Calculus 2 (UMB 565I)

Jan.Eisner@prf.jcu.cz

lecture: Monday 8:00–9:30 AV-P1

excercices: Thursday 8:00–9:30 C2/C3

fix.prf.jcu.cz/~eisner/lock/UMB-565I-CAL2/

Contents

1 Introduction

1.1 What you will know after the lecture

- to operate with infinite sequences
- to sum (some) infinite sequences
- to antiderive (some/a class of) one-variable functions
- to apply antiderivation (e.g. to calculate an area given by a function, a surface of a rotating body, a length of a curve, ...)
- to apply differential calculus in more dimensions
- to locate local extrema of functions of two or more variables
- to locate global extrema of functions of two or more variables with respect to a given domain

Literature for further study:

Introduction to Real Analysis by Prof. W. Trench

(`http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF`)

Sections 2–4 are taken from the Calculus Bible.

Section 6 follows lecture notes of Prof. J. Daněček.

2 INFINITE SERIES

2.1 Sequences

Definition 8.1.1 An infinite sequence (or sequence) is a function, say f, whose domain is the set of all integers greater than or equal to some integer m. If n is an integer greater than or equal to m and $f(n) = a_n$, then we express the sequence by writing its range in any of the following ways:

1. $f(m), f(m+1), f(m+2), \ldots$
2. $a_m, a_{m+1}, a_{m+2}, \ldots$
3. $\{f(n) : n \geq m\}$
4. $\{f(n)\}_{n=m}^{\infty}$
5. $\{a_n\}_{n=m}^{\infty}$

Definition 8.1.2 A sequence $\{a_n\}_{n=m}^{\infty}$ is said to converge to a real number L (or has limit L) if for each $\varepsilon > 0$ there exists some positive integer M such that $|a_n - L| < \varepsilon$ whenever $n \geq M$. We write,

$$\lim_{n\to\infty} a_n = L \quad \text{or} \quad a_n \to L \text{ as } n \to \infty.$$

If the sequence does not converge to a finite number L, we say that it diverges.

Theorem 8.1.1 Suppose that c is a positive real number, $\{a_n\}_{n=m}^{\infty}$ and $\{b_n\}_{n=m}^{\infty}$ are convergent sequences. Then

1. $\lim_{n\to\infty} (ca_n) = c \lim_{n\to\infty} a_n$
2. $\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$
3. $\lim_{n\to\infty} (a_n - b_n) = \lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n$
4. $\lim_{n\to\infty} (a_n b_n) = \lim_{n\to\infty} a_n \lim_{n\to\infty} b_n$
5. $\lim_{n\to\infty} a_n/b_n = \lim_{n\to\infty} a_n / \lim_{n\to\infty} a_n$, if $\lim b_n \neq 0$.
6. $\lim_{n\to\infty} (a_n)^c = (\lim_{n\to\infty} a_n)^c$
7. $\lim_{n\to\infty} (e^{a_n}) = e^{\lim_{n\to\infty} a_n}$
8. Suppose that $a_n \leq b_n \leq c_n$ for all $n \geq m$ and
$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L.$$
Then
$$\lim_{n\to\infty} b_n = L.$$

2.2 Monotone Sequences

Definition 8.2.1 Let $\{a_n\}_{n=m}^{\infty}$ be a given sequence. Then $\{a_n\}_{n=m}^{\infty}$ is said to be

(a) increasing if $a_n < a_{n+1}$ for all $n \geq m$;

(b) decreasing if $a_n + 1 < a_n$ for all $n \geq m$;

(c) nondecreasing if $a_n \leq a_{n+1}$ for all $n \geq m$;

(d) nonincreasing if $a_{n+1} \leq a_n$ for all $n \geq m$;

(e) bounded if $a \leq a_n \leq b$ for some constants a and b and all $n \geq m$;

(f) monotone if $\{a_n\}_{n=m}^{\infty}$ is increasing, decreasing, nondecreasing or nonincreasing.

(g) a Cauchy sequence if for each $\varepsilon > 0$ there exists some M such that $|a_{n_1} - a_{n_2}| < \varepsilon$ whenever $n_1 \geq M$ and $n_2 \geq M$.

Theorem 8.2.1

(a) A monotone sequence converges to some real number if and only if it is a bounded sequence.

(b) A sequence is convergent if and only if it is a Cauchy sequence.

2.3 Infinite Series

Definition 8.3.1 Let $\{a_n\}_{n=1}^{\infty}$ be a given sequence. Let

$$\begin{aligned} s_1 &= a_1, \\ s_2 &= a_1 + a_2, \\ s_3 &= a_1 + a_2 + a_3, \\ &\vdots \\ s_n &= \sum_{k=1}^{n} a_k \end{aligned}$$

for all natural number n. If the sequence $\{s_n\}_{n=1}^{\infty}$ converges to a finite number L, then we write

$$L = a_1 + a_2 + a_3 \cdots = \sum_{k=1}^{\infty} a_k$$

We call $\displaystyle\sum_{k=1}^{\infty} a_k$ an infinite series and write

$$L = a_1 + a_2 + a_3 \cdots = \sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=1}^{n} a_k = L.$$

We say that L is the sum of the series and the series converges to L. If a series does not converge to a finite number, we say that it diverges. The sequence $\{s_n\}_{n=1}^{\infty}$ is called the sequence of the n-th partial sums of the series.

Theorem 8.3.1 Suppose that a and r are real numbers and $a \neq 0$. Then the geometric series

$$a + ar + ar^2 + \ldots = \sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$$

if the quotioent r satisfies $|r| < 1$.

The geometric series diverges if $|r| \geq 1$.

Proof. For each natural number n, let

$$s_n = a + ar + \ldots + ar^{n-1}.$$

On multiplying both sides by r, we get

$$\begin{aligned} rs_n &= ar + ar^2 + \ldots + ar^{n-1} + ar^n \\ s_n - rs_n &= a - ar^n \\ (1-r)s_n &= a(1-r^n) \\ s_n &= \frac{a}{1-r} - \frac{a}{1-r}r^n. \end{aligned}$$

If $|r| < 1$, then

$$\lim_{n\to\infty} s_n = \frac{a}{1-r} - \frac{a}{1-r}\lim_{n\to\infty} r^n = \frac{a}{1-r}.$$

If $|r| > 1$, then $\lim_{n\to\infty} r^n$ is not finite and so the sequence $\{sn\}_{n=1}^{\infty}$ of n-th partial sums diverges.

If $r = 1$, then $s_n = na$ and $\lim_{n\to\infty} na$ is not a finite number.

This completes the proof of the theorem. ■

Theorem 8.3.2 (Divergence Test) If the series $\sum_{k=1}^{\infty} a_k$ converges, then $\lim_{n\to\infty} a_n = 0$. If $\lim_{n\to\infty} a_n \neq 0$, then the series diverges.

Proof. Suppose that the series converges to L. Then

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \left(\sum_{k=1}^{n} a_k \sum_{k=1}^{n-1} a_k \right) = \lim_{n\to\infty} \sum_{k=1}^{n} a_k - \lim_{n\to\infty} \sum_{k=1}^{n-1} a_k = L - L = 0.$$

The rest of the theorem follows from the preceding argument. This completes the proof of this theorem.

■

Theorem 8.3.3 (The Integral Test) Let f be a function that is defined, continuous and decreasing on $[1, \infty)$ such that $f(x) > 0$ for all $x \geq 1$. Then

$$\sum_{n=1}^{\infty} f(n) \quad \text{and} \quad \int_1^{\infty} f(x)\,dx$$

either both converge or both diverge.

Proof. Suppose that f is decreasing and continuous on $[1, \infty)$, and $f(x) > 0$ for all $x \geq 1$. Then for all natural numbers n, we get,

$$\sum_{k=2}^{n+1} f(k) \leq \int_1^{n+1} f(x)\,dx \leq \sum_{k=1}^{n} f(k)$$

It follows that,

$$\sum_{k=2}^{\infty} f(k) \leq \int_1^{\infty} f(x)\,dx \leq \sum_{k=1}^{\infty} f(k)$$

Since $f(1)$ is a finite number, it follows that

$$\sum_{k=1}^{\infty} f(n) \quad \text{and} \quad \int_1^{\infty} f(x)\,dx$$

either both converge or both diverge. This completes the proof of the theorem. $\blacksquare$

Theorem 8.3.4 Suppose that $p > 0$. Then the p-series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$ and diverges if $0 < p \leq 1$. In particular, the harmonic series

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges.

Proof goes via the Integral Test.

2.4 Series with Positive Terms

Theorem 8.4.1 (Algebraic Properties) Suppose that $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ are convergent series and $c > 0$. Then

1. $\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$
2. $\sum_{k=1}^{\infty}(a_k - b_k) = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} b_k$
3. $\sum_{k=1}^{\infty} ca_k = c\sum_{k=1}^{\infty} a_k$
4. If m is any natural number, then the series $\sum_{k=1}^{\infty} c_k$ and $\sum_{k=m}^{\infty} c_k$ either both converge or both diverge.

Theorem 8.4.2 (Comparison Test) Suppose that $0 < a_n \leq b_n$ for all natural numbers $n \geq 1$.

(a) If there exists some M such that $\sum_{k=1}^{n} a_k \leq M$, for all natural numbers n, then $\sum_{k=1}^{\infty} a_k$ converges.

If there exists no such M, then the series diverges.

(b) If $\sum_{k=1}^{\infty} b_k$ converges, then $\sum_{k=1}^{\infty} a_k$ converges.

(c) If $\sum_{k=1}^{\infty} a_k$ diverges, then $\sum_{k=1}^{\infty} b_k$ diverges.

(d) If $c_n > 0$ for all natural numbers n, and

$$\lim_{n\to\infty} \frac{a_n}{c_n} = L, \quad 0 < L < \infty,$$

then the series $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} c_k$ either both converge or both diverge.

Theorem 8.4.3 (Ratio Test) Suppose that $0 < a_n$ for every natural number n and

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = r.$$

Then the series $\displaystyle\sum_{n=1}^{\infty} a_n$

(a) converges if $r < 1$;

(b) diverges if $r > 1$;

(c) may converge or diverge if $r = 1$; the test fails.

Theorem 8.4.4 (Root Test) Suppose that $0 < a_n$ for each natural number n and

$$\lim_{n\to\infty} (a_n)^{1/n} = r.$$

Then the series $\displaystyle\sum_{n=1}^{\infty} a_n$

(a) converges if $r < 1$;

(b) diverges if $r > 1$;

(c) may converge or diverge if $r = 1$; the test fails.

2.5 Alternating Series

Definition 8.5.1 Suppose that for each natural number n, b_n is positive or negative. Then the series $\sum_{k=1}^{\infty} b_k$ is said to converge

(a) absolutely if the series $\sum_{k=1}^{\infty} |b_k|$ converges;

(b) conditionally if the series $\sum_{k=1}^{\infty} b_k$ converges but $\sum_{k=1}^{\infty} |b_k|$ converges diverges.

Theorem 8.5.1 If a series converges absolutely, then it converges.

Definition 8.5.2 Suppose that for each natural number n, $a_n > 0$. Then an alternating series is a series that has one of the following two forms:

(a) $a_1 - a_2 + a_3 - \cdots + (-1)^{n+1} a_n + \cdots = \displaystyle\sum_{k=1}^{\infty} (-1)^{k+1} a_k$

(b) $-a_1 + a_2 - a_3 + \cdots + (-1)^n a_n + \cdots = \displaystyle\sum_{k=1}^{\infty} (-1)^k a_k.$

Theorem 8.5.2 Suppose that $a_n > a_{n+1} > 0$ for all natural numbers m, and $\lim\limits_{n\to\infty} a_n = 0$. Then

(a) $\displaystyle\sum_{n=1}^{\infty} (-1)^n a_n$ and $\displaystyle\sum_{n=1}^{\infty} (-1)^{n+1} a_n$ both converge.

(b) $\left| \displaystyle\sum_{k=1}^{\infty} (-1)^{k+1} a_k - \sum_{k=1}^{n} (-1)^{k+1} a_k \right| < a_{n+1}$, for all n;

(c) $\left| \displaystyle\sum_{k=1}^{\infty} (-1)^{k} a_k - \sum_{k=1}^{n} (-1)^{k} a_k \right| < a_{n+1}$, for all n;

Theorem 8.5.3 Consider a series $\{a_k\}_{k=1}^{\infty}$. Let

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = L, \qquad \lim_{n\to\infty}|a_n|^{1/n} = M.$$

(a) If $L < 1$, then the series $\{a_k\}_{k=1}^{\infty}$ converges absolutely.

(b) If $L > 1$, then the series $\{a_k\}_{k=1}^{\infty}$ does not converge absolutely.

(c) If $M < 1$, then the series $\{a_k\}_{k=1}^{\infty}$ converges absolutely.

(b) If $M > 1$, then the series $\{a_k\}_{k=1}^{\infty}$ does not converge absolutely.

(e) If $L = 1$ or $M = 1$, then the series $\{a_k\}_{k=1}^{\infty}$ may or may not converge absolutely.

2.6 Power Series

Definition 8.6.1 If $a_0, a_1, a_2, \ldots$ is a sequence of real numbers, then the series

$$\sum_{k=0}^{\infty} a_k(x-c)^k$$

is called a *power series* in x.

A real number c is called the *centre* of power series.

A positive number r is called the *radius of convergence* and the interval $(c-r, c+r)$ is called *the interval of convergence* of the power series if the power series converges absolutely for all x in $(c-r, c+r)$ and diverges for all x such that $|x-c| > r$.

The end point $x = c + r$ is included in the interval of convergence if $\sum\limits_{k=0}^{\infty} a_k r^k$ converges.

The end point $x = c - r$ is included in the interval of convergence if the series $\sum\limits_{k=0}^{\infty} (-1)^k a_k r^k$ converges.

If the power series converges only for $x = c$, then the radius of convergence is defined to be zero.

If the power series converges absolutely for all real x, then the radius of convergence is defined to be ∞.

Theorem 8.6.1 If the series $\sum\limits_{n=0}^{\infty} a_n x^n$ converges for $x = r \neq 0$, then the series $\sum\limits_{n=0}^{\infty} a_n x^n$ converges absolutely for all numbers x such that $|x| < |r|$.

Theorem 8.6.2 If the series $\sum_{n=0}^{\infty} a_n(x-c)^n$ converges for some $x-c = r \neq 0$, then the series $\sum_{n=0}^{\infty} a_n(x-c)^n$ converges absolutely for all x such that $|x - c| < |r|$.

Theorem 8.6.3 Let $\sum_{n=0}^{\infty} a_n x^n$ be any power series. Then exactly one of the following three cases is true.

(i) The series converges only for $x = 0$.

(ii) The series converges for all $x \in \mathbb{R}$.

(iii) There exists a number R such that the series converges for all x with $|x| < R$ and diverges for all x with $|x| > R$.

Theorem 8.6.4 Let $\sum_{n=0}^{\infty} a_n (x-c)^n$ be any power series. Then exactly one of the following three cases is true.

(i) The series converges only for $x = c$.

(ii) The series converges for all $x \in \mathbb{R}$.

(iii) There exists a number R such that the series converges for all x with $|x - c| < R$ and diverges for all x with $|x - c| > R$.

Theorem 8.6.5 If $R > 0$ and the series $\sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < R$, then the series $\sum_{n=0}^{\infty} n a_n x^{n-1}$, obtained by term-by-term differentiation of $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for $|x| < R$.

Theorem 8.6.6 If $R > 0$ and the series $\sum_{n=0}^{\infty} a_n (x-c)^n$ converges for all x such that $|x-a| < R$, then the series $\sum_{n=0}^{\infty} a_n (x-c)^n$ may be differentiated with respect to x any number of times and each of the differential series converges for all x such that $|x-a| < R$.

Theorem 8.6.7 Suppose that $R > 0$ and $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and R is radius of convergence of the series $\sum_{n=0}^{\infty} a_n x^n$. Then $f(x)$ is continuous for all x such that $|x| < R$.

Theorem 8.6.8 Suppose that $R > 0$ and $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and R is radius of convergence of the series

$\sum_{n=0}^{\infty} a_n x^n$. For each x such that $|x| < R$, we define

$$F(x) = \int_0^x f(t)dt.$$

Then, for each x such that $|x| < R$, we get

$$F(x) = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1}.$$

Theorem 8.6.9 Suppose that $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for all $|x| < R$, where $R > 0$ is the radius of convergence of the series $\sum_{n=0}^{\infty} a_n x^n$. Then $f(x)$ has continuous derivatives of all orders for $|x| < R$ that are obtained by successive term-by-term differentiations of $\sum_{n=0}^{\infty} a_n x^n$

Definition 8.6.2 The radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ is

(a) zero, if the series converges only for $x = a$;

(b) r, if the series converges absolutely for all x such that $|x-a| < r$ and diverges for all x such that $|x-a| > r$.

(c) ∞, if the series converges absolutely for all real number x.

If the radius of convergence of the power series in $(x-a)$ is r, $0 < r < \infty$, then the interval of convergence of the series is $(a-r, a+r)$. The end points $x = a+r$ or $x = a-r$ are included in the interval of convergence if the corresponding series $\sum_{n=0}^{\infty} a_n r^n$ or $\sum_{n=0}^{\infty} (-1)^n a_n r^n$ converges, respectively.
If $r = \infty$, then the interval of convergence is $(-\infty, \infty)$.

2.7 Taylor Polynomials and Series

Theorem 8.7.1 (Taylor's Theorem) Suppose that $f, f', \cdots, f^{(n+1)}$ are all continuous for all x such that $|x-a| < R$. Then there exists some c between a and x such that

$$f(x) = P_n(x) + R_n(x)$$

where

$$P_n(x) = \sum_{k=0}^{n} f^{(k)}(a)\frac{(x-a)^k}{k!}, \quad R_n(x) = f^{(n+1)}(\xi)\frac{(x-a)^{n+1}}{(n+1)!}$$

The polynomial $P_n(x)$ is called the *n-th degree Taylor polynomial approximation of f*. The term $R_n(x)$ is called the *Lagrange form of the remainder.*

Theorem 8.7.2 (Binomial Series) If m is a real number and $|x| < 1$, then

$$(1+x)^m = 1 + \sum_{k+1}^{\infty} \frac{m(m-1)\cdots(m-k+1)}{k!}x^k = 1 + mx + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \cdots.$$

This series is called the *binomial series.* If we use the notation

$$\binom{m}{k} = \frac{m(m-1)\cdots(m-k+1)}{k!}$$

then $\binom{m}{k}$ is called the *binomial coefficient* and

$$(1+x)^m = 1 + \sum_{k+1}^{\infty} \binom{m}{k} x^k.$$

If m is a natural number, then we get the binomial expansion

$$(1+x)^m = 1 + \sum_{k+1}^{m} \binom{m}{k} x^k.$$

Theorem 8.7.3 The following power series expansions of functions are valid.

1. $(1-x)^{-1} = 1 + \sum_{k=1}^{\infty} x^k$ and $(1+x)^{-1} = 1 + \sum_{k=1}^{\infty} (-1)^k x^k$, $|x| < 1$.

2. $e^x = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!}$ and $e^{-x} = 1 + \sum_{k=1}^{\infty} (-1)^k \frac{x^k}{k!}$, $|x| < \infty$.

3. $\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$, $|x| < \infty$.

4. $\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$, $|x| < \infty$.

5. $\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k-1}}{(2k+1)!}$, $|x| < \infty$.

6. $\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$, $|x| < \infty$.

7. $\ln(1+x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{(k+1)!}$, $-1 < x \leq 1$.

8. $\displaystyle \frac{1}{2}\ln\frac{1+x}{1-x} = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}, \qquad -1 < x < 1.$

9. $\displaystyle \arctan x = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}, \qquad -1 \leq x \leq 1.$

10. $\displaystyle \arcsin x = \sum_{k=0}^{\infty}\binom{-1/2}{k}(-1)^k\frac{x^{2k+1}}{(2k+1)!}, \qquad -1 \leq x \leq 1.$

3 The Definite Integral

nt

3.1 Area Approximation

Example 5.1.1 Find the area bounded by the graph of the function $y = 4$, $y = 0$, $x = 0$, $x = 3$.

[graph]

From geometry, we know that the area is the height 4 times the width 3 of the rectangle.

$Area = 12.$

Example 5.1.2 Find the area bounded by the graphs of $y = 4x$, $y = 0$, $x = 0$, $x = 3$.

[graph]

From geometry, the area of the triangle is 1/2 times the base, 3, times the height, 12.

$Area = 18.$

Example 5.1.3 Find the area bounded by the graphs of $y = 2x$, $y = 0$, $x = 1$, $x = 4$.

[graph]

The required area is covered by a trapezoid. The area of a trapezoid is 1/2 times the sum of the parallel sides times the distance between the parallel sides.

$Area = (2 + 8)(3)/2 = 15.$

Example 5.1.4 Find the area bounded by the curves $y = \sqrt{4 - x^2}$, $y = 0$, $x = -2$, $x = 2$.

[graph]

By inspection, we recognize that this is the area bounded by the upper half of the circle with center at (0, 0) and radius 2. Its equation is

$x^2 + y^2 = 4$ or $y = \sqrt{4 - x^2}$, $-2 \leq x \leq 2$.

Again from geometry, we know that the area of a circle with radius 2 is $\pi r^2 = 4\pi$. The upper half of the circle will have one half of the total area. Therefore, the required area is 2π.

Example 5.1.5 Approximate the area bounded by $y = x^2$, $y = 0$, $x = 0$, and $x = 3$.

Given that the exact area is 9, compute the error of your approximation.

3.2 The Definite Integral

Let f be a function that is continuous on a bounded and closed interval $[a, b]$. Let $p = \{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$ be a partition of $[a, b]$, not necessarily equally spaced. Let

$m_i = \min\{f(x) : x_{i-1} \leq x \leq x_i\}$, $i = 1, 2, \ldots, n$;

$M_i = \max\{f(x) : x_{i-1} \leq x \leq x_i\}$, $i = 1, 2, \ldots, n$;

$\Delta xi = x_i - x_{i-1}$, $i = 1, 2, \ldots, n$;

$\Delta = \max\{\Delta x_i \, : \, i = 1, 2, \ldots, n\}$;

$L(p) = m_1 \Delta x_1 + m_2 \Delta x_2 + \cdots + m_n \Delta x_n$;

$U(p) = M_1 \Delta x_i + M_2 \Delta x_2 + \cdots + M_n \Delta x_n$.

We call $L(p)$ the *lower Riemann sum.* We call $U(p)$ the *upper Riemann sum.*

Clearly $L(p) \leq U(p)$, for every partition.

Let

$$Lf = \inf\{L(p) : p \text{ is a partition of } [a,b]\}, \quad Uf = \sup\{U(p) : p \text{ is a partition of } [a,b]\}.$$

Definition 5.2.1 If f is continuous on $[a,b]$ and $Lf = Uf = I$, then we say that:

(i) f is integrable on $[a,b]$;

(ii) the definite integral of $f(x)$ from $x = a$ to $x = b$ is I;

(iii) I is expressed, in symbols, by the equation

$$I = \int_a^b f(x)\,\mathrm{d}x\,;$$

(iv) the symbol "$\int$" is called the *"integral sign"*; the number "a" is called the *"lower limit"*; the number "b" is called the *"upper limit"*; the function "$f(x)$" is called the *"integrand"*; and the variable "x" is called the *(dummy) "variable of integration"*.

(v) If $f(x) \geq 0$ for each x in $[a,b]$, then the area, A, bounded by the curves $y = f(x)$, $y = 0$, $x = a$ and $x = b$, is defined to be the definite integral of $f(x)$ from $x = a$ to $x = b$. That is,

$$A = \int_a^b f(x)\,\mathrm{d}x\,.$$

(vi) For convenience, we define

$$\int_a^a f(x)\,\mathrm{d}x = 0,$$

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x\,.$$

Theorem 5.2.1 If a function f is continuous on a closed and bounded interval $[a, b]$, then f is integrable on $[a, b]$.

Theorem 5.2.2 (Linearity) Suppose that f and g are continuous on $[a,b]$ and c_1 and c_2 are two arbitrary constants. Then

(i)

$$\int_a^b (f(x)+g(x))\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$$

(ii)

$$\int_a^b (f(x)-g(x))\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x - \int_a^b g(x)\,\mathrm{d}x$$

(iii)

$$\int_a^b c_1 f(x)\,\mathrm{d}x = c_1 \int_a^b f(x)\,\mathrm{d}x\,, \quad \int_a^b c_2 g(x)\,\mathrm{d}x = c_2 \int_a^b g(x)\,\mathrm{d}x\,,$$

$$\int_a^b (c_1 f(x) + c_2 g(x))\,\mathrm{d}x = c_1 \int_a^b f(x)\,\mathrm{d}x + c_2 \int_a^b g(x)\,\mathrm{d}x$$

Theorem 5.2.3 (Additivity) If f is continuous on $[a, b]$ and $a < c < b$, then

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x\,.$$

Theorem 5.2.4 (Order Property) If f and g are continuous on $[a, b]$ and $f(x) \leq g(x)$ for all $x \in [a, b]$, then

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x\,.$$

Theorem 5.2.5 (Mean Value Theorem for Integrals) If f is continuous on $[a, b]$, then there exists some point $c \in [a, b]$ such that

$$\int_a^b f(x)dx = f(c)(b - a).$$

Definition 5.2.2 The number $f(c)$ given in Theorem 5.2.6 is called the *average value* of f on $[a, b]$, denoted $f_{av}[a, b]$. That is

$$f_{av}[a, b] = \frac{1}{b - a}\int_a^b f(x)\,\mathrm{d}x\,.$$

Theorem 5.2.6 (Fundamental Theorem of Calculus, First Form) Suppose that f is continuous on some closed and bounded interval $[a, b]$ and

$$g(x) := \int_a^x f(t)\,\mathrm{d}t$$

for each $x \in [a, b]$. Then $g(x)$ is continuous on $[a, b]$, differentiable on (a, b) and for all $x \in (a, b)$, $g'(x) = f(x)$. That is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x).$$

Theorem 5.2.7 (Fundamental Theorem of Calculus, Second Form) If f and g are continuous on a closed and bounded interval $[a, b]$ and $g'(x) = f(x)$ on (a, b), then

$$\int_a^b f(x)\,\mathrm{d}x = g(b) - g(a).$$

We use the notation:

$$[g(x)]_a^b = g(b) - g(a).$$

Theorem 5.2.8 (Leibniz Rule) If $\alpha(x)$ and $\beta(x)$ are differentiable for all x and f is continuous for all x, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{\alpha(x)}^{\beta(x)} f(t)\,\mathrm{d}t\right) = f(\beta(x)) \cdot \beta'(x) - f(\alpha(x)) \cdot \alpha'(x).$$

Example 5.2.1 [page 201] Compute each of the following definite integrals and sketch the area represented by each integral:

$$(i)\ \int_0^4 x^2\,\mathrm{d}x \qquad\qquad (ii)\ \int_0^{\pi} \sin x\,\mathrm{d}x$$

$$(iii)\ \int_{-\pi/2}^{\pi/2} \cos x\,\mathrm{d}x \qquad\qquad (iv)\ \int_0^{10} e^x\,\mathrm{d}x$$

$$(v)\ \int_0^{\pi/3} \tan x\,\mathrm{d}x \qquad\qquad (vi)\ \int_{\pi/6}^{\pi/2} \cot x\,\mathrm{d}x$$

$$(vii)\ \int_{-\pi/4}^{\pi/4} \sec x\,\mathrm{d}x \qquad\qquad (viii)\ \int_{\pi/4}^{3\pi/4} \csc x\,\mathrm{d}x$$

$$(xi)\ \int_0^{1} \sinh x\,\mathrm{d}x \qquad\qquad (x)\ \int_0^{1} \cosh x\,\mathrm{d}x$$

We note that each of the functions in the integrand is positive on the respective interval of integration, and hence, represents an area. In order to compute these definite integrals, we use the Fundamental Theorem of Calculus, Theorem 5.2.2. As in Chapter 4, we first determine an anti-derivative $g(x)$ of the integrand $f(x)$ and then use

$$\int_a^b f(x)\,\mathrm{d}x = g(b) - g(a) = [g(x)]_a^b.$$

Example 5.2.2 Evaluate each of the following integrals:

$(i)\ \displaystyle\int_1^{10} \frac{1}{x}\,\mathrm{d}x$ $\qquad$ $(ii)\ \displaystyle\int_0^{\pi/2} \sin(2x)\,\mathrm{d}x$

$(iii)\ \displaystyle\int_0^{\pi/6} \cos(3x)\,\mathrm{d}x$ $\qquad$ $(iv)\ \displaystyle\int_0^{2} (x4 - 3x2 + 2x - 1)\,\mathrm{d}x$

$(v)\ \displaystyle\int_0^{3} \sinh(4x)\,\mathrm{d}x$ $\qquad$ $(vi)\ \displaystyle\int_0^{4} \cosh(2x)\,\mathrm{d}x$

Basic List of Indefinite Integrals:

$$\int x^n \, \mathrm{d}x = \frac{x^{n+1}}{n+1} + c \quad \begin{cases} n \in \mathbb{N}, x \neq 0, \\ \\ n \in \mathbb{R}, n \neq -1, x > 0 \end{cases}$$

$$\int \frac{1}{x} \, \mathrm{d}x = \ln|x| + c, \quad x \neq 0$$

$$\int \sin x \, \mathrm{d}x = -\cos x + c$$

$$\int \cos x \, \mathrm{d}x = \sin x + c$$

$$\int \sinh x \, \mathrm{d}x = \cosh x + c$$

$$\int \cosh x \, \mathrm{d}x = \sinh x + c$$

$$\int \frac{1}{1+x^2} \, \mathrm{d}x = \operatorname{arctg} x + c$$

$$\int \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = \arcsin x + c$$

$$\int e^x \, \mathrm{d}x = e^x + c$$

3.3 Integration by Substitution

Many functions are formed by using compositions. In dealing with a composite function it is useful to change variables of integration. It is convenient to use the following differential notation:

If $u = g(x)$, then $\mathrm{d}u = g'(x)\,\mathrm{d}x$.

The symbol "$\mathrm{d}u$" represents the "differential of u," namely, $g(x)\,\mathrm{d}x$.

Theorem 5.3.1 (Change of Variable) If f, g and g' are continuous on an open interval containing $[a, b]$ and $g'(x) \neq 0$ on $[a, b]$, then

$$\int_a^b f(g(x)) \cdot g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u$$

$$\int f(g(x))g'(x)\,\mathrm{d}x = \int f(u)\,\mathrm{d}u\,,$$

where $u = g(x)$ and $\mathrm{d}u = g'(x)dx$.

Remark 18 We say that we have changed the variable from x to u through the substitution $u = g(x)$.

Example 5.3.1 (i)

$$\int_0^2 \sin(3x)\,\mathrm{d}x = \left| \begin{array}{l} u = 3x \\ \mathrm{d}u = 3\,\mathrm{d}x \\ \mathrm{d}x = \dfrac{1}{3}\,\mathrm{d}u \\ 0 \mapsto 0,\ 2 \mapsto 6 \end{array} \right| = \int_0^6 \frac{1}{3}\sin u\,\mathrm{d}u = \frac{1}{3}[-\cos u]_0^6$$

$$= \frac{1}{3}(-\cos 6 - (-1)) = \frac{1}{3}(1 - \cos 6)\,,$$

(ii)

$$\int_0^2 3x\cos(x^2)\,\mathrm{d}x = \left| \begin{array}{l} u = x^2 \\ \mathrm{d}u = 2x\,\mathrm{d}x \\ 3x\,\mathrm{d}x = \dfrac{3}{2}\,\mathrm{d}u \\ 0 \mapsto 0,\ 2 \mapsto 4 \end{array} \right| = \int_0^4 \cos u \left(\frac{3}{2}\,\mathrm{d}u\right) = \frac{3}{2}[\sin u]_0^4 = \frac{3}{2}\sin 4\,,$$

(iii)

$$\int_0^3 x e^{x^2}\,\mathrm{d}x = \left| \begin{array}{l} u = x^2 \\ \mathrm{d}u = 2x\,\mathrm{d}x \\ x\,\mathrm{d}x = \dfrac{1}{2}\,\mathrm{d}u \\ 0 \mapsto 0,\ 3 \mapsto 9 \end{array} \right| = \int_0^9 e^u \frac{1}{2}\,\mathrm{d}u = \frac{1}{2}[e^u]_0^9 = \frac{1}{2}(e^9 - 1)\,.$$

Definition 5.3.1 Suppose that f and g are continuous on $[a, b]$. Then the area bounded by the curves $y = f(x)$, $y = g(x)$, $x = a$ and $x = b$ is defined to be A, where

$$A = \int_a^b |f(x) - g(x)| \, \mathrm{d}x \, .$$

If $f(x) \geq g(x)$ for all $x \in [a, b]$, then

$$A = \int_a^b f(x) - g(x) \, \mathrm{d}x \, .$$

If $g(x) \geq f(x)$ for all $x \in [a, b]$, then

$$A = \int_a^b g(x) - f(x) \, \mathrm{d}x \, .$$

Example 5.3.2 Find the area, A, bounded by the curves $y = \sin x$, $y = \cos x$, $x = 0$ and $x = \pi$. [graph]

[$A = 2\sqrt{2}$]

Example 5.3.3 Find the area, A, bounded by $y = x^2$, $y = x^3$, $x = 0$ and $x = 2$. [graph]

[A=3/2, note that $x^3 \leq x^2$ on $[0, 1]$ and $x^3 \geq x^2$ on $[1, 2]$]

Example 5.3.4 Find the area bounded by $y = x^3$ and $y = x$. To find the interval over which the area is bounded by these curves, we find the points of intersection. [graph] [$A = 1/2$]

3.4 Integration by Parts

The product rule of differentiation yields an integration technique known as *integration by parts*. Let us begin with the product rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(u(x)v(x)\right) = \left(\frac{\mathrm{d}}{\mathrm{d}x}u(x)\right)v(x) + u(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}v(x)\right)$$

On integrating each term with respect to x from $x = a$ to $x = b$, we get

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(u(x)v(x)\right)\mathrm{d}x \;= \int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}x}u(x)\right)v(x)\,\mathrm{d}x \;+ \int_a^b u(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}v(x)\right)\mathrm{d}x$$

By using the differential notation and the fundamental theorem of calculus, we get

$$[u(x)v(x)]_a^b = \int_a^b (u(x)v(x))'\,\mathrm{d}x \;= \int_a^b u'(x)v(x)\,\mathrm{d}x \;+ \int_a^b u(x)v'(x)\,\mathrm{d}x$$

The standard form of this integration by parts formula is written as

$$\int_a^b u(x)v'(x)\,\mathrm{d}x \;= [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)\,\mathrm{d}x$$

$$\int uv'\,\mathrm{d}x \;= uv - \int u'v\,\mathrm{d}x$$

Theorem 5.4.1 (Integration by Parts) If $u(x)$ and $v(x)$ are two functions that are differentiable on some open interval containing $[a, b]$, then

$$\int_a^b u(x)v'(x)\,\mathrm{d}x = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)\,\mathrm{d}x$$

for definite integrals and

$$\int uv'\,\mathrm{d}x = uv - \int u'v\,\mathrm{d}x$$

for indefinite integrals.

Remark 19 The "two parts" of the integrand are "$u(x)$" and "$v'(x)\,\mathrm{d}x$" or "u" and "$\mathrm{d}v$". It becomes necessary to compute $u'(x)$ and $v(x)$ to make the integration by parts step.

Example 5.4.1 Evaluate the following integrals:

$$(i) \int x\sin x\,\mathrm{d}x \qquad (ii) \int xe^{-x}\,\mathrm{d}x \qquad (iii) \int (\ln x)\,\mathrm{d}x$$

$$(iv) \int \arcsin x\,\mathrm{d}x \qquad (v) \int \arccos x\,\mathrm{d}x \qquad (vi) \int x^2e^x\,\mathrm{d}x$$

3.5 The Riemann Integral

In defining the definite integral, we restricted the definition to continuous functions. However, the definite integral as defined for continuous functions is a special case of the general Riemann Integral defined for bounded functions that are not necessarily continuous.

Definition 5.6.1 Let f be a function that is defined and bounded on a closed and bounded interval $[a, b]$. Let

$$P = \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\}$$

be a partition of $[a, b]$. Let

$$C = \{c_i : x_{i-1} \le c_i \le x_i, i = 1, 2, \ldots, n\}$$

be any arbitrary selection of points of $[a, b]$. Then the *Riemann Sum* that is associated with P and C is denoted $R(P)$ and is defined by

$$R(P) = f(c_1)(x_1 - x_0) + f(c_2)(x_2 - x_1) + \ldots + f(c_n)(x_n + x_{n-1}) = \sum_{i=1}^{n} f(c_i)(x_i - x_{i-1}).$$

Let

$$\Delta x_i = x_i - x_{i-1}, i = 1, 2, \ldots, n$$

$$\|\Delta\| = \max_{1 \le i \le n} \{\Delta x_i\}.$$

We write

$$R(P) = \sum_{i=1}^{n} f(c_i)\Delta x_i.$$

We say that

$$\lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i = I$$

if and only if for each $\varepsilon > 0$ there exists some $\delta > 0$ such that

$$\left| \sum_{i=1}^{n} f(c_i)\Delta x_i - I \right| < \varepsilon$$

whenever $\|\Delta\| < \delta$ for all partitions P and all selections C that define the Riemann Sum.
If the limit I exists as a finite number, we say that f *is (Riemann) integrable* and write

$$I = \int_a^b f(x)\,\mathrm{d}x\,.$$

Theorem 5.6.3 If f is continuous on $[a, b]$, then f is (Riemann) integrable and the definite integral and the Riemann integral have the same value.

3.6 Volumes of Revolution

One simple application of the Riemann integral is to define the volume of a solid.

Theorem 5.7.1 Suppose that a solid is bounded by the planes with equations $x = a$ and $x = b$. Let the cross-sectional area perpendicular to the x-axis at x be given by a continuous function $A(x)$. Then the volume V of the solid is given by

$$V = \int_a^b A(x)\,\mathrm{d}x\,.$$

Theorem 5.7.2 Let f be a function that is continuous on $[a, b]$. Let R denote the region bounded by the curves $x = a$, $x = b$, $y = 0$ and $y = f(x)$. Then the volume V obtained by rotating R about the x-axis is given by

$$V = \int_a^b \pi(f(x))^2\,\mathrm{d}x\,.$$

Theorem 5.7.3 Let f and R be defined as in Theorem 5.7.2. Assume that $f(x) > 0$ for all $x \in [a, b]$, either $a \geq 0$ or $b \leq 0$, so that $[a, b]$ does not contain 0. Then the volume V generated by rotating the region R about the y-axis is given by

$$V = \int_a^b (2\pi x f(x))\,\mathrm{d}x\,.$$

Example 5.7.2 Consider the region R bounded by $y = \sin x$, $y = 0$, $x = 0$ and $x = \pi$. Find the volume generated when R rotated about x-axis

Answer: By Theorem 5.7.2, the volume V is given by

$$V = \int_0^\pi \pi \sin^2 x \, \mathrm{d}x = \pi \cdot \left[\frac{1}{2}(x - \sin x \cos x) \right]_0^\pi = \frac{\pi^2}{2} .$$

Example 5.7.3 Consider the region R bounded by the circle $(x-4)^2 + y^2 = 4$. Compute the volume V generated when R is rotated around

(i) $y = 0$ [$32\pi/3$]

(ii) $x = 0$ [$32\pi^2$]

[graph]

(i) Since the area crosses the x-axis, it is sufficient to rotate the top half to get the required solid.

$$V = \int_2^6 \pi y^2 \, \mathrm{d}x = \pi \int_2^6 [4 - (x-4)^2] \, \mathrm{d}x = \pi \left[4x - \frac{1}{3}(x-4)^3 \right]_2^6 = \pi \left[16 - \frac{8}{3} - \frac{8}{3} \right] = \frac{32}{3}\pi .$$

This is the volume of a sphere of radius 2.

3.7 Arc Length and Surface Area

The Riemann integral is useful in computing the length of arcs. Let f and f' be continuous on $[a, b]$. Let C denote the arc

$$C = \{(x, f(x)) : a \leq x \leq b\}.$$

Let

$$P = \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\}$$

be a partition of $[a, b]$. For each $i = 1, 2, \ldots, n$, let

$$\Delta x_i = x_i - x_{i-1}\,, \qquad \Delta y_i = f(x_i) - f(x_{i-1})$$

$$\Delta s_i = \sqrt{(f(x_i) - f(x_{i-1}))^2 + (x_i - x_{i-1})^2}$$

$$\|\Delta\| = \max_{1 \leq i \leq n} \{\Delta x_n\}.$$

Then Δs_i is the length of the line segment joining the two points $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$. Let

$$A(P) = \sum_{i=1}^{n} \Delta s_i.$$

Then $A(P)$ is called the *polygonal approximation of C with respect to the partition P*.

Definition 5.8.1 Let $C = \{(x, f(x)) : x \in [a,b]\}$ where f and f' are continuous on $[a,b]$. Then the arc length L of the arc C is defined by

$$L = \lim_{\|\Delta\|\to 0} A(P) = \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} \sqrt{(f(x_i) - f(x_{i-1}))^2 + (x_i - x_{i-1})^2}.$$

Theorem 5.8.1 The arc length L defined in Definition 5.8.1 is given by

$$L = \int_a^b \sqrt{(f'(x))^2 + 1}\, \mathrm{d}x\,.$$

Example 5.8.1 Let $C = \{(x, \cosh x) : 0 \le x \le 2\}$. Then the arc length L of C is given by

$$L = \int_0^2 \sqrt{1 + \sinh^2 x}\, \mathrm{d}x = \int_0^2 \cosh x\, \mathrm{d}x = [\sinh x]_0^2 = \sinh 2.$$

Example 5.8.2 Let

$$C = \left\{ \left(x, \frac{2}{3} x^{3/2} \right) \,:\, 0 \le x \le 4 \right\}.$$

Then the arc length L of the curve C is given by

$$L = \int_0^4 \sqrt{1 + \left(\frac{2}{3} \cdot \frac{3}{2} x^{1/2} \right)^2}\, \mathrm{d}x = \int_0^4 (1+x)^{1/2}\, \mathrm{d}x = \left[\frac{2}{3}(1+x)^{3/2} \right]_0^4 = \frac{2}{3} \left[5\sqrt{5} - 1 \right].$$

Definition 5.8.2 Let C be defined as in Definition 5.8.1.

(i) The surface area S_x generated by rotating C about the x-axis is given by

$$S_x = \int_a^b 2\pi |f(x)| \sqrt{(f'(x))^2 + 1}\, \mathrm{d}x\,.$$

(ii) The surface area S_y generated by rotating C about the y-axis

$$S_y = \int_a^b 2\pi |x| \sqrt{(f'(x))^2 + 1}\, \mathrm{d}x\,.$$

Example 5.8.3 Let $C = \{(x, \cosh x) : 0 \leq x \leq 4\}$.

(i) Then the surface area S_x generated by rotating C around the x-axis is given by

$$S_x = \int_0^4 2\pi \cosh x \sqrt{1 + \sinh^2 x}\, \mathrm{d}x \; = \cdots = \pi[4 + \sinh 4 \cosh 4].$$

(ii) The surface area S_y generated by rotating the curve C about the y-axis is given by

$$S_y = \int_0^4 2\pi x \sqrt{1 + \sinh^2 x}\, \mathrm{d}x \; = \cdots = 2\pi[x \sinh x - \cosh x]_0^4 = 2\pi[4 \sinh 4 - \cosh 4 + 1]\,.$$

Theorem 5.8.2 Let $C = \{(x(t), y(t)) : a \le t \le b\}$. Suppose that $x'(t)$ and $y'(t)$ are continuous on $[a, b]$.

(i) The arc length L of C is given by

$$L = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}\,\mathrm{d}t\ .$$

(ii) The surface area S_x generated by rotating C about the x-axis is given by

$$S_x = \int_a^b 2\pi|y(t)|\sqrt{(x'(t))^2 + (y'(t))^2}\,\mathrm{d}t\,.$$

(iii) The surface area S_y generated by rotating C about the y-axis is given by

$$S_y = \int_a^b 2\pi|x(t)|\sqrt{(x'(t))^2 + (y'(t))^2}\,\mathrm{d}t\,.$$

Example 5.8.4 Let $C = \Big\{\big(e^t \sin t, e^t \cos t\big) \,:\, 0 \le t \le \pi/2\Big\}$. Then

$$\begin{aligned}
\mathrm{d}s &= \sqrt{(x'(t))^2 + (y'(t))^2}\,\mathrm{d}t \;=\; \sqrt{(e^t(\sin t + \cos t))^2 + (e^t(\cos t - \sin t))^2}\,\mathrm{d}t \\
&= \{e^{2t}(\sin^2 t + \cos^2 t + 2\sin t\cos t + \cos^2 t + \sin^2 t - 2\cos t\sin t)\}^{1/2}\,\mathrm{d}t \\
&= e^t\sqrt{2}\,\mathrm{d}t\ .
\end{aligned}$$

(i) The arc length L of C is given by

$$L = \int_0^{\pi/2} \sqrt{2}e^t\,\mathrm{d}t \;=\; \sqrt{2}\,\big[e^t\big]_0^{\pi/2} = \sqrt{2}(e^{\pi/2} - 1)\,.$$

(ii) The surface area S_x obtained by rotating C about the x-axis is given by

$$S_x = \int_0^{\pi/2} 2\pi(e^t\cos t)(\sqrt{2}e^t\,\mathrm{d}t\,) = \cdots = \frac{2\sqrt{2}\pi}{5}(e^\pi - 2)\,.$$

(iii) The surface area S_y obtained by rotating C about the y-axis is given by

$$S_y = \int_0^{\pi/2} 2\pi(e^t\sin t)(\sqrt{2}e^t\,\mathrm{d}t\,) = \cdots = \frac{2\sqrt{2}\pi}{5}(2e^\pi + 1)\,.$$

4 Techniques of Integration

nt

4.1 Integration by Substitution

Theorem 6.2.1 Let $f(x)$, $g(x)$, $f(g(x))$ and $g'(x)$ be continuous on an interval $[a, b]$. Suppose that $F'(u) = f(u)$ where $u = g(x)$. Then

(i) $\displaystyle\int f(g(x))g'(x)\,\mathrm{d}x = \int f(u)\,\mathrm{d}u = F(g(x)) + C$

(ii) $\displaystyle\int_a^b f(g(x))g'(x)\,\mathrm{d}x = \int_{u=g(a)}^{u=g(b)} f(u)\,\mathrm{d}u = F(g(b)) - F(g(a)).$

4.2 Integration by Parts

Theorem 6.3.1 Let $f(x)$, $g(x)$, $f'(x)$ and $g'(x)$ be continuous on an interval $[a, b]$. Then

(i) $\displaystyle\int f(x)g'(x)\,\mathrm{d}x = f(x)g(x) - \int g(x)f'(x)\,\mathrm{d}x$

(ii) $\displaystyle\int_a^b f(x)g'(x)\,\mathrm{d}x = (f(b)g(b) - f(a)g(a)) - \int_a^b g(x)f'(x)\,\mathrm{d}x$

(iii) $\displaystyle\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u\,.$

where $u = f(x)$ and $\mathrm{d}v = g'(x)\,\mathrm{d}x$ are the parts of the integrand.

4.3 Integration by Partial Fractions

A polynomial with real coefficients can be factored into a product of powers of linear and quadratic factors. This fact can be used to integrate rational functions of the form $P(x)/Q(x)$ where $P(x)$ and $Q(x)$ are polynomials that have no factors in common. If the degree of $P(x)$ is greater than or equal to the degree of $Q(x)$, then by long division we can express the rational function by

$$\frac{P(x)}{Q(x)} = q(x) + \frac{r(x)}{Q(x)}$$

where $q(x)$ is the quotient and $r(x)$ is the remainder whose degree is less than the degree of $Q(x)$. Then $Q(x)$ is factored as a product of powers of linear and quadratic factors. Finally $r(x)/Q(x)$ is split into a sum of fractions of the form

$$\frac{A_1}{ax+b} + \frac{A_2}{(ax+b)^2} + \ldots + \frac{A_n}{(ax+b)^n}$$

and

$$\frac{B_1x+C_1}{ax^2+bx+c} + \frac{B_2x+C_2}{(ax^2+bx+c)^2} + + \ldots + \frac{B_mx+C_m}{(ax^2+bx+c)^m}$$

Many calculators and computer algebra systems, such as Maple or Mathematica, are able to factor polynomials and split rational functions into partial fractions. Once the partial fraction split up is made, the problem of integrating a rational function is reduced to integration by substitution using linear or trigonometric substitutions. It is best to study some examples and do some simple problems by hand.

4.4 Trigonometric substitutions

To integrate

$$\int R(\sin x, \cos x)\,\mathrm{d}x = \int \frac{P(\sin x, \cos x)}{Q(\sin x, \cos x)}\,\mathrm{d}x$$

1. $R(-u, v) = -R(u, v)$ let $\cos x = t$
2. $R(u, -v) = -R(u, v)$ let $\sin x = t$
3. $R(-u, -v) = R(u, v)$ let $\operatorname{tg} x = t$
4. otherwise let $\operatorname{tg} \dfrac{x}{2} = t$

for $\cos x = t$ we have

$$|\sin x| = \sqrt{1 - t^2}\,, \qquad \mathrm{d}x = \frac{-1}{\sqrt{1 - t^2}}\,\mathrm{d}t$$

for $\sin x = t$ we have

$$|\cos x| = \sqrt{1 - t^2}\,, \qquad \mathrm{d}x = \frac{1}{\sqrt{1 - t^2}}\,\mathrm{d}t$$

for $\operatorname{tg} x = t$ we have

$$\cos x = \frac{1}{\sqrt{1 + t^2}}\,, \qquad \sin x = \frac{t}{\sqrt{1 + t^2}}\,, \qquad \mathrm{d}x = \frac{1}{1 + t^2}\,\mathrm{d}t$$

for $\operatorname{tg}\dfrac{x}{2} = t$ we have

$$\cos x = \cos^2\frac{x}{2} - \sin^2\frac{x}{2} = \frac{1-t^2}{1+t^2}\,, \quad \sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = \frac{2t}{1+t^2}\,, \quad \mathrm{d}x = \frac{2}{1+t^2}\,\mathrm{d}t$$

5 Improper Integrals

5.1 Integrals over Unbounded Intervals

Definition 7.1.1 Suppose that a function f is continuous on $(-\infty, \infty)$. Then we define the following improper integrals when the limits exist

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{b\to\infty}\int_a^b f(x)\,\mathrm{d}x \qquad (1) \quad \boxed{\texttt{7.1}}$$

$$\int_{-\infty}^b f(x)\,\mathrm{d}x = \lim_{a\to-\infty}\int_a^b f(x)\,\mathrm{d}x \qquad (2) \quad \boxed{\texttt{7.2}}$$

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \int_{-\infty}^c f(x)\,\mathrm{d}x + \int_c^\infty f(x)\,\mathrm{d}x \qquad (3) \quad \boxed{\texttt{7.3}}$$

provided the integrals on the right hand side exist for some c. If these improper integrals exist, we say that they are convergent; otherwise they are said to be divergent.

Theorem 7.1.2 Suppose that f and g are continuous on $[a, \infty)$ and $0 \le f(x) \le g(x)$ on $[a, \infty)$.

(i) If $\displaystyle\int_a^\infty g(x)\,\mathrm{d}x$ converges, then $\displaystyle\int_a^\infty f(x)\,\mathrm{d}x$ converges.

(ii) If $\displaystyle\int_a^\infty f(x)\,\mathrm{d}x$ diverges, then $\displaystyle\int_a^\infty g(x)\,\mathrm{d}x$ diverges.

Definition 7.1.3 For each $x > 0$, the *Gamma function*, denoted $\Gamma(x)$, is defined by

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\,\mathrm{d}t$$

Theorem 7.1.3 The Gamma function has the following properties:

$$\Gamma(1) = 1$$

$$\Gamma(x+1) = x\Gamma(x)$$

$$\Gamma(n+1) = n!\,, \quad n = \text{ natural number}$$

Theorem 7.1.4 Let f be the normal probability distribution function defined by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\left(\frac{x-\mu}{\sigma\sqrt{2\pi}}\right)^2}$$

where μ is the constant mean of the distribution and σ is the constant standard deviation of the distribution. Then the improper integral

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$$

Let F be the normal distribution function defined by

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t\,.$$

Then $F(b) - F(a)$ represents the percentage of normally distributed data that lies between a and b. This percentage is given by

$$\int_a^b f(x)\,\mathrm{d}x\,.$$

Furthermore,

$$\int_{\mu+a\sigma}^{\mu+b\sigma} f(x)\,\mathrm{d}x = \frac{1}{\sigma\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,\mathrm{d}x$$

5.2 Discontinuities at End Points

Definition 7.2.1 (i) Suppose that f is continuous on $[a, b)$ and

$$\lim_{x\to b-} f(x) = +\infty \quad \text{or} \quad -\infty\,.$$

Then, we define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\xi\to b-} \int_a^\xi f(x)\,\mathrm{d}x\ .$$

If the limit exists, we say that the *improper integral converges*; otherwise we say that *it diverges.*

(ii) Suppose that f is continuous on $(a, b]$ and

$$\lim_{x\to a+} f(x) = +\infty \quad \text{or} \quad -\infty\,.$$

Then, we define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\xi\to a+} \int_\xi^b f(x)\,\mathrm{d}x\ .$$

If the limit exists, we say that the *improper integral converges*; otherwise we say that *it diverges.*

Exercises 7.2

3. Prove that $\displaystyle\int_0^\infty e^{-x}\,\mathrm{d}x = 1$

4. Prove that $\displaystyle\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \frac{\pi}{2}$

5. Prove that $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}\,\mathrm{d}x = \pi$

6. Prove that $\displaystyle\int_1^\infty \frac{1}{x^p}\,\mathrm{d}x = \frac{1}{1-p}$ if and only if $p > 1$

8. Prove that $\displaystyle\int_0^1 \frac{1}{x^p}\,\mathrm{d}x$ converges if and only if $p < 1$.

$\displaystyle\Big($ cf. Integral criterion for series $\displaystyle\sum_{n=1}^\infty \frac{1}{n^p}\Big)$

6 Differential Calculus of Functions of Several Variables

c2

6.1 Introduction

Definition 6.1.1 Let $M \neq \emptyset$. A mapping $\varrho : M \times M \to [0, \infty)$ is a *metric* (distance) on M, if for all $x, y, z \in M$ we have

1. $\varrho(x, y) = 0 \Leftrightarrow x = y$,

2. $\varrho(x, y) = \varrho(y, x)$ (symmetry),

3. $\varrho(x, z) \leq \varrho(x, y) + \varrho(y, x)$ (triangle inequality).

The set M equiped with a metric ϱ is called a *metric space* (M, ϱ).
Example 6.1.1 The set E^1 is a metric space, $d(x, y) = |x - y|$,
The set E^n is a metric space. A mapping $d : E^n \times E^n \to [0, \infty)$ defined by a formula

$$d(X, Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2},$$

where $X = (x_1, x_2, \ldots, x_n)$, $Y = (y_1, y_2, \ldots, y_n) \in E^n$ is called an *Euclidean metric.*

Another exaples of metrices in E^n:

$$d(X,Y) = \sum_{i=1}^{n} |x_i - y_i| = |x_1 - y_1| + |x_2 - y_2| + \ldots + |x_n - y_n| \quad \text{(so called “Postman (Taxicab) metric”)}$$

$$d(X,Y) = \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\} \quad \text{(so called “maximal metric”)}$$

By an Euclidean metric we have also defined a norm on E^n by

$$\|X\| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \varrho(X, 0)$$

Definition 6.1.2

1. A set
$$O_\delta(A) = \{X \in E^n : \|X - A\| < \delta\}$$
is called an open ball centered at $A \in E^n$ with a radius $\delta > 0$.

2. A neighbourhood of a point $A \in E^n$ is an arbitrary set U such that there exists $O_\delta(A) \subset U$.

3. A point $A \in \Omega \subset E^n$ is called an interior point of the set Ω if there is a neighbourhood $O_\delta(A)$ such that $O_\delta(A) \subset \Omega$.

4. We call a set $\Omega \subset E^n$ an open set if each point $x \in \Omega$ is an interior point of the set Ω.

5. By an inerior of the set $\Omega \subset E^n$ we mean a set of all interior points of the set Ω, we denote it by a $Int\Omega$.

6. We call a set $\Omega \subset E^n$ closed if it is a complement to an open set in E^n.

7. A point A is a boundary point of the set $\Omega \subset E^n$ if for any $O_\delta(A)$ we have
$$O_\delta(A) \cap \Omega \neq \emptyset \quad \text{and} \quad O_\delta(A) \cap (E^n \setminus \Omega) \neq \emptyset.$$

8. A boundary of the set Ω is denoted by $\partial\Omega$.

9. A set $\Omega \subset E^n$ is called bounded if there exists $O_\delta(A)$ such that $\Omega \subset O_\delta(A)$.

10. A set $\Omega \subset E^n$ is called convex if for any couple of the points $X, Y \in \Omega$ we have

$$\lambda X + (1 - \lambda)Y \in \Omega$$

for any $\lambda \in [0, 1]$.

11. A set $\Omega \subset E^n$ is called segment-connected if each couple of points from Ω can be connected by a curve which lies completely in Ω.

12. A set $\Omega \subset E^n$ is called a domain if it is open and segment-connected.

Theorem 6.1.1 Let P be a metric space. Then

(i) $\emptyset$ and P are open sets.

(ii) Let $\{A_\alpha\}_{\alpha \in I}$ be a system of open sets. Then $\bigcup\limits_{\alpha \in I} A_\alpha$ is open set.

Theorem 6.1.2 A set Ω is open if and only if it is empty or it is a union of open balls.
Corollary 6.1.1 Open balls are open sets.

Definition 6.1.3 Let $\Omega \subset E^n$, $\Omega \neq \emptyset$. A mapping $f : \Omega \to E$ is called a function of n variables on Ω.

Definition 6.1.4 Let $\Omega \subset E^n$, $\Omega \neq \emptyset$. A vector-valued function of n variables is a mapping $f : \Omega \to E^m$, i.e. $f = (f_1, f_2, \ldots, f_m)$.

Definition 6.1.5 We say, that a sequence of points $\{X_k\}k = 0^\infty \subset E^n$ has a limit $X \in E^n$ if

$$\lim_{k\to\infty} \|X_k - X\| = 0$$

Remark. $X_k = [x_1^k, x_2^k, \ldots, x_n^k]$, $X = [x_1, x_2, \ldots, x_n]$,

$$\lim_{k\to\infty} \|X_k - X\| = 0 \quad \equiv \quad \lim_{k\to\infty} |x_j^k - x_j| = 0 \qquad \forall j = 1, 2, \ldots, n.$$

We will introduce first a notion of a polynomial function of n variables in E^n:

$$P(x_1, x_2, \ldots, x_n) = \sum_{k_1=0}^{m_1} \sum_{k_2=0}^{m_2} \cdots \sum_{k_n=0}^{m_n} a_{k_1 k_2 \ldots k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n},$$

where $n \in N$, $a_{k_1 k_2 \ldots k_n} \in \mathbb{R}$, k_i, m_i are non-negative integers.

Rational function of n variables is then a fraction of two polynomials

$$R(x_1, x_2, \ldots, x_n) = \frac{P(x_1, x_2, \ldots, x_n)}{Q(x_1, x_2, \ldots, x_n)}$$

6.2 Continuity and Limit

6.2.1 Continuity

Definition 6.2.1 (Heine's definition of continuity) Let $M \subset D(f) \subset E^n$. We call a function f continuous at the point $A \in M$ with respect to M if for any sequence $\{X_k\}_{k=0}^{\infty} \subset M$ we have

$$\lim_{k\to\infty} X_k = A \quad \Rightarrow \quad \lim_{k\to\infty} f(X_k) = f(A)$$

Sometimes we write in short

$$X_k \to A \quad \Rightarrow \quad f(X_k) \to f(A).$$

We say that a function f is continuous on a set M if it is continuous w.r.t. M at any point $x \in M$.

Definition 2. Let $M \subset D(f) \subset E^n$. We say that a vector-valued function $f : M \to E^m$ is continuous on the set M if any its coordinate f_i, $i = 1, 2, \ldots, m$, is continuous on M.

Theorem 6.2.1 Let functions f and g are continuous at the point $A \in \Omega \subset E^n$. Then also functions $f + g$, $f - g$, $f \cdot g$, f/g (under the assumption $g(A) \neq 0$) and $|f|$ are continuous at the point A.

Remark Continuity of $f + g$, $f - g$, $f \cdot g$, f/g and $|f|$ on a set.

Theorem 6.2.2 Let a function g be continuous at the point $A \in M \subset D(g) \subset E^n$ w.r.t. M, $g(M) \subset N \subset E^m$ and a function $f : N \to E$ is continuous at the point $B = g(A)$ w.r.t. N. Then a composition $h = f \circ g$ is continuous at the point A w.r.t. M.

6.2.2 Limit

Let us define $\overline{E} := E \cup \{\pm\infty\}$.

Definition 3. Let $\Omega \subset E^n$. We say that a point $A \in E^n$ is a touching point (an accumulation point) of a set Ω if for any ball $B_\delta(A)$ we have

$$B_\delta(A) \cap (\Omega \setminus \{A\}) = \emptyset.$$

Definition 4. Let $\Omega \subset E^n$, $X_0 \in E^n$ be a touching point of a set Ω and $f : \Omega \setminus \{X_0\} \to E$ be a given function. We say that $L \in \overline{E}$ is a limit of the function f at the point X_0 if for any sequence $\{X_k\}_{k=1}^{\infty} \subset \Omega \setminus \{X_0\}$ we have

$$\lim X_k = X_0 \quad \Rightarrow \quad \lim_{k\to\infty} f(X_k) = L.$$

We write

$$\lim_{X\in\Omega, X\to X_0} f(X) = L.$$

Theorem 6.2.3 Let $\Omega \subset E^n$ and a point $X_0 \in E^n$ be an accumulation point of the set Ω. Let$f, g : \Omega \to E$ be given functions. Let us assume that the limits

$$\lim_{X\in\Omega, X\to X_0} f(X), \qquad \lim_{X\in\Omega, X\to X_0} g(X)$$

exist. Then also the limits

$$\lim_{X\in\Omega, X\to X_0} (f(X) \pm g(X)) = \lim_{X\in\Omega, X\to X_0} f(X) \pm \lim_{X\in\Omega, X\to X_0} g(X),$$

$$\lim_{X\in\Omega, X\to X_0} (f(X)g(X)) = \lim_{X\in\Omega, X\to X_0} f(X) \lim_{X\in\Omega, X\to X_0} g(X),$$

$$\lim_{X\in\Omega, X\to X_0} \frac{f(X)}{g(X)} = \frac{\lim_{X\in\Omega, X\to X_0} f(X)}{\lim_{X\in\Omega, X\to X_0} g(X)}$$

exist if the expressions have sence.

Theorem 6.2.4 Let $\Omega \subset E^n$, let a point $X_0 \in E^n$ be an accumulation point of the set Ω and let $g : \Omega \to E$, $f : g(\Omega) \to E$ be given functions. Let the point $g(X_0) \in E^n$ be an accumulation point of the set $g(\Omega)$. Let us assume that

$$\lim_{Y \in g(\Omega), Y \to Y_0} f(Y) = L$$

$$\lim_{X \in \Omega, X \to X_0} g(X) = Y_0,$$

$$g(X_0) = Y_0 \text{ or } g(X) \neq Y_0 \qquad \forall X \in \Omega, X \neq X_0.$$

Then

$$\lim_{X \in \Omega, X \to X_0} f(g(X)) = L$$

Definition 5. Let $\Omega \subset D(f) \subset E^n$. We say that a vector valued function $f : \Omega \to E^m$, $f = (f_1, f_2, \ldots, f_m)$, has a limit at the point $X_0 \in E^n$, if each its component f_i, $i = 1, 2, \ldots, m$, has a limit $L_i \in E$, $i = 1, 2, \ldots, m$, in the point X_0 in the sense of Definition 4.

6.3 Directional derivative, partial derivative, total differential

6.3.1 Directional derivative

Let us denote

$$\varphi_u(t) := f(A + tu)$$

for $A = [a_1, a_2, \ldots, a_n] \in E^n$, $u = (u_1, u_2, \ldots, u_n) \in R^n$ and $t \in R$.

Definition 1. Let f be defined on a neighbourhood $U_\delta(A)$ of the point $A \in E^n$ and let $u \in R$ be a vector. If the limit

$$\lim_{t \to 0} \frac{f(A + tu) - f(A)}{t} = \lim_{t \to 0} \frac{\varphi_u(t) - \varphi_u(0)}{t}$$

exists and is finite, then we say, that function f has at the point A (directional) derivative in the direction u, i.e. φ_u is differentiable at 0.

Number $\varphi'(0)$ is called a derivative of the function f at the point A in the direction u. Directional derivative is denoted by

$$\frac{\partial f(A)}{\partial u}, \quad D_u f(A), \quad d_u f(A), \quad f_u(A), \quad \partial_u f(A).$$

6.3.2 Partial derivative

Definition 2. Let f be defined on a neighbourhood $U_\delta(A)$ of the point $A = [a_1, a_2, \ldots, a_n] \in E^n$ and let $e_1, e_2, \ldots, e_n \in \mathbb{R}^n$ be vectors such that

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots, 0), \\ &\vdots \\ e_n &= (0, 0, \ldots, 0, 1). \end{aligned}$$

If there is a derivative in the direction e_i $(i = 1, 2, \ldots, n)$, i.e. the limit

$$\lim_{t \to 0} \frac{f(A + te_i) - f(A)}{t}$$

exists and is finite, then we say, that f has at the point A a partial derivative with respect to x_i. Partial derivative is denoted by

$$\frac{\partial f(A)}{\partial x_i}, \quad D_i f(A), \quad d_i f(A), \quad f_{x_i}(A), \quad \partial_{x_i} f(A).$$

If $\partial_{x_i} f(A)$ exists for all $i = 1, 2, \ldots, n$, then we say that f is differentiable at A.

Theorem 6.3.1 Let f and g be given functions of $X = [x_1, x_2, \ldots, x_n] \in E^n$. Then

$$\partial_{x_i}(f(X) \pm g(X)) = \partial_{x_i} f(X) \pm \partial_{x_i} g(X),$$

$$\partial_{x_i}(f(X) \cdot g(X)) = \partial_{x_i} f(X) g(X) + f(X) \partial_{x_i} g(X),$$
$$\partial_{x_i}\left(\frac{f(X)}{g(X)}\right) = \frac{\partial_{x_i} f(X) g(X) - f(X) \partial_{x_i} g(X)}{g^2(X)},$$

$i = 1, 2, \ldots, n$, at any point X at which the right hand side has a sense.

Let f be a scalar function of one variable and g be a function of $X = [x_1, x_2, \ldots, x_n] \in E^n$. Then

$$\partial_{x_i}(f(g(X))) = f'(g(X)) \partial_{x_i} g(X), \quad i = 1, 2, \ldots, n,$$

at any point X at which the right hand side has a sense.

Remark: A set of all functions defined on an open set $\Omega \subset E^n$ for which $\partial_{x_i} f \in C(\Omega)$ for all $i = 1, 2, \ldots, n$ we denote by $C^1(\Omega)$.

Remark: Partial derivative of the second order are denoted by

$$\frac{\partial^2 f(A)}{\partial x_i^2}, \quad \frac{\partial^2 f(A)}{\partial x_i \partial x_j}, \quad D_{ij} f(A), \quad d_{ij} f(A), \quad f_{x_i x_j}(A), \quad \partial^2_{x_i x_j} f(A) = \partial_{x_i} \partial_{x_j} f(A).$$

Remark: A set of all functions defined on an open set $\Omega \subset E^n$ for which $\partial_{x_i} f \in C(\Omega)$ for all$i = 1, 2, \ldots, n$ and $\partial_{x_i} \partial_{x_j} f \in C(\Omega)$ for all $i, j = 1, 2, \ldots, n$ we denote by $C^2(\Omega)$.

Theorem 6.3.2 (Schwartz) Let $\Omega \subset E^2$ be open and $A = [x_0, y_0] \in \Omega$. If both second partial derivatives

$$\partial_x \partial_y f, \quad \partial_y \partial_x f$$

exist in a certain neighbourhood of the point A and are continuous then

$$\partial_x \partial_y f(A) = \partial_y \partial_x f(A).$$

6.3.3 (Total) Differential

Theorem 6.3.3 (Riesz Theorem) Let V_n be a space with a scalar product. For any linear functional $L \in V_n'$ (V_n' is a dual space to V_n, $L : V_n \to \mathbb{R}$, L is linear and continuous) there exists a unique vector $x_L \in V_n$ such that

$$L(h) = (x_L|h) \quad \forall h \in V_n.$$

Definition 3. Let f be a function on $\Omega \subset D(f) \subset E^n$ and let a point $A \in \Omega$ be and interior point of the set Ω. e say that a function f is *differentiable at the point* A if there exists a linear mapping $L : \mathbb{R}^n \to \mathbb{R}$ (it depends on the point A) such that

$$\lim_{h \to 0} \frac{|f(A+h) - f(A) - L(h)|}{\|h\|} = 0.$$

Then we call the linear mapping L as a *(total) differential* (or a *tangent mapping*) of the function f at the point A.

We say, that the function f is differentiable on Ω if it is differentiable at any point $A \in \Omega$.

Remark: (Total) differential is denoted by

$$L = df_A = df(A).$$

6.3.4 Gradient

To the diferential of a function f at the point A there is (via a Riesz Representation Theorem) associated a unique vector $\nabla f(A) = \text{grad } f(A)$ from the space $\mathbb{R}^n$ (which is called a *gradient* of the function f at the point A).

We can write

$$df_A(h) = (\nabla f(A)|h) \quad \forall h \in \mathbb{R}^n.$$

Remark:

$$\ker df_A = (\nabla f(A))^{\perp}.$$

Theorem 6.3.4 Let a function $f : \Omega \to E$, $\Omega \subset E^n$ and let A be an interior point of Ω. If f is diferentiable at the point A then

(i) f is continuous at A,

(ii) f has a directional derivative in any direction $u \in \mathbb{R}^n$ and

$$\partial_u f(A) = \frac{\partial f}{\partial u}(A) = df_A(u) \quad \forall u \in \mathbb{R}^n.$$

Corollary 6.3.1 Let the assumptions of the previous Theorem are fulfilled. Then

(i) $df_A : u \mapsto \partial_u f(A)$ is a linear mapping,

(ii) if total differential df_A exists, then it is unique.

Theorem 6.3.5 Let a function $f : B_\delta(A) \to E$, $B_\delta(A) \subset E^n$. Let us assume that all partial derivatives of the first order exist in any point of $B_\delta(A)$ and are continuous at A. Then f is differentiable at the point A.

Cauchy inequality

$$\left|\frac{\partial f}{\partial h}(A)\right| = |(\nabla f(A)|h)| \leq \|h\|\|\nabla f(A)\|$$

and if $\nabla f(A) \neq o$ then the equality in the above expression holds if and only if $h = c\nabla f(A)$ for some $c \in \mathbb{R}$, $c \neq 0$.

Theorem 6.3.1 Let a function $f : \Omega \to E$, $\Omega \subset E^n$ and A be an interior point of Ω. Let us assume that f is differentiable at the point A and $\nabla f(A) \neq o$. Then the direction of the highest increment of the function f at the point A among all possible directions $h \in \mathbb{R}^n$ with $\|h\| = 1$ is

$$h = \frac{\nabla f(A)}{\|\nabla f(A)\|}.$$

The highest decay of the function f is in the direction

$$-h = -\frac{\nabla f(A)}{\|\nabla f(A)\|}.$$

6.3.5 Jacobi matrix

If $f : R^n \to \mathbb{R}^n$ is differerentiable at the point A, then we define a matrix

$$Df(A) = \left(\frac{\partial f}{\partial x_1}(A), \frac{\partial f}{\partial x_2}(A), \cdots, \frac{\partial f}{\partial x_n}(A) \right)$$

which is called *Jacobi matrix* of the function f at the point A.

Since $h = \sum_{i=1}^{n} h_i e_i$, we have due to linearity of df_A that

$$df_A(h) = \sum_{i=1}^{n} h_i \, df_A(e_i) = \sum_{i=1}^{n} h_i \, \frac{\partial f}{\partial x_i}(A) = Df(A)h^T.$$

Theorem 6.3.6 Let $f, g : E^n \to E$ are two differentiable mappings at the interior point A of the set $\Omega \subset E^n$. Then

$$D(f \pm g)(A) = Df(A) \pm Dg(A),$$

$$D(f \cdot g)(A) = Df(A)g(A) + f(A)Dg(A),$$

$$D\left(\frac{f}{g}\right)(A) = \frac{Df(A)g(A) - f(A)Dg(A)}{g^2(A)}, \quad g(A) \neq 0$$

6.3.6 Tangent plane

A graph of a function $f : \Omega \to E$, $\Omega \subset E^n$ is a subset of E^{n+1} defined by

$$Grf = \{[X, y] \in E^n \times E \, : \, X \in \Omega, \, y = f(X)\}.$$

A tangent plane to the graph of f at the point A is a set

$$\{[X, y] \in E^n \times E \, : \, y = f(A) + (\nabla f(A)|(X - A))\}.$$

For $n = 1$, $A = a \in \mathbb{R}$ we have a tangent line

$$y = f(a) + f'(a)(x - a)$$

For $n = 2$, $A = [a, b] \in \mathbb{R}^2$ we have a tangent plane

$$z = f(a, b) + \partial_x f(a, b)(x - a) + \partial_y f(a, b)(y - b) = f(a, b) + df_A((x - a, y - b))$$

6.3.7 Essential Theorems of Differential Calculus

Theorem 6.3.7 (Mean Value Theorem) Let a function $f : \Omega \to E$, $\Omega \subset E^n$ be an open set and the function f has directional derivative in all directions at any point of the set Ω. Let us assume that $A, X \in \Omega$ are such that a segment $\overline{AX} \subset \Omega$ and let $h = X - A$. Then the function $g(t) := f(A + th)$, $t \in [0, 1]$ is defined and differentiable on $[0, 1]$ and we have

$$g'(t) = \frac{\partial f}{\partial h}(A + th), \quad t \in [0, 1]$$

Moreover,

(i) There exists $\xi \in (0, 1)$ such that

$$f(A + h) - f(A) = g(1) - g(0) = \frac{\partial f}{\partial h}(A + \xi h)$$

(ii) If g is continuous on $[0, 1]$ then

$$f(A + h) - f(A) = \int_0^1 \frac{\partial f}{\partial h}(A + th)\, \mathrm{d}t\,.$$

Theorem 6.3.8 (Weierstrass Theorem, Extreme Value Theorem) A function which is continuous on a compact (hence closed and bounded) set $\Omega \subset E^n$ attains both its maximum and a minimum on Ω.
Remark: There are points $M, N \in \Omega$ such that

$$f(M) = \min_{x\in\Omega} f(x), \quad f(N) = \max_{x\in\Omega} f(x).$$

6.3.8 Derivative of a composition of functions

Let $u = g(x)$ have a derivative at the point x_0 and let $y = f(u)$ have a derivative at the point $u_0 = g(x_0)$. Then the composition $y = F(x) = f(g(x))$ has a derivative at the point x_0 and it holds

$$F'(x_0) = f'(u_0)g'(x_0)\,.$$

Theorem 6.8.9 (Derivative of a coposition of functions) Let $u = u(x, y)$ and $v = v(x, y)$ have partial derivatives of the first order at the point (x_0, y_0). Let us call it $u_0 = u(x_0, y_0)$ and $v_0 = v(x_0, y_0)$. Let a function $z = f(u, v)$ be differentiable at the point (u_0, v_0).
Then the composition $z = F(x, y) = f(u(x, y), v(x, y))$ has partial derivatives at the point (x_0, y_0) and it holds

$$\partial_x F(x_0, y_0) = \partial_u f(u_0, v_0)\partial_x u(x_0, y_0) + \partial_v f(u_0, v_0)\partial_x v(x_0, y_0)\,,$$

$$\partial_y F(x_0, y_0) = \partial_u f(u_0, v_0)\partial_y u(x_0, y_0) + \partial_v f(u_0, v_0)\partial_y v(x_0, y_0)\,.$$

6.4 Taylor Theorem

Taylor formula for a function $F \in C^{k+1}((t_0 - \delta, t_0 + \delta))$, $\delta > 0$:

$$\begin{aligned} F(t) &= T_k(t) + R_k(t) \\ &= F(t_0) + F'(t_0)(t - t_0) + \frac{1}{2!}F''(t_0)(t - t_0)^2 + \ldots + \frac{1}{k!}F^{(k)}(t_0)(t - t_0)^k + \frac{1}{(k+1)!}F^{(k+1)}(\xi)(t - t_0)^{k+1} \end{aligned}$$

for any $t \in (t_0 - \delta, t_0 + \delta)$ and some ξ inbetween t a t_0.

Theorem 6.4.1 (Taylor formula with a Lagrange form of the remainder) Let us have a function $f \in C^2(\Omega)$ or $f \in C^3(\Omega)$, respectively, $\Omega \subset E^n$. Let $U_\delta(A) \subset \Omega$ be a neighbourhood of the point $A = [a_1, a_2, \ldots, a_n] \in \Omega$. Then for any vector $u = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ with $A + u \in U_\delta(A)$ we have

$$f(A + u) = T_1(u) + R_1(u) = f(A) + (\nabla f(A)|u) + \frac{1}{2!}(H_f(A + \xi u)u^T|u)$$

for some $\xi \in (0, 1)$ or

$$f(A + u) = T_2(u) + R_2(u) = f(A) + (\nabla f(A)|u) + \frac{1}{2!}(H_f(A)u^T|u) + R_2(u),$$

where $R_1(u)$ or $R_2(u)$ are the reminders with

$$\lim_{\|u\| \to 0} \frac{R_1(u)}{\|u\|} = 0 \quad \text{and} \quad \lim_{\|u\| \to 0} \frac{R_1(u)}{\|u\|^2} = 0\,.$$

6.5 Extremal Points

6.5.1 Introduction

Definition 6.5.1

By a quadratic form on the space E^n we call a function

$$K(x) = K(x_1, \dots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j.$$

We call a quadratic form K

- *positive definite* if $K(x) > 0 \quad \forall x \in E^n$, $x \neq o$,
- *positively semidefinite* if $K(x) \geq 0 \quad \forall x \in E^n$,
- *negative definite* if $K(x) < 0 \quad \forall x \in E^n$, $x \neq o$,
- *negative semidefinite* if $K(x) \leq 0 \quad \forall x \in E^n$,
- *indefinite* if there exist $x, y \in E^n$ such that $K(x) > 0$ and $K(y) < 0$.

Remark: Let us denote

$$D_k = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix}$$

for $k = 1, \dots, n$.

Theorem 6.5.1 (Sylvester criterion) Quadratic form K is

(a) positive definite if and only if $\det D_k > 0$ for all $k = 1, \dots, n$,

(b) negative definite if and only if $(-1)^k \det D_k > 0$ for all $k = 1, \dots, n$,

(c) if $\det D_n \neq 0$ and the form K is not definite then K is indefinite.

6.5.2 Local extrema

Let $\Omega \subset E^n$ and $f : \Omega \to E$ be a given function. A point $A \in \Omega$ such that

$$f(A) \leq f(X) \quad \forall X \in \Omega$$

is called a point of minima (an absolute minimum) of the function f on the set Ω.

Definition 6.5.2 We say that a function $f(X)$ has at the point $A \in \Omega \subset D(f)$ a *local minimum* (*local maximum*) w.r.t. Ω, if there exists a neighbourhood $U_\delta(A)$ such that

$$f(A) \leq f(X) \quad (f(A) \geq f(X))$$

for all $X \in U_\delta(A) \cap \Omega$.

The points local minima (local maxima) are called *extremal points* and values of the function f evaluated at these extremal points are called *extremal values*.

Theorem 6.5.1 If for at least one index $1 \leq j \leq n$ the partial derivative $\partial_{x_j} f(A)$ exists and is non-zero then the function f has no local extrema at the point A.

Theorem 6.5.2 A function f can have sharp local extremal points at most in a countable set.

Theorem 6.5.2 Let a function $f : \Omega \to E$ be diferentiable at the interior point $A \in \Omega$. If A is an extremal point of the function f then

$$df A = 0 \quad (\text{or equivalently } \nabla f(A) = o)$$

Definition 6.5.3 Let Ω be open sobset of E^n and let $f : \Omega \to E$ be diferentiable in Ω. The points $A \in \Omega$ such that

$$df\,A = 0 \quad \text{(or equivalently } \nabla f(A) = o)$$

are called *critical points* (*stationary points*) of the function f in Ω.

Theorem 6.5.3 Let us have a function $f \in C^2(\Omega)$, $\Omega \subset E^n$ be open in E^n and let $A \in \Omega$ be a critical point of the function f. Then

(i) if A is a local minimum then

$$(H_f(A)u^T|u) \geq 0 \quad \forall u \in \mathbb{R}^n,$$

(ii) if

$$(H_f(A)u^T|u) > 0, \quad \forall u \in \mathbb{R}^n, u \neq o$$

then A is a isolated local minimum,

(iii) if $(H_f(A)u^T|u)$ is indefinte then there is no extremal value at the point A

6.5.3 Global extremal points

Usually we look for the *largest (smallest) value* of a given function on the *whole* set Ω. If the set Ω is compact then it follows from Weierstrass Theorem that there is $\min f(\Omega)$ and $\max f(\Omega)$.
Global (absolute) extremal points on the set Ω we obtain by inspection of the points

(1) stationary points of $Int\,\Omega$,

(2) points in $Int\,\Omega$ where some of the (partial) derivatives does not exist,

(3) boundary points of the set Ω (of any dimension less than n).