Advanced Calculus Notes

Michael P. Windham

February 25, 2008

Contents

Introduction

The Notes

These notes are merely an outline of a course I taught several times in several places as "Advanced Calculus". The proofs of theorems and examples were given in class. I present this outline to you for your edification.

References

These books were recommended but not required for the course.

Advanced Calculus, Friedman, (Holt, Rhinehart and Winston)
Advanced Calculus, Fulks, (3rd edition, Wiley)
Advanced Calculus, James, (Wadsworth)
Advanced Calculus, Kaplan, (3rd edition, Addison-Wesley)
Advanced Calculus, Loomis and Sternberg, (Addison-Wesley)
Advanced Calculus, Taylor and Mann, (3rd edition, Wiley)
Advanced Calculus of Several Variables, Edwards, (Academic Press)
Calculus of Vector Functions, Williamson, Crowell, and Trotter, (Prentice Hall)
Calculus on Manifolds, Spivak, (Benjamin)
The Elements of Real Analysis, Bartle, (2nd edition, Wiley)
Mathematical Analysis, Apostol, (2nd edition, Addison-Wesley)
Principles of Mathematical Analysis, Rudin, (3rd edition, McGraw Hill)

Chapter 1

Basics

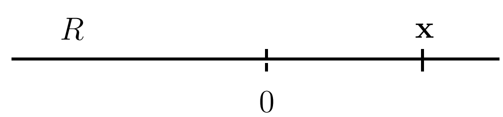
This chapter defines the basic concepts for R^n and functions $R^n \to R^m$.

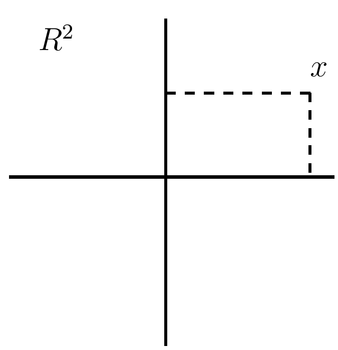
1.1 Euclidean Space R^n

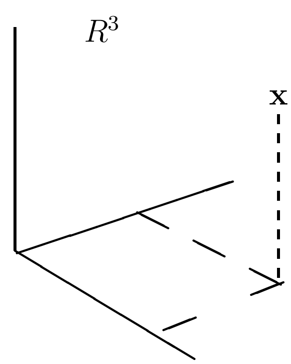
The elements of R^n will be called "vectors" or "points" and will be thought of as ordered n-tuples, $n \times 1$ matrices or geometrical points.

$$\mathbf{x} \in R^n \quad \leftrightarrow \quad (x_1, \ldots, x_n) \quad \leftrightarrow \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Geometry:







R^n is an *inner product space*, that is, it has addition and scalar multiplication operations defined so that it is a vector space, an inner (dot) product and hence a norm. These operations are defined as follows: For

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Addition: $$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

Scalar multiplication: For $r \in R$, $r\mathbf{x} = \begin{bmatrix} rx_1 \\ \vdots \\ rx_n \end{bmatrix}$

Inner product: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$

Norm: $|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_i x_i^2}$

Addition and scalar multiplication allow subtraction to be defined by $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-1)\mathbf{y}$

The basic properties of these operations are as follows.

Addition: for $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z} \in R^n$,

A-1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$

A-2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = x + (\mathbf{y} + \mathbf{z})$

A-3. There is $\mathbf{0} \in R^n$, so that $\mathbf{x} + \mathbf{0} = \mathbf{x}$, for all $\mathbf{x} \in R^n$.

A-4. For each $\mathbf{x} \in R^n$, there is $-\mathbf{x} \in R^n$, so that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

Scalar multiplication: For r and $s \in R$ and $\mathbf{x}$ and $\mathbf{y} \in R^n$,

SM-1. $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$

SM-2. $(r + s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$

SM-3. $(rs)\mathbf{x} = r(s\mathbf{x}) = s(r\mathbf{x})$

SM-4. $1\mathbf{x} = \mathbf{x}$

Inner product: For $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z} \in R^n$ and $r \in R$,

IP-1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

IP-2. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$

IP-3. $\langle r\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, r\mathbf{y}\rangle = r\langle \mathbf{x}, \mathbf{y}\rangle$

IP-4. $\langle \mathbf{x}+\mathbf{y}, \mathbf{z}\rangle = \langle \mathbf{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle$

Norm: For $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z} \in R^n$ and $r \in R$,

N-1. $|\mathbf{x}| \geq 0$, and $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

N-2. $|r\mathbf{x}| = |r||\mathbf{x}|$

N-3. $|\mathbf{x}+\mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$

These properties are by no means the only important properties of the operations on R^n, but they are basic in that the are used to define general structures that act in some sense like the specific concepts in R^n. In particular, any set with operations defined on it that satisfy properties A and SM is called a *vector space*, a vector space with a real valued function that satisfies properties N is called a *normed linear space*, and a vector space with a function defined on it satisfying properties IP is called an *inner product space*. Any inner product space "induces" a norm defined by $|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}$. In a normed linear space the norm induces a *metric*, i.e. a measure of distance between two points given by $|\mathbf{x} - \mathbf{y}|$.

Other properties of note that follow from these basic properties are

1. *Cauchy-Schwartz inequality:* $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq |\mathbf{x}||\mathbf{y}|$

2. $0\mathbf{x} = r\mathbf{0} = \mathbf{0}$

3. $(-1)\mathbf{x} = -\mathbf{x}$

4. $||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}|$

5. And many more.

The *angle between* $\mathbf{x}$ *and* $\mathbf{y}$ is defined by

$$\cos(\theta) = \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{|\mathbf{x}||\mathbf{y}|}$$

1.2 Functions

The symbol $f : R^n \to R^m$ stands for the statement "f is a function from R^n to R^m." For such a function, $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$ and the functions $f_i : R^n \to R$ are called the *component functions* of f

1.2.1 Sets and functions

Associated to a function are various sets that are used to describe or understand the structure. These concepts are also used to associate sets to functions so that the tools of calculus can be used to study the structure of the sets.

- Domain: Dom $f = \{\mathbf{x} \in R^n : f(\mathbf{x}) \text{ exists}\}$
- Range or Image: Im $f = \{\mathbf{y} \in R^m : \mathbf{y} = f(\mathbf{x})$ for some $\mathbf{x} \in$ Dom $f\}$.
- Graph: Graph $f = \{(\mathbf{x}, \mathbf{y}) \in R^{n+m} : \mathbf{x} \in$ Dom f and $\mathbf{y} = f(\mathbf{x})\}$.
- Level sets: For $\mathbf{b} \in R^m$, the *level set* of f at $\mathbf{b} = \{\mathbf{x} \in R^n : f(\mathbf{x}) = \mathbf{b}\}$.

1.2.2 Basic functions

The functions that are used most are built from a few families of *basic* functions most of which are encountered by the end of the first year of calculus.

$R \to R$ Families:

1. Constant functions: For $c \in R$, $f(x) = c$.
2. Power functions: For $r \in R$, $f(x) = x^r$.
3. Trigonometric functions: $\sin, \cos, \tan, \cot, \sec, \csc$, and
   $\arcsin, \arccos, \arctan$, arccot, arcsec, arccsc.
4. Logarithms and exponentials: For $a > 0$, $\log_a$ and $\exp_a$ (also known as a^x).

These functions are combined using the following building processes to give the functions we know and love.

Building processes:

1. Projection: $P_i(\mathbf{x}) = x_i$.
2. Algebraic: For $f, g : R^n \to R^m$
   (a) Addition: $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$.
   (b) Subtraction: $(f - g)(\mathbf{x}) = f(\mathbf{x}) - g(\mathbf{x})$.
   (c) Multiplication:
      i. Inner product: $\langle f, g\rangle(\mathbf{x}) = \langle f(\mathbf{x}), g(\mathbf{x})\rangle$. (If $m = 1$ then this is just multiplication of real-valued functions.)

ii. Scalar product: If $r : R^n \to R$ and $f : R^n \to R^m$ then $(rf)(\mathbf{x}) = r(\mathbf{x})f(\mathbf{x})$. (If $m = 1$ then this is just multiplication of real-valued functions.)

(d) Division: If $m = 1$, $\frac{f}{g}(\mathbf{x}) = \frac{f(\mathbf{x})}{g(\mathbf{x})}$.

3. Composition: For $f : R^n \to R^m$ and $g : R^m \to R^p$, $g \circ f : R^n \to R^p$ is defined by $g \circ f(\mathbf{x}) = g(f(\mathbf{x}))$.

4. Patching: This idea is best illustrated by an example. Define $f : R \to R$ by

$$f(x) = \begin{cases} x^2 & \text{for } x < 0 \\ x + 1 & \text{for } 0 \le x \le 2 \\ e^{-\frac{1}{2}x^2} & \text{for } x > 3 \end{cases}$$

1.3 Exercises

1–1. Prove the following properties of the inner product space R^n.

(a) For each $\mathbf{x} \in R^n$, there is $-\mathbf{x} \in R^n$, so that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

(b) $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$

(c) $|r\mathbf{x}| = |r||\mathbf{x}|$

(d) $0\mathbf{x} = r\mathbf{0} = \mathbf{0}$

(e) $(-1)\mathbf{x} = -\mathbf{x}$

(f) If $\mathbf{x} \ne \mathbf{0}$, then $|\langle \mathbf{x}, \mathbf{y} \rangle| = |\mathbf{x}||\mathbf{y}|$ if and only if $\mathbf{y} = r\mathbf{x}$ for some $r \in R$.

1–2. The vector space R^n can be made into a normed linear space in many ways. In particular, a norm on R^n is any function $\| \bullet \| : R^n \to R$ satisfying

$$\begin{aligned} \|\mathbf{x}\| &\ge 0, \text{ and } \|\mathbf{x}\| = 0 \text{ if and only if } \mathbf{x} = \mathbf{0} \\ \|r\mathbf{x}\| &= |r| \, \|\mathbf{x}\| \\ \|\mathbf{x} + \mathbf{y}\| &\le \|\mathbf{x}\| + \|\mathbf{y}\| \end{aligned}$$

The most common norms defined on R^n are the ℓ_p norms for $1 \le p \le \infty$ defined by

$$\|\mathbf{x}\|_p = \begin{cases} \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} & \text{for } 1 \le p < \infty \\ \max_i |x_i| & \text{for } p = \infty \end{cases}$$

(a) Show that these function are, in fact, norms. (The triangle inequality is quite difficult for $1 < p < \infty$, so prove it for $p = 1$ and $p = \infty$.)

(b) Sketch the unit spheres in R^2 for these norms, i.e. $\{\mathbf{x} : \|\mathbf{x}\|_p = 1\}$, for representative values of p, including $p = 1, 1.5, 2, 4,$ and ∞.

1–3. An inner product on R^n is any function $\langle\!\langle \bullet, \bullet \rangle\!\rangle : R^n \times R^n \to R$, satisfying

$$\begin{aligned} \langle\!\langle \mathbf{x}, \mathbf{x} \rangle\!\rangle &\geq 0, \text{ and } \langle\!\langle \mathbf{x}, \mathbf{x} \rangle\!\rangle = 0 \text{ if and only if } \mathbf{x} = \mathbf{0} \\ \langle\!\langle \mathbf{x}, \mathbf{y} \rangle\!\rangle &= \langle\!\langle \mathbf{y}, \mathbf{x} \rangle\!\rangle \\ r\langle\!\langle \mathbf{x}, \mathbf{y} \rangle\!\rangle &= \langle\!\langle r\mathbf{x}, \mathbf{y} \rangle\!\rangle = \langle\!\langle \mathbf{x}, r\mathbf{y} \rangle\!\rangle \\ \langle\!\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle\!\rangle &= \langle\!\langle \mathbf{x}, \mathbf{z} \rangle\!\rangle + \langle\!\langle \mathbf{y}, \mathbf{z} \rangle\!\rangle \end{aligned}$$

An $n \times n$ symmetric matrix, A, is *positive definite*, if $\mathbf{x}^T A\mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$. Show that, for fixed positive definite symmetric matrix A, $\langle\!\langle \bullet, \bullet \rangle\!\rangle$ defined by $\langle\!\langle \mathbf{x}, \mathbf{y} \rangle\!\rangle = \mathbf{x}^T A\mathbf{y}$ is an inner product.

1–4. Sketch the domain, image, graph and level sets of the function $f(x, y) = \dfrac{1}{xy}$.

1–5. Sketch the following.

(a) The image and graph of $f(t) = (\cos t, \sin t)$.

(b) The image of $f(r, \theta) = (r\cos\theta, r\sin\theta, r)$, for $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$.

(c) The level set of $f(x, y, z) = (x + y + z, x^2 + y^2 + z^2)$ at $(1, 1)$.

(d) The image of $g(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$, for $0 \leq \phi \leq \pi$ and $0 \leq \theta \leq 2\pi$.

1–6. For a function $f : R^n \to R^m$, find

(a) A function whose image is the graph of f.

(b) A function whose level set at $\mathbf{0}$ is the graph of f.

1–7. Recall that a *linear* function $L : R^n \to R^m$ is characterized by matrix multiplication, in that L is linear if and only if there is an $m \times n$ matrix A so that $L(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$. Discuss both the analytical and geometric nature of the domain, image, graph and level sets of a linear function.

1–8. Find the following.

(a) A function $g : R^2 \to R^2$ so that $g(f(x, y)) = (x, y)$, where the function f is defined by $f(x, y) = (3x + 2y, x - y)$.

(b) For the function $F : R^4 \to R^2$ defined by $F(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3 + x_4, 2x_1 - 3x_2 + x_3 + 4x_4)$, let S be the level set of F at $(1, -1)$. Find

  i. A function whose image is S.

  ii. A function whose graph is S.

Chapter 2

Limits and Continuity

2.1 Limits

Definition 2.1.1 *Limit point of a set*
If S is a set in R^n and $\mathbf{x}_0 \in R^n$, then $\mathbf{x}_0$ is a *limit point* of S means for every $\varepsilon > 0$ there is a point $\mathbf{x} \in S$, so that $0 < |\mathbf{x} - \mathbf{x}_0| < \varepsilon$.

Definition 2.1.2 *Limit of $f(\mathbf{x})$ as $\mathbf{x}$ approaches $\mathbf{a}$*
For $f : R^n \to R^m$, and $\mathbf{a}$ a limit point of Dom f, and $\mathbf{b}$ in R^m,

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{b}$$

means for every $\varepsilon > 0$, there is a $\delta > 0$, so that

if $0 < |\mathbf{x} - \mathbf{a}| < \delta$ and $f(\mathbf{x})$ exists, then $|f(\mathbf{x}) - \mathbf{b}| < \varepsilon$

Theorem 2.1.1 $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{b}$, *if and only if* $\lim_{\mathbf{x}\to\mathbf{a}} f_i(\mathbf{x}) = b_i$, *for all* i.

Theorem 2.1.2 ($R \to R$ Families) *If $f(a)$ exists, then* $\lim_{x\to a} f(x) = f(a)$.

Theorem 2.1.3 (Limits and building processes)

1. $\lim_{\mathbf{x}\to\mathbf{a}} P_i(\mathbf{x}) = a_i$
2. Algebraic: *For $f : R^n \to R^m$ and $g : R^n \to R^m$, if* $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{b}$ *and* $\lim_{\mathbf{x}\to\mathbf{a}} g(\mathbf{x}) = \mathbf{c}$, *then*

(a) $\lim_{\mathbf{x}\to\mathbf{a}}(f+g)(\mathbf{x}) = \mathbf{b}+\mathbf{c}$

(b) $\lim_{\mathbf{x}\to\mathbf{a}}(f-g)(\mathbf{x}) = \mathbf{b}-\mathbf{c}$

(c) $\lim_{\mathbf{x}\to\mathbf{a}}|f(\mathbf{x})| = |\mathbf{b}|$

(d) $\lim_{\mathbf{x}\to\mathbf{a}}\langle f,g\rangle(\mathbf{x}) = \langle \mathbf{b},\mathbf{c}\rangle$

(e) *If* $r : R^n \to R$ *and* $\lim_{\mathbf{x}\to\mathbf{a}} r(\mathbf{x}) = s$, *then* $\lim_{\mathbf{x}\to\mathbf{a}} rf(\mathbf{x}) = s\mathbf{b}$.

(f) *If* $m = 1$, *then*

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f}{g}(\mathbf{x}) = \begin{cases} \dfrac{b}{c} & \text{if } c \neq 0 \\ \text{Does not exist} & \text{if } c = 0, \text{ but } b \neq 0 \\ \mathbf{?} & \text{if } c = b = 0 \end{cases}$$

3. *For* $f : R^n \to R^m$, $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{0}$ *if and only if* $\lim_{\mathbf{x}\to\mathbf{a}} |f(\mathbf{x})| = 0$

4. Composition: *For* $f : R^n \to R^m$ *and* $g : R^m \to R^p$, *if* $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{b}$ *and* $\lim_{\mathbf{y}\to\mathbf{b}} g(\mathbf{y}) = g(\mathbf{b})$, *then* $\lim_{\mathbf{x}\to\mathbf{a}} g \circ f(\mathbf{x}) = g(\mathbf{b})$.

5. *For* $f : R^n \to R$, $g : R^n \to R$, *and* $h : R^n \to R$:

   (a) Order: *If* $f(\mathbf{x}) \leq g(\mathbf{x})$ *for* $\mathbf{x}$ *in an open ball centered at* $\mathbf{a}$, *then* $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) \leq \lim_{\mathbf{x}\to\mathbf{a}} g(\mathbf{x})$.

   (b) *Squeeze: If* $f(\mathbf{x}) \leq h(\mathbf{x}) \leq g(\mathbf{x})$ *for* $\mathbf{x}$ *in an open ball centered at* $\mathbf{a}$, *and* $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{a}} g(\mathbf{x}) = b$, *then* $\lim_{\mathbf{x}\to\mathbf{a}} h(\mathbf{x}) = b$.

2.2 Continuity

Definition 2.2.1 *Continuous at a point*
For $f : R^n \to R^m$, $\mathbf{x}_0$ a limit point of Dom f, f is *continuous at* $\mathbf{a}$ means

1. $f(\mathbf{x}_0)$ exists i.e. $\mathbf{x}_0 \in$ Dom f and
2. $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$

Definition 2.2.2 *Continuous on a set*
For $f : R^n \to R^m$, $S \subset$ Dom f, f is *continuous on* S means f is continuous at each point in S.

Theorem 2.2.1 *$f : R^n \to R^m$ is continuous at $\mathbf{x}_0$, if and only if all of its component functions are continuous at $\mathbf{x}_0$.*

Theorem 2.2.2 *If f and $g : R^n \to R^m$ and $r : R^n \to R$ are continuous at $\mathbf{x}_0$, then $f + g, f - g, rf$, and $\langle f, g \rangle$ are continuous at $\mathbf{x}_0$. If $g(\mathbf{x}_0) \neq 0$, then $\frac{f}{g}$ is continuous at $\mathbf{x}_0$.*

Theorem 2.2.3 *The operations of addition, scalar multiplication and inner product are continuous everywhere. The norm is continuous everywhere.*

Corollary 2.2.4 *If the component functions of $f : R^n \to R^m$ are built from $R \to R$ families using a finite number of projections, algebraic operations and compositions, then f is continuous where it is defined.*

2.3 Exercises

2–1. Prove the following.

(a) $\lim_{x\to 1} 3x + 2 = 5$

(b) $\lim_{x\to a} \sqrt{x} = \sqrt{a}$

(c) $f(x) = \frac{1}{x}$ is continuous for all $x \neq 0$.

(d) If sin and cos are continuous at 0, then they are continuous everywhere.

Definition 2.3.1 *Sequence*
A *sequence* of vectors is a list of vectors in R^n indexed by the positive (or nonnegative) integers, e.g. $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_k, \ldots$, which may be abbreviated $\{\mathbf{x}_k\}_{k=1}^{\infty}$.

Definition 2.3.2 *Convergent sequence*
A sequence, $\{\mathbf{x}_k\}_{k=1}^{\infty}$, *converges to* $\mathbf{x} \in R^n$ is denoted

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}$$

and means for every $\varepsilon > 0$, there is an integer K, so that

if $k \geq K$, then $|\mathbf{x}_k - \mathbf{x}| < \varepsilon$

2–2. Prove the following about sequences.

(a) If $\{x_k\}_{k=1}^{\infty} \subset R$ and $\lim_{k\to\infty} x_k = x$, then $\lim_{k\to\infty} x_k^2 = x^2$.

(b) If $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}$ and $\lim_{k\to\infty} \mathbf{y}_k = \mathbf{y}$ then

i. $\lim_{k\to\infty} \mathbf{x}_k + \mathbf{y}_k = \mathbf{x} + \mathbf{y}$

ii. $\lim_{k\to\infty} \langle \mathbf{x}_k, \mathbf{y}_k \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$

(c) Show that $f : R^n \to R^m$ is continuous at $\mathbf{x}$, if and only if for every sequence $\{\mathbf{x}_k\}_{k=1}^{\infty}$ contained in Dom f and converging to $\mathbf{x}$, the sequence $\{f(\mathbf{x}_k)\}_{k=1}^{\infty}$ converges to $f(\mathbf{x})$.

2–3. List any other facts about convergent sequences, similar to the above that you believe to be true.

2–4. Discuss the existence of $\lim_{(x,y)\to(0,0)} f(x, y)$ for the following functions.

(a) $f(x, y) = \dfrac{\sqrt{xy}}{x^2 + y^2}$

(b) $f(x, y) = \dfrac{xy}{x^2 + y^2}$

(c) $f(x, y) = \dfrac{xy}{\sqrt{x^2 + y^2}}$

Chapter 3

Differential Calculus: Basics

3.1 Definitions and Basic Theorems

3.1.1 Affine function

If A is an $m \times n$ matrix and $\mathbf{b}$ is a vector in R^m, then the function $T : R^n \to R^m$ defined by

$$T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

for all $\mathbf{x} \in R^n$ is called an *affine function*.

The sets associated to an affine function are the following. Let $r = \text{rank } A$,

- Domain: R^n
- Image: An r-dimensional plane in R^m thru $\mathbf{b}$, parallel to the column space of A.
- Graph: An n-dimensional plane in R^{n+m} thru $(\mathbf{0}, \mathbf{b})$, parallel to the column space of $\begin{bmatrix} I \\ A \end{bmatrix}$.
- Level set at $\mathbf{c}$: If $\mathbf{c} \in \text{Im } T$, then there is an $\mathbf{x}_0$ so that $T(\mathbf{x}_0) = \mathbf{c}$. The level set at $\mathbf{c}$ is an $(n - r)$-dimensional plane thru $\mathbf{x}_0$, parallel to the kernel (null space) of A. Otherwise, the level set is empty.

3.1.2 Differentiable

Definition 3.1.1 *Differentiable*
A function $f : R^n \to R^m$ is *differentiable* at $\mathbf{x}_0 \in \text{Dom } f$ means there is an affine function of the form $T(\mathbf{x}) = A(\mathbf{x} - \mathbf{x}_0) + f(\mathbf{x}_0)$, so that

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - T(\mathbf{x})}{|\mathbf{x} - \mathbf{x}_0|} = \mathbf{0}$$

The function T is called the *best affine approximation* to f near $\mathbf{x}_0$. The matrix A in T is called the *derivative* of f at $\mathbf{x}_0$, and is denoted $f'(\mathbf{x}_0)$. The function $f' : R^n \to M_{m,n}$ (the set of $m \times n$ matrices) with value at $\mathbf{x}$ given by $f'(\mathbf{x})$ is called the *derivative*.

The geometrical significance of T is described by the following.

- The graph of T is the "tangent" plane to the graph of f at the point $(\mathbf{x}_0, f(\mathbf{x}_0))$.
- (Rank $A = n$) The image of T is the tangent plane to the image of f at $f(\mathbf{x}_0)$.
- (Rank $A = m$) The level set of T at $f(\mathbf{x}_0)$ is the tangent plane to the level set of f at $f(\mathbf{x}_0)$, at $\mathbf{x}_0$

Theorem 3.1.1 *$f : R^n \to R^m$ is differentiable at $\mathbf{x}_0$ if and only if each component function is differentiable at $\mathbf{x}_0$. Moreover,* $f'(\mathbf{x}_0) = \begin{bmatrix} f_1'(\mathbf{x}_0) \\ \vdots \\ f_m'(\mathbf{x}_0) \end{bmatrix}$.

$R \to R$ **Families:** We will assume for the time being that the facts about differentiability and the derivative for real-valued functions of a real variable are "well-known".

Theorem 3.1.2 (Building processes)

1. Algebraic: *If $f, g : R^n \to R^m$ and $r : R^n \to R$ are differentiable at $\mathbf{x}_0$, then*
   (a) *$f + g$ is differentiable at $\mathbf{x}_0$, and $(f+g)'(\mathbf{x}_0) = f'(\mathbf{x}_0) + g'(\mathbf{x}_0)$.*
   (b) *$f - g$ is differentiable at $\mathbf{x}_0$, and $(f-g)'(\mathbf{x}_0) = f'(\mathbf{x}_0) - g'(\mathbf{x}_0)$.*
   (c) *rf is differentiable at $\mathbf{x}_0$, and $(rf)'(\mathbf{x}_0) = r(\mathbf{x}_0)f'(\mathbf{x}_0) + f(\mathbf{x}_0)r'(\mathbf{x}_0)$.*
   (d) *$\langle f, g\rangle$ is differentiable at $\mathbf{x}_0$, and $\langle f, g\rangle'(\mathbf{x}_0) = f^T(\mathbf{x}_0)g'(\mathbf{x}_0) + g^T(\mathbf{x}_0)f'(\mathbf{x}_0)$.*
   (e) *If $m = 1$ and $g(\mathbf{x}_0) \neq 0$, then $\frac{f}{g}$ is differentiable at $\mathbf{x}_0$, and $(\frac{f}{g})'(\mathbf{x}_0) = \frac{g(\mathbf{x}_0)f'(\mathbf{x}_0) - f(\mathbf{x}_0)g'(\mathbf{x}_0)}{g^2(\mathbf{x}_0)}$.*

2. Chain rule (Composition): *If* $f : R^n \to R^m$ *is differentiable at* $\mathbf{x}_0$ *and* $g : R^m \to R^p$ *is differentiable at* $f(\mathbf{x}_0)$, *then* $g \circ f$ *is differentiable at* $\mathbf{x}_0$ *and* $(g \circ f)'(\mathbf{x}_0) = g'(f(\mathbf{x}_0))f'(\mathbf{x}_0)$.

Definition 3.1.2 *Differentiable on a set*
f is *differentiable on a set* S means f is differentiable at each point in S.

3.1.3 Partial derivatives

Definition 3.1.3 *Partial derivative*
The *partial derivative* of f with respect to x_j at $\mathbf{x}_0$ is

$$\begin{aligned}\frac{\partial f}{\partial x_j}(\mathbf{x}_0) = f_{x_j}(\mathbf{x}_0) &\equiv \lim_{x_j \to x_{0j}} \frac{f(x_{01}, \ldots, x_j, \ldots, x_{0n}) - f(x_{01}, \ldots, x_{0j}, \ldots, x_{0n})}{x_j - x_{0j}} \\ &\equiv \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{e_j}) - f(\mathbf{x}_0)}{h}\end{aligned}$$

where $\mathbf{e_j}$ is the j-th standard basis vector in R^n, i.e. the j-th entry is one and the others are zero.

Theorem 3.1.3 *If* f *is differentiable at* $\mathbf{x}_0$, *then all partial derivatives exist at* $\mathbf{x}_0$ *and*

$$f'(\mathbf{x}_0) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)\right]$$

3.1.4 Continuously differentiable

An *open ball* centered at $\mathbf{x}_0$ with radius $r > 0$ is the set $\{\mathbf{x} : |\mathbf{x} - \mathbf{x}_0| < r\}$.

Definition 3.1.4 *Continuously differentiable*
$f : R^n \to R^m$ is *continuously differentiable* at $\mathbf{x}_0$ means all partial derivatives of f are continuous on open ball centered at $\mathbf{x}_0$.
f is *continuously differentiable on a set* S means f is continuously differentiable at each point of S.

Note that if f is continuously differentiable at $\mathbf{x}_0$ then it is continuously differentiable at each point in some ball containing $\mathbf{x}_0$.

Theorem 3.1.4 (Basic functions and building processes) *Theorems 4.1.1 and 4.1.2 are true with "differentiable" replaced with "continuously differentiable."*

Theorem 3.1.5 *If $f : R^n \to R^m$ is continuously differentiable at $\mathbf{x}_0$, then f is differentiable at $\mathbf{x}_0$.*

3.1.5 Directional derivatives and the gradient

Definition 3.1.5 *Directional derivative*
For $f : R^n \to R^m$, $\mathbf{x}_0 \in \text{Dom } f$, and $\mathbf{u} \in R^n$, with $|\mathbf{u}| = 1$, The *directional derivative* of f with respect to $\mathbf{u}$ at $\mathbf{x}_0$ is defined by

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0)}{h}$$

A partial derivative is just the directional derivative in the direction of a standard basis vector.

Theorem 3.1.6 *If f is differentiable at $\mathbf{x}_0$, then for all unit vectors $\mathbf{u} \in R^n$*

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = f'(\mathbf{x}_0)\mathbf{u}.$$

For $f : R^n \to R$, the directional derivative is interpreted as the *rate of change* of f with respect to $\mathbf{x}$ in the "direction" $\mathbf{u}$, at $\mathbf{x}_0$. The direction in which f has the greatest increase is given by the *gradient*

$$\nabla f(\mathbf{x}_0) \equiv (f'(\mathbf{x}_0))^T$$

i.e. the direction in which f increases the fastest is $\mathbf{u} = \frac{1}{|\nabla f(\mathbf{x}_0)|}\nabla f(\mathbf{x}_0)$.

3.2 More Geometry

The columns and rows of the derivative matrix, $f'(\mathbf{x}_0)$ have useful geometrical interpretations.

The j-th column is the partial derivative of f with respect to x_j. These columns are a spanning set for the column space of $f'(\mathbf{x}_0)$ (and a basis if the rank of $f'(\mathbf{x}_0)$ is n). If they are thought of as "arrows" with their tails at $f(\mathbf{x}_0)$ then they are tangent vectors to the image and "generate" the tangent plane.

The transpose of the i-th row is the gradient of the i-th component function. The level set of T at $f(\mathbf{x}_0)$ is the plane tangent at the point $\mathbf{x}_0$ to the level set of f at $f(\mathbf{x}_0)$. The level set of T is the set of $\mathbf{x}$ satisfying the equation

$$f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)+f(\mathbf{x}_0)=f(\mathbf{x}_0)$$

or

$$\mathbf{0}=f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)=\begin{bmatrix} f_1'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \\ \vdots \\ f_m'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \end{bmatrix}$$

but, then for all $i=1,\ldots,m$

$$f_i'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)=\langle \nabla f_i(\mathbf{x}_0),\mathbf{x}-\mathbf{x}_0\rangle=0$$

that is, at the point $\mathbf{x}_0$, the gradients of the components functions are orthogonal to the (tangent plane to the) level set of f.

3.3 Important Relationships

$$\begin{array}{ccccc}
\boxed{\begin{array}{c}\text{Continuously}\\ \text{differentiable}\end{array}} & \begin{array}{c}\Rightarrow\\ \not\Leftarrow\end{array} & \boxed{\text{Differentiable}} & \begin{array}{c}\Rightarrow\\ \not\Leftarrow\end{array} & \boxed{\text{Continuous}} \\
 & & \Downarrow\not\Uparrow & & \not\Downarrow\not\Uparrow \\
 & & \boxed{\begin{array}{c}\text{Directional}\\ \text{derivatives exist}\end{array}} & \begin{array}{c}\Rightarrow\\ \not\Leftarrow\end{array} & \boxed{\text{Derivative exists}}
\end{array}$$

The diagram above depicts the general situation, but *for functions $R \to R$ the situation is much simpler because the derivative, the partial derivatives and the directional derivative are the same and being differentiable is equivalent to the existence of the derivative.*

3.4 Exercises

3–1. For the following functions find the derivative, and at the given point find the best affine approximation. Where meaningful describe the tangent planes to the graph, image and appropriate level sets. Sketches would be wonderful.

(a) $f: R \to R^2$ defined by $f(t)=(\cos t,\sin t)$; point: $t=\frac{\pi}{3}$.

(b) $f: R^2 \to R^3$ defined by $f(r,\theta)=(r\cos\theta, r\sin\theta, r)$; point: $(r,\theta)=(1,\frac{\pi}{4})$.

(c) $f: R^3 \to R^2$ defined by $f(x,y,z)=(x+y+z, x^2+y^2+z^2)$; point: $(x,y,z)=(1,-1,1)$.

(d) $f : R^2 \to R^3$ defined by $f(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$; point: $(\theta, \phi) = (0, \frac{\pi}{2})$.

3–2. What vectors, built using the derivative matrix, are tangent to the graph of a differentiable function? What vectors, built using the derivative matrix, are orthogonal to the graph of a differentiable function?

3–3. Show that the function $f : R^2 \to R$ defined by

$$f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

is continuous but not differentiable at $(0,0)$, although both partial derivatives exist there.

3–4. Show that the function $f : R^2 \to R$ defined by

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

is differentiable everywhere and the partial derivatives are bounded near $(0,0)$, but are discontinuous there.

3–5. Show that the function $f : R^2 \to R$ defined by

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\dfrac{1}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

is differentiable everywhere but has unbounded partial derivatives near $(0,0)$.

3–6. For the function $f : R^2 \to R$ defined by $f(x,y) = xye^{x+y}$,

(a) Find the directional derivative of f in the direction $\mathbf{u} = (\frac{3}{5}, -\frac{4}{5})$ at the point $(1,-1)$.

(b) Find the direction in which f is increasing the fastest at $(1,-1)$

(c) Find the rate of change of f in the direction tangent to the curve that is the image of $g(t) = (t^2, t^3 - 2)$ at the point $g(1)$.

3–7. Show that the function $f : R^2 \to R$ defined by

$$f(x,y) = \begin{cases} \dfrac{x|y|}{\sqrt{x^2+y^2}} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

has a directional derivative in every direction at $(0,0)$, but is not differentiable at $(0,0)$.

Chapter 4

Differential Calculus: Higher Derivatives and Big Theorems

4.1 Higher Order Derivatives and C^k functions

In this section and the next we will consider only real valued functions, i.e. $: R^n \to R$. All of the definitions and theorems can be formulated for general functions but are rarely needed in that form.

Definition 4.1.1 *k-th order derivative*
For $f : R^n \to R$ and non-negative integer k, a *k-th order partial derivative* is of the form

$$\frac{\partial^k f}{\partial x_{j_1} \partial x_{j_2} \ldots \partial x_{j_k}}(\mathbf{x}) = f_{x_{j_k} \ldots x_{j_1}}(\mathbf{x}) \equiv \frac{\partial}{\partial x_{j_1}} \frac{\partial}{\partial x_{j_2}} \ldots \frac{\partial f}{\partial x_{j_k}}(\mathbf{x})$$

Note the reversing of the order of the x's in the two notations for a k-th derivative. In general a function has n^k k-th order derivatives, but the functions one usually sees have fewer because of the following theorem.

Theorem 4.1.1 *If f, f_{x_i}, f_{x_j}, $f_{x_i x_j}$ and $f_{x_j x_i}$ are continuous on an open ball, then $f_{x_i x_j} = f_{x_j x_i}$ there.*

It follows from this theorem that if all the k-th order derivatives are continuous, then the order in which the differentiations are done does not matter so a k-th order derivative is usually denoted

$$\frac{\partial^k f}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}}(\mathbf{x})$$

where $k_1, k_2, \ldots k_n$ are non-negative integers that sum to k, and mean that the function f is differentiated k_j times with respect to x_j.

The single 0-th order derivative of f is simply f, itself.

The second order derivatives are sometimes put into an $n \times n$ matrix called the *hessian* or simply the second derivative matrix

$$f''(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right]$$

Note that if all second order partial derivatives are continuous then the second derivative matrix is symmetric.

Definition 4.1.2 *C^k function spaces*
For a set $S \subset R^n$, $C^k(S)$ is the set of all real-valued functions with all k-th order partial derivatives continuous on S.
$C^\infty(S)$ is the set of functions with continuous partial derivatives of all orders.

$C^k(S)$ is an infinite dimensional vector space. Moreover,

$$C^0(S) \supset C^1(S) \supset C^2(S) \ldots \supset C^k(S) \ldots \supset C^\infty(S)$$

4.2 Taylor's Theorem

Definition 4.2.1 *Taylor Polynomial*
For $f : R^n \to R$, a non-negative integer K and $\mathbf{x}_0$, so that f is C^K in an open ball containing $\mathbf{x}_0$, the *Taylor polynomial* of degree K for f about $\mathbf{x}_0$ is

$$P_K(x) = \sum_{k=0}^{K} \frac{1}{k!} \sum_{k_1+\ldots+k_n=k} \frac{k!}{k_1! \ldots k_n!} \frac{\partial^k f(\mathbf{x}_0)}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} (x_1 - x_{01})^{k_1} \ldots (x_n - x_{0n})^{k_n}$$

Important special cases:

- $P_0(\mathbf{x}) = f(\mathbf{x}_0)$
- $P_1(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$
- $P_2(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T f''(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$

Theorem 4.2.1 *The partial derivatives of f at $\mathbf{x}_0$ are equal to the corresponding partial derivatives of P_K at $\mathbf{x}_0$ for all orders less than or equal to K.*

Theorem 4.2.2 (Taylor's Theorem) *If $f : R^n \to R$ is C^K in an open ball containing $\mathbf{x}_0$ and P_K is the Taylor polynomial of degree K for f about $\mathbf{x}_0$, then*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - P_K(\mathbf{x})}{|\mathbf{x}-\mathbf{x}_0|^K} = 0$$

and P_K is the only polynomial of degree K with this property.

4.3 Inverse and Implicit Function Theorems

4.3.1 Inverse Functions

Definition 4.3.1 *Inverse of a function*
For $f : R^n \to R^n$

- A *(global) inverse* is a function $f^{-1} : R^n \to R^n$, so that

$$f^{-1}(\mathbf{y}) = \mathbf{x} \text{ if and only if } \mathbf{y} = f(\mathbf{x})$$

- For a set $S \subset \text{Dom } f$, f has an *inverse on* S, $f^{-1} : R^n \to R^n$, means for all $\mathbf{x}$ in S

$$f^{-1}(\mathbf{y}) = \mathbf{x} \text{ if and only if } \mathbf{y} = f(\mathbf{x})$$

- For $\mathbf{x}_0 \in \text{Dom } f$, f has a *(local) inverse at* $\mathbf{x}$ means that f has an inverse on an open ball containing $\mathbf{x}$.

It follows that the domain of a global inverse, f^{-1}, is the image of f and the image of f^{-1} is the domain of f. Moreover, $f^{-1}(f(\mathbf{x})) = \mathbf{x}$ and $f(f^{-1}(\mathbf{y})) = \mathbf{y}$.

The existence of an inverse may depend on restrictions on the domain of f, so that a function may have more than one inverse or none at all. For example, $f(x) = x^2$ has $f^{-1}(y) = \sqrt{y}$, on $S = \{x : x \geq 0\}$; $f^{-1}(y) = -\sqrt{y}$ on $S = \{x : x \geq 0\}$; and no inverse if the domain is defined so as to include an open interval containing zero. Moreover, we can say that f has a local inverse at every point in R except zero.

Theorem 4.3.1 (Inverse Function Theorem) *If $f : R^n \to R^n$ is continuously differentiable at $\mathbf{x}_0$ and $f'(\mathbf{x}_0)$ is non-singular, then there is an open ball B containing $\mathbf{x}_0$, so that f, when restricted to B, has a continuously differentiable inverse and for all $\mathbf{x}$ in B*

$$(f^{-1})'(f(\mathbf{x})) = (f'(\mathbf{x}))^{-1}$$

This theorem says that if the best affine approximation to a continuously differentiable f near $\mathbf{x}_0$ has an inverse, then locally so does f.

4.3.2 Implicit Function Theorem

Recall that we have sometimes denoted vectors in R^{n+m} by $(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} \in R^n$ and $\mathbf{y} \in R^m$, we extend this notation to the derivative matrix as follows. For $F : R^{n+m} \to R^m$, the derivative will be partitioned into two pieces

$$F'(\mathbf{x}, \mathbf{y}) = [F_{\mathbf{x}}(\mathbf{x}, \mathbf{y}) \; F_{\mathbf{y}}(\mathbf{x}, \mathbf{y})]$$

where $F_{\mathbf{x}}$ denotes the $m \times n$ matrix whose columns are the partial derivatives of F with respect to the variables in $\mathbf{x}$ and $F_{\mathbf{y}}$ the $m \times m$ matrix whose columns are the partial derivatives of F with respect to the variables in $\mathbf{y}$ variables.

Theorem 4.3.2 (Implicit Function Theorem) *If $F : R^{n+m} \to R^m$ is continuously differentiable at $(\mathbf{x}_0, \mathbf{y}_0)$ and $F_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)$ has an inverse, then there is an open ball B containing $(\mathbf{x}_0, \mathbf{y}_0)$ and a continuously differentiable function $f : R^n \to R^m$ so that, for $(\mathbf{x}, \mathbf{y})$ in B, $F(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}_0, \mathbf{y}_0)$ if and only if $\mathbf{y} = f(\mathbf{x})$. Moreover, for $(\mathbf{x}, f(\mathbf{x})) \in B$,*

$$f'(x) = -(F_{\mathbf{y}}(\mathbf{x}, f(\mathbf{x})))^{-1} F_{\mathbf{x}}(\mathbf{x}, f(\mathbf{x}))$$

This theorem has two important interpretations, one algebraic and one geometric. If T is the best affine approximation to F near $(\mathbf{x}_0, \mathbf{y}_0)$, then

Algebraic: If $T(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}_0, \mathbf{y}_0)$ can be solved for $\mathbf{y}$ in terms of $\mathbf{x}$, then $F(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}_0, \mathbf{y}_0)$ can be solved for $\mathbf{y}$ in terms of $\mathbf{x}$ near $(\mathbf{x}_0, \mathbf{y}_0)$.

Geometric: If the level set of T at $F(\mathbf{x}_0, \mathbf{y}_0)$ is the graph of a function $t : R^n \to R^m$, then the level set of F at $F(\mathbf{x}_0, \mathbf{y}_0)$ is the graph of a function $f : R^n \to R^m$, near $(\mathbf{x}_0, \mathbf{y}_0)$. Moreover, t is the best affine approximation to f near $\mathbf{x}_0$.

The partitioning into $(\mathbf{x}, \mathbf{y})$ was done simply for convenience in precisely stating the theorem. One should more generally interpret the theorem as saying

> For a continuously differentiable function $F : R^N \to R^m$, with $N > m$, if the rank of $F'(\mathbf{z}_0)$ is maximal (i.e. m) then the equation $F(\mathbf{z}) = F(\mathbf{z}_0)$ can be solved for some m variables in terms of the other $N - m$ near $\mathbf{z}_0$.

A particular choice of m variables can be solved for if the partial derivatives at $\mathbf{z}_0$ with respect to those variables are linearly independent.

4.4 Exercises

4–1. If p is a fixed positive integer, what is the largest value of k for which the function $f : R \to R$, defined by

$$f(x) = \begin{cases} x^p \sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is in $C^k(R)$?

4–2. If

$$f(x, y) = \begin{cases} 2xy \, \dfrac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

show that $f_{xy}(0, 0) = -2$ and $f_{yx}(0, 0) = 2$. What does this say about Theorem 5.1.1?

4–3. Taylor polynomials:

(a) Find the Taylor polynomial of degree two for $f(x, y) = xe^{x+y}$ about $(0, 0)$ by using derivatives, then by substitution.

(b) Find the second degree Taylor expansion of $f(x, y, z) = (x^2 + 2xy + y^2)e^z$ about $(1, 2, 0)$.

(c) Compute the second degree Taylor polynomial for $e^{-|\mathbf{x}|^2}$ about $\mathbf{0}$ in R^n.

4–4. For the function $f : R^2 \to R^2$ defined by $f(x, y) = (x^2 - y^2, 2xy)$

(a) Find the points for which f has a local inverse.

(b) Show that f does not have a global inverse.

(c) If f^{-1} is the inverse for f near $(1, 2)$, compute the best affine approximation to f^{-1} near $f(1, 2)$.

(d) What is f^{-1} near $f(1, 2)$?

4–5. For $f : R \to R$, defined by

$$f(x) = \begin{cases} \dfrac{x}{2} + x^2 \sin(\dfrac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

show that the best affine approximation to f near 0 has an inverse but that f does not have an inverse near 0. What does this say about the Inverse Function Theorem?

4–6. Show that the equations

$$\begin{aligned} 2x + y + 2z + u - v &= 1 \\ xy + z - u + 2v &= 1 \\ yz + xz + u^2 + v &= 0 \end{aligned}$$

can be solved near $(1, 1, -1, 1, 1)$ for x, y and z in terms of u and v and find the derivative of the resulting function $f : R^2 \to R^3$ at $(1, 1)$.

4–7. The Inverse Function Theorem can be generalized as follows

If $f : R^n \to R^m$, where $n < m$, is continuously differentiable and T, the best affine approximation to f near $\mathbf{x}_0$, is one-to-one, then there is a open ball B centered at $\mathbf{x}_0$ so that f restricted to B has an inverse.

(a) What is the rank of $f'(\mathbf{x}_0)$ in this situation?

(b) For what points does $f(x, y) = (x + y, (x + y)^2, (x + y)^3)$ have a local inverse?

(c) Prove the generalized Inverse Function Theorem.
[*Hint:* For convenience, let $D = f'(\mathbf{x}_0)$, define $g : R^n \to R^n$ by $g(\mathbf{x}) = (D^T D)^{-1} D^T f(\mathbf{x})$. Justify that this function exists and satisfies the hypotheses of the Inverse Function Theorem at $\mathbf{x}_0$, then make the most of it.]

(d) Can a similar generalization be given for a continuously differentiable function $f : R^n \to R^m$ with $n > m$? Justify your answer.

4–8. For the function $f : R \to R$ given by $f(x) = 1 + x + \sin(x)$, find the second degree Taylor polynomial about $y_0 = 1$ for f^{-1}.

Chapter 5

Applications

5.1 Optimization

A typical optimization problem consists of the following.

> For $f : R^n \to R$ and $S \subset R^n$, find the maximum (or minimum) value of $f(\mathbf{x})$ for $\mathbf{x} \in S$ and the points $\mathbf{x}$ at which the optimum occurs.

5.1.1 Unconstrained Optimization

Definition 5.1.1 *Local optimum*
For $f : R^n \to R$

- $f(\mathbf{x}_0)$ is a *local minimum* for f, if there is an open ball B containing $\mathbf{x}_0$ so that $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ for all $\mathbf{x} \in B$.

- $f(\mathbf{x}_0)$ is a *strict local minimum* for f, if there is an open ball B containing $\mathbf{x}_0$ so that $f(\mathbf{x}) > f(\mathbf{x}_0)$ for all $\mathbf{x} \in B$, $\mathbf{x} \neq \mathbf{x}_0$.

- $f(\mathbf{x}_0)$ is a *local maximum* for f, if there is an open ball B containing $\mathbf{x}_0$ so that $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ for all $\mathbf{x} \in B$.

- $f(\mathbf{x}_0)$ is a *strict local maximum* for f if there is an open ball B containing $\mathbf{x}_0$ so that $f(\mathbf{x}) < f(\mathbf{x}_0)$ for all $\mathbf{x} \in B$, $\mathbf{x} \neq \mathbf{x}_0$.

- A *local optimum (*or *local extremum)* is a value of f that is either a local maximum or local minimum.

Theorem 5.1.1 *If $f : R^n \to R$ has a derivative at $\mathbf{x}_0$ and $f(\mathbf{x}_0)$ is a local optimum, then $f'(\mathbf{x}_0) = \mathbf{0}$.*

Definition 5.1.2 *(Unconstrained) critical point:*
If $f : R^n \to R$ has a derivative at $\mathbf{x}_0$, then $\mathbf{x}_0$ is a *critical point* of f means $f'(\mathbf{x}_0) = \mathbf{0}$.

The theorem can then be rephrased to say that local optima of continuously differentiable functions occur at critical points. Unfortunately, the converse is not true, as is illustrated by $f(x) = x^3$.

The best affine approximation to f at a critical point, $\mathbf{x}_0$, is a constant function, so that the tangent plane to the graph of f at $(\mathbf{x}_0, f(\mathbf{x}_0))$ is "horizontal." Moreover, if $f(\mathbf{x}_0)$ is a local minimum the the graph of f over some ball about $\mathbf{x}_0$ lies "above" its tangent plane, or below for a local maximum. As the example in the previous paragraph points out, the graph of f may cross the tangent plane, in which case the point $(\mathbf{x}_0, f(\mathbf{x}_0))$ is called a *saddle point* or *inflection point* of the graph of f. Which of these possibilities actually occurs can sometimes be determined by looking at the second derivative of f.

Theorem 5.1.2 *If $f : R^n \to R$ is C^2 on an open ball containing a critical point $\mathbf{x}_0$ of f, then*

1. *If $f''(\mathbf{x}_0)$ is positive definite, then $f(\mathbf{x}_0)$ is a strict local minimum.*
2. *If $f''(\mathbf{x}_0)$ is negative definite, then $f(\mathbf{x}_0)$ is a strict local maximum.*
3. *If $f''(\mathbf{x}_0)$ has both positive and negative eigenvalues, then $(\mathbf{x}_0, f(\mathbf{x}_0))$ is a saddle point.*

Note that this theorem does *not* cover all of the possibilities. In any cases not covered by the theorem the second derivative does not, in general, provide any information. Consider the functions $f(x) = x^3$, $g(x) = x^4$ and $h(x) = -x^4$.

This theorem is really stated to make the classification of critical points a computation on the first and second derivatives. The following equivalent version expresses the result in terms of the philosophy of differential calculus.

Theorem 5.1.3 *If $f : R^n \to R$ is C^2 on an open ball containing a point $\mathbf{x}_0$ and P_2 is the Taylor polynomial of degree two for f about $\mathbf{x}_0$, then*

1. *If $P_2(\mathbf{x}_0)$ is a strict global minimum, then $f(\mathbf{x}_0)$ is a strict local minimum.*

2. *If $P_2(\mathbf{x}_0)$ is a strict global maximum, then $f(\mathbf{x}_0)$ is a strict local maximum.*

3. *If $(\mathbf{x}_0, f(\mathbf{x}_0))$ is a saddle point for P_2, then $(\mathbf{x}_0, f(\mathbf{x}_0))$ is a saddle point for f.*

Note that it is not necessary to assume explicitly that $\mathbf{x}_0$ is a critical point.

5.1.2 Constrained Optimization

When the set S over which $f : R^n \to R$ is to be optimized is given by

$$S = \{\mathbf{x} : G(\mathbf{x}) = \mathbf{0}\}$$

for some continuously differentiable function $G : R^n \to R^m$, then the problem is called a *constrained optimization* problem and is often stated

Find the optimum values of $f(\mathbf{x})$ subject to the constraints $G(\mathbf{x}) = \mathbf{0}$.

Definition 5.1.3 *Constrained critical point:*
For $f : R^n \to R$ and $S = \{\mathbf{x} : G(\mathbf{x}) = \mathbf{0}\}$, a point $\mathbf{x}_0 \in S$ is a *constrained critical point* of f on S means $\nabla f(\mathbf{x}_0)$ is orthogonal to S at $\mathbf{x}_0$.

Recall that a vector is orthogonal to a set at a point if the vector is orthogonal to the tangent plane at that point. Since the set S is a level set, If G is sufficiently nice, that is, rank$(G'(\mathbf{x}) = m$, a vector is orthogonal to the tangent plane at $\mathbf{x}$, if it is in the row space of $G'(\mathbf{x})$, i.e. it is a linear combination of the gradients of the component functions of G at $\mathbf{x}$.

Theorem 5.1.4 *If $G : R^n \to R^m$ is continuously differentiable, $n > m$ and $G'(\mathbf{x})$ has rank m at each point of $S = \{\mathbf{x} : G(\mathbf{x}) = \mathbf{0}\}$, then local maxima and minima on S of a differentiable function $f : R^n \to R$ occur at constrained critical points.*

The constrained critical points can be found using so called Lagrange multipliers, λ in the following theorem.

Theorem 5.1.5 (Lagrange) *If $G : R^n \to R^m$ is continuously differentiable, $n > m$ and $G'(\mathbf{x})$ has rank m at each point of $S = \{\mathbf{x} : G(\mathbf{x}) = \mathbf{0}\}$, then the constrained critical points on S of a continuously differentiable function $f : R^n \to R$ are the unconstrained critical points of the function $F : R^{n+m} \to R$ defined by*

$$F(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda^T G(\mathbf{x}).$$

5.2 Surfaces and Tangents

Definition 5.2.1 *Describing surfaces with functions*
A subset S of R^{n+m} is described

- *explicitly* by $f : R^n \to R^m$ means S is the graph of f.
- *implicitly* by $F : R^{n+m} \to R^m$ means there is a $\mathbf{b}$ in R^m so that S is the level set of F at $\mathbf{b}$.
- *parametrically* by $g : R^n \to R^{n+m}$ means S is the image of g.
- A set S described explicitly, implicitly or parametrically is called a *smooth surface* (or *curve*, if $n = 1$), if the describing function is continuously differentiable and its derivative has maximal rank at each point of S.

The terminology for smooth is somewhat misleading. Whether the surface is smooth or not depends on the function used to describe it as well as the set S itself.

If S is the graph of a function $f : R^n \to R^m$ then it can be visualized as an n-dimensional surface in R^{n+m}. It can also be described implicitly and parametrically in the following ways.

- Implicitly: Let $\mathbf{b} = \mathbf{0}$ and $F : R^{n+m} \to R^m$ defined by $F(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}$.
- Parametrically: Let $g : R^n \to R^{n+m}$ be defined by $g(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))$.

It is not generally possible to go from an implicit or parametric description to an explicit description, but the Implicit and Inverse Function Theorems give conditions where explicit descriptions exist locally on smooth surfaces.

Implicit → Explicit: If S is described implicitly by continuously differentiable $F : R^{n+m} \to R^m$ and rank $F'(\mathbf{z}_0) = m$, then the implicit function theorem ensures that near $\mathbf{z}_0$ some m of $z_1, \ldots, z_{n+m}$ can be solved for in terms of the other n variables. The solution $(z_{i_1}, \ldots, z_{i_m}) = f(z_{j_1}, \ldots, z_{j_n})$ provides the explicit description.

Parametric → Explicit: If S is described parametrically by a continuously differentiable $g : R^n \to R^{n+m}$ and rank $g'(\mathbf{t}_0) = n$ then S can be described explicitly as follows. The rank condition says that some n rows of $g'(\mathbf{t}_0)$ are linearly independent. For convenience, suppose the the first n rows are independent. If $g(\mathbf{t})$ is partitioned as follows

$$g(\mathbf{t}) = \begin{bmatrix} G(\mathbf{t}) \\ H(\mathbf{t}) \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}$$

where $G(\mathbf{t}) = \mathbf{x}$ is in R^n and $H(\mathbf{t}) = \mathbf{y}$ is in R^m, then $G'(\mathbf{t}_0)$ consists of the n independent rows of $g'(\mathbf{t}_0)$ and is, therefore, non-singular. So, G has a local inverse and the local explicit description of S (i.e. near $g(\mathbf{t}_0)$) is given by $f(\mathbf{x}) = H(G^{-1}(\mathbf{x}))$.

To summarize, for a function that describes a surface either implicitly or parametrically, if the rank of the derivative of at a point is maximal there is an explicit description near the corresponding point on the surface. Moreover, the existence of the explicit description ensures the existence of the other of the two.

The tangent plane to a smooth surface at a point can be calculated from a describing function, as has been mentioned before. In particular, the tangent plane is described by the best affine approximation in the same way that the surface is described by the function. For example, if the surface is described parametrically by a function g, then the tangent plane to the surface at a point $g(\mathbf{t}_0)$ on the surface is described parametrically by the best affine approximation to g at $\mathbf{t}_0$. For parametric and implicit descriptions, the rank of the derivative matrix must be maximal at a point for tangency at the point to make any geometrical sense.

Tangent and normal (orthogonal) vectors to a smooth surface can be obtained from the derivative matrices of describing functions. In particular,

- Parametric: For $j = 1, \ldots, n$, $\quad g_{t_j}(\mathbf{t})$ is tangent to the surface at at $g(\mathbf{t})$.
- Implicit: For $i = 1, \ldots, m$, $\quad \nabla F_i(\mathbf{x})$ is orthogonal to the surface at $\mathbf{x}$.
- Explicit: For $j = 1, \ldots, n$, $\quad (\mathbf{e}_j, f_{x_j}(\mathbf{x}))$ is tangent to the surface at $(\mathbf{x}, f(\mathbf{x}))$, and for $i = 1, \ldots, m$, $(\nabla f_i(\mathbf{x}), -\mathbf{e}_i)$ is orthogonal at $(\mathbf{x}, f(\mathbf{x}))$.

5.3 Exercises

5–1. Find and classify the critical points of the following functions.

(a) $f(x, y) = x^2 - xy - y^2 + 5y - 1$.

(b) $f(x, y) = (x + y)e^{xy}$.

(c) $f(x, y, z) = xy + xz$.

(d) $f(x, y, z) = x^2 + y^2 + z^2$.

(e) $f(x, y) = (x^2 + y^2)\ln(x^2 + y^2)$.

5–2. Find the point on the image of $f(t) = (\cos t, \sin t, \sin(t/2))$ that is farthest from the origin.

5–3. For the function $f(x, y, z) = x^2 + xy + y^2 + yz + z^2$

(a) Find the maximum value of f, subject to the constraint $x^2+y^2+z^2=1$.

(b) Find the maximum value of f subject to the constraints $x^2+y^2+z^2=1$ and $ax+by+cz=0$, where (a,b,c) is the point at which the maximum is attained in (a).

5–4. Find the minimum distance in R^2 from the ellipse $x^2+2y^2=1$ to the line $x+y=4$.

5–5. The fact that the graph of a function lies "above," "below" or crosses its tangent plane has nothing to do with critical points. For example, we could say that the graph of a differentiable function f lies *above* its tangent plane near the point $(\mathbf{x}_0, f(\mathbf{x}_0))$, means there is an open ball B about $\mathbf{x}_0$ so that

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)$$

for all $\mathbf{x} \in B$. Reversing the inequality defines lying *below*. The point $(\mathbf{x}_0, f(\mathbf{x}_0))$ is a *saddle point* of the graph of f means for any open ball B containing $\mathbf{x}_0$ there are points $\hat{\mathbf{x}}$ and $\tilde{\mathbf{x}} \in B$ so that $f(\hat{\mathbf{x}}) > f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\hat{\mathbf{x}}-\mathbf{x}_0)$ and $f(\tilde{\mathbf{x}}) < f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\tilde{\mathbf{x}}-\mathbf{x}_0)$, i.e. the graph of f crosses its tangent plane.

(a) Prove the following theorem.

Theorem 5.3.1 *If $f : R^n \to R$ is C^2 on an open ball containing $\mathbf{x}_0$, then*

- *If $f''(\mathbf{x}_0)$ is positive definite, the graph of f lies above its tangent plane near the $(\mathbf{x}_0, f(\mathbf{x}_0))$.*
- *If $f''(\mathbf{x}_0)$ is negative definite, the graph of f lies below its tangent plane near the $(\mathbf{x}_0, f(\mathbf{x}_0))$.*
- *If $f''(\mathbf{x}_0)$ has both positive and negative eigenvalues, $(\mathbf{x}_0, f(\mathbf{x}_0))$ is a saddle point.*

(b) Determine whether the graph of the given function lies above, below or crosses its tangent plane at the given point in the domain.

i. $x^2 \sin x$ at $x=1$.

ii. $1/(x-y)$ at $(x,y)=(2,1)$.

iii. x^4+y^4 at $(0,0)$.

iv. $e^{z+w}-x^2-y^2$ at $(0,0,0,0)$.

5–6. Show that for an $n \times n$, symmetric matrix A, if $|\mathbf{x}|=1$, then

$$\lambda_{\text{min}} \leq \mathbf{x}^T A\, \mathbf{x} \leq \lambda_{\text{max}}$$

where λ_{min} and λ_{max} are the smallest and largest eigenvalues of A, respectively.

5–7. Find the tangent plane at $(1, 0, \frac{\sqrt{3}}{2})$ to the surface defined implicitly by $9x^2 + 4y^2 + 36z^2 = 36$. Find tangent and normal vectors to the surface at $(1, 0, \frac{\sqrt{3}}{2})$.

5–8. S is the image of the function defined by $g(r, \theta) = (r \cos\theta, r \sin\theta, \theta)$ for $0 \le r \le 4$ and $0 \le \theta \le 2\pi$.

(a) Sketch the surface.

(b) Find tangent vectors and normal vectors to the surface at $g(2, \pi/2)$.

(c) Show that S can be described explicitly near $(0, 2, \pi/2)$. Find the local explicit description.

5–9. This problem discusses a curve, S, in the plane called the *strophoid*.

(a) S is described parametrically by $g : R \to R^2$ defined by $g(t) = (1 - 2\cos^2 t, \tan t - \sin 2t)$ for $-\frac{\pi}{2} < t < \frac{\pi}{2}$.

i. Sketch the curve.

ii. Find the tangent lines to the curve at $g(0)$, $g(\frac{\pi}{4})$ and $g(-\frac{\pi}{4})$.

(b) Implicit description.

i. Show that the curve is described implicitly by

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

ii. Find the tangent line at the point $(-1, 0)$.

iii. Discuss tangency at $(0, 0)$.

Chapter 6

Topology in R

6.1 The Real Number System

The basic number systems are the *natural numbers*, $N = \{1, 2, 3, \ldots\}$, the *integers*, $Z = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$, the *rational numbers*, $Q = \{\frac{p}{q} : p, q \in Z \text{ and } q \neq 0\}$, and the *real numbers*, R.

The real numbers form a *complete, totally ordered field.* That is, they satisfy the following properties, which are taken as axioms, i.e. assumed without proof.

Field axioms: The operations of addition and multiplication are closed and satisfy the following.

Addition: For $x, y, z \in R$,

A-1. (x+y)+z = x+(y+z)

A-2. There is an additive identity, 0, so that $x + 0 = x$

A-3. For each x there is an additive inverse, $-x$, so that $x + (-x) = 0$

A-4. $x + y = y + x$

Multiplication: For $x, y, z \in R$,

M-1. $(xy)z = x(yz)$

M-2. There is a multiplicative identity, 1, so that $x1 = x$

M-3. For each $x \neq 0$ there is a multiplicative inverse, $\frac{1}{x}$, so that $x\frac{1}{x} = 1$

M-4. $xy = yx$

D-1. $x(y + z) = xy + xz$

Order axioms: There is a non-empty subset, P of R, called the *positive* reals, so that

1. If $x, y \in P$, then $x + y \in P$
2. If $x, y \in P$, then $xy \in P$
3. If $x \in R$, then one and only one of the following is true.
   (a) $x \in P$
   (b) $x = 0$
   (c) $-x \in P$

Archimedean axiom: For each $x \in R$ there is a natural number n, so that $n - x \in P$

Definition 6.1.1 *Cauchy sequence*
A sequence $\{x_k\}_{k=1}^{\infty}$ is *cauchy*, if for every $\varepsilon > 0$ there is K so that if $k, l \geq K$, then $|x_k - x_l| < \varepsilon$.

Example: A convergent sequence is cauchy.

Completeness axiom: Every cauchy sequence converges.

Amazingly enough, any set with two operations defined on its elements that satisfy the four collections of properties "is" the reals, in the sense that there is a one-to-one correspondence between the elements of the set and the real numbers so that arithmetic corresponds, positives correspond, and corresponding convergent sequences converge to corresponding limits.

The following theorems are important tools for studying the topology of R and are all equivalent to the *completeness axiom*, that is, they could replace it as one of the characterizing properties of the reals. In fact, some of them replace both the *completeness axiom* and the *archimedean axiom*, as will be summarized below. Be patient, some of the terminology used in the theorems is not defined until after the statements of all of them.

Theorem 6.1.1 (Bolzano-Weierstrass Property) *Every infinite, bounded set of real numbers has a limit point.*

Theorem 6.1.2 (Cantor Intersection Property) *If $I_0, I_1, \ldots$ is a nested sequence of closed intervals, i.e. $I_k \supset I_{k+1}$, whose lengths go to zero, then $\bigcap_{k=1}^{\infty} I_k = \{x\}$ for some $x \in R$.*

Theorem 6.1.3 (Supremum Property) *A non-empty set that is bounded above has a supremum.*

Theorem 6.1.4 (Monotone Convergence Property) *If $\{x_n\}_{n=1}^{\infty}$ is a monotone increasing sequence, i.e. $x_n \leq x_{n+1}$ for all n, then $\{x_n\}$ converges if and only if it is bounded above.*

Theorem 6.1.5 (Bounded Convergence Property) *Every bounded sequence has a convergent subsequence.*

Definition 6.1.2 *Bounded set*
A set S is *bounded* means there is a number M so that $|x| \leq M$ for all $x \in S$.

A subset of the reals is bounded if it is contained in a closed interval $[a, b]$ for some a and b. A subset, S, of the reals is said to be *bounded above* if there is a b so that $x \leq b$ for all $x \in S$, and b is called an *upper bound* for S. S is *bounded below* if there is an a so that $a \leq x$ for all $x \in S$, and a is called a *lower bound* for S.

Definition 6.1.3 *Subsequence*
For a sequence $\{x_k\}_{k=1}^{\infty}$, a *subsequence* is $\{x_{k_m}\}_{m=1}^{\infty}$, where $k_1 < \ldots < k_m < \ldots$

Definition 6.1.4 *Supremum and infimum*
For a set $S \subset R$,

- b is the *supremum* (or *least upper bound*) of S, if b is an upper bound of S and is less than any other upper bound.
  The supremum of S is denoted $\sup S$.

- a is the *infimum* (or *greatest lower bound*) of S, if a is a lower bound of S and is greater than any other lower bound.
  The infimum of S is denoted $\inf S$.

Corollary 6.1.6 (to Bolzano-Weierstrass) *If S is an infinite bounded set of real numbers, then S has a limit point x so that for any $\varepsilon > 0$, there are infinitely many $y \in S$ so that $0 < |y - x| < \varepsilon$.*

Corollary 6.1.7 (to Cantor Intersection Theorem) *The intersection of a nested sequence of closed intervals is not empty.*

Corollary 6.1.8 (to Supremum Property) *A set that is bounded below has an infimum.*

Corollary 6.1.9 (to Monotone Convergence Theorem) *A monotone decreasing sequence converges if and only if it is bounded below.*

Theorem 6.1.10 (Characterizations of the Reals) *The following are equivalent*

- *Archimedean axiom and completeness axiom*
- *Archimedean axion and Cantor intersection property*
- *Bolzano-Weierstrass property*
- *Monotone convergence property*
- *Bounded convergence property*

6.2 Continuous Functions

Topology is the study of continuous functions and the structures required to define and analyze them. The following theorems are among the most basic facts about continuous functions from the reals to the reals.

Theorem 6.2.1 (Intermediate Value Theorem) *If f is continuous on an interval I and for a and b in I, $f(a) < \xi < f(b)$, then there is a c strictly between a and b so that $f(c) = \xi$.*

Theorem 6.2.2 *If f is continuous on a closed bounded interval $[a, b]$, then f attains a maximum value at some c in the interval.*

Corollary 6.2.3 *If f is continuous on a closed bounded interval $[a, b]$, then f attains a minimum value at some c in the interval.*

Corollary 6.2.4 *The image of a closed bounded interval under a continuous function is a closed bounded interval.*

Theorem 6.2.5 (Rolle's Theorem) *If f is continuous on a closed bounded interval $[a, b]$ and has a derivative on the open interval (a, b) and $f(a) = f(b) = 0$, then there is a c strictly between a and b so that $f'(c) = 0$.*

Corollary 6.2.6 (Mean Value Theorem) *If f is continuous on a closed bounded interval $[a, b]$ and has a derivative on the open interval (a, b), then there is a c strictly between a and b so that*

$$f(b) - f(a) = f'(c)(b - a)$$

Definition 6.2.1 *Uniform continuity*
A function is *uniformly continuous* on a set S means for any $\varepsilon > 0$, there is a $\delta > 0$, so that for all $x, y \in S$,

$$\text{if } |x - y| < \delta, \text{ then } |f(x) - f(y)| < \varepsilon$$

Theorem 6.2.7 *If f is continuous on a closed bounded interval $[a, b]$, then f is uniformly continuous on $[a, b]$.*

6.3 Exercises

6–1. The usual order relations $<$ and $>$ are defined by

- $x < y$ means $y - x \in P$
- $x > y$ means $x - y \in P$

Using what field properties and their consequences you need and the order axioms, prove the following, for a, b, c, d, and d reals.

(a) If $a > b$ and $b > c$, then $a > c$.
(b) Exactly one of the following holds: $a > b$, $a = b$, $a < b$.
(c) If $a \neq 0$, then $a^2 > 0$
(d) $1 > 0$
(e) If $a > b$ and $c > d$, then $a + c > b + d$
(f) If $a > b$ and $c > 0$, then $ac > bc$
(g) If $a > b$ and $c < 0$, then $ac < bc$

6–2. Show that the *Archimedean axiom* is equivalent to $\lim_{k\to\infty} \frac{1}{k} = 0$.

6–3. Show that the *Archimedean axiom* follows from the *supremum property*. (Note that this fact implies that both the *Archimedean axiom* and the *completeness axiom* could be replaced by the *supremum property* in the axiomatic description of R.)

6–4. Suppose $f : D \to R$ has a bounded image and $D_0 \subset D$, show that

$$\inf\{f(x) : x \in D\} \le \inf\{f(x) : x \in D_0\} \le \sup\{f(x) : x \in D_0\} \le \sup\{f(x) : x \in D\}$$

6–5. Suppose f and g are defined on a set D and both have bounded images, show that

$$\begin{aligned} \inf\{f(x) : x \in D\} + \inf\{g(x) : x \in D\} &\le \inf\{f(x) + g(x) : x \in D\} \\ &\le \inf\{f(x) : x \in D\} + \sup\{g(x) : x \in D\} \\ &\le \sup\{f(x) + g(x) : x \in D\} \\ &\le \sup\{f(x) : x \in D\} + \sup\{g(x) : x \in D\} \end{aligned}$$

give examples to show that each inequality can be strict.

6–6. Show that the function $f : R \to R$ defined by

$$f(x) = \begin{cases} x & \text{if } x \text{ is rational} \\ 1 - x & \text{if } x \text{ is irrational} \end{cases}$$

is continuous at $x = \frac{1}{2}$, but is discontinuous elsewhere. (Recall: to show that a function is discontinuous at x, it suffices to find a sequence $\{x_n\}$ converging to x, so that $\{f(x_n)\}$ does not converge to $f(x)$.)

6–7. Show that if f is continuous on $[0, 1]$ and the image of f is contained in $[0, 1]$, then there is a number $x^* \in [0, 1]$ so that $f(x^*) = x^*$. Interpret this result geometrically.

6–8. A set U is *open*, if for each $x \in U$ there is a open ball $B = \{y : |y - x| < \varepsilon\}$, so that $B \subset U$. For a function f and a set B, the *inverse image of* B is $f^{-1}(B) = \{x : f(x) \in B\}$. (Note: the use of the symbol f^{-1} does not necessarily mean that an inverse for f exists.) For a set A, the *image of* A is the set $f(A) = \{y : y = f(x) \text{ for some } x \in A\}$. Show that for a function continuous on an open set, the inverse image of an open set is open, but that the image of an open set may not be open.

6–9. Show that the function $f(x) = 1/(x+1)$ is uniformly continuous on the interval $(0, \infty)$ but not on the interval $(-1, 0)$.

6–10. A function $f : R \to R$ is called *periodic*, if there is a number $p > 0$, so that $f(x+p) = f(x)$, for all $x \in R$. E.g. sin and cos are periodic with $p = 2\pi$. Show that a continuous periodic function is bounded and uniformly continuous on R.

Chapter 7

Topology in R^n

In this chapter we will generalize to R^n some of the important results about topology and continuous functions in R. Many of the definitions given previously make sense in R^n, simply by replacing real numbers by vectors. In particular, the definitions of *bounded set, bounded sequence, subsequence, cauchy sequence, open set* (Exercise 7-8)*, and image and inverse image* of a set under a function will be used in R^n.

The study of R^n does not require the introduction of new axioms. The vector space properties follow from the field axioms in R. There is no "official" generalization of the order and Archimedean axioms, but the following definition is useful. For vectors $\mathbf{a}$ and $\mathbf{b}$ in R^n, $\mathbf{a} < \mathbf{b}$ means that $a_i < b_i$ for $i = 1, \ldots, n$. Similar definitions for $>, \leq$, and $\geq$ will also be used. It should be noted that this "ordering" in R^n does not satisfy all of the usual order properties if $n > 1$. In particular, (a) in Exercise 7-1 holds, but (b) does not.

7.1 Sets

Theorem 7.1.1 *R^n is complete, i.e. every cauchy sequence converges.*

Note that completeness is a theorem (for $n > 1$), not an axiom.

Theorem 7.1.2 (Bolzano-Weierstrass) *Every bounded infinite set in R^n has a limit point.*

Theorem 7.1.3 *Every bounded sequence in R^n has a convergent subsequence.*

The theorems in Chapter 7, usually involved either open intervals or closed bounded intervals. The *open set* is a generalization of an open interval, that is more useful in higher dimensions. Other kinds of sets that capture some of the essential properties of intervals are the following.

Definition 7.1.1 *Sets*

1. A set in R^n is *closed*, means it contains all of its limit points.

2. If $\mathbf{a} \le \mathbf{b}$, the *closed rectangle* with corners $\mathbf{a}$ and $\mathbf{b}$ is $\{\mathbf{x} : \mathbf{a} \le \mathbf{x} \le \mathbf{b}\}$.

3. A set $S \subset R^n$ is *connected*, means whenever S is contained in the union of two disjoint open sets, then its intersection with one of them is empty.

4. A set $S \subset R^n$ is *sequentially compact* means that every sequence of points in S has a subsequence that converges to a point in S.

5. A set $S \subset R^n$ is *compact* means that whenever S is contained in the union of a collection of open sets, S is contained in the union of a finite number of those open sets.

6. A set S is *convex* means for any $\mathbf{x}_1$ and $\mathbf{x}_2$ in S, the set $\{t\mathbf{x}_1 + (1-t)\mathbf{x}_2 : 0 \le t \le 1\}$ is contained in S.

Compactness is usually defined more succinctly in terms of open covers. A *cover* of a set is a collection of sets containing the set in their union. A *subcover* of a cover is a subcollection of the sets in a cover that is also a cover. An *open cover* is a cover consisting of open sets. With these definitions compactness can be described as follows. A set is compact, if every open cover of the set contains a finite subcover.

Definition 7.1.2 *Points*
For a set $S \subset R^n$, the following kinds of points are defined.

1. $\mathbf{x}$ is an *interior point* of S means there is a open ball centered at $\mathbf{x}$ entirely contained in S.

2. $\mathbf{x}$ is a *boundary point* of S means every open ball centered at $\mathbf{x}$ contains a point in S and a point not in S.

3. $\mathbf{x}$ is an *exterior point* of S means there is a open ball centered at $\mathbf{x}$ containing no points of S.

4. $\mathbf{x}$ is an *isolated point* of S means there is a open ball centered at $\mathbf{x}$ whose only point in S is $\mathbf{x}$.

An interior point is always in the set, an exterior point is never in the set and a boundary point may or may not be in the set. An open set contains only interior points and a closed set contains all of its boundary points. A boundary point is either an isolated point in S or a limit point of S and its compliment.

Some important properties and relationships are discussed in the exercises. One of the most famous and useful is the

Theorem 7.1.4 (Heine-Borel) *A set is compact if and only if it is closed and bounded.*

7.2 Continuous Functions

The idea of an open set can be extended to being open relative to a given set. In particular, for $S \subset R^n$, a set is *open in S* means it is the intersection of an open set with S. More technically, $\mathcal{U}$ is open in S means for any $\mathbf{x}_0$ in S there is an $\varepsilon > 0$, so that if $\mathbf{x} \in S$ and $|\mathbf{x} - \mathbf{x}_0| < \varepsilon$, then $\mathbf{x} \in \mathcal{U}$

Theorem 7.2.1 (Global Continuity) *A function $f : R^n \to R^m$ defined on $S \subset R^n$ is continuous on S if and only if the inverse image of an open set in R^m is open in S.*

Theorem 7.2.2 ("Intermediate Value Theorem") *The image of a connected set under a continuous function is connected.*

Theorem 7.2.3 (Maxima and Minima) *If $f : R^n \to R$ is continuous on a compact set C, then f attains a maximum and minimum value on C.*

Theorem 7.2.4 (Uniform Continuity) *If $f : R^n \to R^m$ is continuous on a compact set C, then f is uniformly continuous on C.*

Theorem 7.2.5 (Mean Value Theorem) *If $f : R^n \to R$ is differentiable on a convex set S, then for any $\mathbf{x}$ and $\mathbf{x}_0$ in S, there is a t strictly between 0 and 1, so that for $\mathbf{x}_t = t\mathbf{x} + (1-t)\mathbf{x}_0$,*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_t)\,(\mathbf{x} - \mathbf{x}_0)$$

Note that the Max-Min Theorem and the Mean Value Theorem apply only to real-valued functions.

7.3 Exercises

7–1. Prove Theorem 8.1.3.

7–2. Prove the following.

(a) The union of any collection of open sets is open.
(b) The intersection of a finite number of open sets is open.
(c) The compliment of a closed set is open.
(d) The complement of an open set is closed.
(e) The intersection of any collection of closed sets is closed.
(f) The union of a finite number of closed sets is closed.
(g) The empty set and R^n are both open and closed.
(h) A set is closed if and only if the limit of any convergent sequence in the set is also in the set.

7–3. Prove the following.

(a) The intersection of two connected sets may not be connected.
(b) The union of connected sets may not be connected.
(c) The union of a finite number of compact sets is compact.
(d) The intersection of compact sets is compact.
(e) The intersection of convex sets is convex.
(f) The union of convex sets is not necessarily convex.

7–4. Show that a set is sequentially compact if and only if it is closed and bounded.

7–5. The *diameter* of a set S is defined to be $\mathrm{diam}(S) = \sup\{|\mathbf{x}-\mathbf{y}| : \mathbf{x}, \mathbf{y} \in S\}$, if the sup exists and ∞, otherwise. Show the following.

(a) Show that the diameter of a bounded set is finite.
(b) (A Cantor Intersection Theorem): If $\{C_k\}$ is a nested sequence of non-empty, compact sets with $\lim_{k\to\infty} \mathrm{diam}(C_k) = 0$, then $\cap_{k=1}^{\infty} C_k = \{\mathbf{x}\}$ for some $\mathbf{x} \in R^n$.

7–6. Show that if $f : R^n \to R^m$ is continuous on a compact set C, then $f(C)$ is compact (i.e. the continuous image of a compact set is compact).

7–7. Is Rolle's Theorem true in R^n for $n > 1$? Justify your answer.

7–8. A function $f : R^n \to R^m$ satisfies a *Lipschitz condition* on a set $S \subset R^n$ means there is a constant $L > 0$, so that for all $\mathbf{x}, \mathbf{y} \in S$,

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq L\,|\mathbf{x} - \mathbf{y}|.$$

Prove the following.

(a) If f satisfies a Lipschitz condition on S, then f is uniformly continuous on S.

(b) If f is differentiable on a convex set S and f' is bounded on S, then f is uniformly continuous on S.

Chapter 8

Sequences and Series

8.1 Sequences of Functions

Definition 8.1.1 *Convergence*

1. A sequence of functions $f_k : R^n \to R^m$, for $k = 1, 2, \ldots$, *converges pointwise* to $f : R^n \to R^m$ on $S \subset R^n$ means $S \subset \text{Dom } f_k$, $S \subset \text{Dom } f$ and, for all $\mathbf{x} \in S$, $\lim_{k\to\infty} f_k(\mathbf{x}) = f(\mathbf{x})$. In other words, For each $\mathbf{x} \in S$ and any $\varepsilon > 0$, there is a natural number K, so that
$$\text{if } k \geq K, \text{ then } |f_k(\mathbf{x}) - f(\mathbf{x})| < \varepsilon.$$

2. A sequence of functions $f_k : R^n \to R^m$, for $k = 1, 2, \ldots$, *converges uniformly* to $f : R^n \to R^m$ on $S \subset R^n$ means $S \subset \text{Dom } f_k$, $S \subset \text{Dom } f$ and for any $\varepsilon > 0$, there is a natural number K, so that
$$\text{if } k \geq K, \text{ then } |f_k(\mathbf{x}) - f(\mathbf{x})| < \varepsilon, \text{ for all } \mathbf{x} \in S$$

Clearly, if a sequence of functions converges uniformly it converges pointwise. The converse is not true.

Theorem 8.1.1 *If, for a sequence of functions, $f_k : R^n \to R^m$, $k = 1, 2, \ldots$, each f_k is continuous on $S \subset R^n$ and the sequence converges uniformly to f on S, then f is continuous on S.*

Theorem 8.1.2 *Let $\{f_k\}$ be a sequence of functions from R to R defined on a bounded interval J. Suppose that there is a point x_0 in J at which the sequence $\{f_k(x_0)\}$ converges, f_k is differentiable on J for each k, and the sequence $\{f_k'\}$*

converges uniformly on J to g. Then the sequence $\{f_k\}$ converges uniformly on J to a function f and $f' = g$ on J.

Both of these theorems deal with interchanging the order of two limits. As can be seen this operation is not as simple as one might hope.

The set of functions

$$\mathcal{F}(S) = \{f : R^n \to R^m : f \text{ is defined on } S\}$$

is a vector space. A "norm" can be defined on this vector space by

$$\|f\| = \sup\{|f(\mathbf{x})| : \mathbf{x} \in S\} = \sup_{\mathbf{x}\in S} |f(\mathbf{x})|$$

This "norm" satisfies the usual norm properties, i.e.

$$\begin{aligned} \|f\| &\geq 0, \text{ and } \|f\| = 0 \text{ if and only if } f = \mathbf{0} \\ \|rf\| &= |r|\,\|f\| \text{ for } r \in R \\ \|f+g\| &\leq \|f\| + \|g\| \end{aligned}$$

The only problem is that $\|f\|$ does not exist unless f is bounded on S. If f is not bounded, then $\|f\|$ is said to be $\|f\| = \infty$. On the other hand, for many important subspaces, e.g. the subspace of bounded functions, $\|\cdot\|$ is a norm and can be used to define a topology on the subspace. That is, $\|\cdot\|$ can be used to define open sets, convergent sequences, and continuous functions from one function space to another, just as the Euclidean norm does on R^n. The norm is called the L_∞ (c.f. Exercise 2-2.) or *sup-norm* and is precisely the norm that defines uniform convergence on S. That is, $\|f_k - f\| \to 0$ is equivalent to $\{f_k\}$ converges to f uniformly on S.

This norm is particularly useful when S is compact. In this case, the norm of any continuous function is finite and so the subspaces $C^k(S)$, for $k = 0, 1, 2, \ldots$, are normed linear spaces.

8.2 Series

8.2.1 Series in R^n

Series in R^n are just "infinite" sums of vectors in the same sense that series in elementary calculus are infinite sums of numbers. The basic concepts and results are the same, since sums of vectors are computed component-wise.

Definition 8.2.1 *Sum of a series*
If $\{\mathbf{x}_k\}_{k=1}^{\infty}$ is a sequence of vectors in R^n, then the *series* associated to the sequence is

$$\sum_{k=1}^{\infty} \mathbf{x}_k = \lim_{K\to\infty} \sum_{k=1}^{K} \mathbf{x}_k$$

$s_K = \sum_{k=1}^{K} \mathbf{x}_k$ is called the K–th *partial sum* of the series. The series is said to *converge* or be *summable*, if the limit exists, and to *diverge* if the limit does not exist. If the series converges, the limit is called the *sum* of the series.

Theorem 8.2.1 (Cauchy Criterion) *$\sum_k \mathbf{x}_k$ converges, if and only if, for every $\varepsilon > 0$, there is an M so that*

$$\textit{if } L \geq K \geq M, \textit{then } |\sum_{k=K}^{L} \mathbf{x}_k| < \varepsilon$$

Theorem 8.2.2 *If $\sum_k \mathbf{x}_k$ converges, then $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{0}$.*

Most of the usual limit theorems apply to series. In particular,

Theorem 8.2.3

1. *$\sum_k \mathbf{x}_k$ converges if and only if $\sum_k x_{ik}$ converges for each $i = 1, \ldots, n$, and $\sum_k \mathbf{x}_k = (\sum_k x_{1k}, \ldots, \sum_k x_{nk})$.*
2. *If $\sum_k \mathbf{x}_k$ and $\sum_k \mathbf{y}_k$ converge, then $\sum_k \mathbf{x}_k + \mathbf{y}_k$ converges, and $\sum_k \mathbf{x}_k + \mathbf{y}_k = \sum_k \mathbf{x}_k + \sum_k \mathbf{y}_k$.*
3. *If $\sum_k \mathbf{x}_k$ converges, then $\sum_k c\mathbf{x}_k$ converges for any $c \in R$, and $\sum_k c\mathbf{x}_k = c\sum_k \mathbf{x}_k$.*

Definition 8.2.2 *Absolute and conditional convergence*

1. $\sum_k \mathbf{x}_k$ converges *absolutely* means $\sum_k |\mathbf{x}_k|$ converges.
2. A series converges *conditionally* means the series converges, but does not converge absolutely.

Theorem 8.2.4 *If a series converges absolutely, then it converges.*

Theorem 8.2.5 *If $\sum_k \mathbf{x}_k$ converges absolutely, then any rearrangement of the series converges to the same value.*

$\boxed{\text{Tests for Absolute convergence}}$

1. Comparison Test. If $\{x_k\}$ and $\{y_k\}$ are sequences of *positive real numbers*, and there is a natural number K, so that

$$x_k \leq y_k \text{ for } k \geq K,$$

then

   (a) if $\sum_k y_k$ converges, then $\sum_k x_k$ converges,

   (b) if $\sum_k x_k$ diverges, then $\sum_k y_k$ diverges.

2. Limit Comparison Test. If $\{x_k\}$ and $\{y_k\}$ are sequences of *positive real numbers* and $\lim_k x_k/y_k = L$,

   (a) if $L \neq 0$, then $\sum_k y_k$ converges if and only if $\sum_k x_k$ converges,

   (b) if $L = 0$ and $\sum_k y_k$ converges, then $\sum_k x_k$ converges.

3. Integral Test. If $f : R \to R$ is a positive, decreasing, continuous function on $[M, \infty)$, for some natural number M, then

$$\sum_k f(k) \text{ converges if and only if } \int_M^\infty f(t)dt \text{ converges.}$$

4. Root Test. If $\{\mathbf{x}_k\}$ is a sequence in R^n and

   (a) there is a positive number $r < 1$ and a natural number K so that

$$\text{for } k \geq K, \quad |\mathbf{x}_k|^{1/k} \leq r$$

   then $\sum_k \mathbf{x}_k$ converges absolutely.

   (b) there is a positive number $r > 1$ and a natural number K so that

$$\text{for } k \geq K, \quad |\mathbf{x}_k|^{1/k} \geq r$$

   then $\sum_k \mathbf{x}_k$ diverges.

5. Ratio Test. If $\{\mathbf{x}_k\}$ is a sequence of non-zero vectors in R^n and

   (a) there is a positive number $r < 1$ and a natural number K so that

$$\text{for } k \geq K, \quad \frac{|\mathbf{x}_{k+1}|}{|\mathbf{x}_k|} \leq r$$

   then $\sum_k \mathbf{x}_k$ converges absolutely.

(b) there is a positive number $r > 1$ and a natural number K so that

$$\text{for } k \geq K, \quad \frac{|\mathbf{x}_{k+1}|}{|\mathbf{x}_k|} \geq r$$

then $\sum_k \mathbf{x}_k$ diverges.

$\boxed{\text{Alternating Series Test}}$

If $\{x_k\}$ is a decreasing sequence of *non-negative reals* and $\lim_{k\to\infty} x_k = 0$, then $\sum_k (-1)^k x_k$ converges.

8.2.2 Series of Functions

The notions of series can be extended to summing sequences of functions. In particular, for sequence of functions $f_k : R^n \to R^m$, for $k = 1, 2, \ldots$ defined on $S \subset R^n$, we have

$$\sum_{k=1}^{\infty} f_k = \lim_{K\to\infty} \sum_{k=1}^{K} f_k$$

where the convergence can be pointwise or uniform depending on the context. The major theorems involving sequences of functions and summing series have analogs for series of functions. For example,

Theorem 8.2.6 *If, for a sequence of functions, $f_k : R^n \to R^m$, $k = 1, 2, \ldots$, each f_k is continuous on $S \subset R^n$ and $\sum_{k=1}^{\infty} f_k$ converges uniformly to f on S, then f is continuous on S.*

Theorem 8.2.7 (Cauchy Criterion) *For a sequence of functions, $f_k : R^n \to R^m$, $k = 1, 2, \ldots$, defined on $S \subset R^n$, $\sum_{k=1}^{\infty} f_k$ converges uniformly to f on S if and only if, for $\varepsilon > 0$ there is a natural number M, so that if $L \geq K \geq M$, then $\|\sum_{k=K}^{L} f_k\| < \varepsilon$.*

There are tests for uniform convergence that are specific to series. One of the more useful is

Theorem 8.2.8 (Weierstrass M-Test) *Suppose for a sequence of functions, $f_k : R^n \to R^m$, $k = 1, 2, \ldots$, defined on $S \subset R^n$, $\{M_k\}$ is a sequence of non-negative real numbers such that $\|f_k\| \leq M_k$ for each $k \in N$. If $\sum_k M_k$ converges, then $\sum_k f_k$ converges uniformly on S.*

Two of the most important families of series of functions are

1. Power series: $\sum_{k=0}^{\infty} a_k(x-c)^k$
2. Fourier series: $\sum_{k=0}^{\infty} a_k \sin(kx) + b_k \cos(kx)$.

We examine in detail the important facts concerning power series.

Theorem 8.2.9 *For a power series $\sum_{k=0}^{\infty} a_k(x-c)^k$ there is a number R so that the series converges absolutely for $|x-c| < R$, diverges for $|x-c| > R$, and converges uniformly on any compact subset of $(c-R, c+R)$.*

The number R in the previous theorem is called the *radius of convergence* of the power series and the interval $(c-R, c+R)$ is called the *interval of convergence*. The series may or may not converge at the end points of the interval of convergence.

Theorem 8.2.10 *The limit of a power series is continuous on the interval of convergence, and can be integrated and differentiated term-by-term on the interval of convergence.*

Corollary 8.2.11 *If $f(x) = \sum_k a_k(x-c)^k$ on an open interval, J, containing c, then f is $C^\infty(J)$ and*

$$a_k = \frac{f^{(k)}(c)}{k!}$$

for all k.

Definition 8.2.3 *Analytic function* A function is *analytic* on a open set S means, for each $x_0 \in S$, there is an $R > 0$, $c \in S$, and $\{a_k\}_{k=0}^{\infty}$ so that $|x_0 - c| < R$ and $f(x) = \sum_k a_k(x-c)^k$ for $|x-c| < R$

The set of analytic functions on a given set S is a vector space, and is ,in fact, a proper subspace of $C^\infty(S)$.

The facts about power series and analytic functions on R can easily be generalized to power series on R^n. The form of a power series is just the limit of multivariable polynomials, i.e.

$$\sum_{k=0}^{\infty} \sum_{k_1+\ldots+k_n=k} a_{k_1k_2\ldots k_n}(x_1 - x_{01})^{k_1} \ldots (x_n - x_{0n})^{k_n}$$

8.3 Exercises

8–1. Consider the sequences $\{f_k\}$ defined on $S = \{x \in R : x \geq 0\}$ given by

(a) $\dfrac{x^k}{k}$

(b) $\dfrac{x^k}{1+x^k}$

(c) $\dfrac{x^k}{k+x^k}$

(d) $\dfrac{x^{2k}}{1+x^k}$

Discuss the convergence and uniform convergence of these sequences and the continuity of the limit functions.

8–2. Suppose $\{f_k\}$ and $\{g_k\}$ are sequences of real-valued functions that converge uniformly on $S \subset R^n$ to f and g, respectively, and g is never 0 on S. Prove or give a counterexample.

(a) $\{f_k + g_k\}$ converges uniformly to $f + g$ on S.

(b) $\{f_k g_k\}$ converges uniformly to fg on S.

(c) $\{f_k/g_k\}$ converges uniformly to f/g on S.

8–3. For the functions $f_k : R \to R$, for $k \in N$ defined by

$$f_k(x) = \begin{cases} kx & 0 \leq x \leq 1/k \\ 1/kx & 1/k < x \end{cases}$$

Show that $\{f_k\}$ converges pointwise to 0 on $S = \{x \in R : x \geq 0\}$. Is the convergence uniform on S?

8–4. A sequence of functions, $\{f_k\}$, is *Cauchy* in the sup-norm on $S \subset R^n$ means for $\varepsilon > 0$ there is a $K \in N$, so that if $k, l \geq K$, then $\|f_k - f_l\| < \varepsilon$.

(a) Show that if a sequence of functions is Cauchy in the sup-norm on S, then the sequence converges uniformly on S.

(b) Show that $C^0(S)$ is complete.

8–5. Suppose $f_k(x) = x^{2k+1}/k$ for $k = 1, 2, \ldots$ and $S = [-1, 1]$. Show that the sequence converges uniformly on S to a function f, each f_k is differentiable on S and $\{f_k'\}$ converges pointwise to a function g on S, but $f' \neq g$.

8–6. Show that if $\{x_k\}$ is a sequence of non-negative reals, then $\sum_k x_k$ converges if and only if the partial sums of the sequence are bounded above.

8–7. Show that grouping the terms of a convergent series by introducing parentheses containing a finite number of terms does not destroy convergence or change the value of the limit. Show that grouping terms in a divergent series can produce convergence.

8–8. Show that if a series converges conditionally, then the series of positive terms and the series of negative terms both diverge.

8–9. If $\sum_k x_k$ converges, does $\sum_k x_k^2$?

8–10. Prove the usual version of the ratio test: If $\lim_{k\to\infty} |\mathbf{x}_{k+1}|/|\mathbf{x}_k| = r$, then

- if $r < 1$, then the series converges absolutely,
- if $r > 1$, then the series diverges,
- if $r = 1$, then anything can happen.

8–11. Discuss the convergence and uniform convergence of the series $\sum_k f_k$, where $f_k : R \to R$ is given by

(a) $1/(x^2 + k^2)$

(b) $1/(kx^2)$

(c) $\sin(x/k^2)$

(d) $1/(x^k + 1)$, for $x \geq 0$

8–12. Show that the radius of convergence of a power series $\sum a_k(x-c)^k$ is given by $\lim |a_k|/|a_{k+1}|$ whenever the limit exists.

8–13. Taylor series.

(a) Prove that if $f : R \to R$ is defined for $|x - c| < r$ and there is a constant b so that $|f^{(k)}(x)| \leq b$, for $|x - c| < r$ and any $k = 0, 1, 2, \ldots$, then for $|x - c| < r$

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

(b) Show that

i. $\ln x = \sum_{k=1}^{\infty} (-1)^{k+1}(x-1)^k/k$ for $|x - 1| < 1/2$.

ii. $e^x = \sum_{k=0}^{\infty} x^k/k!$ for all x.

8–14. Show that

(a) $f(x) = e^x$ is analytic on R.

(b) $f(x) = \ln x$ is analytic on $(0, \infty)$

Chapter 9

Riemann Integrals

9.1 Integrals in R

Suppose that $f : R \to R$ is defined on an interval $[a, b]$. There are two approaches to defining the integral of such a function, both of which use the idea of a partition of the interval $[a, b]$.

Definition 9.1.1 *Partition of an interval*

1. A *partition* of an interval $[a, b]$, is a set $\mathcal{P} = \{x_0, \ldots, x_L\}$ satisfying

$$x_0 = a \leq x_1 \leq \ldots \leq x_{l-1} \leq x_l \leq \ldots \leq x_L = b.$$

2. Letting $\Delta x_l = (x_l - x_{l-1})$, the *mesh* of a partition $\mathcal{P}$ is

$$\|\mathcal{P}\| = \max_{l=1,\ldots,L} \Delta x_l$$

3. A *refinement* of a partition $\mathcal{P}$ is a partition $\mathcal{P}'$, so that $\mathcal{P} \subset \mathcal{P}'$.

The integral is defined in terms of limits over partitions that are "finer and finer"

Theorem 9.1.1 *The following are the basic properties of partitions of a given interval $[a, b]$.*

1. *If $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then $\|\mathcal{P}'\| \leq \|\mathcal{P}\|$.*
2. *$\mathcal{P}$ and $\mathcal{Q}$ are partitions, then there is a partition $\mathcal{P}'$ that is a refinement of both of them.*

9.1.1 Riemann Integral

Definition 9.1.2 *Riemann sum*
For a function $f : R \to R$ is defined on an interval $[a, b]$, a partition $\mathcal{P} = \{x_0, \ldots, x_L\}$ of $[a, b]$, and $z_1, \ldots, z_l$, so that $z_l \in [x_{l-1}, x_l]$, for $l = 1, \ldots, L$, the *Riemann sum* is

$$R(f, \mathcal{P}) = \sum_{l=1}^{L} f(z_l) \, \Delta x_l$$

Definition 9.1.3 *Riemann integral*
A function $f : R \to R$ is defined and bounded on an interval $[a, b]$ is *Riemann integrable* on $[a, b]$, means there is a real number I, so that for any $\varepsilon > 0$, there is a partition, $\mathcal{P}_\varepsilon$ of $[a, b]$, so that if $\mathcal{P}$ is any refinement of $\mathcal{P}_\varepsilon$ and $R(f, \mathcal{P})$ is any Riemann sum, then

$$|R(f, \mathcal{P}) - I| < \varepsilon.$$

The number I, if it exists, is called the *Riemann integral* of f on $[a, b]$, and is denoted

$$I = \int_a^b f(x) \, dx.$$

Riemann integrability is often defined slightly differently, i.e. there is a number I so that, for any $\varepsilon > 0$, there is a $\delta > 0$, so that if $\mathcal{P}$ is any partition of $[a, b]$ with $\|\mathcal{P}\| < \delta$ and $R(f, \mathcal{P})$ is any Riemann sum, then $|R(f, \mathcal{P}) - I| < \varepsilon$. These two definitions are equivalent, but refinements are easier to work with, so we will adopt the former as the official definition of Riemann integrability.

Theorem 9.1.2 (Cauchy Criterion) *f is integrable on $[a, b]$ if and only if for any $\varepsilon > 0$, there is a partition, $\mathcal{P}_\varepsilon$ of $[a, b]$, so that if $\mathcal{P}$ and $\mathcal{Q}$ are any refinements of $\mathcal{P}_\varepsilon$, then*

$$|R(f, \mathcal{P}) - R(f, \mathcal{Q})| < \varepsilon.$$

for any Riemann sums $R(f, \mathcal{P})$ and $R(f, \mathcal{Q})$.

9.1.2 Riemann-Darboux Integral

For a function $f : R \to R$ is defined and bounded on an interval $[a, b]$ and a partition $\mathcal{P} = \{x_0, \ldots, x_L\}$ of $[a, b]$,

1. The *upper sum* for f with respect to $\mathcal{P}$ is

$$S(f,\mathcal{P}) = \sum_{l=1}^{L} M_l \, \Delta x_l,$$

   where $M_l = \sup_{[x_{l-1},x_l]} f(x)$, for $l = 1, \ldots, L$.

2. The *lower sum* for f with respect to $\mathcal{P}$ is

$$s(f,\mathcal{P}) = \sum_{l=1}^{L} m_l \, \Delta x_l,$$

   where $m_l = \inf_{[x_{l-1},x_l]} f(x)$, for $l = 1, \ldots, L$.

Definition 9.1.4 *Riemann-Darboux integral*
For a function $f : R \to R$ is defined on an interval $[a,b]$

1. The *upper integral* of f on $[a,b]$ is

$$\overline{\int_a^b} f(x)\,dx = \inf_{\mathcal{P}} S(f,\mathcal{P}).$$

2. The *lower integral* of f on $[a,b]$ is

$$\underline{\int_a^b} f(x)\,dx = \sup_{\mathcal{P}} s(f,\mathcal{P}).$$

3. f is *Riemann-Darboux integrable* on $[a,b]$, means the upper and lower integrals are equal.

Theorem 9.1.3 *For a function $f : R \to R$ is defined and bounded on an interval $[a,b]$, $\overline{\int_a^b} f(x)\,dx$ and $\underline{\int_a^b} f(x)\,dx$ exist, and for $\mathcal{P}$ and $\mathcal{P}'$ partitions of $[a,b]$, with $\mathcal{P} \subset \mathcal{P}'$*

$$s(f,\mathcal{P}) \le s(f,\mathcal{P}') \le \underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx \le S(f,\mathcal{P}') \le S(f,\mathcal{P})$$

Theorem 9.1.4 *A function $f : R \to R$ defined on an interval $[a,b]$ is Riemann integrable if and only if it is Riemann-Darboux integrable. Moreover,*

$$\int_a^b f(x)\,dx = \overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx$$

9.2 Integrals in R^n

The analog in R^n of an interval in R is a closed rectangle as given in Definition 8.1.1.1, i.e. for $\mathbf{a} \le \mathbf{b}$ the set $R = \{\mathbf{x} : \mathbf{a} \le \mathbf{x} \le \mathbf{b}\}$. A partition of a rectangle is obtained from partitions of each of the intervals $[a_i, b_i]$. In particular, if $\mathcal{P}_i = \{x_{i0}, \dots, x_{iL_i}\}$ is a partition of $[a_i, b_i]$, then the corresponding partition $\mathcal{P}$ of the rectangle R consists of the rectangles $R_{l_1,\dots,l_n} = \{\mathbf{x} : x_{i(l_i-1)} \le x_i \le x_{il_i}\}$, for $i = 1, \dots, n$ and $l_i = 1, \dots, L_i$. A refinement of $\mathcal{P}$ is obtained from partitions built from refinements of each of the $\mathcal{P}_i$'s.

Riemann, upper and lower sums are defined analogously, i.e. for $f : R^n \to R$ defined and bounded on a rectangle R,

1. Riemann sum: For $z_{l_1,\dots,l_n} \in R_{l_1,\dots,l_n}$
$$R(f, \mathcal{P}) = \sum_{l_1,\dots,l_n} f(z_{l_1,\dots,l_n})\, \Delta x_{l_1} \dots \Delta x_{l_n}$$

2. Upper sum: For $M_{l_1,\dots,l_n} \ge f(\mathbf{x})$ on $R_{l_1,\dots,l_n}$
$$S(f, \mathcal{P}) = \sum_{l_1,\dots,l_n} M_{l_1,\dots,l_n}\, \Delta x_{l_1} \dots \Delta x_{l_n}$$

3. Lower sum: For $m_{l_1,\dots,l_n} \le f(\mathbf{x})$ on $R_{l_1,\dots,l_n}$
$$s(f, \mathcal{P}) = \sum_{l_1,\dots,l_n} m_{l_1,\dots,l_n}\, \Delta x_{l_1} \dots \Delta x_{l_n}$$

The definitions for Riemann and Riemann-Darboux integrability are the same as for one variable and are again equivalent. The notation for the integral is

$$\int_R f(\mathbf{x})\, dx_1 \dots dx_n.$$

Yet another kind of integral can be defined for $n > 1$, an iterated integral. This integral is based on integrating in the coordinates one at a time, defining a function of one fewer variables, ie.

$$F(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = \int_{a_i}^{b_i} f(x_1, \dots, x_i, \dots, x_n)\, dx_i$$

Definition 9.2.1 *Iterated integrals*
For $f : R^n \to R$ defined and bounded on a rectangle $R = \{\mathbf{x} : \mathbf{a} \le \mathbf{x} \le \mathbf{b}\}$, the *iterated integral* of f over R in the order $i_1, i_2, \dots, i_n$,

$$\int_{a_{i_n}}^{b_{i_n}} \dots \int_{a_{i_1}}^{b_{i_1}} f(\mathbf{x})\, dx_{i_1} \dots dx_{i_n} = \int_{a_{i_n}}^{b_{i_n}} \left(\dots \int_{a_{i_2}}^{b_{i_2}} \left(\int_{a_{i_1}}^{b_{i_1}} f(\mathbf{x})\, dx_{i_1} \right) dx_{i_2} \right) \dots dx_{i_n}$$

Definition 9.2.2 *Integration over bounded sets*
For $f : R^n \to R$ defined and bounded on a bounded set S, let R be any rectangle containing S and $f_S(\mathbf{x}) = f(\mathbf{x})$, for $\mathbf{x} \in S$ and $f_S(\mathbf{x}) = 0$, otherwise, then the *integral of f over S* is defined to be

$$\int_S f(\mathbf{x})\, dx_1 \ldots dx_n = \int_R f_S(\mathbf{x})\, dx_1 \ldots dx_n.$$

9.3 Exercises

9–1. Show that

$$f(x) = \begin{cases} 2 & 0 \le x \le 1 \\ 3 & 1 < x \le 2 \end{cases}$$

is Riemann integrable on $[0, 2]$ using the definition. Do it again using the Cauchy criterion. Do it again using the definition of Riemann-Darboux integrable. Do it again using the definition of Riemann integral based on the mesh of a partition. Don't do it any more.

9–2. Show that $f : R^n \to R$ is Riemann-Darboux integrable on a rectangle R if and only for any $\varepsilon > 0$ there is a partition $\mathcal{P}_\varepsilon$ so that $S(f, \mathcal{P}_\varepsilon) - s(f, \mathcal{P}_\varepsilon) < \varepsilon$.

9–3. Show that a function must be bounded to be integrable by proving the following theorem. If $f : R \to R$ is defined but not bounded above on $[a, b]$, then for any partition $\mathcal{P}$, there is a sequence of Riemann sums, $\{R_N(f, \mathcal{P})\}$, so that $\lim_{N\to\infty} R_N(f, \mathcal{P}) = \infty$.

9–4. What question should one ask about the definition of the integral of a function over a bounded set (Definition 10.2.2)?

9–5. On the rectangle $0 \le x \le 1$ and $0 \le y \le 1$, let $f(x, y) = 1$, if x is rational and $f(x, y) = 2y$, if x is irrational. Show that

$$\int_0^1 (\int_0^1 f(x, y) dy) dx = 1,$$

but f is not integrable on the rectangle.

Chapter 10

Integration Theorems and Improper Integrals

10.1 Integrability and the Basic Properties of the Integral

Definition 10.1.1 *Content zero*
A set $S \subset R^n$ has *content zero* means for any $\varepsilon > 0$, there are a finite number of closed rectangles $R_1, \ldots, R_N$, so that

1. $S \subset R_1 \cup \ldots \cup R_N$ and
2. $\sum_{i=1}^{N} \mathrm{vol}(R_i) < \varepsilon$, where $\mathrm{vol}(R_i)$ is the volume of R_i.

Perhaps, the most important example of a set in R^n, with content zero is a *smooth set*. A set S is *smooth* if it is the image of a compact set in R^m, with $m < n$, under a continuously differentiable function. In particular, a smooth set is a compact, lower dimensional set in R^n. For example, a smooth curve in R^2 is the image of an interval under a continuously differentiable function.

Theorem 10.1.1 (Integrability on Rectangles) *If R is a closed, bounded rectangle and f is continuous except on a set with content zero and bounded on R, then f is integrable on R.*

For a set $S \subset R^n$, the set of all boundary points of S will be denoted ∂S. A set is called *Jordan-measurable* if ∂S has content zero. For example, any set whose boundary is made up of a finite number of smooth sets is Jordan-measurable.

Theorem 10.1.2 (Integrability on Bounded Subsets) *If $f : R^n \to R$ is defined and bounded on a bounded, Jordan-measurable set S and f is continuous on S except on a set with content zero, then f is integrable on S.*

This theorem still does not completely characterized the relationship between integrability and continuity. Such a characterization utilizes the notion of the Lebesgue measure of a set, which is not all that complicated, but requires more time than we have, and is covered in the next course in this subject, Real Analysis. But, for the record, a function is Riemann integrable on a bounded set if, and only if, it is continuous except on a set with Lebesgue measure zero.

Theorem 10.1.3

1. *If R is a closed, bounded rectangle, then $\int_R 1 = \mathrm{vol}(R)$.*
2. *If T is a bounded set and $S \subset T$, then $\int_S f$ exists if, and only if, $\int_T f_S$ exists and they are equal.*
3. *Linearity: If f and g are integrable on S, then for any real numbers a and b, $af + bg$ is integrable on S and*
$$\int_S af + bg = a\int_S f + b\int_S g.$$
4. *Positivity: If f is integrable and $f \geq 0$ on S, then $\int_S f \geq 0$.*
5. *Additivity: If f is integrable on disjoint sets S_1 and S_2, then f is integrable on their union and*
$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f.$$

Theorem 10.1.4 *Suppose I is a function assigning bounded sets and certain real-valued functions to real numbers, $I(S, f)$, satisfying*

1. *If R is a closed, bounded rectangle, then $I(R, 1) = \mathrm{vol}(R)$.*
2. *If T is a bounded set and $S \subset T$, then $I(S, f)$ is defined if, and only if, $I(T, f_S)$ is defined and they are equal.*
3. *If $I(S, f)$ and $I(S, g)$ are defined, then for any real numbers a and b, $I(S, af + bg)$ is defined and*
$$I(S, af + bg) = aI(S, f) + bI(S, g).$$
4. *If $I(S, f)$ is defined and $f \geq 0$ on S, then $I(S, f) \geq 0$.*

Then, if $I(S, f)$ and $\int_S f$ both exist, they are equal.

Theorem 10.1.5 *If f is integrable on S and any iterated integral exists, then the two are equal.*

Theorem 10.1.6 (Interchanging limits with integrals)

1. *If $\{f_k\}$ converges uniformly to f on S and each f_k is integrable on S, then f is integrable on S and*
$$\lim_{k\to\infty}\int_S f_k = \int_S f.$$

2. *If $f : R^{n+m} \to R$ is continuous on $\{(\mathbf{x},\mathbf{y}) : \mathbf{x}\in T \text{ and } \mathbf{y}\in S\}$, where T and S are compact rectangles in R^n and R^m, respectively, then the function*
$$F(\mathbf{y}) = \int_T f(\mathbf{x},\mathbf{y})dx_1\ldots dx_n$$
is continuous on S.

3. *Suppose $f : R^{n+m} \to R$ is defined on $V = \{(\mathbf{x},\mathbf{y}) : \mathbf{x}\in T \text{ and } \mathbf{y}\in S\}$, where T and S are compact rectangles in R^n and R^m, respectively, and the function $F(\mathbf{y}) = \int_T f(\mathbf{x},\mathbf{y})dx_1\ldots dx_n$ exists for each $\mathbf{y}\in S$. If f_{y_j} is continuous on V, then F_{y_j} exists on S and*
$$F_{y_j}(\mathbf{y}) = \int_T f_{y_j}(\mathbf{x},\mathbf{y})dx_1\ldots dx_n$$

Corollary 10.1.7 *If f is continuous on a compact rectangle, then the Riemann integral and all iterated integrals exist and are equal.*

Theorem 10.1.8 (Mean Value Theorem) *If f and g are integrable on a Jordan-measurable set S and $g(\mathbf{x})\ge 0$ on S, then there is a number μ between $\sup_S f(\mathbf{x})$ and $\inf_S f(\mathbf{x})$, so that*
$$\int_S fg = \mu\int_S g.$$
Moreover, if f is continuous and S is compact and connected, then there is a $\mathbf{c}\in S$, so that $\mu = f(\mathbf{c})$, i.e. $\int_S fg = f(\mathbf{c})\int_S g$.

Theorem 10.1.9 (Change of Variables) *If $G : R^n \to R^n$ is continuously differentiable on an open set containing a bounded, Jordan-measurable set S and*

1. *G is one to one on S,*

2. *$\det(G'(\mathbf{x})) = 0$ only on a set with content zero,*

then if $f : R^n \to R$ is bounded and continuous on $G(S)$, then
$$\int_{G(S)} f = \int_S f\circ G|\det(G')|.$$

See *American Mathematical Monthly*: J. Schwartz, "The Formula for Change of Variable in a Multiple Integral," vol. 61, no. 2 (February 1954) and D. E. Varberg, "On Differentiable Transformations in R^n," vol. 73, no. 1, part II, (April 1966).

10.2 Improper Integrals

If S is a set in R^n and $f : R^n \to R$ is defined on S, then a family of sets $\{S_\alpha\}$ *converges* to S, means

1. S_α is a bounded subset of S for each α.
2. If $\alpha < \beta$, then $S_\alpha \subset S_\beta$.
3. $\bigcup_\alpha S_\alpha = S$
4. Every bounded subset of S on which f is bounded is contained in some one of the S_α's.

Theorem 10.2.1 *Suppose f is nonnegative on S and for some particular family of bounded sets $\{S_\alpha\}$ converging to S*

$$\lim_\alpha \int_{S_\alpha} f$$

is finite, then

$$\lim_\alpha \int_{T_\alpha} f = \lim_\alpha \int_{S_\alpha} f$$

for any family of sets $\{T_\alpha\}$ converging to S.

For a function $f : R^n \to R$, functions f^+ and f^- are defined by

1. $f^+(\mathbf{x}) = \max(f(\mathbf{x}), 0) = \frac{1}{2}(|f(\mathbf{x})| + f(\mathbf{x}))$.
2. $f^-(\mathbf{x}) = \max(-f(\mathbf{x}), 0) = \frac{1}{2}(|f(\mathbf{x})| - f(\mathbf{x}))$.

Definition 10.2.1 *Improper integral*

1. If $f : R^n \to R$ is nonnegative, then f is *integrable* on S means for some family of sets converging to S, $\{S_\alpha\}$, f is integrable on each S_α and the $\lim_\alpha \int_{S_\alpha} f$ is finite. The *integral* of f over S is defined to be

$$\int_S f = \lim_\alpha \int_{S_\alpha} f.$$

2. For any $f : R^n \to R$, f is *integrable* on S means f^+ and f^- are integrable on S and the *integral* is defined to be

$$\int_S f = \int_S f^+ - \int_S f^-$$

Theorem 10.2.2 *1. f is integrable on S if, and only if, $|f|$ is integrable on S.*

2. If $|f| \leq g$ and, for $\mathbf{x}_0 \in S$, g is unbounded in any open ball containing $\mathbf{x}_0$ if, and only if, $|f|$ is unbounded in any open ball containing $\mathbf{x}_0$, then if g is integrable on S, then f is integrable on S.

10.3 Exercises

10–1. Show that the following sets have content zero.

(a) The union of a finite number of sets with content zero.

(b) A convergent sequence in R^n.

10–2. Show that following sets are Jordan-measurable.

(a) The unit ball in R^3.

(b) A rectangle in R^n.

10–3. More properties of the integral.

(a) If f is integrable on sets S_1 and S_2, then

$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f - \int_{S_1 \cap S_2} f.$$

(b) Show that if f and g are integrable on S and $f(\mathbf{x}) \geq g(\mathbf{x})$ on S, then $\int_S f \geq \int_S g$.

(c) Show that if f and $|f|$ are integrable on S, then

$$|\int_S f| \leq \int_S |f|.$$

10–4. If $f : R^n \to R^m$, then $\int_S f$ is defined to be

$$\int_S f = (\int_S f_1, \ldots, \int_S f_m).$$

(a) Show that if $f : R^n \to R^m$ and $g : R^n \to R^m$ are integrable over S then for a and b in R,

$$\int_S af + bg = a \int_S f + b \int_S g.$$

(b) If $\mathbf{c} \in R^m$ and $f : R^n \to R^m$ is integrable over S, then

$$\int_S \langle \mathbf{c}, f \rangle = \langle \mathbf{c}, \int_S f \rangle.$$

(c) Show that if $f : R^n \to R^m$ and $|f|$ are integrable on S, then

$$|\int_S f| \leq \int_S |f|.$$

10–5. Compute the following integrals.

(a) For $S = \{(x, y) : 0 < x^2 + y^2 \leq 1\}$,

$$\int_S \frac{1}{\sqrt{x^2 + y^2}} dxdy$$

(b) For $S = \{(x, y) : x^2 + y^2 \geq 1\}$,

$$\int_S \frac{1}{x^2 + y^2} dxdy$$

10–6. Let $G : R^2 \to R^2$ be defined by $G(x, y) = (x^2 - y^2, 2xy)$.

(a) Sketch the image under G of the square with vertices at $(1, 1), (1, \frac{3}{2}), (\frac{3}{2}, 1)$ and $(\frac{3}{2}, \frac{3}{2})$.

(b) Sketch the image of the square in (a) under the best affine approximation to G near $(1, 1)$.

(c) Find the area of the region in (b).

(d) Find the area of the region in (a).

10–7. Compute $\int_S xydxdy$, where $S = \{(x, y) : 1 \leq x^2 - y^2 \leq 2,\ 3 \leq x^2 + y^2 \leq 4,\ x \geq 0,\ y \geq 0\}$.

10–8. For $S = \{\mathbf{x} \in R^n : x_1 + \ldots + x_n \leq 1,\ x_i \geq 0 \text{ for } i = 1, \ldots, n\}$, set up an iterated integral for computing $\int_S f$.

10–9. Compute $\int_{R^2} e^{-\frac{1}{2}(x^2+y^2)} dxdy$. Compute $\int_{R^n} e^{-\frac{1}{2}|\mathbf{x}|^2} dx_1 \ldots dx_n$.

10–10. Find the volume of the unit ball in R^n.

Chapter 11

The Fundamental Theorem of Calculus

11.1 Antiderivatives

11.1.1 The Fundamental Theorem of Calculus in R

Definition 11.1.1 *Extended integral*
For $f : R \to R$ integrable on an interval J and $a, b \in J$,

$$\int_a^b f = \begin{cases} -\int_b^a f & \text{if } a > b \\ 0 & \text{if } a = b \end{cases}$$

Theorem 11.1.1 *For $f : R \to R$ integrable on an interval J,*

1. *For any a, b and $c \in J$,*
$$\int_a^b f = \int_a^c f + \int_c^b f$$

2. *If f is continuous on J, then for any a and $b \in J$, there is a number c between a and b, so that*
$$\int_a^b f = f(c)\,(b-a)$$

Definition 11.1.2 *Antiderivative*
$f : R \to R$ is an *antiderivative* of $F : R \to R$ on $S \subset R$, means $f'(x) = F(x)$ for all $x \in S$.

Theorem 11.1.2 *1. If $f'(x) = 0$ on an interval J, then f is constant on J.*

2. If f and g are antiderivatives of F on an interval J, then there is a constant c so that $g(x) = f(x) + c$ for all $x \in J$.

Theorem 11.1.3 (Fundamental Theorem of Calculus) *If $F : R \to R$ is continuous on an interval J, then*

1. If g is any antiderivative for F on J, then for any a and $b \in J$,

$$\int_a^b F = g(b) - g(a).$$

2. For any $a \in J$, $f : R \to R$, defined by

$$f(x) = \int_a^x F,$$

is an antiderivative for f on J.

Corollary 11.1.4 *If f is continuously differentiable on an interval J, then for any a and b in J,*

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

11.1.2 Antiderivatives in R^n

A *vector field* in R^n is simply a function from R^n to R^n, but the term "vector field" is used for $F : R^n \to R^n$ when $F(\mathbf{x})$ is interpreted as an object with direction and magnitude, rather than as a location, i.e. a point in R^n. The gradient of a function $f : R^n \to R$, is usually called a vector field, since it "points in the direction of maximum increase of f" and its magnitude is the maximum rate of increase.

Definition 11.1.3 *Antiderivative*
For a vector field $F : R^n \to R^n$, a function $f : R^n \to R$ is an *antiderivative* for f on $S \subset R^n$, means

$$\nabla f(\mathbf{x}) = F(\mathbf{x})$$

for all $\mathbf{x} \in S$.

In physics and engineering, a vector field that has an antiderivative is called a *conservative* vector field.

Theorem 11.1.5 *If F is continuously differentiable on an open set S and F has an antiderivative on S, then $F'(\mathbf{x})$ is symmetric for all $\mathbf{x} \in S$.*

Note that if $n = 1$, then any continuous function has an antiderivative, but for $n > 1$, many very simple, very nice vector fields do not have antiderivatives. For example, $F(x, y) = (2y, x)$ does not have an antiderivative on any subset of R^2. Even more disconcerting is the fact that, the converse of this theorem is generally *not* true. In fact, we will see that the existence of an antiderivative is related to the geometry of the set S, as well as the symmetry of the derivative.

11.1.3 Line Integrals and Antiderivatives

Definition 11.1.4 *Piecewise smooth curves*

1. If $c : R \to R^n$ is a continuously differentiable function on a closed bounded interval interval $[a, b] \subset R$ and $c'(t) \neq \mathbf{0}$ for $t \in (a, b)$, then the image of c is called an *(oriented) smooth curve* in R^n, with *initial point* $c(a)$ and *terminal point* $c(b)$.
2. An *(oriented) piecewise smooth curve* is the union of a finite number of smooth curves $\{c_1, \ldots, c_K\}$, so that, for each k, the initial point of c_{k+1} is the terminal point of c_k. The initial point of the piecewise smooth curve is the initial point of c_1 and the terminal point of the curve is the terminal point of c_K.

The term "smooth curve" as defined here appears to refer to a set, but note that the set has an "orientation" given by the initial and terminal points. One should think of a curve as being "drawn" from its initial point to its terminal point by the function parameterizing it. In fact, a oriented smooth curve is sometimes defined as the function c rather than its image, so that the orientation is automatically prescribed by the definition of the function. On the other hand, as will be seen below, our use of oriented smooth curves in this chapter will not depend on the specific function used to parameterize the set, but only on the set and which of its points are the initial point and terminal point, its orientation. To make matters worse, we will, as is typically done, drop the word "oriented" and refer to smooth and piecewise smooth curves, but the reader should not forget that in this chapter a "curve" always means a set with an orientation. Moreover, we will use the same symbol to denote the curve and a function that defines it.

Definition 11.1.5 *Line integral*
For a smooth curve c in R^n and a vector field $F : R^n \to R^n$ continuous on a open set containing c, the *(line) integral of F over c* is

$$\int_c F_1 dx_1 + F_2 dx_2 + \ldots + F_n dx_n = \int_a^b \langle F(c(t)), c'(t) \rangle \, dt.$$

The the *(line) integral* of F over a piecewise smooth curve c is the sum of the integrals of F over the smooth parts of c.

The line integral is sometimes denoted $\int_c F \cdot d\mathbf{x}$, because of the inner product in the definition and an interpretation of integrating the tangential component of F along the curve.

Theorem 11.1.6

1. *The integral of a vector field over an oriented, piecewise smooth curve is independent of the parameterization.*
2. *If the orientation of a piecewise smooth curve is reversed, i.e. the initial point becomes the terminal point and vice versa, then the integral changes sign.*

If c is a piecewise smooth curve, then the the curve given by the same set of points with the orientation reversed is denoted $-c$, and the second part of the previous theorem says that $\int_{-c} F \cdot d\mathbf{x} = -\int_c F \cdot d\mathbf{x}$.

Theorem 11.1.7 *If $f : R^n \to R$ is continuously differentiable on S and c is a piecewise smooth curve in S with initial point $\mathbf{a}$ and terminal point $\mathbf{b}$, then*

$$\int_c \frac{\partial f_1}{\partial x_1} dx_1 + \ldots + \frac{\partial f_n}{\partial x_n} dx_n = f(\mathbf{b}) - f(\mathbf{a}).$$

A *closed* curve is one whose initial and terminal points are the same point.

Theorem 11.1.8 *If $F : R^n \to R^n$ is continuous on a set S, the following are equivalent.*

1. *F has an antiderivative on S.*
2. *If c_1 and c_2 are any piecewise smooth curves in S with the same initial and terminal points, then*
$$\int_{c_1} F_1 dx_1 + \ldots + F_n dx_n = \int_{c_2} F_1 dx_1 + \ldots + F_n dx_n.$$
3. *If c is any piecewise smooth, closed curve in S, then*
$$\int_c F_1 dx_1 + \ldots + F_n dx_n = 0.$$

11.1.4 Green's Theorem and Antiderivatives

A set S in R^2 is said to have a *piecewise smooth boundary*, if the boundary of S consists of a finite number of closed, piecewise smooth curves. The oriented boundary of, denoted ∂S, is "defined" to be the parameterization so that as the boundary is drawn the set S is "on the left."

Theorem 11.1.9 (Green's Theorem) *If $S \subset R^2$ has a piecewise smooth boundary and $F : R^2 \to R^2$ is continuously differentiable on S, then*

$$\int_S \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) = \int_{\partial S} F_1 dx_1 + F_2 dx_2.$$

Corollary 11.1.10 *If every piecewise smooth, closed curve in $S \subset R^2$ is the boundary of a set entirely contained in S, and $F : R^2 \to R^2$ is C^1 and has a symmetric derivative on S, then F has an antiderivative on S.*

11.2 Integrations of Forms on Chains

11.2.1 Differential Forms and Exterior Differentiation

The set of linear functions $l : R^n \to R$ is a vector space. Moreover, for a given linear function l, there is a $1 \times n$ matrix $[\gamma_1 \ldots \gamma_n]$, so that $l(\mathbf{x}) = \sum_i \gamma_i x_i$, for all $\mathbf{x} \in R^n$. This vector space is easily seen to be n-dimensional and in fact the functions $dx_1, \ldots, dx_n$ form a basis, where $dx_i : R^n \to R$ is defined by $dx_i(\mathbf{a}) = a_i$.

For $p = 1, \ldots, n$, we define the set of *alternating, p-linear functions* on R^n, denoted A^p, to be all linear combinations of the following: for $1 \leq i_1, i_2, \ldots, \leq p$, the function $dx_{i_1} \wedge dx_{i_2} \wedge \ldots \wedge dx_{i_p}$ acts on p vectors $\mathbf{a}_1, \ldots, \mathbf{a}_p \in R^n$ by

$$dx_{i_1} \wedge \ldots \wedge dx_{i_p}(\mathbf{a}_1, \ldots, \mathbf{a}_p) = \det \begin{bmatrix} dx_{i_1}(\mathbf{a}_1) & \ldots & dx_{i_1}(\mathbf{a}_p) \\ dx_{i_2}(\mathbf{a}_1) & \ldots & dx_{i_2}(\mathbf{a}_p) \\ & \vdots & \\ dx_{i_p}(\mathbf{a}_1) & \ldots & dx_{i_p}(\mathbf{a}_p) \end{bmatrix}.$$

Note that A^1 is just the space of linear functions. For convenience and completeness, we define $A^0 = R$ and $A^p = \{\mathbf{0}\}$ for $p > n$. A^p is a vector space, of dimension $\begin{pmatrix} n \\ p \end{pmatrix}$, with a basis given by $\{dx_{i_1} \wedge \ldots \wedge dx_{i_p} : i_1 < \ldots < i_p\}$.

Definition 11.2.1 *Differential form*
A differential p-form on a set $S \subset R^n$ is a function $\boldsymbol{\omega} : S \to A^p$, i.e.

$$\boldsymbol{\omega}(\mathbf{x}) = \sum_{i_1 < \ldots < i_p} \omega_{i_1 \ldots i_p}(\mathbf{x}) dx_{i_1} \wedge \ldots \wedge dx_{i_p},$$

where $\omega_{i_1 \ldots i_p} : R^n \to R$.

A form $\boldsymbol{\omega}$ is called a C^k form on S, if all of the coefficients, $\omega_{i_1 \ldots i_p}$ are in $C^k(S)$. The 0-forms on S are just the real valued functions on S.

Definition 11.2.2 *Exterior derivative*
If $\boldsymbol{\omega}$ is a C^1 p-form on S, then the *exterior derivative* of $\boldsymbol{\omega}$ is the $(p+1)$-form, $d\boldsymbol{\omega}$ given by

$$d\boldsymbol{\omega} = \sum_{i_1 < \ldots < i_p} \sum_j \frac{\partial \omega_{i_1 \ldots i_p}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}.$$

Definition 11.2.3 *Closed and exact*

1. A p-form $\boldsymbol{\omega}$ is *closed* means $d\boldsymbol{\omega} = \mathbf{0}$.

2. A p form is *exact* means there is a $(p-1)$-form $\boldsymbol{\tau}$, so that $d\boldsymbol{\tau} = \boldsymbol{\omega}$.

The question one might feel the need to ask at this point is: What is going on here? Suppose you have a function $f : R^n \to R$, then viewing f as a 0-form its exterior derivative is the 1-form $df = f_{x_1} dx_1 + f_{x_2} dx_2 \ldots + f_{x_n} dx_n$ is just another way of writing f' or ∇f. If vector field $F : R^n \to R^n$ is identified with the 1-form $\boldsymbol{\omega} = F_1 dx_1 + \ldots + F_n dx_n$, the saying $\boldsymbol{\omega}$ is closed is equivalent to saying F' is symmetric, and to say that $\boldsymbol{\omega}$ is exact is equivalent to saying that F has an antiderivative. So, the concept of differential forms is designed to generalize the notion of an antiderivative and necessary conditions for the existence of antiderivatives. For example, the following theorem is a generalization of Theorem 12.1.5.

Theorem 11.2.1 *If $\boldsymbol{\omega}$ is a C^2 p-form, then $d^2\boldsymbol{\omega} = d(d\boldsymbol{\omega}) = 0$, i.e. an exact form is closed.*

The exterior derivative also generalizes other well-known operators on vector fields, in particular the curl and the divergence.

A 0-form is just a function. An n-form also has a function naturally associated with it. Namely, an n-form looks like $\boldsymbol{\omega} = \omega \ dx_1 \wedge \ldots \wedge dx_n$, since A^n is one-dimensional. So, the function associated to an n-form is its single coefficient. Moreover, A^1 and A^{n-1} are n-dimensional, so a 1-form and an $(n-1)$-form have n coefficients which can be associated to vector fields. In particular, for a 1-form

$$\boldsymbol{\omega} = \omega_1 dx_1 + \ldots + \omega_n dx_n \leftrightarrow F = (\omega_1, \ldots, \omega_n)$$

and for an $(n-1)$-form, letting $dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n$ denote the basic $(n-1)$-form not containing dx_i,

$$\boldsymbol{\omega} = \sum_i \omega_i dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n \leftrightarrow F = (\omega_1, \ldots, (-1)^{i-1}\omega_i, \ldots, (-1)^{n-1}\omega_n)$$

then we have the following

1. for a 0-form $\boldsymbol{\omega}$, the vector field associated to $d\boldsymbol{\omega}$ is the gradient of the function associated to $\boldsymbol{\omega}$.

2. if F is the vector field associated to an $(n-1)$-form $\boldsymbol{\omega}$, then the function associated to $d\boldsymbol{\omega}$ is the *divergence* of F,

$$\text{div } F = \nabla \cdot F = \sum_i \frac{\partial F_i}{\partial x_i}$$

3. in R^3, if F is the vector field associated to a 1-form $\boldsymbol{\omega}$, then the vector field associated to $d\boldsymbol{\omega}$ is the *curl* of F,

$$\text{curl } F = \nabla \times F = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right)$$

11.2.2 Chains and Integration

A *singular p-cube* in R^n is function $C : R^p \to R^n$ that is continuously differentiable on the unit cube in R^p, $K_p = \{\mathbf{t} : 0 \leq t_i \leq 1\}$. The image of the unit cube under C is also thought of as a singular p-cube. The columns of the derivative matrix are tangent vectors to the image of C. We also define a singular p-cube for $p = 0$, by letting $K_0 = 0$, so that a 0-cube is the single point.

Definition 11.2.4 *p-Chains*
A *p-chain* in R^n is a linear combination of singular p-cubes.

The coefficients of the chain have to do piecing together "smooth" p-cubes to parameterize geometrical objects with "corners" and to "orient" the cubes defining the chain. The geometrical object associated to a chain is the union of the images of the cubes that define the chain. When we say "a p-chain in S" we mean that the images of the p-cubes defining the chain are all contained in the set S.

Definition 11.2.5 *Integration of p-forms on p-chains*
If C is a singular p-cube and $\boldsymbol{\omega}$ is a continuous p-form on (the image of) C, then the *integral of* $\boldsymbol{\omega}$ *over* C is defined to be, for $p = 0$

$$\int_C \boldsymbol{\omega} = \boldsymbol{\omega}(C(0))$$

and

$$\int_C \boldsymbol{\omega} = \int_{K_p} \sum_{i_1<\ldots<i_p} \omega_{i_1\ldots i_p}(C(\mathbf{t}))dx_{i_1} \wedge \ldots \wedge dx_{i_p}(\frac{\partial C}{\partial t_1}(\mathbf{t}), \ldots, \frac{\partial C}{\partial t_p}(\mathbf{t}))dt_1 \ldots dt_p$$

for $p > 0$.
If $\mathbf{c} = \sum_{j=1}^{J} \gamma_j C_j$ is a p-chain and $\boldsymbol{\omega}$ is a continuous p-form on $\mathbf{c}$ (that is on the image of each C_j), then the *integral of* $\boldsymbol{\omega}$ *over* $\mathbf{c}$ is defined to be

$$\int_{\mathbf{c}} \boldsymbol{\omega} = \sum_j \gamma_j \int_{C_j} \boldsymbol{\omega}.$$

This integration generalizes such familiar notions as line and surface integrals and includes the Riemann integral. For $p = 1$ the integral of a 1-form over a 1-chain is just the line integral of the vector field corresponding to the 1-form over the curve given by the 1-chain. For $p = n$, the integral of an n-form over an n-cube is, up to sign, just the Riemann integral of the coefficient of the form over the image of the cube. In R^3, a 2-form $\boldsymbol{\omega} = \omega_1 dx_2 \wedge dx_3 + \omega_2 dx_1 \wedge dx_3 + \omega_3 dx_1 \wedge dx_2$ corresponds to the vector field $F = (\omega_1, -\omega_2, \omega_3)$ and a 2-cube C draws a 2-dimensional surface in R^3. Letting $\mathbf{n} = (C_{t_1} \times C_{t_2})/|C_{t_1} \times C_{t_2}|$, we have

$$\int_C \boldsymbol{\omega} = \int_{K_2} \langle F(C(\mathbf{t})), \mathbf{n}(\mathbf{t})\rangle \, |\frac{\partial C}{\partial t_1}(\mathbf{t}) \times \frac{\partial C}{\partial t_2}(\mathbf{t})| \, dt_1 dt_2$$

$\langle F(C(\mathbf{t})), \mathbf{n}(\mathbf{t})\rangle$ is the normal component of F along the surface and $|C_{t_1}(\mathbf{t}) \times C_{t_2}(\mathbf{t})|dt_1 dt_2$ is the "infinitesimal element of surface area" on the surface.

The $2n$ "faces" of C are singular $(p-1)$ cubes given by,

$$\begin{aligned} C_{i0}(\mathbf{t}) &= C(t_1, t_2, \ldots, t_{i-1}, 0, t_i, \ldots, t_{p-1}) \\ C_{i1}(\mathbf{t}) &= C(t_1, t_2, \ldots, t_{i-1}, 1, t_i, \ldots, t_{p-1}) \end{aligned}$$

for $i = 1, \ldots, n$.

Definition 11.2.6 *Boundary operator*

1. The *boundary* of a singular p-cube is defined to be
$$\partial C = \sum_{i=1}^{p} (-1)^{i-1}(C_{i1} - C_{i0}).$$

2. The *boundary* of a p-chain $\mathbf{c} = \sum_{j=1}^{J} \gamma_j C_j$, where each C_j is a singular p-cube, is defined to be
$$\partial \mathbf{c} = \sum_{j=1}^{J} \gamma_j \partial C_j$$

Definition 11.2.7 *Cycles and boundaries*

1. A *p-cycle* is a p-chain $\mathbf{c}$ with $\partial \mathbf{c} = \mathbf{0}$, that is $\int_{\partial \mathbf{c}} \boldsymbol{\omega} = 0$ for all $\boldsymbol{\omega}$.

2. A *p-boundary* is a p-chain $\mathbf{c}$ for which there is a $(p+1)$-chain $\mathbf{b}$, so that $\partial \mathbf{b} = \mathbf{c}$. Equality means $\int_{\mathbf{c}} \boldsymbol{\omega} = \int_{\partial \mathbf{b}} \boldsymbol{\omega}$ for all $\boldsymbol{\omega}$.

Theorem 11.2.2 *If* $\mathbf{c}$ *is a p-chain, then* $\partial^2 \mathbf{c} = \partial(\partial \mathbf{c}) = 0$, *i.e. a boundary is a cycle.*

11.2.3 THE Fundamental Theorem of Calculus

As was mentioned above, the exterior derivative extends the gradient (curl and divergence), exactness extends the notion of antiderivative, and being closed extends the symmetry of the derivative as a sufficient condition for exactness (having an antiderivative). The final questions involve how integration and (exterior) exterior differentiation are related and when antiderivatives exist. In R these questions are answered by the Fundamental Theorem of Calculus and more generally by the following.

Theorem 11.2.3 (FUNDAMENTAL THEOREM OF CALCULUS)

1. *If* $\boldsymbol{\omega}$ *is an exact,* C^0 *p-form and* $\mathbf{c}$ *is a p-chain, then*
$$\int_{\mathbf{c}} \boldsymbol{\omega} = \int_{\partial \mathbf{c}} \boldsymbol{\tau}$$
for any $(p-1)$-form, $\boldsymbol{\tau}$, *so that* $d\boldsymbol{\tau} = \boldsymbol{\omega}$.

2. *Existence of antiderivatives on open sets. If S is an open set in R^n, then*
   (a) *A C^1 p-form $\boldsymbol{\omega}$ on S is exact if, and only if, $\int_{\mathbf{c}} \boldsymbol{\omega} = 0$ for all p-cycles in S.*
   (b) *Every closed C^1 p-form in S is exact if, and only if, every p-cycle is a boundary.*

Corollary 11.2.4 (Stokes' Theorem) *If $\boldsymbol{\tau}$ is a C^1 p-form and $\mathbf{c}$ is a $(p+1)$-chain, then*

$$\int_{\mathbf{c}} d\boldsymbol{\tau} = \int_{\partial \mathbf{c}} \boldsymbol{\tau}.$$

Stokes' theorem includes Corollary 12.1.4, Theorem 12.1.7 and Green's Theorem. It also includes the Stokes' Theorem and Divergence Theorem from elementary multivariable calculus. The two parts on existence extend the second part of the one-variable theorem of calculus.

Theorem 12.2.3 is the "ultimate" generalization of the Fundamental Theorem of Calculus. The first part relates the integral of a form over a set to the antiderivatives on the boundary of the set. Part 2(a) gives necessary and sufficient conditions for the existence an antiderivative for a particular p-form, and it extends Theorem 12.1.8. Part 2(b) gives necessary and sufficient conditions for every possible p-form to have an antiderivative and generalizes Corollary 12.1.10. It, like Corollary 12.1.10, says something quite surprising, namely that the existence of antiderivatives is intimately related to the geometry, e.g. for S in R^3, every function C^1 on S is the divergence of a vector field on S if, and only if, every closed surface in S is the boundary of a solid in S.

There is one major difference in this version of the Fundamental Theorem and the one-variable version. This version gives is no explicit construction for antiderivatives. This unfortunate fact is really not surprising since the existence is related to the geometry of the set on which existence is desired. So, the construction may be quite different on different sets. Actually, the technique is to construct an antiderivative locally and patch them together in a way determined by the geometry of the set. We will not describe the patching but the local construction is based on the following.

Theorem 11.2.5 *Define for each C^1 form $\boldsymbol{\omega}$ on the open unit ball in R^n, the form $h(\boldsymbol{\omega})$ by*

$$h(\boldsymbol{\omega})(\mathbf{x}) = \sum_{i_1<\ldots<i_p} \left(\int_0^1 t^{p-1}\omega_{i_1\ldots i_p}(t\mathbf{x})dt\right) \sum_{\alpha=1}^{p} (-1)^{\alpha-1} x_{i_\alpha} dx_{i_1} \wedge \ldots \wedge \widehat{dx}_{i_\alpha} \wedge \ldots \wedge dx_{i_p},$$

then

$$\boldsymbol{\omega} = h(d\boldsymbol{\omega}) + d(h(\boldsymbol{\omega})).$$

Corollary 11.2.6 (Poincaré Lemma) *On the open unit ball, every closed form is exact.*

Corollary 11.2.7 *On an open set S in R^n, every closed form is "locally" exact, that is, for p-form, $\boldsymbol{\omega}$, closed on S, for each $\mathbf{x}_0 \in S$, there is an open ball $B \subset S$, centered at $\mathbf{x}_0$ and a $(p-1)$-form $\boldsymbol{\tau}$, so that $d\boldsymbol{\tau} = \boldsymbol{\omega}$ on B.*

11.3 Exercises

11–1. Prove Theorem 12.1.1

11–2. Show that any two antiderivatives of a continuous vector field $F : R^n \to R^n$, on a connected open set S, differ by a constant.

11–3. Suppose $f : R^2 \to R$, $b : R \to R$ and $a : R \to R$ are continuously differentiable. Show that

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,y)\,dy = \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(x,y)\,dy + f(x,b(x))b'(x) - f(x,a(x))a'(x).$$

11–4. Find antiderivatives of the following vector fields on their domains, if they exist.

(a) $F(x,y) = (x-y, x+y)$

(b) $G(x,y,z) = (y,z,x)$

(c) $H(x,y) = (e^x\cos y, -e^x \sin y)$

(d) $L(x,y) = (x/(x^2+y^2), y/(x^2+y^2))$

11–5. Evaluate the following line integrals.

(a) $\int_c e^x \cos y dx + e^x \sin y dy$ where c is the triangle with vertices $(0,0)$, $(1,0)$ and $(1,\pi/2)$ traced counterclockwise.

(b) $\int_c e^x \cos y dx + e^x \sin y dy$ where c is square with corners $(0,0)$, $(1,0)$, $(1,1)$ and $(0,1)$ traced counterclockwise.

(c) $\int_c e^x \cos y dx - e^x \sin y dy$ where c is the triangle with vertices $(0,0)$, $(1,0)$ and $(1,\pi/2)$ traced counterclockwise.

(d) $\int_c (x^2-y^2)dx + (x^2+y^2)dy$ where c is the circle of radius 1 centered at the origin, traced clockwise.

11–6. Show that if a vector field $F : R^2 \to R^2$ is continuously differentiable on an open set containing S, a set with a piecewise smooth boundary, then

$$\int_{\partial S} -F_2 dx + F_1 dy = \int_S \text{div } F.$$

How would you interpret geometrically the line integral on the left?

11–7. Let F be a continuously differentiable vector field defined everywhere but at two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in R^2, and satisfying $F_{1x_2} = F_{2x_1}$. Let c_1 and c_2 be counterclockwise oriented circles centered at $\mathbf{x}_1$ and $\mathbf{x}_2$ with radii less than $|\mathbf{x}_1 - \mathbf{x}_2|$. Suppose

$$\int_{c_k} F \cdot d\mathbf{x} = I_k \qquad \text{for } k = 1, 2.$$

Show that if c is any closed smooth curve that avoids $\mathbf{x}_1$ and $\mathbf{x}_2$, then

$$\int_c F \cdot d\mathbf{x} = n_1 I_1 + n_2 I_2$$

for some integers n_1 and n_2.

11–8. For fixed vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}$, the function $l : R^n \to R$ defined by

$$l(\mathbf{x}) = dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}, \mathbf{x})$$

is linear, so there is a vector $\mathbf{c}$ so that $l(\mathbf{x}) = \mathbf{c}^T\mathbf{x} = \langle \mathbf{c}, \mathbf{x} \rangle$, for all $\mathbf{x} \in R^n$. What is the geometrical relationship between $\mathbf{a}_i$ and $\mathbf{c}$? For $n = 3$ what well-known object is $\mathbf{c}$?

11–9. Show that the 2-form

$$\boldsymbol{\omega} = \frac{x dy \wedge dz + y dz \wedge dx + z dx \wedge dy}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}$$

is closed on $R^3 - \{(0, 0, 0)\}$. Is it exact?

11–10. For the given $\boldsymbol{\omega}$ and $\mathbf{c}$ compute $\int_{\mathbf{c}} \boldsymbol{\omega}$.

(a) $\boldsymbol{\omega} = x^2 dx + y^2 dy$ and $\mathbf{c}$ the straight line from $(1, 1)$ to $(2, 4)$.

(b) $\boldsymbol{\omega} = x^2 dy \wedge dz + y^2 dz \wedge dx + z^2 dx \wedge dy$ and $\mathbf{c}$ the surface defined by $\mathbf{c}(r, \theta) = (r \cos\theta, r \sin\theta, \theta)$ for $1 \le r \le 2$ and $0 \le \theta \le 2\pi$.

(c) $\boldsymbol{\omega}$ the exterior derivative of the form in (b) and $\mathbf{c}$ the unit ball in R^3.

(d) $\boldsymbol{\omega} = x^2 dx + y^2 dy + z^2 dz$ and $\mathbf{c}$ the boundary of the surface in (b).

11–11. The sets of p-forms have some algebraic structure. In particular, if $\boldsymbol{\omega} = \sum_{i_1<\ldots<i_p} \omega_{i_1\ldots i_p} dx_{i_1} \wedge \ldots \wedge dx_{i_p}$ and $\boldsymbol{\tau} = \sum_{i_1<\ldots<i_p} \tau_{i_1\ldots i_p} dx_{i_1} \wedge \ldots \wedge dx_{i_p}$ and f and g are functions, then $f\boldsymbol{\omega}+g\boldsymbol{\tau} = \sum_{i_1<\ldots<i_p}(f\omega_{i_1\ldots i_p}+g\tau_{i_1\ldots i_p})dx_{i_1}\wedge\ldots\wedge dx_{i_p}$. Moreover, if $\boldsymbol{\nu} = \sum_{j_1<\ldots<j_q} \nu_{j_1\ldots j_q} dx_{j_1} \wedge \ldots \wedge dx_{j_q}$ is a q-form, then the *wedge product* of $\boldsymbol{\omega}$ and $\boldsymbol{\nu}$ is defined to be

$$\boldsymbol{\omega} \wedge \boldsymbol{\nu} = \sum_{i_1<\ldots<i_p} \sum_{j_1<\ldots<j_q} \omega_{i_1\ldots i_p} \nu_{j_1\ldots j_q} dx_{i_1} \wedge \ldots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_q}.$$

(a) Compute $\boldsymbol{\omega}\wedge\boldsymbol{\tau}$ where $\boldsymbol{\omega} = x_1dx_1+x_2dx_2+x_3dx_3$ and $\boldsymbol{\tau} = (x_1^2-x_3)dx_1\wedge dx_2 + x_2x_3dx_1 \wedge dx_3 + (x_1+x_2+x_3)dx_2 \wedge dx_3$.

(b) Show that

i. $(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \wedge \boldsymbol{\tau} = \boldsymbol{\omega}_1 \wedge \boldsymbol{\tau} + \boldsymbol{\omega}_2 \wedge \boldsymbol{\tau}$

ii. $(f\boldsymbol{\omega}) \wedge \boldsymbol{\tau} = \boldsymbol{\omega} \wedge (f\boldsymbol{\tau}) = f(\boldsymbol{\omega} \wedge \boldsymbol{\tau})$

(c) How are $\boldsymbol{\omega} \wedge \boldsymbol{\tau}$ and $\boldsymbol{\tau} \wedge \boldsymbol{\omega}$ related?

11–12. **(Integration of functions on p-dimensional surfaces)**: If C is singular p-cube in R^n, and $f : R^n \to R$ is continuous on an open set containing C then the *integral of f over C* is defined to be

$$\int_C f = \int_{K_p} f(C(\mathbf{t}))\sqrt{\det(C'^T(\mathbf{t})C'(\mathbf{t}))}dt_1 \cdots dt_p.$$

(a) Compute the integral of $f(x,y,z) = xyz$ over the unit sphere in R^3.

(b) What does this definition give when $n = p$?

(c) If C is one-to-one, then $\int_C 1$ is called the *volume* of C. Why?

11–13. Prove the following.

(a) **(Stokes' Theorem)** If C is a smooth surface in R^3 (i.e. a singular 2-cube) and F is a continuously differentiable vector field on C, then

$$\int_C \nabla \times F \cdot \mathbf{n}\, d\sigma = \int_{\partial C} F \cdot d\mathbf{x}$$

(b) **(Divergence Theorem)** If C is a smooth set in R^3 (i.e. a singular 3-cube) and F is a continuously differentiable vector field on C, then

$$\int_C \nabla \cdot F = \int_{\partial C} F \cdot \mathbf{n}\, d\sigma$$

11–14. What can you say about the exactness of closed p-forms on the set $S = \{(x,y,z) : 1 < x^2+y^2+z^2 < 4\}$? How about $S = \{(x,y,z) : 1 < x^2+y^2 < 4\}$?

11–15. For $F(x_1,x_2,x_3) = (x_1,-2x_2,x_3)$ find a vector field G, so that curl $G = F$.