) = 0$$

Theorem 1.4.12 (Mean-value theorem)
Suppose $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and f has a derivative at each point of (a,b). Then there exists $\xi \in (a,b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

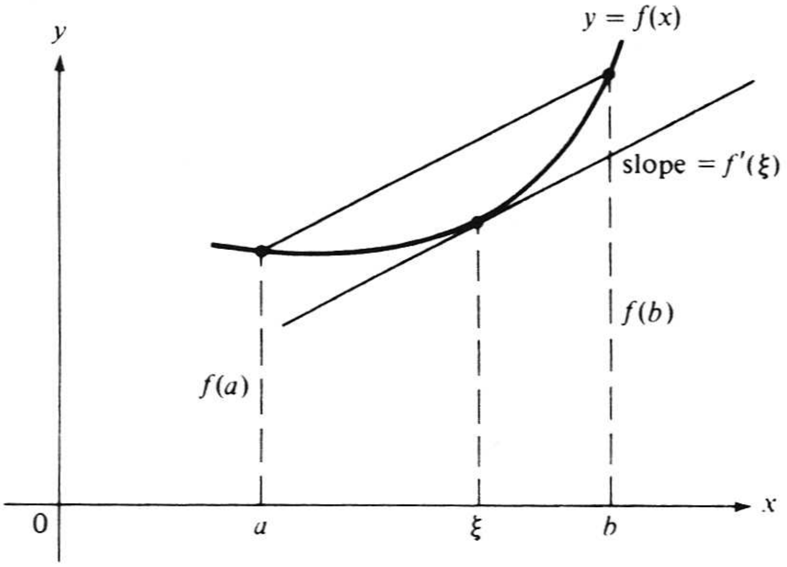

Figure 1.4. Illustrating the Mean-value theorem

Theorem 1.4.13 (Generalised mean value theorem)
If $f, g : [a,b] \to \mathbb{R}$ are continuous functions on $[a,b]$ and differentiable in (a,b), then there exists $\xi \in (a,b)$ such that

$$f'(\xi)\,(g(b) - g(a)) = g'\,(\xi)(f(b) - f(a))$$

Corollary 1.4.14
Assume $f : I \to \mathbb{R}$ is differentiable on I.

i) If $f'(x) \geq 0 \quad \forall x \in I$, then f is increasing.

ii) If $f'(x) > 0 \quad \forall x \in I$, then f is strictly increasing.

iii) If $f'(x) = 0 \quad \forall x \in I$, then f is constant.

Theorem 1.4.15 (l 'Hospital)
Suppose $f, g : [a, b) \to \mathbb{R}$ are differentiable in $[a, b)$, where $a < b \leq \infty$, and $\lim_{x \to b^-} \frac{f'(x)}{g'(x)}$ exists.

i) If $\lim_{x \to b^-} f(x) = \lim_{x \to b^-} g(x) = 0$

ii) or if $\lim_{x \to b^-} g(x) = +\infty$ *(or* $\lim_{x \to b^-} g(x) = -\infty$ *)*

then

$$\lim_{x \to b^-} \frac{f(x)}{g(x)} = \lim_{x \to b^-} \frac{f'(x)}{g'(x)}.$$

Example
i) $\lim_{x \to \infty} \frac{a^k}{x^k} = \infty \quad \forall a > 1, \forall k \in \mathbb{R}$

ii) $\lim_{x \to \infty} \frac{\log x}{x^\alpha} = 0 \quad \forall \alpha \in \mathbb{R}$

iii) $\lim_{x \to \infty} \frac{x^\alpha}{e^x} = 0 \quad \forall \alpha > 0$

Definition 1.4.16 (Higher order derivative)
Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$. If f has a derivative f', and if f' itself is differentiable, we denote the derivative of f' by f''. Continuing this manner, we obtain $f, f', f'', f^{(3)}, \ldots, f^{(n)}$. $f^{(n)}$ is called the nth derivative of f. If f is infinitely differentiable on I we write $f \in C^\infty(I)$.

Theorem 1.4.17 (Taylor)
Suppose $f : [a, b] \to \mathbb{R}$, $f \in C^n[a, b]$, and f^{n+1} exists on (a, b). Let $x_0 \in [a, b]$ and $R_n := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$.
Then there exists $\xi \in (x_0, x)$ such that

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

Remark
i) $n = 0$: $f(x) - f(x_0) = f'(\xi)(x - x_0)$

ii) $n = 1$: $f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(\xi)(x - x_0)$

1.5 Integration

Definition 1.5.1
Let $f : [a, b] \to \mathbb{R}$.

i) By a partition Z *of* $[a, b]$ *is meant a finite set of points* $Z = (x_0, x_1, \ldots, x_n)$ *such that* $a = x_0 < x_1 < x_2 < \ldots < x_n = b$

ii) The mesh of a partition Z*:* $|Z| = \max_{k=1 \ldots n}(x_k - x_{k-1})$

iii) A partition Z' *of* $[a, b]$ *is finer than* Z *(or is a refinement of* Z*) if* $Z' \supset Z$.

iv) Let Z *be a partition of* $[a, b]$ *and* $\xi_k \in [x_{k-1}, x_k]$. *The quantity*

$$S(f, Z, \xi) := \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

is called Riemann sum of f.

v) f *is called Riemann integrable on* $[a, b]$ $(f \in R[a, b])$, *if there exists a number* A *having the following property:*

For every $\epsilon > 0$ *there exists a partition* Z_ϵ *of* $[a, b]$ *such that for every partition* Z *finer than* Z_ϵ *and for every choice of points* ξ_k *in* $[x_{k-1}, x_k]$ *we have* $|S(f, Z, \xi) - A| < \epsilon$.

A *is called the Riemann integral of* f.

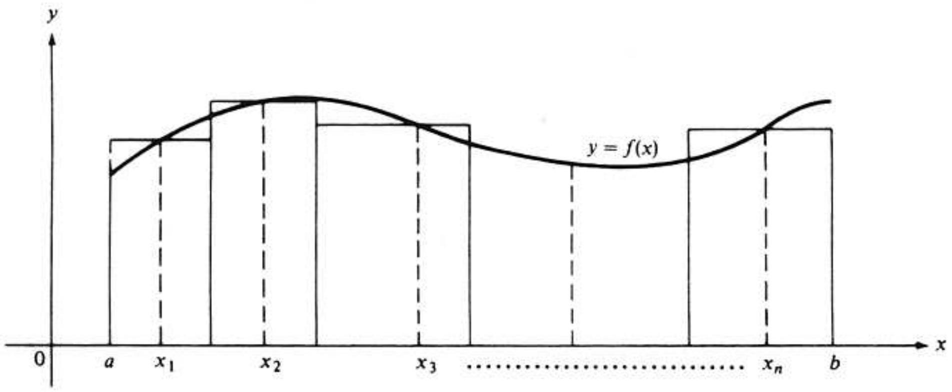

Figure 1.5. A Riemann sum

Theorem 1.5.2
If f *is Riemann integrable on* $[a, b]$, *then* f *is bounded.*

Theorem 1.5.3 (Properties)

i) $f \in R[a,b], g \in R[a,b], c \in \mathbb{R} \Rightarrow cf \in R[a,b], f+g \in R[a,b]$ *and*

$$\int_a^b cf(x)dx = c\int_a^b f(x)dx$$

$$\int_a^b (f(x)+g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

ii) $f \in R[a,b]$ *if and only if* $f \in R[a,c]$ *and* $f \in R[c,b]$ $(a < c < b)$ *and*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

iii) Let $f, g \in R[a,b]$ *such that* $f(x) \le g(x) \quad \forall x \in [a,b]$*. Then*

$$\int_a^b f(x)dx \le \int_a^b g(x)dx$$

iv) $\inf_{x\in[a,b]} f(x)(b-a) \le \int_a^b f(x)dx \le \sup_{x\in[a,b]} f(x)(b-a)$

v) If $f \in R[a,b]$*, then* $|f| \in R[a,b]$ *and*

$$\left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx$$

Theorem 1.5.4
Let $f : [a,b] \to \mathbb{R}$*. If* f *is continuous on* $[a,b]$*, then* f *is Riemann integrable on* $[a,b]$*.*

Theorem 1.5.5
Suppose that $f : [a,b] \to \mathbb{R}$ *is Riemann integrable on* $[a,b]$*. Then* $F(x) = \int_a^x f(t)dt$ *is continuous on* $[a,b]$*.*

Theorem 1.5.6 (The fundamental theorem of calculus – first form)
If $f \in R[a,b]$ *and if there is a differentiable function* F *on* $[a,b]$ *such that* $F' = f$*, then*

$$\int_a^b f(x)dx = F(b) - F(a) =: F|_a^b$$

Theorem 1.5.7 (The fundamental theorem of calculus – second form)
Let $f : [a,b] \to \mathbb{R}$ *be continuous on* $[a,b]$*. Let* $F(x) = \int_a^x f(t)dt$*. Then* F *is differentiable on [a,b] and*

$$F'(x) = f(x) \quad \forall x \in [a,b]$$

Theorem 1.5.8 (Integration by parts)
Suppose $f, g : [a, b] \to \mathbb{R}$ are differentiable on $[a, b]$ and f', g' are continuous on $[a, b]$. Then

$$\int_a^b f'(x)g(x)dx = fg|_a^b - \int_a^b f(x)g'(x)dx$$

Theorem 1.5.9 (Taylor)
Suppose $f : [a, b] \to \mathbb{R}$ having $n+1$ continuous derivatives on $[a, b]$. Then

$$R_n = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)(x-t)^n dt \quad \forall x \in [a, b], x_0 \in [a, b]$$

Corollary 1.5.10
If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, then

$$\left(\frac{d}{dx}\right)^{n+1} \left(\frac{1}{n!}\int_{x_0}^{x} f(t)(x-t)^n dt\right) = f(x)$$

Theorem 1.5.11 (Change of variable theorem)
Assume $f \circ g$ is defined, f is continuous and g has a continuous derivative g' on $[a, b]$. Then

$$\int_a^b f(g(t))g'(t)dt = \int_{g(a)}^{g(b)} f(x)dx$$

Theorem 1.5.12 (First mean value theorem for integration)
Let $f, g \in R[a, b]$ and g be nonnegative on $[a, b]$.

i) Then there exists a number $\mu \in [\inf_{x\in[a,b]} f, \sup_{x\in[a,b]} f]$

$$\int_a^b f(x)g(x)dx = \mu \int_a^b g(x)dx$$

ii) If f is continuous on $[a, b]$, then there exists $\xi \in (a, b)$

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx$$

Theorem 1.5.13 (Second mean value theorem for integration)
Let $f \in R[a, b]$, $g \in C^1[a, b]$ and g' be nonnegative on $[a, b]$. Then there is $\xi \in [a, b]$

$$\int_a^b f(x)g(x)dx = g(a)\int_a^{\xi} f(x)dx + g(b)\int_{\xi}^{b} f(x)dx$$

Theorem 1.5.14 (Integration of sequences)
If $f_n : [a,b] \to \mathbb{R}$ is a sequence of Riemann integrable functions on $[a,b]$ and f_n converges uniformly to f on $[a,b]$, then f is Riemann integrable on $[a,b]$ and

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx = \int_a^b \lim_{n\to\infty} f_n(x)dx$$

Corollary 1.5.15

i) $\sum_{k=0}^{\infty} \int_a^b f_k(x)dx = \int_a^b \sum_{k=0}^{\infty} f_k(x)dx$ *if* $f_n \in R[a,b]$ *and if* $\sum_{k=0}^{\infty} f_k$ *converges uniformly on* $[a,b]$.

ii) If $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ *has radius of convergence* $R > 0$, *then*

$$\int_{x_0}^{x} \sum_{k=0}^{\infty} a_k(t-x_0)^k dt = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-x_0)^{k+1} \quad \forall x : |x-x_0| < R$$

Definition 1.5.16 (Improper integral)
Assume that $f : [a,b) \to \mathbb{R}$, $b \in \mathbb{R} \cup \{\infty\}$, and assume that $f \in R[a,c]$ $\forall a \le c < b$. Define $\int_a^b f(x)dx = \lim_{c\to b^-} \int_a^c f(x)dx$ if this limit exists. In that case, we say that the integral on the left converges.

Theorem 1.5.17 (Comparison test)
Let $f, g : [a,b) \to \mathbb{R}$. Assume that $f, g \in R[a,c]$ $\forall a \le c < b$. If $|f(x)| \le |g(x)|$ $\forall x \in [a,b)$ and if $\int_a^b g(x)dx$ converges, then $\int_a^b f(x)dx$ converges.

Theorem 1.5.18 (Integral test)
Assume that f is monotone on $[1,\infty)$. Suppose that $\sum_{k=1}^{\infty} u_k$ is a series with $u_k = f(k)$, $\quad k \in \mathbb{N}$. Then

i) $\sum_{k=1}^{\infty} u_k$ *converges if* $\int_1^{\infty} f(x)dx$ *converges and*

ii) $\sum_{k=1}^{\infty} u_k$ *diverges if* $\int_1^{\infty} f(x)dx$ *diverges.*

Definition 1.5.19 (The Gamma function)
For $0 < x < \infty$ the Gamma function, denoted by $\Gamma(x)$ is defined by the formula

$$\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}dt.$$

Theorem 1.5.20

i) functional equation:

$$\Gamma(x+1) = x\Gamma(x) \quad \forall x > 0$$

ii) $\Gamma(n+1) = n! \quad \forall n \in \mathbb{N}_0$

iii) $\Gamma(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\dots(x+n)} \quad \forall x > 0$

iv) $\int_0^\infty e^{-at}t^{x-1}dt = \frac{\Gamma(x)}{a^x}$

v) $\Gamma^{(n)}(x) = \int_0^\infty (\log t)^n t^{x-1}e^{-t}dt$

Theorem 1.5.21 (The Euler summation formula)
Assume f has a continuous derivative on $[0, n]$. Then

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)dx + \frac{1}{2}(f(0) + f(n)) + \int_0^n (x - [x] - 1/2)f'(x)dx$$

Theorem 1.5.22 (Stirling's formula)
Let $n \in \mathbb{N}$. Then there is a number θ between 0 and 1 such that

$$n! = \sqrt{2\pi n}n^n e^{-n} e^{\frac{\theta}{12n}}$$

Theorem 1.5.23 (Weierstrass approximation theorem)
Let $[a, b]$ be a closed interval, and let f be a continuous function on $[a, b]$. Then f can be uniformly approximated by polynomials on $[a, b]$.

With other words there exists a polynomial $P(x) = \sum_{k=0}^{n} a_k x^k \quad (n \in \mathbb{N}, a_k \in \mathbb{R})$ such that

$$|f(x) - P(x)| \leq \epsilon \quad \forall x \in [a, b]$$

A more general theorem which includes the Weierstrass approximation theorem as a special case was done by M. H. Stone.

Chapter 2

Linear Algebra

2.1 Linear Spaces

Definition 2.1.1 (Linear space)
Let K be a given field and let V denote a nonempty set of objects together with two operations $+, \cdot$ which assigns any $x, y \in V$ a sum $x + y \in V$ and any $x \in V, c \in K$ a product $cx \in V$. Then $(V, +, \cdot)$ is called a linear space over K or vector space over K if the following axioms hold:

i) $x, y \in V \Rightarrow x + y \in V$

ii) $c \in \mathbb{R}, x \in V \Rightarrow cx \in V$

iii) $(x + y) + z = x + (y + z) \quad \forall x, y, z \in V$

iv) $x + 0 = x \quad \forall x \in V$

v) $x + (-x) = 0 \quad \forall x \in V$

vi) $x + y = y + x \quad \forall x, y \in V$

vii) $(ab)x = a(bx) \quad \forall x \in V, \forall a, b \in K$

viii) $a(x + y) = ax + ay \quad \forall x, y \in V, \forall a \in K$

ix) $(a + b)x = ax + bx \quad \forall a, b \in K, \forall x \in V$

x) $1 \cdot x = x \quad 1 \in K, \forall x \in V$

Linear spaces over the field $\mathbb{R}$ are also called real linear spaces to emphasise that we are multiplying the elements of V by real numbers. The numbers used as multipliers are also called scalars.

Example

i) $V = \mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$ *with addition and multiplication by scalars defined in terms of components is a vector space over $\mathbb{R}$.*

ii) Let K be an arbitrary field and let X be an nonempty set. The set V of all functions from X into K together with the sum of two functions defined by $(f+g)(x) = f(x) + g(x)$ and the product of a scalar $c \in K$ and a function $f \in V$ defined by $(cf)(x) = cf(x)$ is a linear space over K.

Definition 2.1.2 (Linear combinations)
Let V be a vector space over K. If $x_1, \dots, x_k \in V$ and $c_1, \dots, c_k$ are scalars, then the vector $c_1x_1 + \dots + c_kx_k$ is called a finite linear combination of $x_1, \dots, x_k$.

Definition 2.1.3 (Linear independence, linear dependence)
Let V be a vector space over K.

i) Let $x_1, \dots, x_k \in V$. The set of vectors $\{x_1, \dots, x_k\}$ is said to be linearly independent over K, if for any scalars $c_1, \dots, c_k \in K$ the relation $c_1x_1 + \dots + c_kx_k = 0 \Rightarrow c_1 = \dots = c_k = 0$ holds,

ii) otherwise $\{x_1, \dots, x_k\}$ is said to be linearly dependent. If this is the case, there exists $c_1, \dots, c_k \in K$, not all of them 0, such that $c_1x_1 + \dots + c_kx_k = 0$.

Definition 2.1.4 (Subspaces)
Let $S \neq \emptyset$ be a subset of a linear space V over a field K. If S is also a linear space over K with the same operation of addition and multiplication by scalars, then S is called a subspace.

Example
Let V be the linear space of all functions from an interval $[a,b]$ into the real field. Then the set of all continuous functions $C[a,b]$ in V is a subspace of V.

Theorem 2.1.5
S is a subspace of V if and only if

i) $S \neq \emptyset$.

ii) $x, y \in S$ implies $x + y \in S$.

iii) $x \in S$ implies $cx \in S$ for every $c \in K$.

Corollary 2.1.6
The intersection of any number of subspaces of a vector space V is a subspace of V.

Definition 2.1.7 (Linear spans)
Let $S \neq \emptyset$ be a subset of V. The set of all finite linear combinations of elements in S, denoted by $L(S)$, is a subspace of V. We call this the subspace spanned by S, or the linear span of S.

Theorem 2.1.8
$L(S)$ is the smallest subspace of V containing S.

Theorem 2.1.9
Let $S = \{x_1, \dots, x_k\} \subset V$ be an independent set, and $L(S)$ be the subspace spanned by S. Then every set of $k+1$ elements in $L(S)$ is dependent.

Definition 2.1.10 (Dimension, Basis)
A vector space V is said to be of finite dimension $n \in \mathbb{N}$, written $\dim V = n$, if there exists linearly independent vectors $\{e_1, e_2, \dots, e_n\}$ which span V. The set $\{e_1, e_2, \dots, e_n\}$ is then called a basis of V.

Theorem 2.1.11
Suppose the set $\{v_1, \dots, v_n\}$ is a basis of a vector space V. If $\{w_1, \dots, w_m\}$ is linearly independent, then $m \leq n$ and $\{w_1, \dots, w_m, v_{i_1}, \dots, v_{i_{n-m}}\}$ is a basis of V.

Corollary 2.1.12
If V has a finite basis then every finite basis for V has the same numbers of elements.

Theorem 2.1.13
Let $\dim V = n \in \mathbb{N}$. Then

i) Any set of $n+1$ or more vectors is dependent.

ii) Any independent set is part of a basis, i.e. it can be extended to a basis.

iii) An independent set with n elements is a basis.

Definition 2.1.14
Let V be a vector space. The sum $U + W = \{u + w : u \in U, w \in W\}$ of the subspaces U and W of V is a subspace of V.

Theorem 2.1.15
Let U and W be finite dimensional subspaces of a vector space V. Then $U + W$ has finite dimension and

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Definition 2.1.16 (Inner product)
A linear space V over K is said to have an inner product, if to each pair of vectors $x, y \in V$ there is assigned a scalar $(x, y) \in K$ satisfying the following axioms:

i) $(x, y) = \overline{(y, x)} \quad \forall x, y \in V$

ii) $(x, y+z) = (x, y) + (x, z) \quad \forall x, y, z \in V$

iii) $c(x, y) = (cx, y) \quad \forall x, y \in V, \forall c \in K$

iv) $(x, x) > 0$ *if* $x \neq 0 \quad \forall x \in V$

v) $(x, x) = 0$ *if* $x = 0$

A linear space over $\mathbb{R}$ *with an inner product is called a real Euclidean space.*

Example
The dot product $(x, y) := \sum_{i=1}^{n} x_i y_i$ *in* $\mathbb{R}^n$ *is an inner product.*

Definition 2.1.17 (Norm)
The mapping $||.|| : V \to \mathbb{R}$ *which has the following properties*

i) $||x|| = 0 \Leftrightarrow x = 0 \quad \forall x \in V$

ii) $||x|| > 0 \Leftrightarrow x \neq 0 \quad \forall x \in V$

iii) $||cx|| = |c| ||x|| \quad \forall x \in V, c \in K$

iv) $||x + y|| \leq ||x|| + ||y|| \quad \forall x, y \in V$

is called norm.

Example
In an Euclidean space V $||x|| = \sqrt{(x, x)}$ *defines a norm.*

Theorem 2.1.18 (Cauchy-Schwartz inequality)
Let V *be an Euclidean space. Every inner product satisfies the Cauchy-Schwarz inequality*

$$|(x, y)| = ||x|| ||y||.$$

Definition 2.1.19
A collection of vectors $\{x_1, \dots, x_m\}$ *in* V *is called orthogonal, if* $(x_i, x_j) = 0 \quad (i \neq j)$. *In particular the set is said to be orthonormal, if* $(x_i, x_j) = \delta_{ij}$.

Example
In the real linear space $C[0, 2\pi]$ *with the inner product* $(f, g) = \int_0^{2\pi} f(x) g(x) dx$ *the set* $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots\}$ *is an orthogonal set. Dividing each element by its norm, we obtain the orthonormal set* $\{\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots\}$

Theorem 2.1.20 (Gram-Schmidt)
Let V *be a finite dimensional vector space with an inner product. From any given basis* $\{x_1, \dots, x_n\}$ *it is possible to construct an orthogonal basis* $\{y_1, \dots, y_n\}$ *by setting*

i) $y_1 = x_1$ *and*

ii) $y_m = x_m - \sum_{k=1}^{m} \frac{(x_m, x_k)}{||x_k||^2} x_k \quad m = 2, \dots, n.$

2.2 Matrices

Definition 2.2.1 (Matrix)
Let K be an arbitrary field.

i) A rectangular array consisting of m rows and n columns

$$A = (a_{ij})_{m\times n} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

where the a_{ij} are scalars in K is called $m \times n$ matrix.

ii) The elements a_{ij} may be usually real or complex numbers.

iii) Notation:

$A \in K^{m,n}$

$$A = (a_1, \dots, a_n) = \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix} \quad \text{with} \quad a_k \in K^m, a^k \in K^n$$

iv) If $A = (a_{ij})_{m\times n}$, then the transpose of A is the $n \times m$ matrix A^T whose i, j entry is a_{ji}.

v) The matrix

$$E = E_{n\times n} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

is called the identity matrix.

Definition 2.2.2
Let $A = (a_{ij}), B = (b_{ij})$ two $m \times n$ matrices, $c \in \mathbb{R}$. We define

i) the sum of two matrices: $(A + B) = (a_{ij} + b_{ij})$

ii) the product of a matrix and scalar: $cA = (ca_{ij})$

Definition 2.2.3 (Matrix multiplication)
Let $A = (a_{ij})_{m\times n}$ be any $m \times n$ matrix, and let $B = (b_{ij})_{n\times k}$ be any $n \times k$ matrix. Then the product AB is defined by the matrix $C = (c_{ij})_{m\times k}$

$$c_{ij} := \sum_{\nu=1}^{n} a_{i\nu} b_{\nu j}$$

Example
Let $A \in K^{m,n}, B \in K^{n,k}$.

i) $AB = (Ab_1, Ab_2, \dots, Ab_k)$

ii) $Ax = \sum_{\nu=1}^{n} a_\nu x_\nu \quad \forall x \in K^n$

Theorem 2.2.4 (Properties)

i) The matrix product is not commutative.

ii) associative law: $(AB)C = A(BC)$

iii) distributive law: $A(B + C) = AB + BC$

iv) $AE = EA = A$

v) $(AB)^T = B^T A^T$

Definition 2.2.5 (Inverse matrix)
The $n \times n$ *square matrix* A *is said to be invertible, if there exists a matrix* B *with the property*

$$AB = BA = E.$$

Such a matrix B *is unique. We call* B *the inverse of* A *and denote it by* A^{-1}*. The matrix* A^{-1} *is called the inverse of* A*.*

2.3 Linear Systems of Equations

Let $A = (a_{ij})$ be an $m \times n$ matrix of real numbers, $b \in \mathbb{R}^m$. A set of m equations of the form

$$\sum_{k=1}^{n} a_{ik}x_k = b_i \quad i = 1, 2, \dots, m$$

is called a system of m linear equations in n unknowns. This system can be written more simply as

$$Ax = b.$$

With each linear system, we can associate the system $Ax = 0$, obtained by replacing b by 0. This is called homogeneous system. If $b \neq 0$, it is called nonhomogeneous system.

In solving a system of linear equations by elimination of unknowns we use the Gauss Jordan elimination method.

Definition 2.3.1 (Gauss Jordan method)
The Gauss Jordan method consists of applying three basic types of operations on the equations of a linear system:

(1) Interchanging two equations.

(2) Multiplying all the terms of an equation by an nonzero scalar.

(3) Adding to one equation a multiple of another equation.

To use matrices to solve systems of linear equations, we will focus on these row operations as they apply to matrices.

Definition 2.3.2 (Elementary Row operations)
(R1) Interchange two rows of A.

(R2) Multiply a row of A by a scalar $\neq 0$.

(R3) Add a scalar multiple of one row of A to another row of A.

Definition 2.3.3 (Elementary matrices)
An elementary matrix is a matrix formed by performing an elementary row operation on an identity matrix E.

Remark
Elementary row operations can be achieved by multiplying A on the left by elementary matrices.

Definition 2.3.4 (Echelon matrix)
An echelon matrix is an echelon matrix, if the number of zeros preceding the first nonzero entry of a row increases row by row until only zeros remain. The leading entry's of any nonzero rows are called the distinguished elements.

Example (Echelon matrix)

$$\begin{pmatrix} 0 & 1 & 3 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Definition 2.3.5 (Reduced matrix)
In particular, an echelon matrix is a reduced matrix, if the distinguished elements are

i) the only entries in their respective column.

ii) each equal to 1.

Example (Reduced matrix)

$$\begin{pmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Definition 2.3.6 (Row equivalent)
A matrix A is said to be row equivalent to a matrix B, if B can be obtained from A by performing a finite sequence of elementary row operations.

Theorem 2.3.7
Let A be an $m \times n$ matrix. Then $A \neq 0$ is row equivalent to an echelon matrix, in other words for all matrices $A \neq 0$ we have $m \times m$ elementary matrices $G_1, \ldots, G_s$ such that $B := G_s \cdots G_1 A$ is an echelon matrix.

Definition 2.3.8 (Row space, Row rank)
Let A be an $m \times n$ matrix.

i) The subspace of R^n consisting of all linear combinations of row vectors of A is the row space of A.

ii) The subspace of R^m consisting of all linear combinations of column vectors of A is the column space of A.

iii) The dimension of the row space and of the column space are called, respectively, the row rank and column rank.

Theorem 2.3.9
Let A be an $m \times n$ matrix.

i) Elementary row operations does not change the row rank and the column rank of A.

ii) If A has an echelon matrix with r nonzero rows, then the row rank equals r.

Theorem 2.3.10
The row rank and the column rank are equal.

Definition 2.3.11 (Rank of a matrix)
The rank of a matrix A, written $\operatorname{rank}(A)$, is the common value of its row rank and column rank.

Definition 2.3.12 (Augmented matrix)
Consider the linear system $Ax = b$. The matrix

$$(A|b) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

is called the augmented matrix.

Theorem 2.3.13 (Existence of a solution)
The nonhomogeneous system $Ax = b$ has a solution if and only if the coefficient matrix A and the augmented matrix $(A|b)$ have the same rank.

$$\operatorname{rank}(A) = \operatorname{rank}(A|b)$$

Theorem 2.3.14 (General solution of $\boldsymbol{Ax = b}$)
Let $Ax = b$, $A \in \mathbb{R}^{m,n}, b \in \mathbb{R}^m$. Suppose x_0 is a particular solution of the nonhomogeneous system and $X_{A,0}$ is the general solution of the associated homogeneous system. Then

i) $X_{A,0}$ is a subspace of $\mathbb{R}^m$ with dimension $\dim X_{A,0} = n - \operatorname{rank}(A)$.

ii) $x_0 + X_{A,0} = \{x_0 + x_1 : x_1 \in X_{A,0}\}$ is the general solution of the nonhomogeneous system.

Corollary 2.3.15
The homogeneous system has a nontrivial solution if and only if $\operatorname{rank}(A) < m$. In that case the echelon form of A has $n - \operatorname{rank}(A)$ free variables , say $x_{i_1}, x_{i_2}, \dots , x_{i_{n-r}}$. Let v_j be the solution by setting $x_{i_j} = 1$ and all other free variables $= 0$. Then the solutions $v_1, \dots , v_{n-r}$ are independent and so form a basis of the solution space.

Procedure for solving $Ax = b$

- Reduce $(A|b)$ to obtain the corresponding augmented matrix $(A'|b')$. Solutions of $Ax = b$ are the same as solutions of $A'x = b'$.
- If $\text{rank}(A) \neq \text{rank}(A|b)$, then the system has no solution.
- Identify the $n - \text{rank}(A)$ dependent unknowns.
- Write each dependent x_i in terms of independent unknowns and b_i. The independent unknowns can be assigned any values, and are arbitrary constants in this general solution.

Example (A system with a unique solution)
Consider the system

$$\begin{pmatrix} 2 & -5 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \\ 10 \end{pmatrix}$$

By performing the Gauss Jordan method on the augmented matrix, we get the corresponding augmented matrix

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right)$$

Solving these equations backwards give us the solution

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 124 \\ 75 \\ 31 \end{pmatrix}$$

Example (A system with more than one solution)
Consider a system $Ax = b$ with

$$\begin{pmatrix} 2 & -5 & 4 & 1 & -1 \\ 1 & -2 & 1 & -1 & 1 \\ 1 & -4 & 6 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \\ 10 \end{pmatrix}$$

and the corresponding augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & -16 & 19 & 124 \\ 0 & 1 & 0 & -9 & 11 & 75 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right)$$

If we let $x_4 = t$, $x_5 = s$, the vector

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 124 \\ 75 \\ 81 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 16 \\ 9 \\ 3 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} -19 \\ -11 \\ -4 \\ 0 \\ 1 \end{pmatrix}$$

gives the general solution of the system.

Example (A system with no solution)
Applying the row operations on the system

$$\begin{aligned} 2x_1 - 5x_2 + 4x_3 &= -3 \\ 1x_1 - 2x_2 + 1x_3 &= 5 \\ 1x_1 - 4x_2 + 5x_3 &= 10 \end{aligned}$$

we arrive at the corresponding augmented matrix

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 0 & 31 \end{array}\right)$$

Because $\mathrm{rank}(A) = 2 \neq 3 = \mathrm{rank}(A|b)$ *the system has no solution.*

2.4 Linear Transforms

Definition 2.4.1 (Linear Transform)
Let V, W be linear spaces over the same field K. A mapping $T : V \to W$ is called a linear transformation if

i) $T(x + y) = T(x) + T(y) \quad \forall x, y \in V$

ii) $T(cx) = cT(x) \quad \forall x \in V, c \in K$

$L(V, W)$ denotes the linear space of all linear transformations of V into W.

Remark
By induction, we have $T(\sum_{i=1}^{n} c_i x_i) = \sum_{i=1}^{n} c_i T(x_i)$ if and only if $T \in L(V, W)$.

Definition 2.4.2 (Image, Kernel)
Let $T \in L(V, W)$.

i) The image of T, written $\operatorname{img}(T)$ *is the set* $\operatorname{img}(T) = \{w \in W : T(v) = w \text{ for some } v \in V\}$

ii) The kernel of T, written $\operatorname{kern}(T)$ *is the set* $\operatorname{kern}(T) = T^{-1}(\{0\}) = \{v \in V : T(v) = 0\}$

Corollary 2.4.3
Let $T \in L(V, W)$.

i) $\operatorname{img}(T) \subset W$ *is a subspace of W.*

ii) $\operatorname{kern}(T) \subset V$ *is a subspace of V.*

Theorem 2.4.4
Let V be of finite dimension and let $T \in L(V, W)$. Then

$$\dim \operatorname{kern}(T) + \dim \operatorname{img}(T) = \dim V.$$

Theorem 2.4.5
Let $T \in L(V, W)$. Then the following statements are equivalent:

i) T is one-to-one on V.

ii) T is invertible and $T^{-1} : T(V) \to V$ is linear.

iii) $\operatorname{kern}(T) = \{0\}$.

Theorem 2.4.6
Let $T \in L(V, W)$ and $\dim V = n$. Then the following statements are equivalent:

i) T *is one-to-one on* V.

ii) *If* $\{e_1, \dots, e_r\}$ *are independent in* V, *then* $\{T(e_1), \dots, T(e_r)\}$ *are independent in* $T(V)$.

iii) $\dim T(V) = n$.

iv) *If* $\{e_1, \dots, e_n\}$ *is a basis for* V, *then* $\{T(e_1), \dots, T(e_n)\}$ *is a basis for* $T(V)$.

Theorem 2.4.7
Let $\{e_1, \dots, e_n\}$ *be a basis for* V *and* $\{u_1, \dots, u_n\}$ *arbitrary elements in* W. *Then there is one and only linear transform* $T : V \to W$ *such that*

$$T(e_k) = u_k \quad k = 1, 2, \dots, n.$$

(T is completely determined by its action on a given set of basis elements.)

Corollary 2.4.8
This T *maps an arbitrary element* x *as follows:*
If $x = \sum_{k=1}^{n} x_k e_k$, *then* $T(x) = \sum_{k=1}^{n} x_k u_k$.

Definition 2.4.9
If $(w_1, \dots, w_m)$ *is a ordered basis for* W, *then* $T(e_k)$ *can be expressed uniquely as* $T(e_k) = \sum_{i=1}^{m} t_{ik} w_i \quad k = 1, \dots, n$. *The matrix* $(t_{ik})_{m \times n}$ *is called the matrix representation of* T *relative to the given choice of ordered bases* $(e_1, \dots, e_n)$ *for* V *and* $(w_1, \dots, w_m)$ *for* W.

Corollary 2.4.10
Let $T \in L(V, W)$, $(e_1, \dots, e_n)$ *and* $(w_1, \dots, w_m)$ *ordered bases for* V *and* W, *and* (t_{ik}) *the matrix representation of* T, $x = \sum_{k=1}^{n} x_k e_k$ *and* $T(x) = \sum_{i=1}^{m} y_i w_i$. *Then*

$$y_i = \sum_{k=1}^{n} t_{ik} x_k \quad i = 1, 2, \dots, m.$$

Definition 2.4.11 (Isomorphism)
A linear transform $T \in L(V, W)$ *is called an isomorphism if it is one-to-one. The linear spaces* V, W *are said to be isomorphic if there is an isomorphism of* V *onto* W.

Theorem 2.4.12 (Isomorphism theorem)
Let $\dim V = n, \dim W = m$, $(e_1, \dots, e_n)$ *a basis for* V, $(w_1, \dots, w_m)$ *a basis for* W *and* $T \in L(V, W)$. *Then the transformation* $\phi : L(V, W) \to K^{m,n}$ $\phi(T) = (t_{ik})_{m \times n}$ *is linear and one-to-one.*

2.5 Determinants

Definition 2.5.1 (Permutation)
Let $n \in \mathbb{N}$.

i) *A one-to-one mapping* $\sigma : \{1, 2, , \dots , n\} \to \{1, 2, \dots , n\}$ *is called a permutation of the integers* 1 *to* n.

ii) S_n *denotes the set* $S_n = \{\sigma : \sigma \text{ permutation of } \{1, 2, \dots , n\}\}$.

iii) *We denote the* $\sigma \in S_n$ *by*

$$\sigma = \begin{pmatrix} 1, & 2, & \dots, & n \\ \sigma(1), & \sigma(2), & \dots, & \sigma(n) \end{pmatrix} = (\sigma_1, \sigma_2, \dots , \sigma_n)$$

iv) *We say* σ *is even or odd according to whether there is an even or odd number of pairs* (i, k) *for which* $i > k$, *but* i *precedes* k *in* σ.

v) *The signum or parity of* $\sigma \in S_n$ *is defined by*

$$\operatorname{sign}(\sigma) = \prod_{1 \le i < j \le n} \operatorname{sign}(\sigma(j) - \sigma(i))$$

Example
$\sigma = (2, 4, 3, 1), \tau = (4, 2, 3, 1)$

$$\sigma \circ \tau = (1, 4, 3, 2) \neq \tau \circ \sigma = (2, 1, 3, 4)$$

$$\sigma^{-1} = (4, 1, 3, 2) \quad \operatorname{sign}(\sigma^{-1}) = 1$$

Theorem 2.5.2

i) $\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau) \quad \forall \sigma, \tau \in S_n$

ii) $\forall \sigma \in S_n \quad \exists \sigma^{-1} \in S_n : \quad \sigma \circ \sigma^{-1} = \sigma^{-1}\sigma = id$

iii) $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$

iv) $\tau \in S_n \Rightarrow S_n = \{\sigma \circ \tau : \sigma \in S_n\}$

v) $S_n = \{\sigma^{-1} : \sigma \in S_n\}$

vi) $|S_n| = n!$

Definition 2.5.3 (Determinant)
Let $A = (a_{ij})_{n \times n}$ *be a square matrix. The number*

$$\det A := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma_1} \cdot \ldots \cdot a_{n\sigma_n}$$

is called the determinant of A.

Example

i)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\Rightarrow detA = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

ii)

$$\det \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} = a_{11} \cdot \ldots \cdot a_{nn}$$

Theorem 2.5.4 (Properties)
Let A be an n square matrix, E the identity matrix.

i) $\det A = \det(A^T)$

ii) $\det E = 1$

iii) $\det(a_1, \dots, a_k, \dots, a_l, \dots, a_n) = -\det(a_1, \dots, a_l, \dots, a_k, \dots, a_n)$

iv) $\det(a_1, \dots, \lambda a_k, \dots, a_n)) = \lambda \det(a_1, \dots, a_n)$ *for all scalars* λ

v) $\det(a_1, \dots, a'_k + a''_k, \dots, a_n) = \det(a_1, \dots, a'_k, \dots, a_n) + \det(a_1, \dots, a''_k, \dots, a_n)$

Corollary 2.5.5
Let A be an n square matrix.

i) $\det(a_1, \dots, a_n) = 0$ *if* $a_k = 0$ *for some* k.

ii) $\det(a_1, \dots, a_n) = 0$ *if* $a_i = a_j$ *for some* $i \neq j$.

iii) $\det A = 0$ *if rows (columns) are dependent.*

Theorem 2.5.6 (Product formula for determinants)
The determinant is a multiplicative function, for any $n \times n$ matrices A and B we have

$$\det(AB) = \det A \det B.$$

Definition 2.5.7 (Minor, Cofactor)
Let A be an n square matrix. Let M_{ij} denote the $(n-1)$ square sub-matrix of A obtained by deleting its ith row and jth column.

i) $\det(M_{ij})$ *is called the minor of the element a_{ij} of A.*

ii) $A_{ij} = (-1)^{i+j} \det M_{ij}$ *is called the cofactor of* a_{ij}.

Theorem 2.5.8 (Expansion formula)
Let A_{kj} *denote the matrix obtained from* A *by replacing the* j*th column by the unit coordinate vector* e_k. *For any* $n \times n$ *matrix* A *we have*

i) (expansion by row i*)*

$$\sum_{\nu=1}^{n} a_{i\nu} A_{i\nu} = \det A$$

ii) (expansion by column i*)*

$$\sum_{\nu=1}^{n} a_{\nu i} A_{\nu i} = \det A$$

Definition 2.5.9 (Adjoint matrix)
The transpose of the matrix of cofactors of elements a_{ij} *of* A, *denoted by* $\mathrm{adj}A$, *is called the adjoint of* A.

$$\mathrm{adj}A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix}$$

Theorem 2.5.10
For any square matrix A

$$A(\mathrm{adj}A) = (\mathrm{adj}A)A = (\det A)E.$$

Corollary 2.5.11
If $\det A \neq 0$ *then* $A^{-1} = \frac{1}{\det A}(\mathrm{adj}A)$.

Theorem 2.5.12 (Cramer's rule)
Let A *be an* $n \times n$ *matrix with* $\det A \neq 0$. *Then there is a unique solution for the system* $Ax = b \quad \forall b \in \mathbb{R}^n$ *given by* $x = A^{-1}b$, *i. e.*

$$x_k = \frac{1}{\det A} \det(a_1, \dots, a_{k-1}, b, a_{k+1}, \dots, a_n)$$

2.6 Eigenvalues and Eigenvectors

Definition 2.6.1 (Eigenvalue, Eigenvector)
A scalar λ is an eigenvalue of A if and only if there is $x \neq 0$:

$$Ax = \lambda x \quad x \neq 0.$$

x is called eigenvector of A belonging to λ. The scalar λ is called an eigenvalue corresponding to x.

Theorem 2.6.2
Let A be an $n \times n$ matrix.

i) λ is an eigenvalue of A if and only if $\det(\lambda E - A) = 0$.

ii) If λ is an eigenvalue of A, then any solution $x \neq 0$ of $(\lambda E - A)x = 0$ is a corresponding eigenvalue.

Definition 2.6.3 (Characteristic polynomial)
Let A be an $n \times n$ matrix. Then $c_a(\lambda) := \det(\lambda E - A)$ is called the characteristic polynomial of A.

Theorem 2.6.4 (Gerschgorin)
Let $A = (a_{ij})_{n \times n}$. Let $C_i, \tilde{C}_k$ be the circles

$$C_i := \{\lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq \sum_{k \neq i}^{n} |a_{ik}|\} \quad i = 1, \dots, n$$
$$\tilde{C}_k := \{\lambda \in \mathbb{C} : |\lambda - a_{kk}| \leq \sum_{i \neq k}^{n} |a_{ik}|\} \quad k = 1, \dots, n$$

Then each eigenvalue of A lies in $\cup_{i=1}^{n} C_i$, $\cup_{K=1}^{n} \tilde{C}_k$. The circles $C_i, \tilde{C}_k$ are called Gerschgorin circles for A.

Theorem 2.6.5 (Cayley-Hamilton)
Let A be an $n \times n$ matrix and $c_A(\lambda) = \det(\lambda E - A)$ the characteristic polynomial. Then $c_A(A) = 0 \in K^{n,n}$, i. e. every matrix is a zero of its characteristic polynomial.

Definition 2.6.6 (Diagonalisable)
Let A be an $n \times n$ matrix. A is diagonalisable if and only if there exists an invertible $n \times n$ matrix P such that $P^{-1}AP$ is a diagonal matrix.

Corollary 2.6.7
Let A be a $n \times n$ matrix. If A is diagonalisable, then there exists an independent set $x_1, \dots, x_n$ such that

$$Ax_k = \lambda_k x_k, \quad x_k \neq 0 \quad k = 1, \dots, n.$$

Theorem 2.6.8 (Condition for diagonalizability)
Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. If there is an independent set $\{x_1, \ldots, x_n\}$ of eigenvectors, then A is diagonalisable.

Suppose x_i is an eigenvector associated with λ_i, and let $P := (x_1, \ldots, x_n)$, then P is invertible and diagonalises A.

Theorem 2.6.9
Let $\{x_1, \ldots, x_k\}$ be a set of Eigenvectors of A and assume that the corresponding eigenvalues $\lambda_1, \ldots, \lambda_k$ are distinct. Then $\{x_1, \ldots, x_k\}$ are independent.

Corollary 2.6.10
An $n \times n$ matrix with n distinct eigenvalues is diagonalisable.

Definition 2.6.11 (Orthogonal matrix, symmetric matrix)
Let A be an $n \times n$ matrix.

i) A is called an orthogonal matrix if $AA^T = A^T A = E$.

ii) A is symmetric if $A = A^T$.

Corollary 2.6.12
A is orthogonal if and only if the row vectors of A are orthonormal.

Corollary 2.6.13
If A is an orthogonal matrix, then $\det A = \pm 1$.

Theorem 2.6.14 (Principal axis theorem)
Let A be a symmetric matrix. Then there exists an orthogonal matrix that diagonalises A.

Definition 2.6.15 (Quadratic form)
An expression $\sum_{j,k=1}^{n} a_{jk} \bar{z}_j z_k$, in which $a_{jk} \in \mathbb{C}, z_j \in \mathbb{C}$ is called quadratic form. If $a_{jk} \in \mathbb{R}, z_j \in \mathbb{R}$, then $\sum_{j,k=1}^{n} a_{jk} x_j x_k$ is a real quadratic form.

Remark
A real quadratic form can be written as $x^T A x$ with a symmetric matrix A.

Theorem 2.6.16 (Principal axis theorem)
Let $A = (a_{ij})_{n \times n}$ be a symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Let Q be an orthogonal matrix that diagonalises A. Then the change of coordinates $x = Qy$ transforms $\sum_{i,j}^{n} a_{ij} x_i x_j$ to $\sum_{i=1}^{n} \lambda_i y_i^2$.

Example
Consider the equation

$$2x_1^2 + 4x_1x_2 + 5x_2^2 + 4x_1 + 13x_2 - 1/4 = 0 \qquad (2.1)$$

We write as (2.1) as

$$x^T A x + b^T x - 1/4 = 0$$

where $A = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}, B = \begin{pmatrix} 4 \\ 13 \end{pmatrix}$. *A has the eigenvalues* $\lambda_1 = 1, \lambda_2 = 6$ *and an orthonormal set of eigenvectors* $u_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, u_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. *An orthogonal diagonalising matrix is* $Q = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$. *By rotation* $y = Qx$ *(2.1) reduces to*

$$y_1^2 + 6y_2^2 - \sqrt{5}y_1 + 6\sqrt{5}y_2 - 1/4 = 0$$

By completing the squares in y_1 *and* y_2

$$(y_1 - \frac{1}{2}\sqrt{5})^2 + 6(y_2 + \frac{1}{2}\sqrt{5})^2 = 9$$

This is the equation of an ellipse with its center at $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ *in the* $y_1 y_2$*-system.*

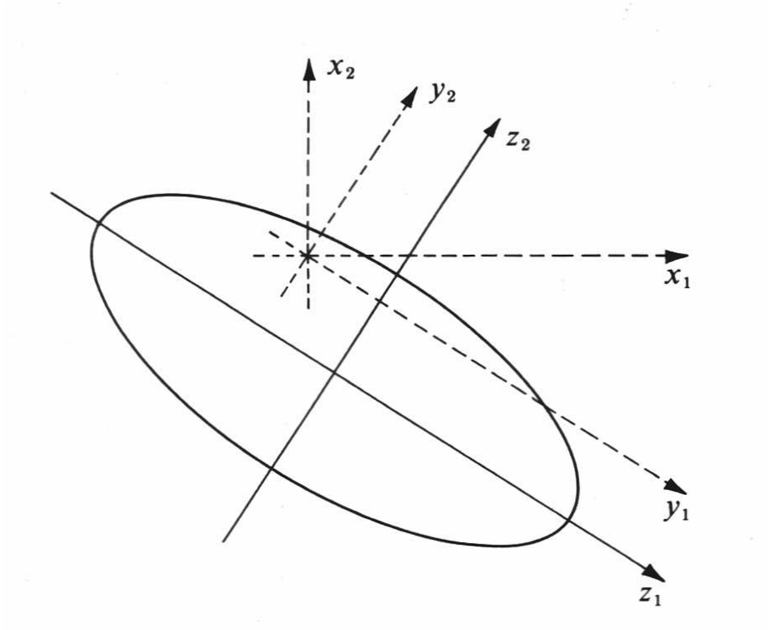

Figure 2.1.

Bibliography

[1] Tom Apostel. *Calculus, One-Variable Calculus, with an Introduction to Linear Algebra*, volume 1. Blaisdell Publishing Company, second edition, 1967.

[2] Tom Apostel. *Calculus, Multi-Variable Calculus and Linear Algebra, with Applications to Differential Equations and Probaility*, volume 2. Xerox College Publishing, second edition, 1969. ISBN 0-536-00008-5.

[3] Anthony Edward Labarre. *Intermediate Mathematical Analysis.* Holt, Rinehart and Winston, Inc., 1968.

[4] Serge Lang. *Undergraduate Analysis.* Springer-Verlag, 1983. ISBN 3-540-90800-5.

[5] Walter Rudin. *Principles of Mathematical Analysis.* Holt, Rinehart and Winston, Inc., third edition, 1976. ISBN 0-07-085613-3.

[6] Karl Robert Stromberg. *An Introduction to Classical Real Analysis.* Wadsworth International Group, 1981. ISBN 0-534-98012-0.