BADAR MUNIR

Virtual University of Pakistan

Real Analysis II (MTH631)

Salman Amin Malik

Virtual University
Learning Management System

To my unknown students

About the instructor

Dr. Malik did his MS and PhD (Mathematics) from University of La Rochelle, La Rochelle, France in 2009 and 2012, respectively. Prior to MS and PhD, Dr. Malik completed his MPhil and MSc (Mathematics) from Department of Mathematics, University of the Punjab, Lahore, Pakistan. He has been a liated with several universities in Pakistan and abroad. He has the experience of teaching a wide range of mathematics courses at undergraduate and graduate level.

Dr. Malik has published several research articles in international journals and conferences. His area of research includes the study of di erential equations with nonlocal operators and their applications to image processing. He is also interested in inverse problems related to reaction-di usion equations with nonlocal integro-di erential operators and boundary conditions. These models have numerous applications in anomalous di usion/transport, biomedical imaging and non-destructive testing.

About the handouts

The books followed during this course are: W. Rudin, Principles of Mathematical Analysis, Third Edition, McGraw-Hill, 1976. ISBN: 9780070542358. and W. F. Trench, Introduction to Real Analysis, Pearson Education, 2013. Consequently, the most of the examples considered in these notes are from the above mentioned books and their exercises, but not restricted to those books only. If you nd any typing error in the text kindly report to me by writing an email to salman.amin.malik@gmail.com.

Course Information

Title and Course Code: Real Analysis II (MTH631)

Number of Credit Hours: 3 credits

Course Objective: Real Analysis II is the follow up course of Real Analysis I and in general an advanced course related to mathematical analysis. The topics of the Real Analysis II are linked with its rst course namely Real Analysis I, indeed, we will extend the ideas of Real Analysis I to Euclidean space R^n, we will discuss sequences and series of functions, limits and continuity of functions of several variables, partial derivatives their applications, multiple integrals etc. Upon completion of this course students will be able to

- Understand the convergence of sequence of functions (LO1).
- Understand the pointwise convergence, uniform convergence, several tests for convergence (LO2).
- Apply the interchange of limit and integration, derivative of sequence of functions (LO3).
- Understand the in nite series of functions, convergence, Weierstrass's test and some other results about the convergence (LO4).
- Apply Dirichlet's test for uniform convergence, series of product of two functions, interchange of sum and intgeration (LO5).
- Represent and study the function which could be written as power series,term by term integral and derivative of a power series, (LO6). item Understand the concept of equicontinuous function, The Stone-Weierstrass Theorem (LO7).
- Understand and nd the Fourier series, Fourier coe cients, convergence of Fourier series (LO8).
- Apply the best approximation theorem and understand the Euler gamma function and the beta function and their properties (LO9)
- Understand the functions of several variables, Heine-Borel Theorem, limits and continuity of functions of several variables (LO10)
- Vector valued functions and their calculus, Bounded functions and several results about vector valued functions (LO11)
- Di erentiablity in R^n, Di erentials, Directional derivatives, Partial derivatives, Maxima and minima (LO12)
- Improper integrals, Multiple integrals, Functions of bounded variation (LO15)

Prerequisites: Real Analysis I (MTH621)

The textbooks for this course:

[1] W. Rudin, Principles of Mathematical Analysis, Third Edition, McGraw-Hill, 1976. ISBN: 9780070542358.

[2] W. F. Trench, Introduction to Real Analysis, Pearson Education, 2013.

[3] S. Ponnusamy, Foundations of Mathematical Analysis, Birkhauser, 2012.

Reference books:

[4] A. N. Kolmogorov and S. V. Fomin, Introductory Real Analysis, Revised English Edition Translated and Edited by R. A. Silverman, Dover Publication, Inc. New York.

[5] R. G. Bartle and D. R. Sherbert, Introduction to Real Analysis, Third Edition, 2000, John Wiley & Sons Inc.

- Sequences and Series of functions
- Functions of several variables
- Vector valued functions
- Integral Calculus

Contents

1 Sequences and Series of Functions 2

1.1 Informal way 2

1.2 Pointwise Convergence 3

1.3 Norm De ned Over a Set 5

1.4 Uniform Convergence 5

1.5 Cauchy's Uniform Convergence Criterion 8

1.6 Properties Preserved by Uniform Convergence 10

1.6.1 Continuity of the Limit Function at a Point 10

1.6.2 Interchange of Limit and Integration 12

1.6.3 Under What Conditions We May Have $F' = \lim_{n\to\infty} F_n'$ . . . 14

1.7 Series of Functions 15

1.8 Convergence of Series of Functions 15

1.8.1 Cauchy's criterion for functional series 17

1.8.2 Dominated Series of Real Numbers for Series of Functions 18

1.8.3 Weierstrass M-test/dominated Convergence Test 18

1.9 Power Series 30

1.10 The Taylor's Series 36

1.11 The Abel's Theorem 44

1.12 Pointwise and Uniform Bounded Functions 47

1.13 Equicontinuous Functions on a Set 48

1.14 The Stone-Weierstrass Theorem 51

1.15 Fourier Series 53

1.15.1 Periodic Functions 53

1.15.2 Periodic Extension 55

1.15.3 Trigonometric Polynomials 56

1.16 The space E 57

1.16.1 Fourier Series of Even and Odd Functions 61

1.17 Fourier Series for Arbitrary Periodic Function 66

1.18 Best Approximation Theorem 69

2 Functions of Several Variables 71

2.1 Euclidean Spaces 71

2.2 Schwarz's Inequality 72

2.2.1 Line Segment in $\mathbb{R}^n$ 75

2.3 Neighbourhoods and Open Sets in $\mathbb{R}^n$ 75

2.4 Convergence of a Sequence in $\mathbb{R}^n$ 77

2.5 Principle of nested sets 78

2.6 Heine-Borel Theorem 79

2.7 Connected Sets in $\mathbb{R}^n$ 81

2.7.1 Polygonal Path ... 81
2.8 Polygonally Connected Set ... 81
2.9 Sequences in R^n ... 84
2.10 Domain of Function of n Variable ... 85
2.11 Limit at a Point of a Function of n Variables ... 85
2.12 In nite Limits and Limits at $\mathbf{X} \to \infty$ ... 88
2.12.1 Limit at In nity ... 90
2.13 Continuity ... 92
2.14 Vector Valued Functions ... 93
2.14.1 Composite Function ... 93
2.15 Bounded Functions ... 95
2.16 Directional Derivative ... 96
2.16.1 Partial Derivative ... 97
2.16.2 Equality of Mixed Partial Derivatives ... 100
2.16.3 Generalization of Equality of Mixed Partial Derivative ... 101
2.17 Di erentiability of Functions of Several Variables ... 102
2.17.1 Linear Function ... 104
2.17.2 Di erential ... 105
2.17.3 A su cient Condition for Di erentiability ... 107
2.17.4 Continuously Di erentiable Function ... 108
2.17.5 Geometric Interpretation of Di erentiability ... 110
2.18 Maxima and Minima ... 113
2.19 Di erentiable Vector Valued Function ... 114
2.20 The Chain Rule ... 115
2.21 Higher derivatives of composite functions ... 118
2.22 rth Di erential ... 122
2.23 Taylor's Theorem for Functions of **n** Variables ... 124
2.23.1 Positive De nite ... 126

3 Integral Calculus 129
3.1 Locally Integrable Functions ... 129
3.1.1 The Comparison Test ... 134
3.2 Absolute integrability ... 137
3.3 Nonoscillatory and Oscillatory Functions ... 138
3.4 Conditional convergence ... 138
3.5 Dirichlet's Test ... 139
3.6 Rectangles in R^n ... 140
3.7 Riemann Sum in R^n ... 142
3.8 Riemann Integral in R^n ... 143
3.9 Upper and Lower Integrals ... 146
3.10 Sets with Zero Content ... 151
3.11 Integral Over Bounded Set ... 152
3.12 Di erentiable Surfaces ... 153
3.13 Iterated Integrals ... 156

List of Figures

1.1 Plot of $F_n(x) = \left(\frac{1}{n+x}\right)$ $n \geq 1$, for $n = 1, 2, 4, 8, 100$ 2
1.2 Plot of $F_n(x) = \frac{x}{n+x}$, $n \geq 1$, for $n = 1, 2, 4, 8, 20, 100$ 3
1.3 Uniform convergence graphically 5
1.4 Periodic functions 53
1.5 Geometric proof 54
1.6 Periodic extension example 1 55
1.7 Periodic extension example 2 56
1.8 At discontinuous points 64
1.9 Example 64
1.10 Example 65

2.1 The set S 76
2.2 Heine-Borel Theorem for $n = 2$ 79
2.3 Disconnected set which is not a region 83
2.4 A connected set which is not a region 83
2.5 A region 84
2.6 Domain of the function 86
2.7 Composite of vector valued functions 94
2.8 Domain of the function 110
2.9 Geometric interpretation of di erentiability 111

3.1 Rectangle in $\mathbb{R}^2$ 141
3.2 Rectangular parallelepiped in $\mathbb{R}^3$ 141
3.3 Partitioning of a rectangle in $\mathbb{R}^2$ 142
3.4 S_1 and S_2 156

Chapter 1

Sequences and Series of Functions

1.1 Informal way

$D \subset R$

If $F_k, F_{k+1}, \ldots, F_n, \ldots$ are real-valued functions de ned on a subset D of the real numbers, we say that $\{F_n\}$ is an in nite sequence or (simply a sequence) of functions on D. For each $x_0 \in D$, we have a sequence of real numbers and we can talk about the convergence of that sequence of real numbers.

If the sequence of values $\{F_n(x)\}$ converges for each x in some subset S of D, then $\{F_n\}$ de nes a limit function on S.

Example: The functions

$$F_n(x) = \left(\frac{1}{n+x}\right), \quad n \geq 1,$$

de ne a sequence on $D = [0, \infty)$.

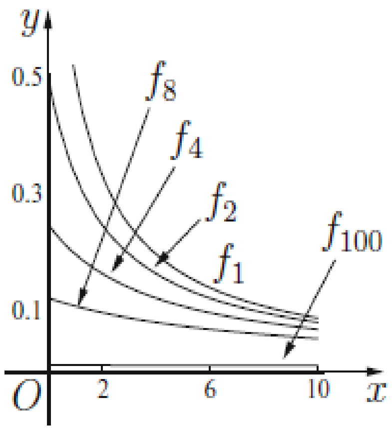

Figure 1.1: Plot of $F_n(x) = \left(\frac{1}{n+x}\right), \quad n \geq 1$, for $n = 1, 2, 4, 8, 100$

Example: The functions

$$F_n(x) = \left(\frac{x}{n+x}\right), \quad n \geq 1,$$

de ne a sequence on $D = [0, \infty)$.

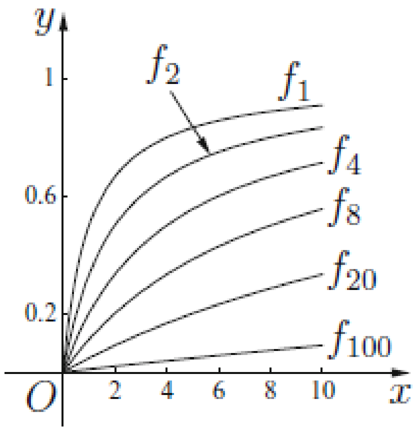

Figure 1.2: Plot of $F_n(x) = \left(\frac{x}{n+x}\right), \quad n \geq 1$, for $n = 1, 2, 4, 8, 20, 100$

Example: The functions

$$F_n(x) = \left(1 - \frac{nx}{n+1}\right)^{n/2}, \quad n \geq 1,$$

de ne a sequence on $D = (-\infty, 1]$.

1.2 Pointwise Convergence

Suppose that $\{F_n\}$ is a sequence of functions on D and the sequence of values $\{F_n(x)\}$ converges for each x in some subset S of D. Then we say that $\{F_n\}$ converges pointwise on S to the limit function F, de ned by

$$F(x) = \lim_{n\to\infty} F_n(x), \quad x \in S.$$

Example: The sequence of functions de ned by

$$\lim_{n\to\infty} F_n(x) = \begin{cases} \infty, & x < 0, \\ 1, & x = 0, \\ 0, & 0 < x \leq 1. \end{cases}$$

Therefore, $\{F_n\}$ converges pointwise on $S = [0, 1]$ to the limit function F de ned by

$$F(x) = \begin{cases} 1, & x = 0, \\ 0, & 0 < x \leq 1. \end{cases}$$

Example: Consider the functions

$$F_n(x) = x^n e^{-nx}, \quad x \geq 0, \quad n \geq 1.$$

Equating the derivative

$$F_n'(x) = nx^{n-1}e^{-nx}(1 - x)$$

to zero shows that the maximum value of $F_n(x)$ on $[0, \infty)$ is e^{-n}, attained at $x = 1$.
Therefore,

$$|F_n(x)| \leq e^{-n}, \quad x \geq 0,$$

so $\lim_{n\to\infty} F_n(x) = 0$ for all $x \geq 0$. The limit function in this case is identically zero on $[0, \infty)$.

Example: For $n \geq 1$, let F_n be de ned on $(-\infty, \infty)$ by

$$F_n(x) = \begin{cases} 0, & x < -\frac{2}{n}, \\ -n(2 + nx), & -\frac{2}{n} \leq x < -\frac{1}{n}, \\ n^2 x, & -\frac{1}{n} \leq x < \frac{1}{n}, \\ n(2 - nx), & \frac{1}{n} \leq x < \frac{2}{n}, \\ 0, & x \geq \frac{2}{n} \end{cases}$$

Since $F_n(0) = 0$ for all n, $\lim_{n\to\infty} F_n(0) = 0$. If $x \neq 0$, then $F_n(x) = 0$ if $n \geq 2/|x|$. Therefore,

$$\lim_{n\to\infty} F_n(x) = 0, \quad -\infty < x < \infty,$$

so the limit function is identically zero on $(-\infty, \infty)$.

Example: Show that the sequence of functions

$$F_n(x) = \left(\frac{1}{n + x}\right), \quad n \geq 1,$$

de ne a sequence on $D = [0, \infty)$, converges to 0.

Example: For each positive integer n, let S_n be the set of numbers of the form $x = p/q$, where p and q are integers with no common factors and $1 \leq q \leq n$.
De ne

$$F_n(x) = \begin{cases} 1, & x \in S_n, \\ 0, & x \notin S_n. \end{cases}$$

If x is irrational, then $x \notin S_n$ for any n, so $F_n(x) = 0$, $n \geq 1$. If x is rational, then $x \in S_n$ and $F_n(x) = 1$ for all su ciently large n.
Therefore,

$$\lim_{n\to\infty} F_n(x) = F(x) = \begin{cases} 1 & \text{if } x \text{ is rational}, \\ 0 & \text{if } x \text{ is irrational}. \end{cases}$$

1.3 Norm De ned Over a Set

Let us introduce the notation

$$\|g\|_S = \sup_{x \in S} |g(x)|.$$

Lemma: If g and h are de ned on S, then

$$\begin{aligned} \|g+h\|_S &\leq \|g\|_S + \|h\|_S \\ \|gh\|_S &\leq \|g\|_S \|h\|_S. \end{aligned}$$

Moreover, if either g or h is bounded on S, then

$$\|g-h\|_S \geq |\|g\|_S - \|h\|_S|.$$

1.4 Uniform Convergence

A sequence $\{F_n\}$ of functions de ned on a set S converges uniformly to the limit function F on S if

$$\lim_{n\to\infty} ||F_n - F||_S = 0.$$

Thus, $\{F_n\}$ converges uniformly to F on S if for each $\varepsilon > 0$ there is an integer N such that

$$||F_n - F||_S < \varepsilon \quad \text{if} \quad n \geq N. \tag{1.1}$$

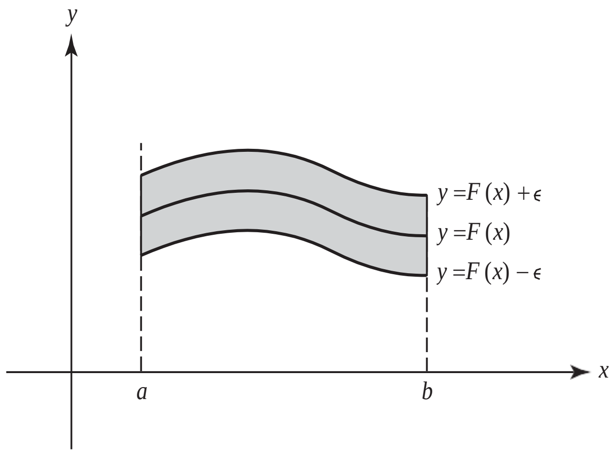

Figure 1.3: Uniform convergence graphically

A sequence $\{F_n\}$ of functions de ned on a set S converges uniformly to the limit function F on S if

$$\lim_{n\to\infty} ||F_n - F||_S = 0.$$

Thus, $\{F_n\}$ converges uniformly to F on S if for each $\varepsilon > 0$ there is an integer N such that

$$\|F_n - F\|_S < \varepsilon \quad \text{if} \quad n \geq N. \tag{1.2}$$

If $S = [a, b]$ and F is the function with graph shown in then (1.2) implies that the graph of

$$y = F_n(x), \qquad a \leq x \leq b,$$

lies in the shaded band

$$F(x) - \varepsilon < y < F(x) + \varepsilon, \quad a \leq x \leq b, \qquad n \geq N$$

.

Example: The sequence $\{F_n\}$ de ned by

$$F_n(x) = x^n e^{-nx}, \qquad n \geq 1,$$

converges uniformly to $F \equiv 0$.

We have

$$\|F_n - F\|_S = \|F_n\|_S = e^{-n},$$

so

$$\|F_n - F\|_S < \varepsilon$$

if $n > -\log \varepsilon$. For these values of n, the graph of

$$y = F_n(x), \qquad 0 \leq x < \infty,$$

lies in the strip

$$-\varepsilon \leq y \leq \varepsilon, \quad x \geq 0$$

Theorem: Let $\{F_n\}$ be de ned on S. Then

1. $\{F_n\}$ converges pointwise to F on S if and only if there is, for each $\varepsilon > 0$ and $x \in S$, an integer N (which may depend on x as well as ε) such that

$$|F_n(x) - F(x)| < \varepsilon \quad \text{if} \quad n \geq N(\varepsilon, x).$$

2. $\{F_n\}$ converges uniformly to F on S if and only if there is for each $\varepsilon > 0$ an integer N (which depends only on ε and not on any particular x in S) such that

$$|F_n(x) - F(x)| < \varepsilon \quad \text{for all } x \text{ in } S \text{ if } n \geq N(\varepsilon).$$

Theorem: If $\{F_n\}$ converges uniformly to F on S, then $\{F_n\}$ converges pointwise to F on S.

The converse is false; that is, pointwise convergence does not imply uniform convergence.

Counter example: For $n \geq 1$, let F_n be de ned on $(-\infty, \infty)$ by

$$F_n(x) = \begin{cases} 0, & x < -\frac{2}{n}, \\ -n(2+nx), & -\frac{2}{n} \leq x < -\frac{1}{n}, \\ n^2x, & -\frac{1}{n} \leq x < \frac{1}{n}, \\ n(2-nx), & \frac{1}{n} \leq x < \frac{2}{n}, \\ 0, & x \geq \frac{2}{n} \end{cases}$$

The sequence $\{F_n\}$ of converges pointwise to $F \equiv 0$ on $(-\infty, \infty)$, but not uniformly. Because

$$\|F_n - F\|_{(-\infty,\infty)} = F_n\left(\frac{1}{n}\right) = \left|F_n\left(\frac{-1}{n}\right)\right| = n,$$

so

$$\lim_{n\to\infty} \|F_n - F\|_{(-\infty,\infty)} = \infty.$$

Counter example: For $n \geq 1$, let F_n be de ned on $(-\infty, \infty)$ by

$$F_n(x) = \begin{cases} 0, & x < -\frac{2}{n}, \\ -n(2+nx), & -\frac{2}{n} \leq x < -\frac{1}{n}, \\ n^2x, & -\frac{1}{n} \leq x < \frac{1}{n}, \\ n(2-nx), & \frac{1}{n} \leq x < \frac{2}{n}, \\ 0, & x \geq \frac{2}{n} \end{cases}$$

However, the convergence is uniform on

$$S_\rho = (-\infty, \rho] \cup [\rho, \infty)$$

for any $\rho > 0$, since

$$\|F_n - F\|_{S_\rho} = 0 \quad \text{if} \quad n > \frac{2}{\rho}.$$

How to show that a sequence of functions is not uniformly convergent?

Suppose that a sequence of function F_n is point wise convergent on the set S. Then the convergence of F_n is not uniform, if there exists an $\varepsilon > 0$ such that to each integer N there correspond and integer $n > N$ and a point $x_n \in S$ for which we have

$$|F_n(x_n) - F(x_n)| \geq \varepsilon.$$

Example: If $F_n(x) = x^n$, $n \geq 1$, then $\{F_n\}$ converges pointwise on $S = [0, 1]$ to

$$F(x) = \begin{cases} 1, & x = 1, \\ 0, & 0 \leq x < 1. \end{cases}$$

The convergence is not uniform on S. To see this, suppose that $0 < \varepsilon < 1$. Then

$$|F_n(x) - F(x)| > 1 - \varepsilon \quad \text{if} \quad (1-\varepsilon)^{1/n} < x < 1.$$

Therefore,

$$1 - \varepsilon \leq \|F_n - F\|_S \leq 1$$

for all $n \geq 1$. Since ε can be arbitrarily small, it follows that

$$\|F_n - F\|_S = 1 \quad \text{for all} \quad n \geq 1.$$

Example: If $F_n(x) = x^n$, $n \geq 1$, then $\{F_n\}$ converges pointwise on $S = [0, 1]$ to

$$F(x) = \begin{cases} 1, & x = 1, \\ 0, & 0 \leq x < 1. \end{cases}$$

However, the convergence is uniform on $[0, \rho]$ if $0 < \rho < 1$, since then

$$\|F_n - F\|_{[0,\rho]} = \rho^n$$

and $\lim_{n\to\infty} \rho^n = 0$. Another way to say the same thing: $\{F_n\}$ converges uniformly on every closed subset of $[0, 1)$.

1.5 Cauchy's Uniform Convergence Criterion

Theorem: A sequence of functions $\{F_n\}$ converges uniformly on a set S if and only if for each $\varepsilon > 0$ there is an integer N such that

$$\|F_n - F_m\|_S < \varepsilon \quad \text{if} \quad n, m \geq N. \tag{1.3}$$

Proof: For necessity, suppose that $\{F_n\}$ converges uniformly to F on S.
Then, if $\varepsilon > 0$, there is an integer N such that

$$\|F_k - F\|_S < \frac{\varepsilon}{2} \quad \text{if} \quad k \geq N.$$

Therefore,

$$\begin{aligned} \|F_n - F_m\|_S &= \|(F_n - F) + (F - F_m)\|_S \\ &\leq \|F_n - F\|_S + \|F - F_m\|_S \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{if} \quad m, n \geq N. \end{aligned}$$

For su ciency, we rst observe that (1.17) implies that

$$|F_n(x) - F_m(x)| < \varepsilon \quad \text{if} \quad n, m \geq N,$$

for any xed x in S.

Therefore, Cauchy's convergence criterion for sequences of constants implies that $\{F_n(x)\}$ converges for each x in S; that is, $\{F_n\}$ converges pointwise to a limit function F on S.

To see that the convergence is uniform, we write

$$\begin{aligned}|F_m(x) - F(x)| &= |[F_m(x) - F_n(x)] + [F_n(x) - F(x)]| \\ &\le |F_m(x) - F_n(x)| + |F_n(x) - F(x)| \\ &\le \|F_m - F_n\|_S + |F_n(x) - F(x)|.\end{aligned}$$

This and (1.17) imply that

$$|F_m(x) - F(x)| < \epsilon + |F_n(x) - F(x)| \quad \text{if} \quad n, m \ge N. \tag{1.4}$$

Since $\lim_{n\to\infty} F_n(x) = F(x)$,

$$|F_n(x) - F(x)| < \epsilon$$

for some $n \ge N$, so (1.4) implies that

$$|F_m(x) - F(x)| < 2\epsilon \quad \text{if} \quad m \ge N.$$

But this inequality holds for all x in S, so

$$\|F_m - F\|_S \le 2\epsilon \quad \text{if} \quad m \ge N.$$

Since ϵ is an arbitrary positive number, this implies that $\{F_n\}$ converges uniformly to F on S.

Example: Suppose that g is di erentiable on $S = (-\infty, \infty)$ and

$$|g'(x)| \le r < 1, \qquad -\infty < x < \infty. \tag{1.5}$$

Let F_0 be bounded on S and de ne

$$F_n(x) = g(F_{n-1}(x)), \qquad n \ge 1. \tag{1.6}$$

Show that $\{F_n\}$ converges uniformly on S.

Solution: We rst note that if u and v are any two real numbers, then (1.5) and the mean value theorem imply that

$$|g(u) - g(v)| \le r|u - v|. \tag{1.7}$$

Recalling (1.6) and applying this inequality with $u = F_{n-1}(x)$ and $v = 0$ shows that

$$\begin{aligned}|F_n(x)| &= |g(0) + (g(F_{n-1}(x)) - g(0))| \\ &\le |g(0)| + |g(F_{n-1}(x)) - g(0)| \\ &\le |g(0)| + r|F_{n-1}(x)|.\end{aligned}$$

Therefore, since F_0 is bounded on S, it follows by induction that F_n is bounded on S for $n \geq 1$.

Moreover, if $n \geq 1$, then (1.6) and (1.7) with $u = F_n(x)$ and $v = F_{n-1}(x)$ imply that

$$\begin{aligned} |F_{n+1}(x) - F_n(x)| &= |g(F_n(x)) - g(F_{n-1}(x))| \\ &\leq r|F_n(x) - F_{n-1}(x)|, \quad -\infty < x < \infty, \end{aligned}$$

so

$$||F_{n+1} - F_n||_S \leq r||F_n - F_{n-1}||_S.$$

By induction, this implies that

$$||F_{n+1} - F_n||_S \leq r^n||F_1 - F_0||_S. \tag{1.8}$$

If $n > m$, then

$$\begin{aligned} ||F_n - F_m||_S &= ||(F_n - F_{n-1}) + (F_{n-1} - F_{n-2}) + \cdots \\ &\quad +(F_{m+1} - F_m)||_S \\ &\leq ||F_n - F_{n-1}||_S + ||F_{n-1} - F_{n-2}||_S + \cdots \\ &\quad +||F_{m+1} - F_m||_S. \end{aligned}$$

Now (1.8) implies that

$$\begin{aligned} ||F_n - F_m||_S &\leq ||F_1 - F_0||_S(1 + r + r^2 + \cdots + r^{n-m-1})r^m \\ &< ||F_1 - F_0||_S \frac{r^m}{1-r}. \end{aligned}$$

Therefore, if $$||F_1 - F_0||_S \frac{r^N}{1-r} < \epsilon,$$

then $||F_n - F_m||_S < \epsilon$ if $n, m \geq N$.

1.6 Properties Preserved by Uniform Convergence

1.6.1 Continuity of the Limit Function at a Point

Theorem: If $\{F_n\}$ converges uniformly to F on S and each F_n is continuous at a point x_0 in S, then so is F. Similar statements hold for continuity from the right and left.

Proof: Suppose that each F_n is continuous at x_0. If $x \in S$ and $n \geq 1$, then

$$\begin{aligned} |F(x) - F(x_0)| &\leq |F(x) - F_n(x)| + |F_n(x) - F_n(x_0)| + |F_n(x_0) - F(x_0)| \\ &\leq |F_n(x) - F_n(x_0)| + 2||F_n - F||_S. \end{aligned} \tag{1.9}$$

Suppose that $\varepsilon > 0$. Since $\{F_n\}$ converges uniformly to F on S, we can choose n so that $\|F_n - F\|_S < \varepsilon$. For this xed n, (1.9) implies that

$$|F(x) - F(x_0)| < |F_n(x) - F_n(x_0)| + 2\varepsilon, \quad x \in S. \tag{1.10}$$

Since F_n is continuous at x_0, there is a $\delta > 0$ such that

$$|F_n(x) - F_n(x_0)| < \varepsilon \quad \text{if} \quad |x - x_0| < \delta.$$

So, from (1.10),

$$|F(x) - F(x_0)| < 3\varepsilon, \quad \text{if} \quad |x - x_0| < \delta.$$

Therefore, F is continuous at x_0.

Similar arguments apply to the assertions on continuity from the right and left.

Corollary: If $\{F_n\}$ converges uniformly to F on S and each F_n is continuous on S, then so is F; that is, a uniform limit of continuous functions is continuous.

Proof: See video lectures.

Remark: If $\{F_n\}$ converges uniformly to F on S. Is the following

$$\int_a^b F(x)\,dx = \lim_{n\to\infty} \int_a^b F_n(x)\,dx,$$

is true?

Example: $\int_a^b F(x)\,dx = \lim_{n\to\infty} \int_a^b F_n(x)\,dx$, is not true generally.

Consider the sequence of functions de ned on $S = [0, 1]$

$$F_n(x) = \begin{cases} 0, & x = 0, \\ n, & 0 \le x \le \frac{1}{n}, \\ 0, & \frac{1}{n} < x < 1. \end{cases}$$

Then the sequence $\{F_n\}$ converges pointwise to $F(x) = 0$ on $[0, 1]$ and it is not uniformly convergent. We have

$$\int_0^1 F_n(x)\,dx = \int_0^{1/n} n\,dx + \int_{1/n}^1 0\,dx = 1 \qquad \text{But} \qquad \int_0^1 F(x)\,dx = 0$$

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

1.6.2 Interchange of Limit and Integration

Theorem: Suppose that $\{F_n\}$ converges uniformly to F on $S = [a, b]$. Assume that F and all F_n are integrable on $[a, b]$.
Then
$$\int_a^b F(x)\,dx = \lim_{n\to\infty} \int_a^b F_n(x)\,dx. \tag{1.11}$$

Proof: Consider
$$\left| \int_a^b F_n(x)\,dx - \int_a^b F(x)\,dx \right|.$$

$$\begin{aligned}\left| \int_a^b F_n(x)\,dx - \int_a^b F(x)\,dx \right| &\le \int_a^b |F_n(x) - F(x)|\,dx \\ &\le (b-a)\|F_n - F\|_S\end{aligned}$$

and $\lim_{n\to\infty} \|F_n - F\|_S = 0$, the conclusion follows.

Remark: Recall the theorem we have just proved; i.e.,

Theorem: Suppose that $\{F_n\}$ converges uniformly to F on $S = [a, b]$. Assume that F and all F_n are integrable on $[a, b]$.
Then
$$\int_a^b F(x)\,dx = \lim_{n\to\infty} \int_a^b F_n(x)\,dx.$$

The hypotheses of Theorem are stronger than necessary.

Theorem: Suppose that $\{F_n\}$ converges pointwise to F and each F_n is integrable on $[a, b]$.

1. If the convergence is uniform, then F is integrable on $[a, b]$ and
$$\int_a^b F(x)\,dx = \lim_{n\to\infty} \int_a^b F_n(x)\,dx.$$
holds.

2. If the sequence $\{\|F_n\|_{[a,b]}\}$ is bounded and F is integrable on $[a, b]$, then
$$\int_a^b F(x)\,dx = \lim_{n\to\infty} \int_a^b F_n(x)\,dx.$$
holds.

Remark: Part (1) of this theorem shows that it is not necessary to assume that F is integrable on $[a, b]$, since this follows from the uniform convergence. Part (2) is known as the bounded convergence theorem. Neither of the assumptions of (2) can be omitted.

Example (Unbounded sequence of functions): For $n \geq 1$, let F_n be de ned on $(-\infty, \infty)$ by

$$F_n(x) = \begin{cases} 0, & x < -\frac{2}{n}, \\ -n(2+nx), & -\frac{2}{n} \leq x < -\frac{1}{n}, \\ n^2 x, & -\frac{1}{n} \leq x < \frac{1}{n}, \\ n(2-nx), & \frac{1}{n} \leq x < \frac{2}{n}, \\ 0, & x \geq \frac{2}{n} \end{cases}$$

$\{\|F_n\|_{[0,1]}\}$ is unbounded while F is integrable on $[0, 1]$,

$$\int_0^1 F_n(x)\, dx = 1, \quad n \geq 1, \quad \text{but} \quad \int_0^1 F(x)\, dx = 0.$$

Example (Bounded sequence of functions but limit is not integrable): For each positive integer n, let S_n be the set of numbers of the form $x = p/q$, where p and q are integers with no common factors and $1 \leq q \leq n$.

De ne

$$F_n(x) = \begin{cases} 1, & x \in S_n, \\ 0, & x \notin S_n. \end{cases}$$

If x is irrational, then $x \notin S_n$ for any n, so $F_n(x) = 0$, $n \geq 1$. If x is rational, then $x \in S_n$ and $F_n(x) = 1$ for all su ciently large n. Therefore,

$$\lim_{n\to\infty} F_n(x) = F(x) = \begin{cases} 1 & \text{if } x \text{ is rational}, \\ 0 & \text{if } x \text{ is irrational}. \end{cases}$$

In this example it is clear that $\|F_n\|_{[a,b]} = 1$ for every nite interval $[a, b]$, F_n is integrable for all $n \geq 1$, and F is nonintegrable on every interval.

Example: The sequence $\{F_n\}$ de ned by

$$F_n(x) = x^n \sin \frac{1}{x^{n-1}}.$$

The sequence of functions converges $\{F_n\}$ converges uniformly to $F \equiv 0$ on $[r_1, r_2]$ if $0 < r_1 < r_2 < 1$ (or, equivalently, on every compact subset of $(0, 1)$).

However,

$$F_n'(x) = nx^{n-1} \sin \frac{1}{x^{n-1}} - (n-1) \cos \frac{1}{x^{n-1}},$$

so $\{F_n'(x)\}$ does not converge for any x in $(0, 1)$.

1.6.3 Under What Conditions We May Have $F' = \lim_{n\to\infty} F_n'$

Theorem: Suppose that F_n' is continuous on $[a, b]$ for all $n \geq 1$ and $\{F_n'\}$ converges uniformly on $[a, b]$. Suppose also that $\{F_n(x_0)\}$ converges for some x_0 in $[a, b]$.

Then $\{F_n\}$ converges uniformly on $[a, b]$ to a di erentiable limit function F, and

$$F'(x) = \lim_{n\to\infty} F_n'(x), \quad a < x < b, \tag{1.12}$$

while

$$F_+'(a) = \lim_{n\to\infty} F_n'(a+) \quad \text{and} \quad F_-'(b) = \lim_{n\to\infty} F_n'(b-). \tag{1.13}$$

Proof: Since F_n' is continuous on $[a, b]$, due to fundamental theorem of calculus, we can write

$$F_n(x) = F_n(x_0) + \int_{x_0}^{x} F_n'(t)\, dt, \quad a \leq x \leq b. \tag{1.14}$$

$$\text{Let} \quad L = \lim_{n\to\infty} F_n(x_0), \qquad G(x) = \lim_{n\to\infty} F_n'(x). \tag{1.15}$$

Since F_n' is continuous and $\{F_n'\}$ converges uniformly to G on $[a, b]$, G is continuous on $[a, b]$.

Therefore, (1.14) and using the fact we have proved $\int_a^b F(x)\, dx = \lim_{n\to\infty} \int_a^b F_n(x)\, dx$ (with F and F_n replaced by G and F_n') imply that $\{F_n\}$ converges pointwise on $[a, b]$ to the limit function

$$F(x) = L + \int_{x_0}^{x} G(t)\, dt.$$

$$F(x) = L + \int_{x_0}^{x} G(t)\, dt. \tag{1.16}$$

The convergence is actually uniform on $[a, b]$, since subtracting (1.14) from (1.16) yields

$$\begin{aligned} |F(x) - F_n(x)| &\leq |L - F_n(x_0)| + \left| \int_{x_0}^{x} |G(t) - F_n'(t)|\, dt \right| \\ &\leq |L - F_n(x_0)| + |x - x_0|\, \|G - F_n'\|_{[a,b]}. \end{aligned}$$

Consequently,

$$\|F - F_n\|_{[a,b]} \leq |L - F_n(x_0)| + (b - a)\|G - F_n'\|_{[a,b]},$$

where the right side approaches zero as $n \to \infty$.

Since G is continuous on $[a, b]$, (1.15), (1.16), De nition ??, and Theorem ?? imply (1.12) and (1.13).

1.7 Series of Functions

If $\{f_j\}_k^\infty$ is a sequence of real-valued functions defined on a set D of real numbers, then $\sum_{j=k}^\infty f_j$ is an infinite series (or simply a series) of functions on D.

The partial sums of , $\sum_{j=k}^\infty f_j$ are defined by

$$F_n = \sum_{j=k}^{n} f_j, \quad n \geq k.$$

If $\{F_n\}_k^\infty$ converges pointwise to a function F on a subset S of D, we say that $\sum_{j=k}^\infty f_j$ converges pointwise to the sum F on S, and write

$$F = \sum_{j=k}^{\infty} f_j, \quad x \in S.$$

If $\{F_n\}$ converges uniformly to F on S, we say that $\sum_{j=k}^\infty f_j$ converges uniformly to F on S.

1.8 Convergence of Series of Functions

The infinite series of functions $\sum_{j=k}^\infty f_j$ on D is said to be uniformly convergent if the sequence of partial sum $\{F_n\}$ defined by

$$F_n = \sum_{j=k}^{n} f_j, \quad n \geq k.$$

converges uniformly to $F(x)$ on D.

Example: For the functions

$$f_j(x) = x^j, \quad j \geq 0,$$

define the infinite series of functions

$$\sum_{j=0}^{\infty} x^j$$

on $D = (-\infty, \infty)$.

Pointwise convergence: The nth partial sum of the series is

$$F_n(x) = 1 + x + x^2 + \cdots + x^n,$$

or, in closed form,

$$F_n(x) = \begin{cases} \dfrac{1-x^{n+1}}{1-x}, & x \neq 1, \\ n+1, & x = 1. \end{cases}$$

We have seen earlier that $\{F_n\}$ converges pointwise to

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

if $|x| < 1$ and diverges if $|x| \geq 1$.

Hence, we write

$$\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}, \quad -1 < x < 1.$$

Since the di erence

$$F(x) - F_n(x) = \frac{x^{n+1}}{1-x}$$

can be made arbitrarily large by taking x close to 1,

$$\|F - F_n\|_{(-1,1)} = \infty,$$

so the convergence is not uniform on $(-1, 1)$.

We have seen earlier that $\{F_n\}$ converges pointwise to

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

if $|x| < 1$ and diverges if $|x| \geq 1$.

Neither is it uniform on any interval $(-1, r]$ with $-1 < r < 1$, since

$$\|F - F_n\|_{(-1,r)} \geq \frac{1}{2}$$

for every n on every such interval.

Example: For the functions $f_j(x) = x^j$, $j \geq 0$, discuss the uniform convergence of the in nite series of functions $\sum_{j=0} f_j(x)$.

Uniform convergence: The series does converge uniformly on any interval $[-r, r]$ with $0 < r < 1$, since

$$\|F - F_n\|_{[-r,r]} = \frac{r^{n+1}}{1-r}$$

and $\lim_{n\to\infty} r^n = 0$. Put another way, the series converges uniformly on closed subsets of $(-1, 1)$.

Uniform convergence (using ε): See video lectures.

Remark: A necessary condition for $\sum_{j=0}^{\infty} f_j(x)$ to converge on S is that $f_j(x) \to 0$ for each $x \in S$.

Remark: As for series of constants, the convergence, pointwise or uniform, of a series of functions is not changed by altering or omitting nitely many terms. This justi es adopting the convention that we used for series of constants: when we are interested only in whether a series of functions converges, and not in its sum, we will omit the limits on the summation sign and write simply $\sum f_n$.

1.8.1 Cauchy's criterion for functional series

Recall the following Theorem knows as Cauchy's convergence criterion

Theorem: A sequence of functions $\{F_n\}$ converges uniformly on a set S if and only if for each $\varepsilon > 0$ there is an integer N such that

$$\|F_n - F_m\|_S < \varepsilon \quad \text{if} \quad n, m \geq N. \tag{1.17}$$

Theorem: A series $\sum f_n$ converges uniformly on a set S if and only if for each $\varepsilon > 0$ there is an integer N such that

$$\|f_n + f_{n+1} + \cdots + f_m\|_S < \varepsilon \quad \text{if} \quad m \geq n \geq N. \tag{1.18}$$

Proof: Apply Cauchy's convergence criterion to the partial sums of $\sum f_n$, observing that

$$f_n + f_{n+1} + \cdots + f_m = F_m - F_{n-1}.$$

Theorem: A series $\sum f_n$ converges uniformly on a set S if and only if for each $\varepsilon > 0$ there is an integer N such that

$$\|f_n + f_{n+1} + \cdots + f_m\|_S < \varepsilon \quad \text{if} \quad m \geq n \geq N. \tag{1.19}$$

Corollary: If $\sum f_n$ converges uniformly on S, then $\lim_{n\to\infty} \|f_n\|_S = 0$. Setting $m = n$.

Remark: The above conditions is necessary but not su cient.

Example: We have proved that the series $\sum_{j=0} f_j(x)$, where

$$f_j(x) = x^j, \quad j \geq 0,$$

is uniformly convergent on any compact subset of $(-1, 1)$ say $[-r, r]$, where $0 < r < 1$.

Let us apply Cauchy's criterion for functional series, recall that we have

$$F_n(x) = 1 + x + x^2 + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Consider

$$\begin{aligned} |F_m - F_n| &= \left|\frac{1 - x^{n+1}}{1 - x} - \frac{1 - x^{m+1}}{1 - x}\right| = \left|\frac{x^{m+1} - x^{n+1}}{1 - x}\right| \\ &\leq \frac{2|x^{n+1}|}{1 - |x|} \\ &\leq \frac{2|r^{n+1}|}{1 - |r|}. \end{aligned}$$

We have

$$\|F_m - F_n\|_{[-r,r]} \le \frac{2|r^{n+1}|}{1-|r|}.$$

Since

$$\frac{2|r^{n+1}|}{1-|r|} \to 0 \quad \text{as} \quad n \to \infty,$$

there is an integer $N(\varepsilon)$ can be found for which

$$\frac{2|r^{n+1}|}{1-|r|} < \varepsilon, \quad \text{when} \quad n > N(\varepsilon).$$

We have

$$\|F_m - F_n\|_{[-r,r]} \le \varepsilon,$$

hence by Cauchy's criterion the series $\sum_{j=0}^{\infty} x^j$, is uniformly convergent on $[-r, r]$.

1.8.2 Dominated Series of Real Numbers for Series of Functions

Let $\{M_n\}$ be a sequence of nonnegative real numbers, and $\{F_n(x)\}$ a sequence of functions defined on the set S such that

$$|F_n(x)| \le M_n, \quad \forall x \in S \quad \text{and} \quad n \in \mathbb{N}.$$

Then the series of functions $\sum_{n=1}^{\infty} F_n(x)$ is said to be dominated on S by the series $\sum_{n=1}^{\infty} M_n$.

Example: Consider $F_n = \frac{1}{x^2+n^2}$ and the series of functions $\sum_{n=1}^{\infty} F_n$ is dominated by the series $\sum 1/n^2$ because

$$|F_n| < \frac{1}{n^2} =: M_n.$$

We know that $\sum 1/n^2 < \infty$.

1.8.3 Weierstrass M-test/dominated Convergence Test

Theorem The series $\sum f_n$ converges uniformly on S if

$$\|f_n\|_S \le M_n, \quad n \ge k, \tag{1.20}$$

where $\sum M_n < \infty$.

Proof: From Cauchy's convergence criterion for series of constants, there is for each $\varepsilon > 0$ an integer N such that

$$M_n + M_{n+1} + \cdots + M_m < \varepsilon \quad \text{if} \quad m \ge n \ge N.$$

which, because of (1.20), implies that

$$\|f_n\|_S + \|f_{n+1}\|_S + \cdots + \|f_m\|_S < \varepsilon \quad \text{if} \quad m, n \ge N.$$

$$\|f_n + f_{n+1} + \cdots + f_m\|_S < \varepsilon \quad \text{if} \quad m, n \geq N.$$

Due to Cauchy's criterion, we conclude that $\sum f_n$ converges uniformly on S.

Recall the following necessary condition for uniform convergence:

If $\sum f_n$ converges uniformly on S, then $\lim_{n\to\infty} \|f_n\|_S = 0$.

Example: Check the uniform convergence of the following series of functions

- $\sum \frac{1}{x^2+n^2}$.
- $\sum \frac{\sin nx}{n^2}$

Solution: We have

$$\frac{1}{x^2+n^2} \leq \frac{1}{n^2}, \qquad \left|\frac{\sin nx}{n^2}\right| \leq \frac{1}{n^2}.$$

Taking $M_n = 1/n^2$ and recalling that

$$\sum \frac{1}{n^2} < \infty.$$

Due to Weierstrass M-test, we can conclude

$$\sum \frac{1}{x^2+n^2} \quad \text{and} \quad \sum \frac{\sin nx}{n^2}$$

converge uniformly on $(-\infty, \infty)$.

Example: Check the uniform convergence of the series

$$\sum f_n(x) = \sum \left(\frac{x}{1+x}\right)^n.$$

Solution: The given series converges uniformly on any set S such that

$$\left|\frac{x}{1+x}\right| \leq r < 1, \quad x \in S. \tag{1.21}$$

For such a set S, we have $\|f_n\|_S \leq r^n$.
By Weierstrass's test applies, with $\sum M_n = \sum r^n < \infty$.
Since (1.21) is equivalent to

$$\frac{-r}{1+r} \leq x \leq \frac{r}{1-r}, \quad x \in S,$$

this means that the series converges uniformly on any compact subset of $(-1/2, \infty)$.

Example: Check the uniform convergence of the series

$$\sum f_n(x) = \sum \left(\frac{x}{1+x}\right)^n.$$

Solution: See the solution in video lecture.

Recall: If $\sum f_n$ converges uniformly on S, then $\lim_{n\to\infty} \|f_n\|_S = 0$. The series does not converge uniformly on $S = (-1/2, b)$ with $b < \infty$ or on $S = [a, \infty)$ with $a > -1/2$, because in these cases $\|f_n\|_S = 1$ for all n.

Absolute convergence: A series of functions $\sum f_n$ is said to converge absolutely on S if $\sum |f_n|$ converges pointwise on S, and converges absolutely uniformly on S if $\sum |f_n|$ converges uniformly on S.

Remarks:

- The condition of absolutely convergence (pointwise or uniform) is stronger than the usual convergence (pointwise or uniform)
- In our proof of Weierstrass's M-test, we actually proved that $\sum f_n$ converges absolutely uniformly on S.
- Show that if a series converges absolutely uniformly on S, then it converges uniformly on S.

Theorem: The series

$$\sum_{n=k}^{\infty} f_n g_n$$

converges uniformly on S if $\{f_n\}$ converges uniformly to zero on S, $\sum (f_{n+1} - f_n)$ converges absolutely uniformly on S, and

$$\|g_k + g_{k+1} + \cdots + g_n\|_S \leq M, \quad n \geq k, \tag{1.22}$$

for some constant M.

Proof: Let

$$G_n = g_k + g_{k+1} + \cdots + g_n,$$

and consider the partial sums of $\sum_{n=k}^{\infty} f_n g_n$:

$$H_n = f_k g_k + f_{k+1} g_{k+1} + \cdots + f_n g_n. \tag{1.23}$$

By substituting $g_k = G_k$ and $g_n = G_n - G_{n-1}, \quad n \geq k+1,$

into (1.23), we obtain

$$H_n = f_k G_k + f_{k+1}(G_{k+1} - G_k) + \cdots + f_n(G_n - G_{n-1}).$$

Which we rewrite as

$$H_n = (f_k - f_{k+1})G_k + (f_{k+1} - f_{k+2})G_{k+1} + \cdots + (f_{n-1} - f_n)G_{n-1} + f_n G_n,$$

or

$$H_n = J_{n-1} + f_n G_n, \tag{1.24}$$

where

$$J_{n-1} = (f_k - f_{k+1})G_k + (f_{k+1} - f_{k+2})G_{k+1} + \cdots + (f_{n-1} - f_n)G_{n-1}. \tag{1.25}$$

That is, $\{J_n\}$ is the sequence of partial sums of the series

$$\sum_{j=k}^{\infty} (f_j - f_{j+1})G_j. \tag{1.26}$$

From (1.22) and the de nition of G_j,

$$\left|\sum_{j=n}^{m} [f_j(x) - f_{j+1}(x)]G_j(x)\right| \le M \sum_{j=n}^{m} |f_j(x) - f_{j+1}(x)|, \quad x \in S,$$

so

$$\left\|\sum_{j=n}^{m} (f_j - f_{j+1})G_j\right\|_S \le M \left\|\sum_{j=n}^{m} |f_j - f_{j+1}|\right\|_S.$$

Now suppose that $\varepsilon > 0$. Since $\sum (f_j - f_{j+1})$ converges absolutely uniformly on S, Cauchy's convergence criterion implies that there is an integer N such that the right side of the last inequality is less than ε if $m \ge n \ge N$. The same is then true of the left side, so Cauchy's convergence criterion implies that (1.26) converges uniformly on S.

We have now shown that $\{J_n\}$ as de ned in (1.25) converges uniformly to a limit function J on S. Returning to (1.24), we see that

$$H_n - J = J_{n-1} - J + f_n G_n.$$

Hence, we have

$$\begin{aligned} \|H_n - J\|_S &\le \|J_{n-1} - J\|_S + \|f_n\|_S \|G_n\|_S \\ &\le \|J_{n-1} - J\|_S + M\|f_n\|_S. \end{aligned}$$

Since $\{J_{n-1} - J\}$ and $\{f_n\}$ converge uniformly to zero on S, it now follows that $\lim_{n\to\infty} \|H_n - J\|_S = 0$. Therefore, $\{H_n\}$ converges uniformly on S.

Corollary: The series $\sum_{n=k}^{\infty} f_n g_n$ converges uniformly on S if

$$f_{n+1}(x) \le f_n(x), \quad x \in S, \quad n \ge k,$$

$\{f_n\}$ converges uniformly to zero on S, and

$$\|g_k + g_{k+1} + \cdots + g_n\|_S \le M, \quad n \ge k,$$

for some constant M.

Example: Consider the series

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

with $f_n = 1/n$ (constant), $g_n(x) = \sin nx$, and

$$G_n(x) = \sin x + \sin 2x + \cdots + \sin nx.$$

We have

$$|G_n(x)| \leq \frac{1}{|\sin(x/2)|}, \quad n \geq 1, \quad x \neq 2k\pi \qquad (k = \text{integer}).$$

Therefore, $\{\|G_n\|_S\}$ is bounded, and the series converges uniformly on any set S on which $\sin x/2$ is bounded away from zero.

Example: For example, if $0 < \delta < \pi$, then

$$\left|\sin \frac{x}{2}\right| \geq \sin \frac{\delta}{2}$$

if x is at least δ away from any multiple of 2π; hence, the series converges uniformly on

$$S = \bigcup_{k=-\infty}^{\infty} [2k\pi + \delta, 2(k+1)\pi - \delta].$$

Since

$$\sum \left|\frac{\sin nx}{n}\right| = \infty, \quad x \neq k\pi.$$

This result cannot be obtained from Weierstrass's test.

Example: The series

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

satisfies the hypotheses of Corollary on $(-\infty, \infty)$, with

$$f_n(x) = \frac{1}{n + x^2}, \quad g_n = (-1)^n, \quad G_{2m} = 0, \quad \text{and} \quad G_{2m+1} = -1.$$

Therefore, the series converges uniformly on $(-\infty, \infty)$. This result cannot be obtained by Weierstrass's test, since

$$\sum \frac{1}{n + x^2} = \infty$$

for all x.

Recall the following result:

Theorem: If $\{F_n\}$ converges uniformly to F on S and each F_n is continuous at a point x_0 in S, then so is F. Similar statements hold for continuity from the right and left.

Theorem: If $\sum_{n=k} f_n$ converges uniformly to F on S and each f_n is continuous at a point x_0 in S, then so is F. Similar statements hold for continuity from the right and left.

Proof: See Lecture.

Recall the following: Theorem: If $\sum_{n=k} f_n$ converges uniformly to F on S and each f_n is continuous at a point x_0 in S, then so is F. Similar statements hold for continuity from the right and left.

Example: Recall, we have proved that the series

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

converges uniformly on every compact subset of $(-1/2, \infty)$.

Since the terms of the series are continuous on every such subset, implies that F is also.

In fact, we can state a stronger result: F is continuous on $(-1/2, \infty)$, since every point in $(-1/2, \infty)$ lies in a compact subinterval of $(-1/2, \infty)$.

Example: Show that the function

$$G(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n}$$

is continuous except perhaps at $x_k = 2k\pi$ (k = integer).

We have seen that the series $\sum_{n=1}^{\infty} \frac{\sin nx}{n}$ is uniformly convergent by applying Dirichlet's Test for Uniform Convergence except at $x_k = 2k\pi$ (k = integer).

Example: The function

$$H(x) = \sum_{n=1}^{\infty} (-1)^n \frac{1}{n+x^2}$$

is continuous for all x.

Theorem: Suppose that $\{F_n\}$ is a sequence of Riemann integrable functions de ned on an interval $[a, b]$. If $\{F_n\}$ converges uniformly on $[a, b]$ to F, then F is Riemann integrable on $[a, b]$, and

$$\lim_{n\to\infty} \int_a^b F_n(x)dx = \int_a^b F(x)dx.$$

For each $t \in [a, b]$

$$\int_a^t F_n(x)dx,$$

converges uniformly on $[a, b]$ to

$$\int_a^t F(x)dx.$$

Proof: We need to show that the function F is integrable on $[a, b]$. Observe that the following statements holds:

- F_n is bounded, because each F_n is integrable on $[a, b]$.
- F is bounded, because

$$|F(x)| \leq |F_n(x) - F(x)| + |F_n(x)| \leq \delta_n + |F_n(x)|,$$

where $\delta_n = \sup_{x\in[a,b]} |F_n(x) - F(x)|$.

- Since F_n converges uniformly to F, for every $\varepsilon > 0$, there exists an N such that

$$|F_n(x) - F(x)| < \frac{\varepsilon}{3(b-a)}, \quad \text{for all} \quad x \in [a, b], n > N.$$

Also, F_n is integrable, there exists a partition P of $[a, b]$ such that

$$S(P, F_n) - s(P, F_n) < \frac{\varepsilon}{3}.$$

For each $x \in [a, b]$ with $n = N$

$$|F_n(x) - F(x)| < \frac{\varepsilon}{3(b-a)}, \quad \text{for all} \quad x \in [a, b], n > N,$$

implies that

$$F_n(x) - \frac{\varepsilon}{3(b-a)} < F(x) < F_n(x) + \frac{\varepsilon}{3(b-a)}.$$

Therefore,

$$s(P, F_n) - \frac{\varepsilon}{3} < s(P, F) \leq S(P, F) < S(P, F_n) + \frac{\varepsilon}{3}$$

Hence F is integrable. Finally, for $n \geq N$ and for each $t \in [a, b]$, we have

$$\begin{aligned} \left|\int_a^t F_n(x)dx - \int_a^t F(x)dx\right| &\leq \int_a^t |F_n(x) - F(x)|dx \\ &\leq \frac{\varepsilon(b-a)}{3(b-a)}, \quad \text{for all} \quad x \in [a, b], n > N. \end{aligned}$$

Remark: The limit of a uniformly convergent series of integrable functions is integrable, and so term-by-term integration is permissible for such a series.

Theorem: Suppose that $\sum_{n=k} f_n$ converges uniformly to F on $S = [a, b]$. Assume that F and f_n, $n \geq k$, are integrable on $[a, b]$. Then

$$\int_a^b F(x)\,dx = \sum_{n=k}^{\infty} \int_a^b f_n(x)\,dx.$$

We say in this case that $\sum_{n=k}^{\infty} f_n$ can be integrated term by term over $[a, b]$.

Example: Consider the $\{F_n\}$ de ned by

$$F_n(x) = \frac{x}{1 + n_{x^2}}, \qquad x \in [a, b] \subset \mathbb{R}.$$

Then Weieretrass's M-test shows that $\sum F_n$ converges uniformly on $[a, b]$

Consequently, term-by-term integration is permissible in this series.

Example: Consider the following

$$\frac{1}{1 - x} = \sum_{n=0}^{\infty} x^n, \qquad -1 < x < 1.$$

The series converges uniformly, and the limit function is integrable on any closed subinterval $[a, b]$ of $(-1, 1)$.

Hence,

$$\int_a^b \frac{dx}{1 - x} = \sum_{n=0}^{\infty} \int_a^b x^n\,dx.$$

Consequently,

$$\log(1 - a) - \log(1 - b) = \sum_{n=0}^{\infty} \frac{b^{n+1} - a^{n+1}}{n + 1}.$$

Remark: We have seen that

$$\log(1 - a) - \log(1 - b) = \sum_{n=0}^{\infty} \frac{b^{n+1} - a^{n+1}}{n + 1}.$$

Letting $a = 0$ and $b = x$ yields

$$\log(1 - x) = -\sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1}, \quad -1 < x < 1.$$

Example: Evaluate the

$$\sum_{n=1}^{\infty} \int_0^1 x(e^x - 1)e^{-nx}dx.$$

Solution: The sequence of partial sum is

$$F_n = \sum_{k=1}^{n} x(e^x - 1)e^{-kx}dx.$$

Observe that $F_n(0) = 0$ and for $x > 0$

$$F_n(x) = x(e^x - 1)\frac{e^{-x}(1 - e^{-nx})}{1 - e^{-x}},$$

$$F_n(x) = x(1 - e^{-nx}).$$

Example: Evaluate the

$$\sum_{n=1}^{\infty} \int_0^1 x(e^x - 1)e^{-nx}dx.$$

Solution: For the function xe^{-nx}, we have seen that it attains its maximum at $x = 1/n$, we have

$$\|F_n(x) - x\| = \sup_{x\geq 0} |F_n(x) - x|$$

$$\|F_n(x) - x\| = \sup_{x\geq 0} |xe^{-nx}| = \frac{1}{en}.$$

So, as $n \to \infty$, we have $\|F_n(x) - x\| \to 0$.

Example: Evaluate the

$$\sum_{n=1}^{\infty} \int_0^1 x(e^x - 1)e^{-nx}dx.$$

Solution: The series of functions

$$\sum_{n=1}^{\infty} x(e^x - 1)e^{-nx}dx,$$

converges uniformly to $F(x) = x$.

Applying the theorem of interchange of sum and integral sign, we can conclude that

$$\sum_{n=1}^{\infty} \int_0^1 x(e^x - 1)e^{-nx}dx = \int_0^1 \sum_{n=1}^{\infty} x(e^x - 1)e^{-nx}dx = \int_0^1 xdx.$$

Example: Consider

$$F_n(x) = \frac{x}{1 + nx^2}, \qquad x \in \mathbb{R}.$$

$$|F_n(x)| = \left|\frac{x}{1 + nx^2}\right| \leq \frac{|x|}{2\sqrt{n}|x|} = \frac{1}{2\sqrt{n}}.$$

$F_n(x)$ is uniformly convergent to $F(x) = 0$ on $\mathbb{R}$. We have

$$F_n'(x) = \frac{1 - nx^2}{(1 + nx^2)^2}$$

When $x = 0$, we have $\lim_{n\to\infty} F_n'(x) = 0$ and for x 0 $\lim_{n\to\infty} F_n'(x) = 1$.

Remark: What we have observed in this example is:

- We have a sequence of di erentiable functions $\{F_n\}$ de ned on S.
- F_n converges uniformly to F on S.
- F is di erentiable on S.
- There exists $x \in S$ with $F'(x) \neq \lim_{n\to\infty} F_n'(x)$, because $F_n'(0) \to 1 \neq F'(0)$.

Thus, even if the limit of a uniformly convergent sequence (respectively series) of di erentiable functions on S is di erentiable on S, it may happen that the derivative of the limit is not the limit of the sequence (respectively sequence of partial sums) of derivatives of the di erentiable functions.

Theorem: Suppose that f_n is a sequence of functions such that:

- f_n is continuously di erentiable on $[a, b]$ for each $n \geq k$, i.e., $f_n \in C^1[a, b]$.
- $\sum_{n=k}^{\infty} f_n(x_0)$ converges for some x_0 in $[a, b]$.

- $\sum_{n=k}^{\infty} f_n'$ converges uniformly on $[a, b]$.

  Then $\sum_{n=k}^{\infty} f_n$ converges uniformly on $[a, b]$ to a di erentiable function F, such that

$$F'(x) = \sum_{n=k}^{\infty} f_n'(x), \quad a < x < b,$$

$$\text{while} \quad F'(a+) = \sum_{n=k}^{\infty} f_n'(a+) \quad \text{and} \quad F'(b-) = \sum_{n=k}^{\infty} f_n'(b-).$$

Proof: Since f_n' is uniformly convergent to g on any closed interval contained in $[a, b]$, say in an interval with endpoints x_0 and x, $x \in [a, b]$. Thus, for all $x \in [a, b]$, we have

$$\int_{x_0}^{x} g(t)dt = \lim_{n\to\infty} \int_{x_0}^{x} f_n'(t)dt.$$

Recall the fundamental theorem of calculus, we have

$$\int_{x_0}^{x} g(t)dt = \lim_{n\to\infty} (f_n(x) - f_n(x_0)).$$

Recall the $\lim_{n\to\infty} f_n(x_0)$ exists (given hypothesis), we can obtain

$$\int_{x_0}^{x} g(t)dt + \lim_{n\to\infty} f_n(x_0) = \lim_{n\to\infty} f_n(x), \quad \text{on} \quad [a, b].$$

The above convergence is uniform. By setting $F(x) = \lim_{n\to\infty} f_n(x)$, we have

$$\int_{x_0}^{x} g(t)dt + \lim_{n\to\infty} f_n(x_0) = F(x), \quad \text{on} \quad [a, b].$$

Now, g, being the limit of a uniformly convergent sequence of continuous functions on $[a, b]$, is continuous on $[a, b]$.

Recall the second fundamental theorem of calculus with $G(x) = \int_{x_0}^{x} g(t)dt$ is di erentiable and $G'(x) = g(x)$ on $[a, b]$.

Therefore, we have

$$F'(x) = g(x), \qquad F'(x) = \lim_{n\to\infty} f_n'(x), \quad \text{on} \quad [a, b].$$

Remark: The series $\sum_{n=k}^{\infty} f_n$ can be di erentiated term by term on $[a, b]$.

How to apply this result?

- We first verify that $\sum_{n=k}^{\infty} f_n(x_0)$ converges for some x_0 in $[a, b]$.
- Then differentiate $\sum_{n=k}^{\infty} f_n$ term by term. If the resulting series converges uniformly. Then term by term differentiation was legitimate.

Example: The series

$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n} \cos \frac{x}{n} \tag{1.27}$$

converges at $x_0 = 0$. Differentiating term by term yields the series

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2} \sin \frac{x}{n} \tag{1.28}$$

of continuous functions. This series converges uniformly on $(-\infty, \infty)$, by Weierstrass's test. Consequently, the series (1.27) converges uniformly on every finite interval to the differentiable function

$$F(x) = \sum_{n=1}^{\infty} (-1)^n \frac{1}{n} \cos \frac{x}{n}, \quad -\infty < x < \infty,$$

$$F'(x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2} \sin \frac{x}{n}, \quad -\infty < x < \infty.$$

Example: Consider the series

$$E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots. \tag{1.29}$$

The series converges uniformly on every interval $[-r, r]$ by Weierstrass's test, because

$$\frac{|x|^n}{n!} \le \frac{r^n}{n!}, \quad |x| \le r,$$

$$\sum \frac{r^n}{n!} < \infty$$

for all r, by the ratio test.

Differentiating the right side of (1.30) term by term yields the series

$$\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

which is the same as (1.30).

Example: Consider the series

$$E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots. \tag{1.30}$$

Therefore, the di erentiated series is also uniformly convergent on $[-r, r]$ for every r, so the term by term di erentiation is legitimate and

$$E'(x) = E(x), \quad -\infty < x < \infty.$$

This is not surprising if you recognize that $E(x) = e^x$.

Remark: Failure to verify that the given series converges at some point can lead to erroneous conclusions.

Example: For example, di erentiating

$$\sum_{n=1}^{\infty} \cos \frac{x}{n} \tag{1.31}$$

term by term.

We have

$$\sum_{n=1}^{\infty} \frac{1}{n} \sin \frac{x}{n}.$$

Since

$$\left| \frac{1}{n} \sin \frac{x}{n} \right| \le \frac{|x|}{n^2} \le \frac{r}{n^2}, \qquad |x| \le r,$$

and $\sum 1/n^2 < \infty$. which converges uniformly on $[-r, r]$ for every r,

We cannot conclude from this that (1.31) converges uniformly on $[-r, r]$. In fact, it diverges for every x.

1.9 Power Series

An in nite series of the form

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n, \tag{1.32}$$

where x_0 and $a_0, a_1, \ldots,$ are constants, is called a power series in $x - x_0$. If $x_0 = 0$ then power series becomes

$$\sum_{n=0}^{\infty} a_n x^n.$$

Theorem: The radius of convergence of $\sum a_n (x - x_0)^n$ is given by

$$\frac{1}{R} = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|.$$

if the limit exists in the extended real number system.

Theorem: For the power series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$, define R in the extended real numbers by

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n}. \tag{1.33}$$

Theorem: A power series

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

with positive radius of convergence R is continuous and differentiable in its interval of convergence, and its derivative can be obtained by differentiating term by term; that is,

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}, \tag{1.34}$$

which can also be written as

$$f'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n. \tag{1.35}$$

This series also has radius of convergence R.

Proof: Since

$$\begin{aligned}
\limsup_{n\to\infty}((n+1)|a_n|)^{1/n} &= \limsup_{n\to\infty}(n+1)^{1/n}|a_n|^{1/n} \\
&= \left(\lim_{n\to\infty} (n+1)^{1/n}\right)\left(\limsup_{n\to\infty} |a_n|^{1/n}\right) \\
&= \left[\lim_{n\to\infty} \exp\left(\frac{\log(n+1)}{n}\right)\right] \\
&\quad \left(\limsup_{n\to\infty} |a_n|^{1/n}\right) = \frac{e^0}{R} = \frac{1}{R},
\end{aligned}$$

the radius of convergence of the power series obtained by term by term differentiation is R. Therefore, the power series in

$$f'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n,$$

converges uniformly in every interval $[x_0 - r, x_0 + r]$ such that $0 < r < R$.

The term by term differentiation is valid for the power series and the series

$$f'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n,$$

converges uniformly for all x in $(x_0 - R, x_0 + R)$.

Theorem: A power series

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

with positive radius of convergence R has derivatives of all orders in its interval of convergence, which can be obtained by repeated term by term differentiation. That is,

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n(x-x_0)^{n-k}. \tag{1.36}$$

The radius of convergence of each of these series is R.

Proof:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n(x-x_0)^{n-k}.$$

The proof is by induction. The assertion is true for $k = 1$, by the Theorem we proved in previous module.

Suppose that it is true for some $k \geq 1$. By shifting the index of summation, we can write

$$f^{(k)}(x) = \sum_{n=0}^{\infty} (n+k)(n+k-1)\cdots(n+1)a_{n+k}(x-x_0)^n, \qquad |x-x_0| < R.$$

Defining

$$b_n = (n+k)(n+k-1)\cdots(n+1)a_{n+k}. \tag{1.37}$$

We rewrite this as

$$f^{(k)}(x) = \sum_{n=0}^{\infty} b_n(x-x_0)^n, \qquad |x-x_0| < R.$$

By Theorem of term by term differentiation of power series, we can differentiate this series term by term to obtain

$$f^{(k+1)}(x) = \sum_{n=1}^{\infty} nb_n(x-x_0)^{n-1}, \qquad |x-x_0| < R.$$

Substituting from (1.37) for b_n for $|x-x_0| < R$ yields

$$f^{(k+1)}(x) = \sum_{n=1}^{\infty} (n+k)(n+k-1)\cdots(n+1)na_{n+k}(x-x_0)^{n-1}.$$

Shifting the summation index yields

$$f^{(k+1)}(x) = \sum_{n=k+1}^{\infty} n(n-1)\cdots(n-k)a_n(x-x_0)^{n-k-1}, \qquad |x-x_0| < R,$$

which is (1.42) with k replaced by $k+1$. This completes the induction.

Example: We have proved that

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \qquad |x| < 1.$$

Repeated di erentiation yields

$$\frac{k!}{(1-x)^{k+1}} = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)x^{n-k}$$

$$= \sum_{n=0}^{\infty} (n+k)(n+k-1)\cdots(n+1)x^n, \qquad |x| < 1,$$

$$\frac{1}{(1-x)^{k+1}} = \sum_{n=0}^{\infty} \binom{n+k}{k} x^n, \qquad |x| < 1.$$

Example: Show that the series

$$S(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad \text{and} \quad C(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

converges for all x.

Di erentiating yields

$$S'(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{(2n)!} = C(x)$$

and

$$C'(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n-1}}{(2n-1)!} = -\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = -S(x).$$

These results should not surprise you if you recall that

$$S(x) = \sin x \quad \text{and} \quad C(x) = \cos x.$$

Theorem: If

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n, \qquad |x-x_0| < R,$$

then

$$a_n = \frac{f^{(n)}(x_0)}{n!}.$$

Proof: We have

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n(x-x_0)^{n-k}.$$

Setting $x = x_0$ in the above equation yields

$$f^{(k)}(x_0) = k!a_k.$$

Theorem: If

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=0}^{\infty} b_n(x-x_0)^n \tag{1.38}$$

for all x in some interval $(x_0 - r, x_0 + r)$, then

$$a_n = b_n, \quad n \geq 0. \tag{1.39}$$

Proof: Let $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ and $g(x) = \sum_{n=0}^{\infty} b_n(x-x_0)^n$.
From previous result, we have

$$a_n = \frac{f^{(n)}(x_0)}{n!} \quad \text{and} \quad b_n = \frac{g^{(n)}(x_0)}{n!}. \tag{1.40}$$

From (1.38), $f = g$ in $(x_0 - r, x_0 + r)$. Therefore,

$$f^{(n)}(x_0) = g^{(n)}(x_0), \quad n \geq 0.$$

This and (1.40) imply (1.39).

Theorem (Recall the following): For the power series, de ne R in the extended real numbers by

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n}. \tag{1.41}$$

In particular, $R = 0$ if $\limsup_{n\to\infty} |a_n|^{1/n} = \infty$, and $R = \infty$ if $\limsup_{n\to\infty} |a_n|^{1/n} = 0$.

Then the power series converges

1. only for $x = x_0$ if $R = 0$;

2. for all x if $R = \infty$, and absolutely uniformly in every bounded set;

3. for x in $(x_0 - R, x_0 + R)$ if $0 < R < \infty$, and absolutely uniformly in every closed subset of this interval.

Remark: The series diverges if $|x - x_0| > R$. No general statement can be made concerning convergence at the endpoints $x = x_0 + R$ and $x = x_0 - R$: the series may converge absolutely or conditionally at both, converge conditionally at one and diverge at the other, or diverge at both.

Theorem (Recall the following): Suppose that $\sum_{=k} f_n$ converges uniformly to F on $S = [a, b]$. Assume that F and f_n, $n \geq k$, are integrable on $[a, b]$.

Then
$$\int_a^b F(x)\,dx = \sum_{n=k}^{\infty} \int_a^b f_n(x)\,dx.$$

Theorem: If x_1 and x_2 are in the interval of convergence of
$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n,$$
then
$$\int_{x_1}^{x_2} f(x)\,dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} \left[(x_2 - x_0)^{n+1} - (x_1 - x_0)^{n+1} \right];$$
that is, a power series may be integrated term by term between any two points in its interval of convergence.

Proof: See Lecture.

Some questions related to Power Series.

[illegible]s.

- We discussed, what are the properties of its sum.

- What properties guarantee that a given function f can be represented as the sum of a convergent power series in $x - x_0$?

Recall the following:

Theorem: A power series
$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$
with positive radius of convergence R has derivatives of all orders in its interval of convergence, which can be obtained by repeated term by term di erentiation; thus,
$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n(x - x_0)^{n-k}. \tag{1.42}$$

The radius of convergence of each of these series is R. If

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n, \quad |x - x_0| < R,$$

then

$$a_n = \frac{f^{(n)}(x_0)}{n!}.$$

1.10 The Taylor's Series

The only power series in $x - x_0$ that can possibly converge to f in such a neighborhood is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \tag{1.43}$$

This is called the Taylor series of f about x_0 (also, the Maclaurin series of f, if $x_0 = 0$). The mth partial sum of (1.43) is the Taylor polynomial

$$T_m(x) = \sum_{n=0}^{m} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

Remark: The Taylor series of an infinitely differentiable function f may converge to a sum different from f.

Example: Consider the function

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

the function f is infinitely times differentiable on $(-\infty, \infty)$ and $f^{(n)}(0) = 0$ for $n \geq 0$. So its Maclaurin series is identically zero.

Taylor's theorem: If f is infinitely differentiable on (a, b) and x and x_0 are in (a, b) then, for every integer $n \geq 0$,

$$f(x) - T_n(x) = \frac{f^{(n+1)}(c_n)}{(n+1)!}(x - x_0)^{n+1}, \tag{1.44}$$

where c_n is between x and x_0.

Therefore,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

for an x in (a, b) if and only if

$$\lim_{n\to\infty} \frac{f^{(n+1)}(c_n)}{(n+1)!}(x - x_0)^{n+1} = 0.$$

Remark: It is not always easy to check this condition, because the sequence $\{c_n\}$ is usually not precisely known, or even uniquely de ned; however, the next theorem is su ciently general to be useful.

Theorem: Suppose that f is in nitely di erentiable on an interval I and

$$\lim_{n\to\infty} \frac{r^n}{n!} \|f^{(n)}\|_I = 0. \tag{1.45}$$

Then, if $x_0 \in I^0$, the Taylor series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

converges uniformly to f on

$$I_r = I \cap [x_0 - r, x_0 + r].$$

Proof: We know that

$$f(x) - T_n(x) = \frac{f^{(n+1)}(c_n)}{(n+1)!}(x - x_0)^{n-1},$$

$$\|f - T_n\|_{I_r} \le \frac{r^{n+1}}{(n+1)!}\|f^{(n+1)}\|_{I_r} \le \frac{r^{n+1}}{(n+1)!}\|f^{(n+1)}\|_I,$$

so (1.45) implies the conclusion.

Example:

$$= 0, \quad 0 < r < \infty$$

holds for all r.

Since

Apply the previous theorem, with $I = (-\infty, \infty)$, $x_0 = 0$, and r arbitrary. We have the the well known series expansion of $\sin x$, that is,

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad -\infty < x < \infty,$$

and the convergence is uniform on bounded sets.

Example: A similar argument shows that

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad -\infty < x < \infty,$$

with uniform convergence on bounded sets.

Example: If $f(x) = e^x$, then $f^{(k)}(x) = e^x$ and $\|f^{(k)}\|_I = e^r$, $k \geq 0$, if $I = [-r, r]$.
Since

$$\lim_{n\to\infty} \frac{r^n}{n!} e^r = 0.$$

we conclude that

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad -\infty < x < \infty,$$

with uniform convergence on bounded sets.

Example: If $f(x) = (1+x)^q$, then

$$\frac{f^{(n)}(x)}{n!} = \binom{q}{n} (1+x)^{q-n}, \quad \text{so} \quad \frac{f^{(n)}(0)}{n!} = \binom{q}{n}. \tag{1.46}$$

The Maclaurin series

$$\sum_{n=0}^{\infty} \binom{q}{n} x^n$$

is called the binomial series. We saw in Analysis I that this series equals $(1+x)^q$ for all x if q is a nonnegative integer.

Example: We will now show that if q is an arbitrary real number, then

$$\sum_{n=0}^{\infty} \binom{q}{n} x^n = f(x) = (1+x)^q, \quad 0 \leq x < 1. \tag{1.47}$$

Since

$$\lim_{n\to\infty} \left| \binom{q}{n+1} \Big/ \binom{q}{n} \right| = \lim_{n\to\infty} \left| \frac{q-n}{n+1} \right| = 1,$$

the radius of convergence of the series in (1.47) is 1.
From (1.46),

$$\frac{\|f^{(n)}\|_{[0,1]}}{n!} \leq [\max(1, 2^q)] \left| \binom{q}{n} \right|, \quad n \geq 0.$$

Example: Therefore, if $0 < r < 1$,

$$\limsup_{n\to\infty} \frac{r^n}{n!} \|f^{(n)}\|_{[0,1]} \leq [\max(1, 2^q)] \lim_{n\to\infty} \left| \binom{q}{n} \right| r^n = 0,$$

where the last equality follows from the absolute convergence of the series in (1.47) on $(-1, 1)$.

Theorem: If

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad |x - x_0| < R_1, \tag{1.48}$$

$$g(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n, \quad |x - x_0| < R_2, \tag{1.49}$$

and α and β are constants, then

$$\alpha f(x) + \beta g(x) = \sum_{n=0}^{\infty} (\alpha a_n + \beta b_n)(x - x_0)^n, \quad |x - x_0| < R,$$

where $R \geq \min\{R_1, R_2\}$.

Proof: See the video lectures.

Recall the following theorem:

Theorem: If $\sum_{n=0}^{\infty}$ [illegible] and $\sum$ [illegible] the Cauchy product of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ converges absolutely to AB.

Theorem: If f and g are given by power series

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad |x - x_0| < R_1,$$

$$g(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n, \quad |x - x_0| < R_2,$$

then

$$f(x)g(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n, \quad |x - x_0| < R, \tag{1.50}$$

$$c_n = \sum_{r=0}^{n} a_r b_{n-r} = \sum_{r=0}^{n} a_{n-r} b_r$$

and $R \geq \min\{R_1, R_2\}$.

Proof: Suppose that $R_1 \leq R_2$.
Since the series

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad |x - x_0| < R_1,$$

$$g(x) = \sum_{n=0}^{\infty} b_n(x - x_0)^n, \quad |x - x_0| < R_2,$$

converge absolutely to $f(x)$ and $g(x)$.

If $|x - x_0| < R_1$, their Cauchy product converges to $f(x)g(x)$ if $|x - x_0| < R_1$, by product of series.

The nth term of this product is

$$\sum_{r=0}^{n} a_r(x - x_0)^r b_{n-r}(x - x_0)^{n-r} = \left(\sum_{r=0}^{n} a_r b_{n-r}\right)(x - x_0)^n = c_n(x - x_0)^n.$$

Example: If

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \quad |x| < 1,$$
$$g(x) = \sum_{n=0}^{\infty} b_n x^n, \quad |x| < R,$$
$$\frac{g(x)}{1-x} = \sum_{n=0}^{\infty} s_n x^n, \quad |x| < \min\{1, R\},$$

where

$$\begin{aligned} s_n &= (1)b_0 + (1)b_1 + \cdots + (1)b_n \\ &= b_0 + b_1 + \cdots + b_n. \end{aligned}$$

Example: We have already discussed

$$(1 + x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n, \quad |x| < 1.$$

Also

$$(1 + x)^q = \sum_{n=0}^{\infty} \binom{q}{n} x^n, \quad |x| < 1.$$

Since

$$(1 + x)^p(1 + x)^q = (1 + x)^{p+q} = \sum_{n=0}^{\infty} \binom{p+q}{n} x^n,$$

while the Cauchy product is $\sum_{n=0}^{\infty} c_n x^n$, with

$$c_n = \sum_{r=0}^{n} \binom{p}{r}\binom{q}{n-r}.$$

Join VU WhatsApp Group:
https://chat.whatsapp.com/JURZWfAElue6qixpzkfduE

Product of power series implies that

$$c_n = \binom{p+q}{n}.$$

This yields the identity

$$\binom{p+q}{n} = \sum_{r=0}^{n} \binom{p}{r}\binom{q}{n-r},$$

valid for all p and q. The quotient

$$f(x) = \frac{h(x)}{g(x)} \tag{1.51}$$

of two power series

$$\begin{aligned} h(x) &= \sum_{n=0}^{\infty} c_n(x-x_0)^n, \quad |x-x_0| < R_1, \\ g(x) &= \sum_{n=0}^{\infty} b_n(x-x_0)^n, \quad |x-x_0| < R_2, \end{aligned}$$

can be represented as a power series

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n \tag{1.52}$$

with a positive radius of convergence, provided that

$$b_0 = g(x_0) \neq 0.$$

This is surely plausible. Since $g(x_0) \neq 0$ and g is continuous near x_0, the denominator of (1.51) di ers from zero on an interval about x_0. Therefore, f has derivatives of all orders on this interval, because g and h do.

Since

$$f(x)g(x) = h(x),$$

The result about the product of Power series implies that

$$\sum_{r=0}^{n} a_r b_{n-r} = c_n, \quad n \geq 0.$$

Solving these equations successively yields

$$\begin{aligned} a_0 &= \frac{c_0}{b_0}, \\ a_n &= \frac{1}{b_0}\left(c_n - \sum_{r=0}^{n-1} b_{n-r}a_r\right), \quad n \geq 1. \end{aligned}$$

Remark: It is not worthwhile to memorize these formulas. Rather, it is usually better to view the procedure as follows: Multiply the series f (with unknown coe cients) and g according to the procedure of Theorem ??, equate the resulting coe cients with those of h, and solve the resulting equations successively for a_0, a_1, $\dots$.

Example: Suppose that we wish to nd the coe cients in the Maclaurin series

$$\tan x = a_0 + a_1 x + a_2 x^2 + \cdots .$$

We rst observe that since $\tan x$ is an odd function, its derivatives of even order vanish at $x_0 = 0$, so $a_{2m} = 0$, $m \geq 0$. Therefore,

$$\tan x = a_1 x + a_3 x^3 + a_5 x^5 + \cdots .$$

Since

$$\tan x = \frac{\sin x}{\cos x},$$

it follows from series of $\sin x$ and $\cos x$ that

$$a_1 x + a_3 x^3 + a_5 x^5 + \cdots = \frac{x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots}{1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots}$$

so

$$(a_1 x + a_3 x^3 + a_5 x^5 + \cdots)\left(1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots\right) = x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots,$$

or

$$a_1 x + \left(a_3 - \frac{a_1}{2}\right) x^3 + \left(a_5 - \frac{a_3}{2} + \frac{a_1}{24}\right) x^5 + \cdots = x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots .$$

Comparing coe cients of like powers of x on the two sides of this equation must be equal; hence,

$$a_1 = 1, \quad a_3 - \frac{a_1}{2} = -\frac{1}{6}, \quad a_5 - \frac{a_3}{2} + \frac{a_1}{24} = \frac{1}{120}$$

$$a_1 = 1, \quad a_3 = -\frac{1}{6} + \frac{1}{2}(1) = \frac{1}{3}, \quad a_5 = \frac{1}{120} + \frac{1}{2}\left(\frac{1}{3}\right) - \frac{1}{24}(1) = \frac{2}{15}.$$

Therefore,

$$\tan x = x + \frac{x^3}{3} + \frac{2}{15} x^5 + \cdots .$$

Example: To nd the reciprocal of the power series

$$g(x) = 1 + e^x = 2 + \sum_{n=1}^{\infty} \frac{x^n}{n!},$$

we let $h = 1$ in (1.51). If

$$\frac{1}{g(x)} = \sum_{n=0}^{\infty} a_n x^n,$$

then

$$\begin{aligned}
1 &= (a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots)\left(2 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots\right) \\
&= 2a_0 + (a_0 + 2a_1)x + \left(\frac{a_0}{2} + a_1 + 2a_2\right)x^2 \\
&\quad + \left(\frac{a_0}{6} + \frac{a_1}{2} + a_2 + 2a_3\right)x^3 + \cdots .
\end{aligned}$$

From Corollary,

$$\begin{aligned}
2a_0 &= 1, \\
a_0 + 2a_1 &= 0, \\
\frac{a_0}{2} + a_1 + 2a_2 &= 0, \\
\frac{a_0}{6} + \frac{a_1}{2} + a_2 + 2a_3 &= 0.
\end{aligned}$$

Solving these equations successively yields

$$\begin{aligned}
a_0 &= \frac{1}{2}, \\
a_1 &= -\frac{a_0}{2} = -\frac{1}{4}, \\
a_2 &= -\frac{1}{2}\left(\frac{a_0}{2} + a_1\right) = -\frac{1}{2}\left(\frac{1}{4} - \frac{1}{4}\right) = 0, \\
a_3 &= -\frac{1}{2}\left(\frac{a_0}{6} + \frac{a_1}{2} + a_2\right) = -\frac{1}{2}\left(\frac{1}{12} - \frac{1}{8} + 0\right) = \frac{1}{48},
\end{aligned}$$

so

$$\frac{1}{1+e^x} = \frac{1}{2} - \frac{x}{4} + \frac{x^3}{48} + \cdots .$$

Example: To find the reciprocal of

$$g(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad (1.53)$$

we again let $h = 1$ in (1.51). If

$$(e^x)^{-1} = \sum_{n=0}^{\infty} a_n x^n,$$

then

$$1 = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right) = \sum_{n=0}^{\infty} c_n x^n,$$

where

$$c_n = \sum_{r=0}^{n} \frac{a_r}{(n-r)!}.$$

We have, $c_0 = a_0 = 1$ and $c_n = 0$ if $n \geq 1$; hence,

$$a_n = -\sum_{r=0}^{n-1} \frac{a_r}{(n-r)!}, \quad n \geq 1. \tag{1.54}$$

Solving these equations successively for $a_0, a_1, \ldots$ yields

$$\begin{aligned}
a_1 &= -\frac{1}{1!}(1.32) = -1, \\
a_2 &= -\left[\frac{1}{2!}(1) + \frac{1}{1!}(-1)\right] = \frac{1}{2}, \\
a_3 &= -\left[\frac{1}{3!}(1) + \frac{1}{2!}(-1) + \frac{1}{1!}\left(\frac{1}{2}\right)\right] = -\frac{1}{6} \\
a_4 &= -\left[\frac{1}{4!}(1) + \frac{1}{3!}(-1) + \frac{1}{2!}\left(\frac{1}{2}\right) + \frac{1}{1!}\left(-\frac{1}{6}\right)\right] = \frac{1}{24}
\end{aligned}$$

From this, we see that

$$a_k = \frac{(-1)^k}{k!}$$

for $0 \leq k \leq 4$ and are led to conjecture that this holds for all k. To prove this by induction, we assume that it is so for $0 \leq k \leq n-1$ and compute from (1.54):

$$\begin{aligned}
a_n &= -\sum_{r=0}^{n-1} \frac{1}{(n-r)!} \frac{(-1)^r}{r!} \\
&= -\frac{1}{n!} \sum_{r=0}^{n-1} (-1)^r \binom{n}{r} \\
&= \frac{(-1)^n}{n!}.
\end{aligned}$$

Thus, we have shown that

$$(e^x)^{-1} = \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{n!}.$$

Since this is precisely the series that results if x is replaced by $-x$ in (1.53), we have veri ed a fundamental property of the exponential function: that

$$(e^x)^{-1} = e^{-x}.$$

This also follows from Example ??.

1.11 The Abel's Theorem

Theorem: Let f be de ned by a power series with nite radius of convergence R.

[illegible]

$$\lim_{x\to(x_0+R)-} f(x) = \quad a_nR^n.$$

[illegible]

$$\lim_{x\to(x_0-R)+} f(x) = \text{[illegible]}$$

Proof: Let

$$g(y) = \sum_{n=0}^{\infty} b_n y^n, \quad \sum_{n=0}^{\infty} b_n = s \quad \text{(finite)}.$$

We will show that

$$\lim_{y\to 1-} g(y) = s. \tag{1.55}$$

We have

$$g(y) = (1-y)\sum_{n=0}^{\infty} s_n y^n, \tag{1.56}$$

where

$$s_n = b_0 + b_1 + \cdots + b_n.$$

Since

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n \quad \text{and therefore} \quad 1 = (1-y)\sum_{n=0}^{\infty} y^n, \quad |y| < 1, \tag{1.57}$$

we can multiply through by s and write

$$s = (1-y)\sum_{n=0}^{\infty} sy^n, \quad |y| < 1.$$

Subtracting this from (1.56) yields

$$g(y) - s = (1-y)\sum_{n=0}^{\infty} (s_n - s)y^n, \quad |y| < 1.$$

If $\varepsilon > 0$, choose N so that

$$|s_n - s| < \varepsilon \quad \text{if} \quad n \geq N+1.$$

Then, if $0 < y < 1$,

$$\begin{aligned}
|g(y) - s| &\leq (1-y)\sum_{n=0}^{N} |s_n - s|y^n + (1-y)\sum_{n=N+1}^{\infty} |s_n - s|y^n \\
&< (1-y)\sum_{n=0}^{N} |s_n - s|y^n + (1-y)\varepsilon y^{N+1}\sum_{n=0}^{\infty} y^n \\
&< (1-y)\sum_{n=0}^{N} |s_n - s| + \varepsilon,
\end{aligned}$$

because of the second equality in (1.57).

Therefore,

$$|g(y) - s| < 2\varepsilon$$

if

$$(1 - y)\sum_{n=0}^{N} |s_n - s| < \varepsilon.$$

To obtain rst part of the theorem from this, let $b_n = a_n R^n$ and $g(y) = f(x_0 + Ry)$; to obtain second part, let $b_n = (-1)^n a_n R^n$ and $g(y) = f(x_0 - Ry)$.

Example: The series

$$f(x) = \frac{1}{1+x} =$$

This shows that the converse of Abel's theorem is false.

Integrating the series term by term yields

$$\log(1 + x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}, \quad |x| < 1,$$

where the power series converges at $x = 1$. The Abel's theorem implies that

$$\log 2 = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{n+1}.$$

Example: If $q \geq 0$, the binomial series

$$\sum_{n=0}^{\infty} \binom{q}{n} x^n$$

converges absolutely for $x = \pm 1$. This is obvious if q is a nonnegative integer, and it follows from Raabe's test for other positive values of q, since

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\binom{q}{n+1} \Big/ \binom{q}{n}\right| = \frac{n-q}{n+1}, \quad n > q,$$

and

$$\lim_{n\to\infty} n\left(\left|\frac{a_{n+1}}{a_n}\right| - 1\right) = \lim_{n\to\infty} n\left(\frac{n-q}{n+1} - 1\right) = \lim_{n\to\infty} \frac{n}{n+1}(-q-1) = -q-1.$$

Therefore, Abel's theorem imply that

$$\sum_{n=0}^{\infty} \binom{q}{n} = 2^q \quad \text{and} \quad \sum_{n=0}^{\infty} (-1)^n \binom{q}{n} = 0, \quad q \geq 0.$$

1.12 Pointwise and Uniform Bounded Functions

A sequence of functions $\{F_n\}$ on the set S is said to be pointwise bounded on S if the sequence of functions is bounded for every $x \in S$, that is, if there exists a nite valued function $\phi(x)$ de ned on S such that

$$|F_n(x)| < \phi(x), \quad x \in S, n = 1, 2, 3, \ldots.$$

[illegible] that

[illegible], ...

Remark: If $\{F_n\}$ is pointwise bounded on S and S_1 is countable subset of S, it is always possible to nd a subsequence $\{F_{n_k}\}$ such that subsequence is convergent. However, even if $\{F_n\}$ is uniformly bounded sequence of continuous functions on a compact set S, there need not exist a subsequence which converges pointwise on S.

Example: Consider the sequence of functions

$$F_n(x) = \sin nx, \qquad x \in [0, 2\pi].$$

Suppose there exists a sequence $\{n_k\}$ such that $\{\sin n_k x\}$ converges, for every $x \in [0, 2\pi]$. Then we must have

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x) = 0, \qquad x \in [0, 2\pi].$$

Hence

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x)^2 = 0, \qquad x \in [0, 2\pi].$$

By Lebesgue's theorem concerning integration of bounded convergent sequences, we have

$$\lim_{k\to\infty} \int_0^{2\pi} \lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x)^2 = 0.$$

But we have

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x)^2 = 2\pi.$$

which is a contradiction.

Example: Consider the sequence of functions

$$F_n(x) = \frac{x^2}{x^2 + (1 - nx)^2}, \qquad S = [0, 1].$$

Then $|F_n| \leq 1$, so that $\{F_n(x)\}$ is uniformly bounded on $[0,1]$. Also

$$\lim_{n\to\infty} F_n(x) = 0, \qquad x \in [0,1].$$

But

$$F_n(\frac{1}{n}) = 1,$$

so no subsequence can converge uniformly on $[0,1]$.

1.13 Equicontinuous Functions on a Set

A family of functions F de ned on the set S is equicontinuous if for all $f \in$ F and for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(x_1) - f(x_2)| \leq \varepsilon \quad \text{if} \quad x_1, x_2 \in [a,b], \quad |x_1 - x_2| < \delta. \tag{1.58}$$

Remark: It is clear that every member of F is uniformly continuous.

Theorem: If $\{F_n\}$ is a pointwise bounded sequence of functions on a countable set S, then $\{F_n\}$ has a subsequence $\{F_{n_k}\}$ such that subsequence converges for all $x \in S$.

Proof: Let $\{x_i\}$, $i = 1, 2, 3, \ldots$ be the points of S arranged in a sequence.

Since $\{F_n(x_i)\}$ is bounded, there exists a subsequence, which we shall denote by $\{F_{i,k}\}$, such that $\{F_{i,k}(x_i)\}$ converges as $k \to \infty$.

Consider the sequences $S_1, S_2, \ldots,$ de ned by

$$S_1 : F_{1,1}\ F_{1,2}\ F_{1,3}\ F_{1,4}\ \ldots$$

$$S_2 : F_{2,1}\ F_{2,2}\ F_{2,3}\ F_{2,4}\ \ldots$$

$$S_3 : F_{3,1}\ F_{3,2}\ F_{3,3}\ F_{3,4}\ \ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Consider the sequences $S_1, S_2, \ldots,$ de ned by

$$S_1 : F_{1,1}\ F_{1,2}\ F_{1,3}\ F_{1,4}\ \ldots$$

$$S_2 : F_{2,1}\ F_{2,2}\ F_{2,3}\ F_{2,4}\ \ldots$$

$$S_3 : F_{3,1}\ F_{3,2}\ F_{3,3}\ F_{3,4}\ \ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

The sequence has the following properties

- S_n is a subsequence of S_{n-1}, for $n = 2, 3, 4, \ldots$

- Due to the boundedness of $\{F_n(x_n)\}$, we can say that $F_{n,k}(x_n)$ converges, as $k \to \infty$.
- The order in which the functions appear is the same in each sequence, i.e., if one function precedes another in S_1, they are in the same relation in every S_n, until one or the other is deleted. Hence, when going from one row in the above array to the next below, functions may move to the left but never to the right.

We consider the sequence

$$E : F_{1,1}\ F_{2,2}\ F_{3,3}\ \ldots$$

By (3) property E is a subsequence of S_n, for $n = 1, 2, 3,$ The order in which the functions appear is the same in each sequence, i.e., if one function precedes another in S_1, they are in the same relation in every S_n, until one or the other is deleted.

Hence, when going from one row in the above array to the next below, functions may move to the left but never to the right. The (2) property of the sequence ensures that $\{F_{n,n}(x_i)\}$ converges as $n \to \infty$ for every $x \in S$.

Theorem: If K is a compact subset and if $\{F_n\}$ is a sequence of continuous functions de ned on K and $\{F_n\}$ converges uniformly then $\{F_n\}$ is equicontinuous on K.

Proof: Since the sequence of functions $\{F_n\}$ is uniformly convergent, for every $\varepsilon > 0$, there is an integer N such that

$$\|F_n - F_N\|_K < \varepsilon, \qquad n > N.$$

We know that continuous functions on compact sets are uniformly continuous, there is a $\delta > 0$ such that

$$|F_i(x) - F_i(y)| < \varepsilon, \quad |x - y| < \delta, 1 \le i \le N.$$

Theorem: If K is a compact subset and if $\{F_n\}$ is a sequence of continuous functions de ned on K and $\{F_n\}$ converges uniformly then $\{F_n\}$ is equicontinuous on K.

For $n > N$ and $|x - y| < \delta$, we have

$$\begin{aligned} |F_n(x) - F_n(y)| &\le |F_n(x) - F_N(x)| + |F_N(x) - F_N(y)| \\ &\quad + |F_N(y) - F_n(y)| \\ &< 3\varepsilon. \end{aligned}$$

Theorem: If $\{F_n\}$ is a sequence of continuous functions de ned on a compact set S and if $\{F_n\}$ is a pointwise bounded and equicontinuous on S, then

[illegible]

[illegible]

Proof: Since $\{F_n\}$ is equicontinuous then by de nition for every $\varepsilon > 0$, we have

$$|F_n(x) - F_n(y)| < \varepsilon, \qquad |x - y| < \delta.$$

From Analysis I, we know that S is compact then there are nitely many points $p_1, p_2, ...p_r$ in S such that to every $x \in S$ corresponds at least one p_1 such that $|x - p_1| < \delta$.

Since $\{F_n\}$ is pointwise bounded, there exists $M_i < \infty$ such that

$$|F_n(p_i)| < M_i, n \in \mathbb{N}.$$

If we take

$$M = \max\{M_1, ..., M_r\},$$

then $|F_n(x)| < M + \varepsilon$ for every $x \in S$. This proves the rst part of the theorem.

Theorem: If $\{F_n\}$ is a sequence of continuous functions de ned on a compact set S and if $\{F_n\}$ is a pointwise bounded and equicontinuous on S, then

[illegible]

Proof: Let E be a countable dense subset of S. Then from previous theorem we have a subsequence $\{F_{n_i}(x)\}$ such that the subsequence $\{F_{n_i}(x)\}$ converges for every $x \in E$.

Fix the notation $F_{n_i}(x) = g_i$, we shall prove that $\{g_i\}$ converges uniformly on S.

Let $\varepsilon > 0$, and choose δ as before. Let $V(x, \delta)$ be the set of all $y \in S$ such that $|x - y| < \delta$.

Since E is dense in S, and S is compact, there are nitely many points $x_1, ..., x_m$ in E such that

$$S \subset V(x_1, \delta) \cup ... \cup V(x_m, \delta) \qquad (*).$$

Since $\{g_i(x)\}$ converges for every $x \in E$, there is an integer N such that

$$|g_i(x_s) - g_j(x_s)| < \varepsilon, \quad \text{whenever} \quad i, j \geq N, 1 \leq s \leq m.$$

If $x \in S$, from (*) shows that $x \in V(x_s, \delta)$ for some s, so that

$$|g_i(x) - g_i(x_s)| < \varepsilon$$

for every i.

If $i \geq N$ and $j \geq N$, it follows that

$$|g_i(x) - g_j(x)| \leq |g_i(x) - g_i(x_s)| + |g_i(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)|$$

$$|g_i(x) - g_j(x)| \leq 3\varepsilon.$$

1.14 The Stone-Weierstrass Theorem

Theorem: If f is continuous function on $[a, b]$, there exists a sequence of polynomials P_n such that

$$\lim_{n\to\infty} P_n(x) = f(x),$$

uniformly on $[a, b]$.

Proof: Without any loss of generality, we may assume that $[a, b] = [0, 1]$.
We may also assume that $f(0) = f(1) = 0$. As we can consider

$$g(x) = f(x) - f(0) - c[f(1) - f(0)], \qquad x \in [0, 1].$$

If g can be obtained as the limit of uniformly convergent sequence of polynomials, it is clear that the same is true for f, since $f - g$ is a polynomial.

Furthermore, we de ne $f(x)$ to be zero for x outside $[0, 1]$. Then f is uniformly continuous on the whole line.

We take

$$Q_n(x) = c_n(1 - x^2)^n, \qquad n = 1, 2, \ldots,$$

where c_n is chosen so that

$$\int_{-1}^{1} Q_n(x)dx = 1, \qquad n = 1, 2, \ldots.$$

Consider the function

$$(1 - x^2)^n - 1 + nx^2,$$

which is zero at $x = 0$ and whose derivative is positive in $(0, 1)$.

Since

$$\begin{aligned}
\int_{-1}^{1} (1 - x^2)^n dx &= 2\int_{0}^{1} (1 - x^2)^n dx \\
&\geq 2\int_{0}^{1/\sqrt{n}} (1 - x^2)^n dx \\
&\geq 2\int_{0}^{1/\sqrt{n}} (1 - nx^2)dx \\
&= \frac{4}{3\sqrt{n}} \\
&= \frac{1}{\sqrt{n}}
\end{aligned}$$

It follows from

$$\int_{-1}^{1} Q_n(x)dx = 1, \qquad n = 1, 2, \ldots.$$

that $c_n < \sqrt{n}$.

For any $\delta > 0$, we have

$$Q_n(x) \leq \sqrt{n}(1-\delta^2)^n, \qquad \delta \leq |x| \leq 1.$$

So that $Q_n \to 0$ uniformly in $\delta \leq |x| \leq 1$.

Now set

$$P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt, \qquad x \in [0,1].$$

By change of variable and assumption on f implies that

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_{0}^{1} f(t)Q_n(t-x)dt,$$

and the last integral is clearly a polynomial in x.

Thus $\{P_n\}$ is a sequence of polynomials.

Given $\varepsilon > 0$, we chose $\delta > 0$ such that $|y - x| < \delta$ implies

$$|f(y) - f(x)| < \frac{\varepsilon}{3}.$$

Let $M = \sup|f(x)|$, we see that for $x \in [a,b]$, we have

$$\begin{aligned}
|P_n(x) - f(x)| &= \left|\int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)dt\right| \\
&\leq \int_{-1}^{1} |f(x+t) - f(x)|Q_n(t)dt \\
&\leq 2M\int_{-1}^{-\delta} Q_n(t)dt + \frac{\varepsilon}{2}\int_{-\delta}^{\delta} Q_n(t)dt \\
&\quad +2M\int_{\delta}^{1} Q_n(t)dt \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2}
\end{aligned}$$

1.15 Fourier Series

One of the fundamental methods of solving many problems in engineering elds is to represent the behavior of a system by a combination of simple behaviors. Mathematically, this is related to representing a function f(x) in the form of a functional series

$$f(x) = \sum_{k=1} c_k \phi_k(x).$$

Here the functions $\phi_k(x)$ are suitable elementary functions, also called the base set of functions, and the c_k are called the coe cients of the expansion.

For the Taylor series

$$f(x) =$$

the set $\{1, x, \ldots, x^n, \ldots\}$ is a base set of functions

Fourier Series: A Fourier series expansion of a function is a representation of the function as a linear combination of sines and cosines, that is, the base set of the representation is

$$\{1, \cos nx, \sin nx\}_{n=1}^{\infty}.$$

1.15.1 Periodic Functions

A function $f : \Omega \subset \mathbb{R} \to \mathbb{R}$ is said to be periodic if there exists a nonzero real number ω such that

$$f(x) = f(x + \omega), \qquad x \in \Omega.$$

The simplest examples of periodic functions from R into R include the well known sine and cosine functions, since for each $k \in \mathbb{Z}\backslash\{0\}$.

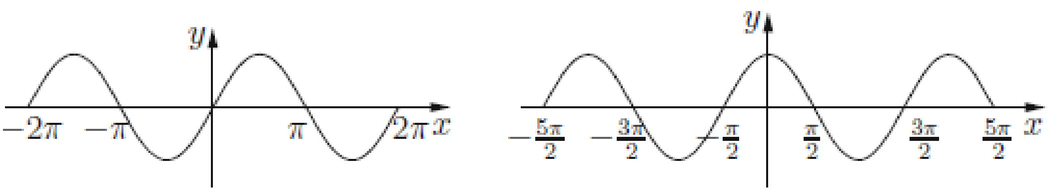

Figure 1.4: Periodic functions

Remark: If ω_1 and ω_2 are such that

$$f(x + \omega_1) = f(x), \qquad f(x + \omega_2) = f(x).$$

Then so is $\omega_1 \pm \omega_2$.

$$f(x + (\omega_1 \pm \omega_2))$$

There is a smallest positive value ω of a periodic function f called the primitive period (or the basic period or the fundamental period) of $f(x)$.
The reciprocal of the primitive period is called the frequency of the periodic function.

Lemma: [illegible] $f(cx)$ is ω/c. If $f(x)$ and $g(x)$ are periodic with the same period ω, then $h(x) = af(x) + bg(x)$ is also a periodic function with period ω. Here ω is not necessarily a primitive period.

Proof: Let $\phi(x) = f(cx)$, then

$$\phi(x) = f(cx) = f(cx + \omega) = f(c(x + \omega/c)) = \phi(x + \omega/c), \quad x \in \mathrm{R}.$$

This shows that ω/c is a period.

For the second part, we consider

$$h(x + \omega) = af(x + \omega) + bg(x + \omega) = af(x) + bg(x) = h(x).$$

Example: $\sin(cx)$ and $\cos(cx)$ are periodic functions with period $2\pi/c$.

The function

$$[illegible] (a_n \cos nx + b_n \sin nx),$$

is a periodic function with period 2π.

Although, individual functions, $\cos x, \cos 2x, \cos 3x, \ldots,$ have periods $2\pi, \pi, 2\pi/3, \ldots,$ respectively.

Lemma: If $f(x)$ is a periodic function with period ω, then

$$\int_c^{c+\omega} f(x)dx = \int_0^{\omega} f(x)dx,$$

whenever f is integrable on $[0, \omega]$.

Proof: Geometrically, it is obvious

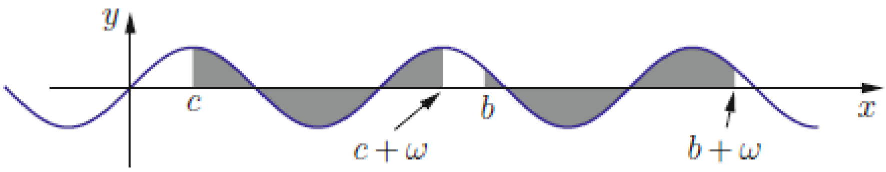

Figure 1.5: Geometric proof

Consider

$$
\begin{aligned}
\int_c^{c+\omega} f(x)dx &= \int_c^0 f(x)dx + \int_0^\omega f(x)dx + \int_\omega^{c+\omega} f(x)dx \\
&= -\int_0^c f(x)dx + \int_0^\omega f(x)dx + \int_0^c f(s)ds \\
&= \int_0^\omega f(x)dx.
\end{aligned}
$$

showing that the integral of a periodic function with period ω taken over an arbitrary interval of length ω always has the same value.

1.15.2 Periodic Extension

Suppose that f is a function de ned on $[a, a+\omega]$. Then the periodic extension of f [illegible]

$$\{ \text{[illegible]}$$

where n is an integer.

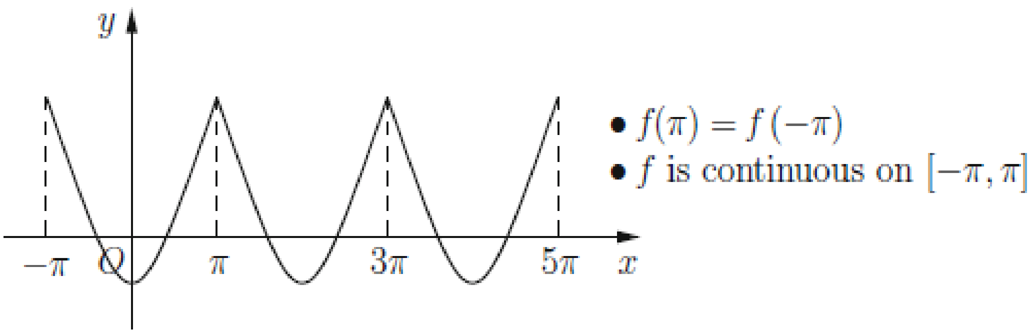

Figure 1.6: Periodic extension example 1

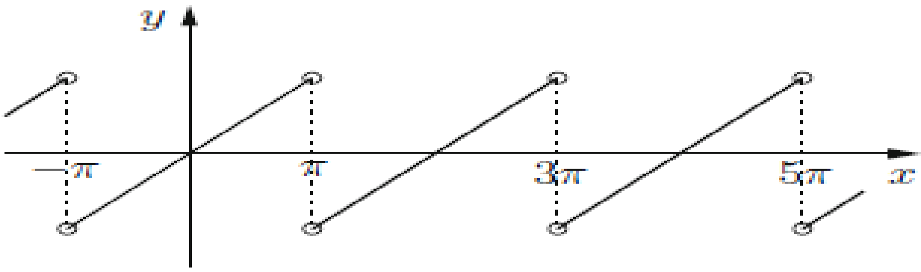

Figure 1.7: Periodic extension example 2

1.15.3 Trigonometric Polynomials

Any linear combination of the trigonometric functions $\sin kx$, $\cos kx$, given by

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}\left(a_k \cos kx + b_k \sin kx\right), \qquad x \in \mathbb{R},$$

where a_k and b_k are real numbers, is known as trigonometric polynomials.

Recall the Stone and Weierstrass theorem stating that the trigonometric polynomials are dense in $C[a, b]$ for any closed interval $[a, b]$, provided that $b - a < 2\pi$.

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}\left(a_k \cos kx + b_k \sin kx\right), \qquad x \in \mathbb{R},$$

The sequence $\{s_n\}$, converges on a set E, then we may de ne a function $f : E \to \mathbb{R}$ by

$$f(x) = \lim_{n\to\infty} s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\left(a_k \cos kx + b_k \sin kx\right), \qquad x \in E.$$

The series on the right is called a trigonometric series. The constants a_0, a_k, b_k $(k \in \mathbb{N})$ are called coe cients of the trigonometric series.

We have taken the constant term in series as $a_0/2$ rather than a_0 so that we can make $a_0/2$ t in a general formula later.

We observe that if the series on the right converges for all real $t[0, 2\pi]$, then the sum f must satisfy

$$f(x) = f(x + 2\pi), \qquad x \in \mathbb{R}.$$

Vector Space: A vector space is a nonempty set V of objects, called vectors, on which are de ned two operations, called addition and multiplication by scalars (real numbers), subject to the ten axioms (or rules). The axioms must hold for all vectors u, v, and w in V and for all scalars c and d.

1. The sum of u and v, denoted by $\mathrm{u} + \mathrm{v}$, is in V.

2. $u + v = v + u$.
3. $(u + v) + w = u + (v + w)$
4. There is a zero vector 0 in V such that $u + 0 = u$.
5. For each u in V, there is a vector $-u$ in V such that $u + (-u) = 0$.
6. The scalar multiple of u by c, denoted by cu, is in V.
7. $c(u + v) = cu + cv$.
8. $(c + d)u = cu + du$.
9. $c(du) = (cd)u$.
10. $1u = u$.

Remark: Using only these axioms, one can show that the zero vector in Axiom 4 is unique, and the vector $-u$, called the negative of u, in Axiom 5 is unique for each u in V.

The Inner Product: Let u, v, and w be vectors in vector space V, and let c be a scalar. Then an inner product is a function $<.,.>: V \times V \rightarrow \mathbb{F}$ such that

1. $< v, u >=< u, v >$
2. $< (v + u), w >=< v, w > + < u, w >$
3. $< cu, v >=< u, cv >= c < v, u >$
4. $< u, u >\geq 0$, and $< u, u >= 0$ if and only if $u = 0$.

1.16 The space E

Let us de ne the space E be the set of all real valued piecewise de ned periodic function f on the interval $[-\pi, \pi]$.

Theorem: The space E is a linear space, that is, a vector space. Moreover, E an inner product space with respect to the inner product

$$< f, g >= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx.$$

The trigonometric functions: The set of functions

$$\Phi = \{\frac{1}{\sqrt{2}}, \cos(nx), \sin(nx) : n \in \mathbb{N}\}$$

is an infinite orthonormal system in E with respect to the inner product defined

$$<f, g>= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx.$$

Let $\Phi = \{\phi_1, \phi_2, ..., \phi_n, ...\}$ be an orthonormal basis of an infinite dimensional inner product space X, and let $f \in X$. Then the infinite series

$$\sum_{k=1}^{\infty} <f, \phi_k> \phi_k(x) := \sum_{k=1}^{\infty} c_k\phi_k(x),$$

is called the Fourier series of f (relative to Φ), and the coefficients $c_k =< f, \phi_k >$ are called the kth Fourier coefficient of f.

We introduce

$$||f||^2 =< f, f >= \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx.$$

Suppose that we are given a trigonometric series of the form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\left(a_k \cos kx + b_k \sin kx\right), \qquad x \in E.$$

Clearly, since each term of the series has period 2π, if it converges to a function $f(x)$, then $f(x)$ must be a periodic function with period 2π.

Thus, only 2π-periodic functions are expected to have trigonometric series of the above form.

Problem: Suppose that f is a 2π-periodic function. Under what conditions does the function have a representation of the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos nx + b_n \sin nx\right).$$

When it does, what should be a_n, b_n?

Assume for the moment that the series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos nx + b_n \sin nx\right), \qquad (*)$$

converges uniformly on R. This is the case if

$$\frac{|a_0|}{2} + \sum_{n=1}^{\infty}\left(|a_n| + |b_n|\right)$$

converges, so that the series (*) is dominated by the convergent series in R.

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{1}{\pi}\int_{-\pi}^{\pi}\left\{\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n\cos nx + b_n \sin nx)\right\}$$

$$= \frac{a_0}{2}\left\{\frac{1}{\pi}\int_{-\pi}^{\pi} dx\right\} + \sum_{n=1}^{\infty}\left\{\frac{a_n}{\pi}\int_{-\pi}^{\pi}\cos nxdx + \frac{b_n}{\pi}\int_{-\pi}^{\pi}\sin nx\right)$$

$$= a_0$$

Recall:

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos nx\cos kxdx = \delta_{nk} = \frac{1}{\pi}\int_{-\pi}^{\pi}\sin nx\sin kxdx$$

$$\int_{-\pi}^{\pi}\cos nx\sin kxdx = 0$$

and

$$2\cos\alpha\cos\beta = \cos(\alpha+\beta) + \cos(\alpha-\beta)$$

$$2\sin\alpha\sin\beta = \cos(\alpha-\beta) - \cos(\alpha+\beta)$$

$$2\sin\alpha\cos\beta = \sin(\alpha+\beta) + \sin(\alpha-\beta).$$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos nx + b_n\sin nx\right), \qquad (*)$$

Multiply by $\cos kx$ and the series for $f(x)\cos kx$ can be integrated term by term for each xed k, we can determine a_k and b_n.

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kxdx = \frac{a_0}{2}\frac{1}{\pi}\int_{-\pi}^{\pi}\cos kxdx + \sum_{n=1}\left(a_n\int_{-\pi}^{\pi}\cos kx\cos nxdx + b_n\int_{-\pi}^{\pi}\sin nx\cos kxdx\right).$$

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kxdx.$$

Multiply by $\sin kx$ and the series for $f(x)\sin kx$ can be integrated term by term for each xed k, we can determine b_k and b_n.

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kxdx = \frac{a_0}{2}\frac{1}{\pi}\int_{-\pi}^{\pi}\sin kxdx + \sum_{n=1}\left(a_n\int_{-\pi}^{\pi}\sin kx\cos nxdx + b_n\int_{-\pi}^{\pi}\sin nx\sin kxdx\right).$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx dx.$$

[illegible]

de ned by

[illegible]

are called the Fourier coe cients of f. The corresponding trigonometric series

$$[illegible] + [illegible]\Big(a_k \cos kx + b_k \sin kx\Big),$$

is called the Fourier series of f. We express this association by writing

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty}\Big(a_k \cos kx + b_k \sin kx\Big),$$

to indicate that the Fourier series on the right may or may not converge to f at some point $t \in [-\pi, \pi]$.

Theorem: If the trigonometric series of the form

$$[illegible] + [illegible]\Big(a_k \cos kx + b_k \sin kx\Big), \qquad (*)$$

[illegible]

More precisely, if the trigonometric series (*) converges uniformly to f on $[-\pi, \pi]$, then the a_k and b_k are given by

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx dx,\ k \geq 0,\quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx dx,\ k \geq 1.$$

Remark: We have no idea what happens if the series

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}\Big(a_k \cos kx + b_k \sin kx\Big), \qquad (*)$$

doesn't converge uniformly on $[-\pi, \pi]$.

However, since

$$|a_k \cos kx + b_k \sin kx| \leq |a_k| + |b_k|,$$

Weierstrass M-test shows that the trigonometric series (*) converges absolutely and uniformly on every closed interval $[a, b]$ whenever

$$\sum_{k=1}^{\infty}(|a_k| + |b_k|)$$

is convergent.

1.16.1 Fourier Series of Even and Odd Functions

Even and odd functions possess certain simple but useful properties:

- The product of two even (or odd) functions is an even function.
- The sum of two even (or odd) functions is an even (or odd) function.
- The product of an even and an odd function is an odd function.
- For a Riemann integrable function f de ned on $[-c, c]$ $(c > 0)$, it is evident that

$$\int_{-c}^{c} f(x)dx = 2\int_{0}^{c} f(x)dx, \qquad \text{if } f \text{ is even}$$

$$\int_{-c}^{c} f(x)dx = 0, \qquad \text{if } f \text{ is odd}$$

Fourier series of even function: Suppose that $f(x)$ is a periodic function of period 2π. Let us further assume that f is even on $(-\pi, \pi)$, i.e., $f(x) = f(-x)$ for all $x \in (-\pi, \pi)$.

Then the product function $f(x)\sin kx$ is odd, which means that $b_k = 0$ for all $k \geq 1$, and hence we have the Fourier cosine series

$$f(x) \simeq \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx, \; a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx dx.$$

Fourier series of odd function: Suppose that $f(x)$ is a periodic function of period 2π. Let us further assume that f is odd on $(-\pi, \pi)$, i.e., $f(x) = -f(-x)$ for all $x \in (-\pi, \pi)$.

Then the product function $f(x)\cos kx$ is odd, which means that $a_k = 0$ for all $k \geq 0$, and hence we have the Fourier cosine series

$$f(x) \cong \sum_{k=1}^{\infty} b_k \sin kx, \; b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx dx.$$

Example: Consider $f(x) = |x|$ on $[-\pi, \pi]$.

Then f is even and continuous on $[-\pi, \pi]$.

$$a_n = -\frac{2(1-(-1)^n)}{n^2\pi}.$$

We have

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

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

Remark: Note that the Fourier series here converges uniformly to $|x|$ on $[-\pi, \pi]$ but not on the whole interval $(-\infty, \infty)$, and so outside the interval $(-\infty, \infty)$, $f(x)$ is determined by the periodicity condition $f(x) = f(x + 2\pi)$.

we can make use of this series to nd the values of some numerical series. For instant $x = 0$ gives

$$\frac{\pi^2}{8} = \sum_{k=1}^{\infty} \frac{1}{(2k+1)^2}.$$

Some natural questions arise:

- For what values of x does the Fourier series of f converge? Does it converge for all x in $[-\pi, \pi]$? If it converges on $[-\pi, \pi]$ but not to f, what will be its sum?
- If the Fourier series of f converges at x, does it converge to f?
- If the Fourier series of f converges to f on $[-\pi, \pi]$, does it converge uniformly to f on $[-\pi, \pi]$?

Is the continuity of f is su cient to guarantee convergence of the Fourier series of f on $[-\pi, \pi]$?

In 1876, Paul du Bois-Reymond constructed a continuous function $f : [-\pi, \pi] \to \mathbb{R}$ whose Fourier series failed to converge to f at each point in a dense subset of $[-\pi, \pi]$.

Indeed, the following are true statements

- There exists a continuous function whose Fourier series diverges at a point.
- There exists a continuous function whose Fourier series converges everywhere on $[-\pi, \pi]$, but not uniformly.
- There exists a continuous function whose Fourier series diverges for points in some set S and converges on $(-\pi, \pi) \setminus S$.

The space E: Let us de ne the space E be the set of all real valued piecewise de ned periodic function f on the interval $[-\pi, \pi]$.

De ne

$$E' = \left\{ f \in E : \begin{array}{l} \lim_{h\to 0+} \dfrac{f(x+h) - f(x+)}{h} \text{ exists } x \in [-\pi, \pi) \\ \lim_{h\to 0-} \dfrac{f(x+h) - f(x-)}{h} \text{ exists } x \in (-\pi, \pi] \end{array} \right\}$$

Theorem: Let $f \in E'$. Then for each $x \in (-\pi, \pi)$, the Fourier series of $f(x)$ converges to the value

$$\frac{f(x-) + f(x+)}{2}.$$

At the end points $x = \pm\pi$, the series converges to

$$\frac{f(\pi-) + f(-\pi+)}{2}.$$

Remark: If $f \in E'$ is continuous at x, then $f(x-) = f(x+) = f(x)$, and so at such points

$$\frac{f(x-) + f(x+)}{2} = f(x).$$

Thus, the Fourier series of f converges to f(x) at the point x where it is continuous.

At the point of discontinuity x, the Fourier series of f assumes the mean of the one-sided limits of f.

Corollary: If $f : [-\pi, \pi] \to \mathbb{R}$ is continuous, and if $f(-\pi) = f(\pi)$, $f'(x)$ exists and is piecewise continuous on $[-\pi, \pi]$, then the Fourier series of f converges to $f(x)$ at every point $x \in [-\pi, \pi]$.

Theorem: Suppose that $f : [-\pi, \pi] \to \mathbb{R}$ is piecewise continuous on $[-\pi, \pi]$ and piecewise monotone,that is, there exists a partition $P = x_0, x_1, \ldots, x_n$ of $[-\pi, \pi]$ such that the restriction $f|_{[x_{k-1}, x_k]}, k = 1, 2 \ldots, n,$ is either increasing or decreasing.

Let $f(x)$ be de ned for other values of x by the periodicity condition $f(x) = f(x + 2\pi)$. Then the Fourier series of f on $[-\pi, \pi]$ converges to

- $f(x)$ if f is continuous at $x \in (-\pi, \pi)$.
- $(f(x+) + f(x-))/2$ if f is discontinuous at x.
- $(f(\pi-) + f((-\pi)-))/2$ if f is discontinuous at $x = \pm\pi$.

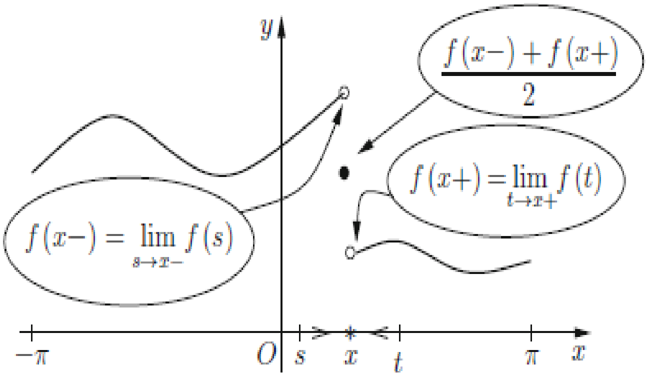

Figure 1.8: At discontinuous points

Example:If $f(x) = x$ on $[-\pi, \pi)$ and $f(\pi) = -\pi$. Find the Fourier sine series of f.

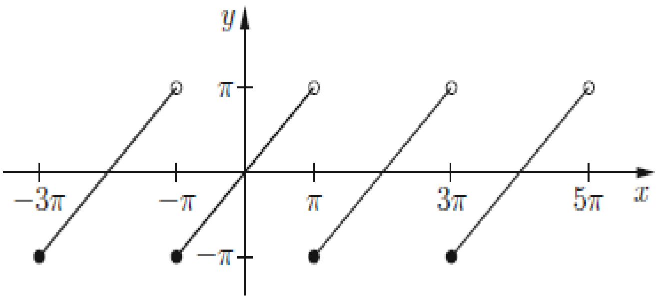

Figure 1.9: Example

- f is odd function, hence $a_n = 0$.
- $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin nx dx = \frac{2}{\pi}\int_0^{\pi} x \sin nx dx = \frac{2(-1)^{n-1}}{n}.$

$$x \cong 2 \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sin kx.$$

Remarks: Note that the Fourier series does not necessarily agree with $f(x) = x$ at every point in $[-\pi, \pi]$.

The Fourier series vanishes at both endpoints $x = \pm\pi$, whereas the function does not vanish at either endpoint.

However, the Dirichlet's theorem states that series converges to $f(x)$ at every interior point of $(-\pi, \pi)$.

For example at $x = \pi/2$ the symbol $\cong$ could be replaced by $=$ and so

$$\frac{\pi}{2} = 2\left(1 - \frac{0}{2} + \frac{(-1)}{3} - \frac{0}{4} + \frac{1}{5} + \ldots\right).$$

$$x \cong 2 \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sin kx.$$

Remarks: Finally, we remark that at the endpoints $x = \pm\pi$, the series converges to

$$\frac{f(\pi-) + f((-\pi)-)}{2} = \frac{\pi + (-\pi)}{2} = 0.$$

we could also consider f as follows: $f(x) = x$ on $(-\pi, \pi)$ and $f(-\pi) = f(\pi) = 0$.

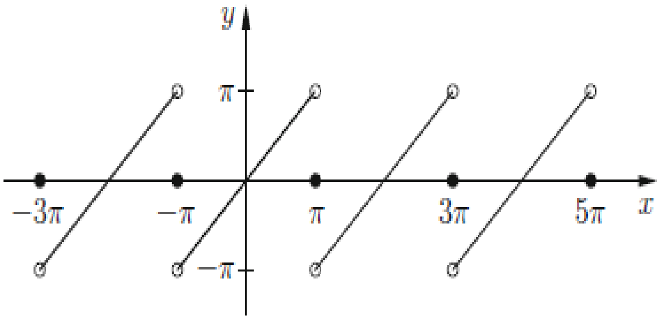

Figure 1.10: Example

Example: If $f(x) = e^x$ on $[-\pi, \pi)$ and $f(x + 2\pi) = f(x)$ for $x \in \mathbb{R}$. Determine the Fourier series of the function f.

Some facts about complex numbers.

Example: If $f(x) = e^x$ on $[-\pi, \pi)$ and $f(x + 2\pi) = f(x)$ for $x \in \mathbb{R}$. Determine the Fourier series of the function f.

$$\int e^{inx} = \int \cos nx dx + i \int \sin nx dx.$$

According to this, the Fourier coe cients are easy to derive quickly by writing

$$\begin{aligned} a_n - ib_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} e^{-inx} e^x dx \\ &= \frac{1}{\pi} \frac{e^{(1-in)x}}{1 - in} \Big|_{-\pi}^{\pi} \\ &= \frac{1}{\pi} \left(\frac{e^{(1-in)\pi} - e^{-(1-in)\pi}}{1 - in} \right) \\ &= \frac{(-1)^n (e^{\pi} - e^{-\pi})}{\pi(1 - in)} \end{aligned}$$

$$a_n = \frac{2(-1)^n \sinh \pi}{\pi(1 + n^2)}, \qquad b_n = \frac{2(-1)^{n-1} n \sinh \pi}{\pi(1 + n^2)}.$$

We have

$$e^x \cong \frac{\sinh\pi}{\pi} + \frac{2\sinh\pi}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{(1+n^2)}\cos nx + \frac{2\sinh\pi}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}n}{(1+n^2)}\sin nx.$$

Remark: In particular, at the point of continuity $x = 0$, it follows that

$$1 = \frac{\sinh\pi}{\pi} + \frac{2\sinh\pi}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{(1+n^2)}.$$

Which can be written as

$$\frac{\pi\csc\pi - 1}{2} = \sum_{n=1}^{\infty}\frac{(-1)^n}{(1+n^2)}.$$

Remark: According to Dirichlet's theorem, at the endpoint $x = \pi$, we have

$$\frac{e^{\pi}+e^{-\pi}}{2} = \frac{\sinh\pi}{\pi} + \frac{2\sinh\pi}{\pi}\sum_{n=1}^{\infty}\frac{1}{(1+n^2)},$$

$$\pi\coth\pi = 1 + 2\sum_{n=1}^{\infty}\frac{1}{(1+n^2)}.$$

Which reduces to

$$\frac{\pi\coth\pi - 1}{2} = \sum_{n=1}^{\infty}\frac{1}{(1+n^2)}.$$

1.17 Fourier Series for Arbitrary Periodic Function

Suppose that f is a $2L$-periodic and Riemann integrable function. The function $f(at)$ has period $2L/a$.

In particular, $f((L/\pi)t)$ is 2π-periodic, and so the Fourier series expansion has the following in terms of the variable t:

$$f\left(\frac{L}{\pi}t\right) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos nt + b_n\sin nt\right), \qquad t\in[-\pi,\pi],$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{L}{\pi}t\right)\cos nt\,dt = \frac{1}{L}\int_{-L}^{L} f(x)\cos\left(\frac{k\pi}{L}x\right)dx.$$

$$f\left(\frac{L}{\pi}t\right) \cong \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos nt + b_n\sin nt\right), \qquad t\in[-\pi,\pi],$$

where
$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\frac{L}{\pi}t)\cos nt dt = \frac{1}{L}\int_{-L}^{L} f(x)\cos(\frac{k\pi}{L}x)dx,$$
and similarly,
$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(\frac{k\pi}{L}x)dx.$$
We remark that the interval of integration in the last two formulas for the Fourier coe cients can be replaced with an arbitrary interval $[c, c + 2L]$, of length $2L$. Changing the variable t, by setting $t = (\pi/L)x$.

Theorem: Let f be a periodic function with period $2L$. Then the Fourier expansion of f is given by
$$f(x) \cong \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos(\frac{n\pi}{L}x) + b_n\sin(\frac{n\pi}{L}x)\right), \qquad x \in [-L, L],$$
where
$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(\frac{n\pi}{L}x)dx,$$
and
$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(\frac{n\pi}{L}x)dx.$$

Remark: The interval of integration in the last formulas for the Fourier coe cients can be replaced with the interval $[c, c + 2L]$, where c is any real number; we usually let $c = -L$. Notice that
$$\cos(\frac{n\pi}{L}(x + 2L)) = \cos(\frac{n\pi}{L}x)$$
$$\sin(\frac{n\pi}{L}(x + 2L)) = \sin(\frac{n\pi}{L}x).$$

Corollary: The Fourier series of an even function f with period $2L$ is a Fourier cosine series
$$f(x) \cong \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(\frac{k\pi}{L}x), \qquad x \in [c, c + 2L],$$
where
$$a_n = \frac{1}{L}\int_{c}^{c+2L} f(x)\cos(\frac{n\pi}{L}x)dx.$$
and the Fourier series of an odd function f with period $2L$ is a Fourier sine series
$$f(x) \cong \sum_{n=1}^{\infty} b_n\sin(\frac{n\pi}{L}x), \qquad x \in [c, c + 2L],$$

where

$$b_n = \frac{1}{L}\int_c^{c+2L} f(x)\sin(\frac{n\pi}{L}x)dx$$

where c is any real number.

Example: Consider the function

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

Here, we have $L = 2$, and the function is even. We have

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(\frac{k\pi}{L}x)dx,$$

and

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(\frac{k\pi}{L}x)dx.$$

we obtain

$$a_0 = \frac{1}{2}, \qquad a_n = 0.$$

$$b_n = \frac{1 + (-1)^{n-1}}{n\pi}, \qquad n \ge 1.$$

Example: Consider the function $f(x) = |\sin x|$. The function is de ned for all x and the function has period π.

Clearly, f represents a continuous, piecewise smooth, even function of period π, and therefore it is everywhere equal to its Fourier series, consisting of cosine terms only.

We have $c = 0$, and $L = \pi/2$, then we have

$$\begin{aligned} a_k &= \frac{2}{\pi}\int_0^{\pi} f(x)\cos(2kx)dx \\ &= \frac{2}{\pi}\int_0^{\pi} \sin x\cos(2kx)dx \\ &= \frac{1}{\pi}\int_0^{\pi}\Big[\sin(1+2k)x - \sin(2k-1)x\Big]dx \end{aligned}$$

$$\begin{aligned} &= \frac{1}{\pi}\int_0^{\pi}\Big[\sin(1+2k)x - \sin(2k-1)x\Big]dx \\ &= \frac{1}{\pi}\left(-\frac{\cos(1+2k)x}{2k+1} + \frac{\cos(2k-1)x}{2k-1}\right)\Bigg|_0^{\pi} \\ &= -\frac{1}{\pi}\left(\frac{(-1)^{2k+1}-1}{2k+1} - \frac{(-1)^{2k-1}-1}{2k-1}\right) \\ a_k &= -\frac{4}{\pi(4k^2-1)}. \end{aligned}$$

Thus, the Fourier series expansion of $|\sin x|$ is

$$|\sin x| = \frac{2}{\pi} - \frac{4}{\pi}\sum_{k=1} \frac{\cos 2kx}{4k^2 - 1}, \qquad x \in [-\pi, \pi].$$

1.18 Best Approximation Theorem

Theorem: Let $\Phi = \phi_1, \dots, \phi_n$ be an orthonormal set of functions in the inner product space E, and let c_k be the Fourier coe cients of f relative to ϕ_k:

$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\phi_k(x)dx := < f, \phi_k > .$$

If $T_n(x)$ is an arbitrary Fourier polynomial relative to ϕ_k, that is, $T_n(x) = \sum_{k=1}^{n} d_k\phi_k(x)$ for some constants $d_1, \dots, d_n$, then we have

$$\left\| f - \sum_{k=1}^{n} c_k\phi_k(x) \right\|^2 \le \|f - T_n\|^2,$$

with equality if and only if $c_k = d_k$ for each $k = 1, \dots, n$. Moreover,

$$\sum_{k=1}^{n} |c_k|^2 \le \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx.$$

Proof: Setting $S_n = \sum_{k=1}^{n} c_k\phi_k(x)$. Then we have

$$\begin{aligned}
\|f - T_n\|^2 &= \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x) - T_n(x)|^2 dx \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx + \frac{1}{\pi}\int_{-\pi}^{\pi} |T_n(x)|^2 dx \\
&\quad - 2\frac{1}{\pi}\sum d_k \int_{-\pi}^{\pi} f(x)\phi_k(x)dx \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx + \sum_{k=1}^{n} |d_k|^2 - 2\sum_{k=1}^{n} c_k d_k \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx + \sum_{k=1}^{n} |d_k|^2 - 2\sum_{k=1}^{n} c_k d_k \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx + \sum_{k=1}^{n} |c_k - d_k|^2 - 2\sum_{k=1}^{n} c_k \\
&= \|f - S_n\|^2 + \sum_{k=1}^{n} |c_k - d_k|^2.
\end{aligned}$$

Therefore,

$$\|f - T_n\|^2 \geq \|f - S_n\|^2,$$

with equality if and only if $c_k = d_k$ for each $k = 1, \ldots, n$.

$$\|f - T_n\|^2 \geq \|f - S_n\|^2,$$

Note that f and ϕ_k are xed, while the d_k are allowed to vary.

In particular, setting $d_k = c_k$, shows that the minimum value of $\|f - T_n\|^2 \geq \|f - S_n\|^2$, is given by

$$\min_{T_n} \|f - T_n\|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx - \sum_{k=1}^{n} |c_k|^2 = \|f\|^2 - \sum_{k=1} |c_k|^2,$$

which has to be nonnegative. This gives

$$\sum_{k=1} |c_k|^2 \leq \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx \quad \text{for all } n.$$

Chapter 2

Functions of Several Variables

2.1 Euclidean Spaces

The vector sum of

$$\mathbf{X} = (x_1, x_2, \dots, x_n) \quad \text{and} \quad \mathbf{Y} = (y_1, y_2, \dots, y_n)$$

is

$$\mathbf{X} + \mathbf{Y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n). \tag{2.1}$$

If a is a real number, the scalar multiple of $\mathbf{X}$ by a is

$$a\mathbf{X} = (ax_1, ax_2, \dots, ax_n). \tag{2.2}$$

In R^4, let

$$\mathbf{X} = (1, -2, 6, 5) \quad \text{and} \quad \mathbf{Y} = \left(3, -5, 4, \tfrac{1}{2}\right).$$

Then

$$\mathbf{X} + \mathbf{Y} = \left(4, -7, 10, \tfrac{11}{2}\right)$$

and

$$6\mathbf{X} = (6, -12, 36, 30).$$

Theorem: If $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are in R^n and a and b are real numbers, then

- $\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$ (vector addition is commutative).
- $(\mathbf{X} + \mathbf{Y}) + \mathbf{Z} = \mathbf{X} + (\mathbf{Y} + \mathbf{Z})$ (vector addition is associative).
- There is a unique vector $\mathbf{0}$, called the zero vector, such that $\mathbf{X} + \mathbf{0} = \mathbf{X}$ for all $\mathbf{X}$ in R^n.
- For each $\mathbf{X}$ in R^n there is a unique vector $-\mathbf{X}$ such that $\mathbf{X} + (-\mathbf{X}) = \mathbf{0}$.
- $a(b\mathbf{X}) = (ab)\mathbf{X}$.
- $(a + b)\mathbf{X} = a\mathbf{X} + b\mathbf{X}$.
- $a(\mathbf{X} + \mathbf{Y}) = a\mathbf{X} + a\mathbf{Y}$.
- $1\mathbf{X} = \mathbf{X}$.

Remark: Clearly, $\mathbf{0} = (0, 0, \ldots, 0)$ and, if $\mathbf{X} = (x_1, x_2, \ldots, x_n)$, then

$$-\mathbf{X} = (-x_1, -x_2, \ldots, -x_n).$$

We write $\mathbf{X} + (-\mathbf{Y})$ as $\mathbf{X} - \mathbf{Y}$. The point $\mathbf{0}$ is called the origin.

Length, distance: The length of the vector $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ is

$$|\mathbf{X}| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}.$$

The distance between points $\mathbf{X}$ and $\mathbf{Y}$ is $|\mathbf{X} - \mathbf{Y}|$.

In particular, $|\mathbf{X}|$ is the distance between $\mathbf{X}$ and the origin. If $|\mathbf{X}| = 1$, then $\mathbf{X}$ is a unit vector.

Example: The lengths of the vectors

$$\mathbf{X} = (1, -2, 6, 5) \quad \text{and} \quad \mathbf{Y} = \left(3, -5, 4, \tfrac{1}{2}\right)$$

are

$$|\mathbf{X}| = (1^2 + (-2)^2 + 6^2 + 5^2)^{1/2} = \sqrt{66}$$

and

$$|\mathbf{Y}| = \left(3^2 + (-5)^2 + 4^2 + \left(\tfrac{1}{2}\right)^2\right)^{1/2} = \frac{\sqrt{201}}{2}.$$

The distance between $\mathbf{X}$ and $\mathbf{Y}$ is

$$|\mathbf{X} - \mathbf{Y}| = \left((1-3)^2 + (-2+5)^2 + (6-4)^2 + \left(5 - \tfrac{1}{2}\right)^2\right)^{1/2} = \frac{\sqrt{149}}{2}.$$

The inner product $\mathbf{X} \cdot \mathbf{Y}$ of $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{Y} = (y_1, y_2, \ldots, y_n)$ is

$$\mathbf{X} \cdot \mathbf{Y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

2.2 Schwarz's Inequality

Lemma: If $\mathbf{X}$ and $\mathbf{Y}$ are any two vectors in $\mathbb{R}^n$, then

$$|\mathbf{X} \cdot \mathbf{Y}| \leq |\mathbf{X}|\,|\mathbf{Y}|, \tag{2.3}$$

with equality if and only if one of the vectors is a scalar multiple of the other.

Proof: Suppose that $\mathbf{Y} \neq \mathbf{0}$ and t is any real number. Then

$$\begin{aligned} 0 &\leq \sum_{i=1}^n (x_i - ty_i)^2 \\ &= \sum_{i=1}^n x_i^2 - 2t\sum_{i=1}^n x_iy_i + t^2\sum_{i=1}^n y_i^2 \\ &= |\mathbf{X}|^2 - 2(\mathbf{X} \cdot \mathbf{Y})t + t^2|\mathbf{Y}|^2. \end{aligned} \tag{2.4}$$

The last expression is a second-degree polynomial p in t. From the quadratic formula, the zeros of p are

$$t = \frac{(\mathbf{X}\cdot\mathbf{Y}) \pm \sqrt{(\mathbf{X}\cdot\mathbf{Y})^2 - |\mathbf{X}|^2|\mathbf{Y}|^2}}{|\mathbf{Y}|^2}$$

Hence,

$$(\mathbf{X}\cdot\mathbf{Y})^2 \leq |\mathbf{X}|^2|\mathbf{Y}|^2. \tag{2.5}$$

because if not, then p would have two distinct real zeros and therefore be negative between them, contradicting the inequality (2.4).

Proof:

$$(\mathbf{X}\cdot\mathbf{Y})^2 \leq |\mathbf{X}|^2|\mathbf{Y}|^2, \tag{2.6}$$

Taking square roots in (2.6) yields (2.3) if $\mathbf{Y} \neq \mathbf{0}$. If $\mathbf{X} = t\mathbf{Y}$, then $|\mathbf{X}\cdot\mathbf{Y}| = |\mathbf{X}||\mathbf{Y}| = |t||\mathbf{Y}|^2$ (verify), so equality holds in (2.3).

Conversely, if equality holds in (2.3), then p has the real zero $t_0 = (\mathbf{X}\cdot\mathbf{Y})/|\mathbf{Y}|^2$, and

$$\sum_{i=1}^{n}(x_i - t_0 y_i)^2 = 0$$

from (2.4); therefore, $\mathbf{X} = t_0\mathbf{Y}$.

Theorem: If $\mathbf{X}$ and $\mathbf{Y}$ are in $\mathbb{R}^n$, then

$$|\mathbf{X}+\mathbf{Y}| \leq |\mathbf{X}| + |\mathbf{Y}|, \tag{2.7}$$

with equality if and only if one of the vectors is a nonnegative multiple of the other.

Proof: By definition,

$$\begin{aligned}
|\mathbf{X}+\mathbf{Y}|^2 &= \sum_{i=1}^{n}(x_i+y_i)^2 \\
&= \sum_{i=1}^{n}x_i^2 + 2\sum_{i=1}^{n}x_iy_i + \sum_{i=1}^{n}y_i^2 \\
&= |\mathbf{X}|^2 + 2(\mathbf{X}\cdot\mathbf{Y}) + |\mathbf{Y}|^2 \qquad (2.8)\\
&\leq |\mathbf{X}|^2 + 2|\mathbf{X}|\,|\mathbf{Y}| + |\mathbf{Y}|^2 \quad \text{(by Schwarz's inequality)}\\
&= (|\mathbf{X}| + |\mathbf{Y}|)^2.
\end{aligned}$$

Hence,

$$|\mathbf{X}+\mathbf{Y}|^2 \leq (|\mathbf{X}| + |\mathbf{Y}|)^2.$$

Taking square roots yields (2.7).

From the third line of (2.8), equality holds in (2.7) if and only if $\mathbf{X}\cdot\mathbf{Y} = |\mathbf{X}||\mathbf{Y}|$, which is true if and only if one of the vectors $\mathbf{X}$ and $\mathbf{Y}$ is a nonnegative scalar multiple of the other.

Corollary: If $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are in R^n, then

$$|\mathbf{X} - \mathbf{Z}| \leq |\mathbf{X} - \mathbf{Y}| + |\mathbf{Y} - \mathbf{Z}|.$$

Proof: Write

$$\mathbf{X} - \mathbf{Z} = (\mathbf{X} - \mathbf{Y}) + (\mathbf{Y} - \mathbf{Z}),$$

and apply triangle inequality with $\mathbf{X}$ and $\mathbf{Y}$ replaced by $\mathbf{X} - \mathbf{Y}$ and $\mathbf{Y} - \mathbf{Z}$.

Corollary: If $\mathbf{X}$ and $\mathbf{Y}$ are in R^n, then

$$|\mathbf{X} - \mathbf{Y}| \geq ||\mathbf{X}| - |\mathbf{Y}||.$$

Proof: Since

$$\mathbf{X} = \mathbf{Y} + (\mathbf{X} - \mathbf{Y}),$$

Triangle inequality implies that

$$|\mathbf{X}| \leq |\mathbf{Y}| + |\mathbf{X} - \mathbf{Y}|,$$

which is equivalent to $|\mathbf{X}| - |\mathbf{Y}| \leq |\mathbf{X} - \mathbf{Y}|$.

Interchanging $\mathbf{X}$ and $\mathbf{Y}$ yields

$$|\mathbf{Y}| - |\mathbf{X}| \leq |\mathbf{Y} - \mathbf{X}|.$$

Since $|\mathbf{X} - \mathbf{Y}| = |\mathbf{Y} - \mathbf{X}|$, the last two inequalities imply the stated conclusion.

Theorem: If $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are members of R^n and a is a scalar, then

- $|a\mathbf{X}| = |a|\,|\mathbf{X}|$.
- $|\mathbf{X}| \geq 0$, with equality if and only if $\mathbf{X} = \mathbf{0}$.
- $|\mathbf{X} - \mathbf{Y}| \geq 0$, with equality if and only if $\mathbf{X} = \mathbf{Y}$.
- $\mathbf{X} \cdot \mathbf{Y} = \mathbf{Y} \cdot \mathbf{X}$.
- $\mathbf{X} \cdot (\mathbf{Y} + \mathbf{Z}) = \mathbf{X} \cdot \mathbf{Y} + \mathbf{X} \cdot \mathbf{Z}$.
- $(c\mathbf{X}) \cdot \mathbf{Y} = \mathbf{X} \cdot (c\mathbf{Y}) = c(\mathbf{X} \cdot \mathbf{Y})$.

2.2.1 Line Segment in R^n

The equation of a line through a point $\mathbf{X}_0 = (x_0, y_0, z_0)$ in R^3 can be written parametrically as

$$x = x_0 + u_1 t, \quad y = y_0 + u_2 t, \quad z = z_0 + u_3 t, \quad -\infty < t < \infty,$$

where u_1, u_2, and u_3 are not all zero. We write this in vector form as

$$\mathbf{X} = \mathbf{X}_0 + t\mathbf{U}, \quad -\infty < t < \infty, \tag{2.9}$$

with $\mathbf{U} = (u_1, u_2, u_3)$, and we say that the line is through $\mathbf{X}_0$ in the direction of $\mathbf{U}$. There are many ways to represent a given line parametrically.

For example,

$$\mathbf{X} = \mathbf{X}_0 + s\mathbf{V}, \quad -\infty < s < \infty, \tag{2.10}$$

represents the same line as (2.9) if and only if $\mathbf{V} = a\mathbf{U}$ for some nonzero real number a.

Then the line is traversed in the same direction as s and t vary from $-\infty$ to ∞ if $a > 0$, or in opposite directions if $a < 0$. To write the parametric equation of a line through two points $\mathbf{X}_0$ and $\mathbf{X}_1$ in R^3.

We take $\mathbf{U} = \mathbf{X}_1 - \mathbf{X}_0$ in (2.9), which yields

$$\mathbf{X} = \mathbf{X}_0 + t(\mathbf{X}_1 - \mathbf{X}_0) = t\mathbf{X}_1 + (1 - t)\mathbf{X}_0, \quad -\infty < t < \infty.$$

The line segment from $\mathbf{X}_0$ to $\mathbf{X}_1$ consists of those points for which $0 \le t \le 1$.

Suppose that $\mathbf{X}_0$ and $\mathbf{U}$ are in R^n and $\mathbf{U} \neq \mathbf{0}$.

Then the line through $\mathbf{X}_0$ in the direction of $\mathbf{U}$ is the set of all points in R^n of the form

$$\mathbf{X} = \mathbf{X}_0 + t\mathbf{U}, \quad -\infty < t < \infty.$$

A set of points of the form

$$\mathbf{X} = \mathbf{X}_0 + t\mathbf{U}, \quad t_1 \le t \le t_2,$$

is called a line segment. The line segment from $\mathbf{X}_0$ to $\mathbf{X}_1$ is the set of points of the form

$$\mathbf{X} = \mathbf{X}_0 + t(\mathbf{X}_1 - \mathbf{X}_0) = t\mathbf{X}_1 + (1 - t)\mathbf{X}_0, \quad 0 \le t \le 1.$$

2.3 Neighbourhoods and Open Sets in R^n

If $\varepsilon > 0$, the ε-neighborhood of a point $\mathbf{X}_0$ in R^n is the set

$$N_\varepsilon(\mathbf{X}_0) = \{\mathbf{X} \,|\, |\mathbf{X} - \mathbf{X}_0| < \varepsilon\}.$$

$N_\varepsilon(\mathbf{X}_0)$ in R^2

We are going to de ne neighborhood, interior point, interior of a set, open set, closed set,limit point, boundary point, boundary of a set, closure of a set, isolated point, exterior point, and exterior of a set.

Example: Let S be the set of points in $\mathbb{R}^2$ in the square bounded by the lines $x = \pm 1$, $y = \pm 1$, except for the origin and the points on the vertical lines $x = \pm 1$ thus,

$$S = \{(x, y) : (x, y) \neq (0, 0),\ -1 < x < 1,\ \ -1 \leq y \leq 1\}.$$

Every point of S not on the lines $y = \pm 1$ is an interior point.

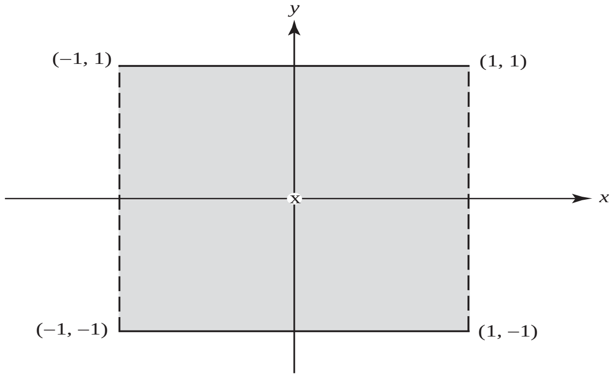

Figure 2.1: The set S

$$S^0 = \{(x, y) : (x, y) \quad (0, 0),\ \ -1 < x, y < 1\}.$$

S is a deleted neighborhood of $(0, 0)$ and is neither open nor closed.

The closure of S is

$$\overline{S} = \{(x, y) : -1 \leq x, y \leq 1\},$$

and every point of $\overline{S}$ is a limit point of S.

The origin and the perimeter of S form ∂S, the boundary of S. The exterior of S consists of all points (x, y) such that $|x| > 1$ or $|y| > 1$. The origin is an isolated point of S^c.

Example: If $\mathbf{X}_0$ is a point in $\mathbb{R}^n$ and r is a positive number, the open n-ball of radius r about $\mathbf{X}_0$ is the set

$$B_r(\mathbf{X}_0) = \{\mathbf{X} : |\mathbf{X} - \mathbf{X}_0| < r\}.$$

Thus, ε-neighborhoods are open n-balls. If $\mathbf{X}_1$ is in $S_r(\mathbf{X}_0)$ and

$$|\mathbf{X} - \mathbf{X}_1| < \varepsilon = r - |\mathbf{X} - \mathbf{X}_0|,$$

then $\mathbf{X}$ is in $S_r(\mathbf{X}_0)$. Thus, $S_r(\mathbf{X}_0)$ contains an ε-neighborhood of each of its points, and is therefore open.

We can show that the closure of $B_r(\mathbf{X}_0)$ is the closed n-ball of radius r about $\mathbf{X}_0$, de ned by

$$\overline{S}_r(\mathbf{X}_0) = \{\mathbf{X} : |\mathbf{X} - \mathbf{X}_0| \le r\}.$$

Remark: Open and closed n-balls are generalizations to $\mathbb{R}^n$ of open and closed intervals.

Lemma: If $\mathbf{X}_1$ and $\mathbf{X}_2$ are in $S_r(\mathbf{X}_0)$ for some $r > 0$, then so is every point on the line segment from $\mathbf{X}_1$ to $\mathbf{X}_2$.

Proof: The line segment is given by

$$\mathbf{X} = t\mathbf{X}_2 + (1 - t)\mathbf{X}_1, \quad 0 < t < 1.$$

Suppose that $r > 0$. If

$$|\mathbf{X}_1 - \mathbf{X}_0| < r, \quad |\mathbf{X}_2 - \mathbf{X}_0| < r,$$

and $0 < t < 1$, then

$$\begin{aligned} |\mathbf{X} - \mathbf{X}_0| &= |t\mathbf{X}_2 + (1 - t)\mathbf{X}_1 - t\mathbf{X}_0 - (1 - t)\mathbf{X}_0| \\ &= |t(\mathbf{X}_2 - \mathbf{X}_0) + (1 - t)\mathbf{X}_1 - \mathbf{X}_0)| \\ &< tr + (1 - t)r = r. \end{aligned}$$

2.4 Convergence of a Sequence in $\mathbb{R}^n$

A sequence of points $\{\mathbf{X}_r\}$ in $\mathbb{R}^n$ converges to the limit $\overline{\mathbf{X}}$ if

$$\lim_{r\to\infty} |\mathbf{X}_r - \overline{\mathbf{X}}| = 0.$$

In this case we write

$$\lim_{r\to\infty} \mathbf{X}_r = \overline{\mathbf{X}}.$$

Theorem: Let

$$\overline{\mathbf{X}} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n) \quad \text{and} \quad \mathbf{X}_r = (x_{1r}, x_{2r}, \ldots, x_{nr}), \quad r \ge 1.$$

Then $\lim_{r\to\infty} \mathbf{X}_r = \overline{\mathbf{X}}$ if and only if

$$\lim_{r\to\infty} x_{ir} = \overline{x}_i, \quad 1 \le i \le n;$$

that is, a sequence $\{\mathbf{X}_r\}$ of points in $\mathbb{R}^n$ converges to a limit $\overline{\mathbf{X}}$ if and only if the sequences of components of $\{\mathbf{X}_r\}$ converge to the respective components of $\overline{\mathbf{X}}$.

Theorem (Cauchy's Convergence Criterion): A sequence $\{\mathbf{X}_r\}$ in $\mathbb{R}^n$ converges if and only if for each $\varepsilon > 0$ there is an integer K such that

$$|\mathbf{X}_r - \mathbf{X}_s| < \varepsilon \quad \text{if} \quad r, s \geq K.$$

Diameter of a Set: If S is a nonempty subset of $\mathbb{R}^n$, then

$$d(S) = \sup\{|\mathbf{X} - \mathbf{Y}| : \mathbf{X}, \mathbf{Y} \in S\}$$

is the diameter of S.

If $d(S) < \infty$, S is bounded; if $d(S) = \infty$, S is unbounded.

2.5 Principle of nested sets

Theorem: If $S_1, S_2, \ldots$ are closed nonempty subsets of $\mathbb{R}^n$such that

$$S_1 \supset S_2 \supset \cdots \supset S_r \supset \cdots \tag{2.11}$$

and

$$\lim_{r\to\infty} d(S_r) = 0, \tag{2.12}$$

then the intersection

$$I = \bigcap_{r=1}^{\infty} S_r$$

contains exactly one point.

Proof: Let $\{\mathbf{X}_r\}$ be a sequence such that $\mathbf{X}_r \in S_r$ $(r \geq 1)$.

Because of $S_1 \supset S_2 \supset \cdots \supset S_r \supset \cdots$, $\mathbf{X}_r \in S_k$ if $r \geq k$, so

$$|\mathbf{X}_r - \mathbf{X}_s| < d(S_k) \quad \text{if} \quad r, s \geq k.$$

From $\lim_{r\to\infty} d(S_r) = 0$ and Cauchy's convergence theorem, $\mathbf{X}_r$ converges to a limit $\overline{\mathbf{X}}$. Since $\overline{\mathbf{X}}$ is a limit point of every S_k and every S_k is closed, $\overline{\mathbf{X}}$ is in every S_k (A set is closed if and only if it contains all its limit points). Therefore, $\overline{\mathbf{X}} \in I$, so $I \neq \emptyset$. Moreover, $\overline{\mathbf{X}}$ is the only point in I, since if $\mathbf{Y} \in I$, then

$$|\overline{\mathbf{X}} - \mathbf{Y}| \leq d(S_k), \quad k \geq 1,$$

and (2.12) implies that $\mathbf{Y} = \overline{\mathbf{X}}$.

2.6 Heine-Borel Theorem

We are going to state and prove the Heine-Borel theorem for $\mathbb{R}^n$.

This theorem concerns compact sets. As in $\mathbb{R}$, a compact set in $\mathbb{R}^n$ is a closed and bounded set.

Recall that a collection H of open sets is an open covering of a set S if

$$S \subset \cup\{H : H \in \mathrm{H}\}.$$

Theorem: If H is an open covering of a compact subset S, then S can be covered by nitely many sets from H.

Proof: The proof is by contradiction. We rst consider the case where $n = 2$, so that you can visualize the method.

Suppose that there is a covering H for S from which it is impossible to select a nite subcovering.

Since S is bounded, S is contained in a closed square

$$T = \{(x, y) | a_1 \leq x \leq a_1 + L, a_2 \leq x \leq a_2 + L\}$$

with sides of length L

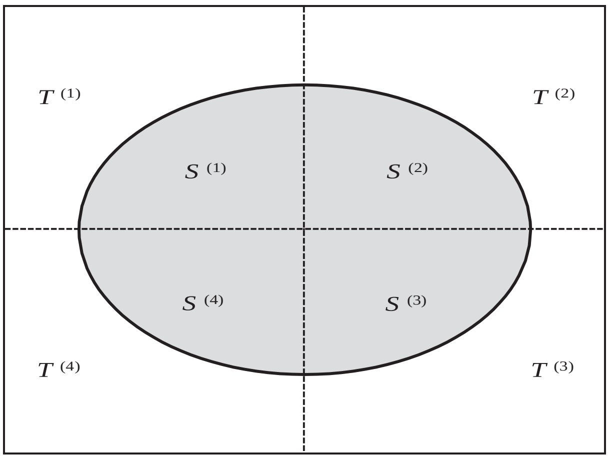

Figure 2.2: Heine-Borel Theorem for $n = 2$

Bisecting the sides of T leads to four closed squares, $T^{(1)}, T^{(2)}, T^{(3)}$, and $T^{(4)}$, with sides of length $L/2$. Let

$$S^{(i)} = S \cap T^{(i)}, \quad 1 \leq i \leq 4.$$

Each $S^{(i)}$, being the intersection of closed sets, is closed, and

$$S = \bigcup_{i=1}^{4} S^{(i)}.$$

Moreover, H covers each $S^{(i)}$, but at least one $S^{(i)}$ cannot be covered by any nite subcollection of H, since if all the $S^{(i)}$ could be, then so could S. Let S_1 be a set with this property, chosen from $S^{(1)}$, $S^{(2)}$, $S^{(3)}$, and $S^{(4)}$.

We are now back to the situation we started from: a compact set S_1 covered by H, but not by any nite subcollection of H. However, S_1 is contained in a square T_1 with sides of length $L/2$ instead of L. Bisecting the sides of T_1 and repeating the argument, we obtain a subset S_2 of S_1 that has the same properties as S, except that it is contained in a square with sides of length $L/4$. Continuing in this way produces a sequence of nonempty closed sets S_0 $(= S)$, S_1, S_2, . . . , such that $S_k \supset S_{k+1}$ and $d(S_k) \leq L/2^{k-1/2}$ $(k \geq 0)$.

From Principle of Nested Sets Theorem, there is a point $\overline{\mathbf{X}}$ in $\bigcap_{k=1}^{\infty} S_k$.

Since $\overline{\mathbf{X}} \in S$, there is an open set H in H that contains $\overline{\mathbf{X}}$, and this H must also contain some ε-neighborhood of $\overline{\mathbf{X}}$. Since every $\mathbf{X}$ in S_k satis es the inequality

$$|\mathbf{X} - \overline{\mathbf{X}}| \leq 2^{-k+1/2}L,$$

it follows that $S_k \subset H$ for k su ciently large.

This contradicts our assumption on H, which led us to believe that no S_k could be covered by a nite number of sets from H.

Consequently, this assumption must be false: H must have a nite subcollection that covers S. This completes the proof for $n = 2$.

The idea of the proof is the same for $n > 2$. The counterpart of the square T is the hypercube with sides of length L:

$$T = \{(x_1, x_2, \ldots, x_n) : a_i \leq x_i \leq a_i + L, i = 1, 2, \ldots, n\}.$$

Halving the intervals of variation of the n coordinates x_1, x_2, . . . , x_n divides T into 2^n closed hypercubes with sides of length $L/2$:

$$T^{(i)} = \{(x_1, x_2, \ldots, x_n) : b_i \leq x_i \leq b_i + L/2, 1 \leq i \leq n\},$$

where $b_i = a_i$ or $b_i = a_i + L/2$. If no nite subcollection of H covers S, then at least one of these smaller hypercubes must contain a subset of S that is not covered by any nite subcollection of S. Now the proof proceeds as for $n = 2$.

Remark: The Bolzano Weierstrass theorem is valid in $\mathbb{R}^n$; its proof is the same as in R.

2.7 Connected Sets in $\mathbb{R}^n$

A subset S of $\mathbb{R}^n$ is connected if it is impossible to represent S as the union of two disjoint nonempty sets such that neither contains a limit point of the other.

If S cannot be expressed as $S = A \cup B$, where

$$A \neq \emptyset, \quad B \neq \emptyset, \quad \overline{A} \cap B = \emptyset, \quad \text{and} \quad A \cap \overline{B} = \emptyset. \tag{2.13}$$

If S can be expressed in this way, then S is disconnected.

Example: The empty set and singleton sets are connected, because they cannot be represented as the union of two disjoint nonempty sets.

Example: The space $\mathbb{R}^n$ is connected.

If $\mathbb{R}^n = A \cup B$ with $\overline{A} \cap B = \emptyset$ and $A \cap \overline{B} = \emptyset$, then $\overline{A} \subset A$ and $\overline{B} \subset B$.
That is, A and B are both closed and therefore are both open.
Since the only nonempty subset of $\mathbb{R}^n$ that is both open and closed is $\mathbb{R}^n$ itself, one of A and B is $\mathbb{R}^n$ and the other is empty.

2.7.1 Polygonal Path

If $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k$ are points in $\mathbb{R}^n$

Let L_i is the line segment from $\mathbf{X}_i$ to $\mathbf{X}_{i+1}$, $1 \leq i \leq k-1$, we say that L_1, L_2, $\ldots$, L_{k-1} form a polygonal path from $\mathbf{X}_1$ to $\mathbf{X}_k$.

We say that $\mathbf{X}_1$ and $\mathbf{X}_k$ are connected by the polygonal path.

2.8 Polygonally Connected Set

A set S is polygonally connected if every pair of points in S can be connected by a polygonal path lying entirely in S.

Theorem: An open set S in $\mathbb{R}^n$ is connected if and only if it is polygonally connected.

Proof: For su ciency, we will show that if S is disconnected, then S is not polygonally connected.

Let $S = A \cup B$, where A and B satisfy

$$A \neq \emptyset, \quad B \neq \emptyset, \quad \overline{A} \cap B = \emptyset, \quad \text{and} \quad A \cap \overline{B} = \emptyset.$$

.

Suppose that $\mathbf{X}_1 \in A$ and $\mathbf{X}_2 \in B$, and assume that there is a polygonal path in S connecting $\mathbf{X}_1$ to $\mathbf{X}_2$. Then some line segment L in this path must contain a point $\mathbf{Y}_1$ in A and a point $\mathbf{Y}_2$ in B.

The line segment

$$\mathbf{X} = t\mathbf{Y}_2 + (1-t)\mathbf{Y}_1, \quad 0 \le t \le 1,$$

is part of L and therefore in S. Now de ne

$$\rho = \sup\{\tau : tY_2 + (1-t)\mathbf{Y}_1 \in A,\ 0 \le t \le \tau \le 1\}.$$

Let $\mathbf{X}_\rho = \rho\mathbf{Y}_2 + (1-\rho)\mathbf{Y}_1$. Then $\mathbf{X}_\rho \in \overline{A} \cap \overline{B}$.

However, since $\mathbf{X}_\rho \in A \cup B$ and $\overline{A} \cap B = A \cap \overline{B} = \emptyset$, this is impossible.

Therefore, the assumption that there is a polygonal path in S from $\mathbf{X}_1$ to $\mathbf{X}_2$ must be false.

For necessity, suppose that S is a connected open set and $\mathbf{X}_0 \in S$. Let A be the set consisting of $\mathbf{X}_0$ and the points in S can be connected to $\mathbf{X}_0$ by polygonal paths in S. Let B be set of points in S that cannot be connected to $\mathbf{X}_0$ by polygonal paths. If $\mathbf{Y}_0 \in S$, then S contains an ε-neighborhood $N_\varepsilon(\mathbf{Y}_0)$ of $\mathbf{Y}_0$, since S is open. Any point $\mathbf{Y}_1$ in $N_\varepsilon(\mathbf{Y}_0$ can be connected to $\mathbf{Y}_0$ by the line segment

$$\mathbf{X} = t\mathbf{Y}_1 + (1-t)\mathbf{Y}_0, \quad 0 \le t \le 1,$$

which lies in $N_\varepsilon(\mathbf{Y}_0)$ and therefore in S. This implies that $\mathbf{Y}_0$ can be connected to $\mathbf{X}_0$ by a polygonal path in S if and only if every member of $N_\varepsilon(\mathbf{Y}_0)$ can also. Thus, $N_\varepsilon(\mathbf{Y}_0) \subset A$ if $\mathbf{Y}_0 \in A$, and $N_\varepsilon(\mathbf{Y}_0) \in B$ if $\mathbf{Y}_0 \in B$. Therefore, A and B are open. Since $A \cap B = \emptyset$, this implies that $A \cap \overline{B} = \overline{A} \cap B = \emptyset$. Since A is nonempty ($\mathbf{X}_0 \in A$), it now follows that $B = \emptyset$, since if $B \neq \emptyset$, S would be disconnected. Therefore, $A = S$, which completes the proof of necessity.

Remark: Any polygonally connected set, open or not, is connected. The converse is false. A set (not open) may be connected but not polygonally connected.

Regions in $\mathbb{R}^n$: A region S in $\mathbb{R}^n$ is the union of an open connected set with some, all, or none of its boundary; thus, S^0 is connected, and every point of S is a limit point of S^0.

Example: Intervals are the only regions in $\mathbb{R}$. The n-ball $B_r(\mathbf{X}_0)$ is a region in $\mathbb{R}^n$, as is its closure $\overline{S}_r(\mathbf{X}_0)$. The set $S = \{(x, y) : x^2 + y^2 \le 1 \text{ or } x^2 + y^2 \ge 4\}$ is not a region in $\mathbb{R}^2$, since it is not connected.

The set S_1 obtained by adding the line segment

$$L_1 : \quad \mathbf{X} = t(0, 2) + (1-t)(0, 1), \quad 0 < t < 1,$$

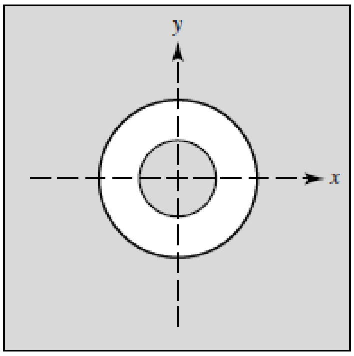

Figure 2.3: Disconnected set which is not a region

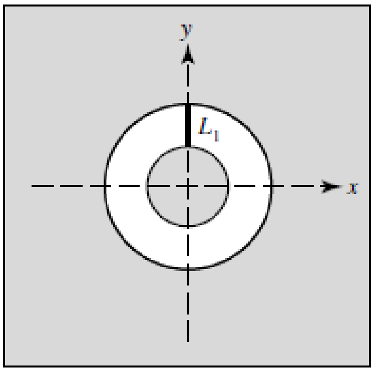

Figure 2.4: A connected set which is not a region

to S is connected but is not a region, since points on the line segment are not limit points of S_1^0. The set S_2 obtained by adding to S_1 the points in the rst quadrant bounded by the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$ and the line segments L_1 and

$$L_2: \quad X = t(2,0) + (1-t)(1,0), \quad 0 < t < 1,$$

is a region.

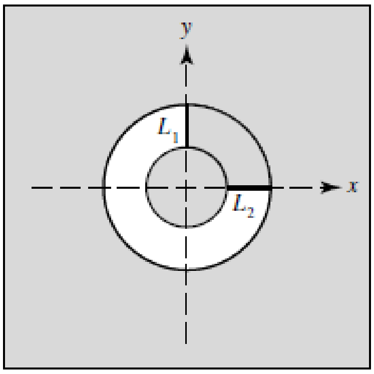

Figure 2.5: A region

2.9 Sequences in $\mathbb{R}^n$

A sequence $\{\mathbf{X}_r\}$ of points in $\mathbb{R}^n$ converges to a limit $\overline{\mathbf{X}}$ if and only if for every $\varepsilon > 0$ there is an integer K such that

$$|\mathbf{X}_r - \overline{\mathbf{X}}| < \varepsilon \quad \text{if} \quad r \geq K.$$

The $\mathbb{R}^n$ de nitions of divergence, boundedness, subsequence, and sums, di erences, and constant multiples of sequences are analogous to those we discussed in Analysis I.

Since $\mathbb{R}^n$ is not ordered for $n > 1$, monotonicity, limits inferior and superior of sequences in $\mathbb{R}^n$, and divergence to $\pm\infty$ are unde ned for $n > 1$.

Products and quotients of members of $\mathbb{R}^n$ are also unde ned if $n > 1$.

Several theorems from Analysis I remain valid for sequences in $\mathbb{R}^n$, with proofs unchanged, provided that $|\quad|$ is interpreted as distance in $\mathbb{R}^n$.

1. uniqueness of the limit.

2. Boundedness of a convergent sequence.

3. Concerning limits of sums, differences, and constant multiples of convergent sequences.

4. Every subsequence of a convergent sequence converges to the limit of the sequence.

2.10 Domain of Function of n Variable

We denote the domain of a function f by D_f and the value of f at a point $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ by $f(\mathbf{X})$ or $f(x_1, x_2, \ldots, x_n)$.

If a function is defined by a formula such as

$$f(\mathbf{X}) = \left(1 - x_1^2 - x_2^2 - \cdots - x_n^2\right)^{1/2} \tag{2.14}$$

$$g(\mathbf{X}) = \left(1 - x_1^2 - x_2^2 - \cdots - x_n^2\right)^{-1} \tag{2.15}$$

without specification of its domain, it is to be understood that its domain is the largest subset of $\mathbb{R}^n$ for which the formula defines a unique real number.

2.11 Limit at a Point of a Function of n Variables

A function $f(\mathbf{X})$ approaches the limit L as $\mathbf{X}$ approaches $\mathbf{X}_0$ and write

$$\lim_{\mathbf{X}\to\mathbf{X}_0} f(\mathbf{X}) = L,$$

if $\mathbf{X}_0$ is a limit point of D_f and, for every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|f(\mathbf{X}) - L| < \varepsilon$$

for all $\mathbf{X}$ in D_f such that

$$0 < |\mathbf{X} - \mathbf{X}_0| < \delta.$$

Example: If $g(x, y) = 1 - x^2 - 2y^2$, then

$$\lim_{(x,y)\to(x_0,y_0)} g(x, y) = 1 - x_0^2 - 2y_0^2 \tag{2.16}$$

for every (x_0, y_0).

To see this, we write

$$\begin{aligned} |g(x, y) - g(x_0 - y_0)| &= |(1 - x^2 - 2y^2) - (1 - x_0^2 - 2y_0^2)| \\ &\leq |x^2 - x_0^2| + 2|y^2 - y_0^2| \\ &= |(x + x_0)(x - x_0)| \\ &\quad +2|(y + y_0)(y - y_0)| \\ &\leq |\mathbf{X} - \mathbf{X}_0|(|x + x_0| + 2|y + y_0|), \end{aligned} \tag{2.17}$$

since

$$|x - x_0| \leq |\mathbf{X} - \mathbf{X}_0| \quad \text{and} \quad |y - y_0| \leq |\mathbf{X} - \mathbf{X}_0|.$$

If $|\mathbf{X} - \mathbf{X}_0| < 1$, then $|x| < |x_0| + 1$ and $|y| < |y_0| + 1$.

This and (2.17) imply that

$$|g(x, y) - g(x_0 - y_0)| < K|\mathbf{X} - \mathbf{X}_0| \quad \text{if} \quad |\mathbf{X} - \mathbf{X}_0| < 1,$$

where

$$K = (2|x_0| + 1) + 2(2|y_0| + 1).$$

Therefore, if $\varepsilon > 0$ and

$$|\mathbf{X} - \mathbf{X}_0| < \delta = \min\{1, \varepsilon/K\},$$

then

$$\left| g(x, y) - (1 - x_0^2 - 2y_0^2) \right| < \varepsilon.$$

Example: The function

$$h(x, y) = \frac{\sin \sqrt{1 - x^2 - 2y^2}}{\sqrt{1 - x^2 - 2y^2}}$$

is de ned only on the interior of the region bounded by the ellipse

$$x^2 + 2y^2 = 1.$$

It is not de ned at any point of the ellipse itself or on any deleted neighborhood of

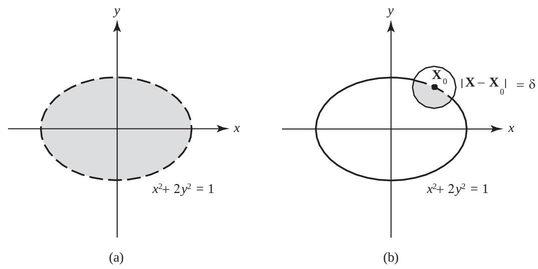

Figure 2.6: Domain of the function

such a point. Nevertheless,

$$\lim_{(x,y)\to(x_0,y_0)} h(x, y) = 1 \tag{2.18}$$

if

$$x_0^2 + 2y_0^2 = 1. \tag{2.19}$$

To see this, let

$$u(x, y) = \sqrt{1 - x^2 - 2y^2}.$$

Then

$$h(x, y) = \frac{\sin u(x, y)}{u(x, y)}. \tag{2.20}$$

Recall that

$$\lim_{r \to 0} \frac{\sin r}{r} = 1.$$

Therefore, if $\varepsilon > 0$, there is a $\delta_1 > 0$ such that

$$\left| \frac{\sin u}{u} - 1 \right| < \varepsilon \quad \text{if} \quad 0 < |u| < \delta_1. \tag{2.21}$$

From previous example, we have

$$\lim_{(x,y) \to (x_0, y_0)} (1 - x^2 - 2y^2) = 0.$$

If (2.19) holds, so there is a $\delta > 0$ such that

$$0 < u^2(x, y) = (1 - x^2 - 2y^2) < \delta_1^2,$$

if $\mathbf{X} = (x, y)$ is in the interior of the ellipse and $|\mathbf{X} - \mathbf{X}_0| < \delta$; that is, if $\mathbf{X}$ is in the shaded region.

Therefore,

$$0 < u = \sqrt{1 - x^2 - 2y^2} < \delta_1 \tag{2.22}$$

if $\mathbf{X}$ is in the interior of the ellipse and $|\mathbf{X} - \mathbf{X}_0| < \delta$; that is, if $\mathbf{X}$ is in the shaded region. This, (2.20), and (2.21) imply that

$$|h(x, y) - 1| < \varepsilon$$

for such $\mathbf{X}$, which is the required result.

Theorem: If $\lim_{\mathbf{X} \to \mathbf{X}_0} f(\mathbf{X})$ exists, then it is unique.

Proof: See lecture.

Example: The function

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

is de ned everywhere in $\mathbb{R}^2$ except at $(0, 0)$. Does $\lim_{(x,y) \to (0,0)} f(x, y)$ exist?

If we try to answer this question by letting (x, y) approach $(0, 0)$ along the line $y = x$, we see the functional values

$$f(x, x) = \frac{x^2}{2x^2} = \frac{1}{2}$$

and conclude that the limit is $1/2$.

However, if we let (x, y) approach $(0, 0)$ along the line $y = -x$, we see the functional values

$$f(x, -x) = -\frac{x^2}{2x^2} = -\frac{1}{2}$$

and conclude that the limit equals $-1/2$.

In fact, they are both incorrect. What we have shown is that

$$\lim_{x\to 0} f(x, x) = \frac{1}{2} \quad \text{and} \quad \lim_{x\to 0} f(x, -x) = -\frac{1}{2}.$$

Since $\lim_{x\to 0} f(x, x)$ and $\lim_{x\to 0} f(x, -x)$ must both equal $\lim_{(x,y)\to(0,0)} f(x, y)$.

Theorem: Suppose that f and g are de ned on a set D, $\mathbf{X}_0$ is a limit point of D, and

$$\lim_{\mathbf{X}\to\mathbf{X}_0} f(\mathbf{X}) = L_1, \quad \lim_{\mathbf{X}\to\mathbf{X}_0} g(\mathbf{X}) = L_2.$$

Then

$$\lim_{\mathbf{X}\to\mathbf{X}_0} (f + g)(\mathbf{X}) = L_1 + L_2, \tag{2.23}$$

$$\lim_{\mathbf{X}\to\mathbf{X}_0} (f - g)(\mathbf{X}) = L_1 - L_2, \tag{2.24}$$

$$\lim_{\mathbf{X}\to\mathbf{X}_0} (fg)(\mathbf{X}) = L_1 L_2, \tag{2.25}$$

$$if\,(L_2) = 0,$$

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \frac{f}{g}(\mathbf{X}) = \frac{L_1}{L_2}. \tag{2.26}$$

2.12 In nite Limits and Limits at $\mathbf{X} \to \infty$

We say that $f(\mathbf{X})$ approaches ∞ as $\mathbf{X}$ approaches $\mathbf{X}_0$

$$\lim_{\mathbf{X}\to\mathbf{X}_0} f(\mathbf{X}) = \infty$$

if $\mathbf{X}_0$ is a limit point of D_f and, M, there is a $\delta > 0$ such that

$$f(\mathbf{X}) > M \quad \text{whenever} \quad 0 < |\mathbf{X} - \mathbf{X}_0| < \delta \quad \text{and} \quad \mathbf{X} \in D_f.$$

We say that

$$\lim_{\mathbf{X}\to\mathbf{X}_0} f(\mathbf{X}) = -\infty$$

$$if$$

$$\lim_{\mathbf{X}\to\mathbf{X}_0} (-f)(\mathbf{X}) = \infty.$$

Example: If

$$f(\mathbf{X}) = (1 - x_1^2 - x_2^2 - \cdots - x_n^2)^{-1/2},$$

then

$$\lim_{\mathbf{X}\to\mathbf{X}_0} f(\mathbf{X}) = \infty$$

if $|\mathbf{X}_0| = 1$, because

$$f(\mathbf{X}) = \frac{1}{|\mathbf{X} - \mathbf{X}_0|},$$

so

$$f(\mathbf{X}) > M \quad \text{if} \quad 0 < |\mathbf{X} - \mathbf{X}_0| < \delta = \frac{1}{M}.$$

Example: If

$$f(x, y) = \frac{1}{x + 2y + 1},$$

then $\lim_{(x,y)\to(1,-1)} f(x, y)$ does not exist (why not?).
But

$$\lim_{(x,y)\to(1,-1)} |f(x, y)| = \infty.$$

To see this, we observe that

$$\begin{aligned} |x + 2y + 1| &= |(x - 1) + 2(y + 1)| \\ &\le \sqrt{5}|\mathbf{X} - \mathbf{X}_0| \quad \text{(by Schwarz's inequality)}, \end{aligned}$$

where $\mathbf{X}_0 = (1, -1)$. So

$$|f(x, y)| = \frac{1}{|x + 2y + 1|} \ge \frac{1}{\sqrt{5}|\mathbf{X} - \mathbf{X}_0|}.$$

Therefore,

$$|f(x, y)| > M \quad \text{if} \quad 0 < |\mathbf{X} - \mathbf{X}_0| < \frac{1}{M\sqrt{5}}.$$

Example: The function

$$f(x, y, z) = \frac{\sin\left(\dfrac{1}{x^2 + y^2 + z^2}\right)}{x^2 + y^2 + z^2}$$

assumes arbitrarily large values in every neighborhood of $(0, 0, 0)$.
For example, if $\mathbf{X}_k = (x_k, y_k, z_k)$, where

$$x_k = y_k = z_k = \frac{1}{\sqrt{3\left(k + \frac{1}{2}\right)\pi}},$$

then

$$f(\mathbf{X}_k) = \left(k + \frac{1}{2}\right)\pi.$$

However, this does not imply that $\lim_{\mathbf{X}\to\mathbf{0}} f(\mathbf{X}) = \infty$. Since, for example, every neighborhood of $(0, 0, 0)$ also contains points

$$\mathbf{X}_k = \left(\frac{1}{\sqrt{3k\pi}}, \frac{1}{\sqrt{3k\pi}}, \frac{1}{\sqrt{3k\pi}} \right).$$

For which $f(\overline{\mathbf{X}_k}) = 0$.

2.12.1 Limit at In nity

If D_f is unbounded, we say that

$$\lim_{|\mathbf{X}|\to\infty} f(\mathbf{X}) = L \quad (\text{ nite})$$

if for every $\varepsilon > 0$, there is a number R such that

$$|f(\mathbf{X}) - L| < \varepsilon \quad \text{whenever} \quad |\mathbf{X}| \geq R \quad \text{and} \quad \mathbf{X} \in D_f.$$

Example: If

$$f(x, y, z) = \cos\left(\frac{1}{x^2 + 2y^2 + z^2} \right),$$

then

$$\lim_{|\mathbf{X}|\to\infty} f(\mathbf{X}) = 1. \tag{2.27}$$

To see this, we recall that the continuity of $\cos u$ at $u = 0$ implies that for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|\cos u - 1| < \varepsilon \quad \text{if} \quad |u| < \delta.$$

Since

$$\frac{1}{x^2 + 2y^2 + z^2} \leq \frac{1}{|\mathbf{X}|^2}.$$

It follows that if $|\mathbf{X}| > 1/\sqrt{\delta}$, then

$$\frac{1}{x^2 + 2y^2 + z^2} < \delta.$$

Therefore,

$$|f(\mathbf{X}) - 1| < \varepsilon.$$

Example: Consider the function de ned only on the domain

$$D = \{(x, y) : 0 < y \leq ax\}, \quad 0 < a < 1,$$

by

$$f(x, y) = \frac{1}{x - y}.$$

We will show that

$$\lim_{|\mathbf{X}| \to \infty} f(x, y) = 0. \tag{2.28}$$

It is important to keep in mind that we need only consider (x, y) in D, since f is not de ned elsewhere.

In D,

$$x - y \geq x(1 - a) \tag{2.29}$$

and

$$|\mathbf{X}|^2 = x^2 + y^2 \leq x^2(1 + a^2).$$

So

$$x \geq \frac{|\mathbf{X}|}{\sqrt{1 + a^2}}.$$

This and (2.29) imply that

$$x - y \geq \frac{1 - a}{\sqrt{1 + a^2}}|\mathbf{X}|, \quad \mathbf{X} \in D.$$

So

$$|f(x, y)| \leq \frac{\sqrt{1 + a^2}}{1 - a} \frac{1}{|\mathbf{X}|}, \quad \mathbf{X} \in D.$$

This and (2.29) imply that

$$x - y \geq \frac{1 - a}{\sqrt{1 + a^2}}|\mathbf{X}|, \quad \mathbf{X} \in D.$$

So

$$|f(x, y)| \leq \frac{\sqrt{1 + a^2}}{1 - a} \frac{1}{|\mathbf{X}|}, \quad \mathbf{X} \in D.$$

Therefore,

$$|f(x, y)| < \varepsilon$$

if $\mathbf{X} \in D$ and

$$|\mathbf{X}| > \frac{\sqrt{1 + a^2}}{1 - a} \frac{1}{\varepsilon}.$$

Remarks: In the same manner we can de ne $\lim_{|\mathbf{X}| \to \infty} f(\mathbf{X}) = \infty$ and $\lim_{|\mathbf{X}| \to \infty} f(\mathbf{X}) = -\infty$. We will have the following notion $\lim_{\mathbf{X} \to \mathbf{X}_0} f(\mathbf{X})$ exists means that $\lim_{\mathbf{X} \to \mathbf{X}_0} f(\mathbf{X}) = L$, where L is nite; to leave open the possibility that $L = \pm\infty$.

We will say that $\lim_{\mathbf{X} \to \mathbf{X}_0} f(\mathbf{X})$ exists in the extended reals. A similar convention applies to limits as $|\mathbf{X}| \to \infty$.

2.13 Continuity

If $\mathbf{X}_0$ is in D_f and is a limit point of D_f, then we say that f is continuous at $\mathbf{X}_0$ if

$$\lim_{\mathbf{X}\to\mathbf{X}_0} f(\mathbf{X}) = f(\mathbf{X}_0).$$

Theorem: Suppose that $\mathbf{X}_0$ is in D_f and is a limit point of D_f. Then f is continuous at $\mathbf{X}_0$ if and only if for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(\mathbf{X}) - f(\mathbf{X}_0)| < \varepsilon$$

whenever

$$|\mathbf{X} - \mathbf{X}_0| < \delta \quad \text{and} \quad \mathbf{X} \in D_f.$$

Example: The function

$$f(x, y) = 1 - x^2 - 2y^2$$

is continuous on $\mathbb{R}^2$.

Solution: See lecture.

Example: Consider the function

$$h(x, y) = \begin{cases} \dfrac{\sin\sqrt{1-x^2-2y^2}}{\sqrt{1-x^2-2y^2}}, & x^2 + 2y^2 < 1, \\ 1, & x^2 + 2y^2 = 1, \end{cases}$$

then it follows from the example we have discussed that h is continuous on the ellipse

$$x^2 + 2y^2 = 1.$$

Example: Can we rede ne the function

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

to make it continuous at $(0, 0)$.

The limit

$$\lim_{(x,y)\to(0,0)} f(x, y)$$

does not exist.

Consequently, it is impossible to de ne the function at origin to make it continuous.

Theorem: If f and g are continuous on a set S in $\mathbb{R}^n$, then so are $f + g$, $f - g$, and fg. Also, f/g is continuous at each $\mathbf{X}_0$ in S such that $g(\mathbf{X}_0) \neq 0$.

End 26 lec 93

2.14 Vector Valued Functions

27 Lecture

Suppose that $g_1, g_2, \ldots, g_n$ are real-valued functions de ned on a subset T of $\mathbb{R}^m$.
We de ne the vector-valued function $\mathbf{G}$ on T by

$$\mathbf{G}(\mathbf{U}) = (g_1(\mathbf{U}), g_2(\mathbf{U}), \ldots, g_n(\mathbf{U})), \quad \mathbf{U} \in T.$$

Then $g_1, g_2, \ldots, g_n$ are the component functions of $\mathbf{G} = (g_1, g_2, \ldots, g_n)$. We say that

$$\lim_{\mathbf{U}\to\mathbf{U}_0} \mathbf{G}(\mathbf{U}) = \mathbf{L} = (L_1, L_2, \ldots, L_n)$$

if

$$\lim_{\mathbf{U}\to\mathbf{U}_0} g_i(\mathbf{U}) = L_i, \quad 1 \le i \le n,$$

and that $\mathbf{G}$ is continuous at $\mathbf{U}_0$ if $g_1, g_2, \ldots, g_n$ are each continuous at $\mathbf{U}_0$.

Theorem: For a vector-valued function $\mathbf{G}$,

$$\lim_{\mathbf{U}\to\mathbf{U}_0} \mathbf{G}(\mathbf{U}) = \mathbf{L}$$

if and only if for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|\mathbf{G}(\mathbf{U}) - \mathbf{L}| < \varepsilon \quad \text{whenever} \quad 0 < |\mathbf{U} - \mathbf{U}_0| < \delta \quad \text{and} \quad \mathbf{U} \in D_{\mathbf{G}}.$$

Similarly, $\mathbf{G}$ is continuous at $\mathbf{U}_0$ if and only if for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{U}_0)| < \varepsilon \quad \text{whenever} \quad |\mathbf{U} - \mathbf{U}_0| < \delta \quad \text{and} \quad \mathbf{U} \in D_{\mathbf{G}}.$$

2.14.1 Composite Function

Let f be a real-valued function de ned on a subset of $\mathbb{R}^n$, and let the vector-valued function $\mathbf{G} = (g_1, g_2, \ldots, g_n)$ be de ned on a domain $D_{\mathbf{G}}$ in $\mathbb{R}^m$.
Let the set

$$T = \{\mathbf{U} : \mathbf{U} \in D_{\mathbf{G}} \quad \text{and} \quad \mathbf{G}(\mathbf{U}) \in D_f\},$$

be nonempty.
Composite function: De ne the real-valued composite function

$$h = f \circ \mathbf{G}$$

on T by

$$h(\mathbf{U}) = f(\mathbf{G}(\mathbf{U})), \quad \mathbf{U} \in T.$$

$$T = \{\mathbf{U} : \mathbf{U} \in D_{\mathbf{G}} \quad \text{and} \quad \mathbf{G}(\mathbf{U}) \in D_f\},$$

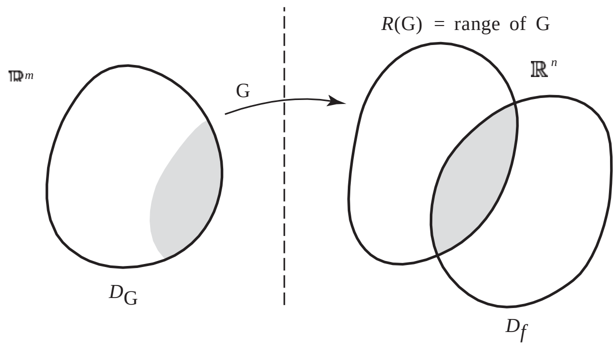

Figure 2.7: Composite of vector valued functions

Theorem: Suppose that $\mathbf{U}_0$ is in T and is a limit point of T, $\mathbf{G}$ is continuous at $\mathbf{U}_0$, and f is continuous at $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$. Then $h = f \circ \mathbf{G}$ is continuous at $\mathbf{U}_0$.

Proof: Suppose that $\varepsilon > 0$. Since f is continuous at $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$, there is an $\varepsilon_1 > 0$ such that

$$|f(\mathbf{X}) - f(\mathbf{G}(\mathbf{U}_0))| < \varepsilon \tag{2.30}$$

if

$$|\mathbf{X} - \mathbf{G}(\mathbf{U}_0)| < \varepsilon_1 \quad \text{and} \quad \mathbf{X} \in D_f. \tag{2.31}$$

Since $\mathbf{G}$ is continuous at $\mathbf{U}_0$, there is a $\delta > 0$ such that

$$|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{U}_0)| < \varepsilon_1 \quad \text{if} \quad |\mathbf{U} - \mathbf{U}_0| < \delta \quad \text{and} \quad \mathbf{U} \in D_{\mathbf{G}}.$$

By taking $\mathbf{X} = \mathbf{G}(\mathbf{U})$ in (2.30) and (2.31), we see that

$$|h(\mathbf{U}) - h(\mathbf{U}_0)| = |f(\mathbf{G}(\mathbf{U}) - f(\mathbf{G}(\mathbf{U}_0))| < \varepsilon$$

if

$$|\mathbf{U} - \mathbf{U}_0| < \delta \quad \text{and} \quad \mathbf{U} \in T.$$

Example: If

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

and

$$g(x, y) = 1 - x^2 - 2y^2,$$

then $D_f = [0, \infty]$, $D_g = \mathrm{R}^2$, and

$$T = \{(x, y) : x^2 + 2y^2 \le 1\}.$$

We have proved that g is continuous on R^2.

We can obtain the same conclusion by observing that the functions $p_1(x, y) = x$ and $p_2(x, y) = y$ are continuous on R^2.

Theorem: Suppose that $\mathbf{U}_0$ is in T and is a limit point of T, $\mathbf{G}$ is continuous at $\mathbf{U}_0$, and f is continuous at $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$.
Then $h = f \circ \mathbf{G}$ is continuous at $\mathbf{U}_0$. Since f is continuous on D_f, the function

$$h(x, y) = f(g(x, y)) = \sqrt{1 - x^2 - 2y^2}$$

is continuous on T.

Example: If

$$g(x, y) = \sqrt{1 - x^2 - 2y^2}$$

and

$$f(s) = \begin{cases} \frac{\sin s}{s}, & s \neq 0, \\ 1, & s = 0, \end{cases}$$

then $D_f = (-\infty, \infty)$ and

$$D_g = T = \{(x, y) : x^2 + 2y^2 \leq 1\}.$$

We have proved that g is continuous on T. Since f is continuous on D_f, the composite function $h = f \circ g$ defined by

$$h(x, y) = \begin{cases} \frac{\sin\sqrt{1-x^2-2y^2}}{\sqrt{1-x^2-2y^2}}, & x^2 + 2y^2 < 1, \\ 1, & x^2 + 2y^2 = 1, \end{cases}$$

is continuous on

$$D_g = T = \{(x, y) : x^2 + 2y^2 \leq 1\}.$$

2.15 Bounded Functions

The definitions of bounded above, bounded below, and bounded on a set S are the same for functions of n variables as for functions of one variable, as are the definitions of supremum and infimum of a function on a set S.

Theorem: If f is continuous on a compact set S in $\mathbb{R}^n$, then f is bounded on S.

Theorem: Let f be continuous on a compact set S in $\mathbb{R}^n$ and

$$\alpha = \inf_{\mathbf{X} \in S} f(\mathbf{X}), \quad \beta = \sup_{\mathbf{X} \in S} f(\mathbf{X}).$$

Then

$$f(\mathbf{X}_1) = \alpha \quad \text{and} \quad f(\mathbf{X}_2) = \beta$$

for some $\mathbf{X}_1$ and $\mathbf{X}_2$ in S.

Proof: See lecture.

Theorem: Let f be continuous on a region S in R^n.

Suppose that $\mathbf{A}$ and $\mathbf{B}$ are in S and

$$f(\mathbf{A}) < u < f(\mathbf{B}).$$

Then $f(\mathbf{C}) = u$ for some $\mathbf{C}$ in S.

Proof: If there is no such $\mathbf{C}$, then $S = R \cup T$, where

$$\begin{aligned} R &= \{\mathbf{X} : \mathbf{X} \in S \text{ and } f(\mathbf{X}) < u\} \\ T &= \{\mathbf{X} : \mathbf{X} \in S \text{ and } f(\mathbf{X}) > u\}. \end{aligned}$$

If $\mathbf{X}_0 \in R$, the continuity of f implies that there is a $\delta > 0$ such that

$$f(\mathbf{X}) < u \quad \text{if} \quad |\mathbf{X} - \mathbf{X}_0| < \delta$$

and $\mathbf{X} \in S$.

This means that $\mathbf{X}_0 \notin \overline{T}$. Therefore, $R \cap \overline{T} = \emptyset$. Similarly, $\overline{R} \cap T = \emptyset$. Therefore, S is disconnected, which contradicts the assumption that S is a region. Hence, we conclude that $f(\mathbf{C}) = u$ for some $\mathbf{C}$ in S.

Theorem: A function f is uniformly continuous on a subset S of its domain in R^n if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$|f(\mathbf{X}) - f(\mathbf{X}')| < \epsilon$$

whenever

$$|\mathbf{X} - \mathbf{X}'| < \delta$$

and $\mathbf{X}, \mathbf{X}' \in S$.

Remark: We emphasize that δ must depend only on ϵ and S, and not on the particular points $\mathbf{X}$ and $\mathbf{X}'$.

Theorem: If f is continuous on a compact set S in R^n, then f is uniformly continuous on S.

Proof: See lecture.

2.16 Directional Derivative

Let $\mathbf{\Phi}$ be a unit vector and $\mathbf{X}$ a point in R^n.

The directional derivative of f at $\mathbf{X}$ in the direction of $\mathbf{\Phi}$ is de ned by

$$\frac{\partial f(\mathbf{X})}{\partial \mathbf{\Phi}} = \lim_{t \to 0} \frac{f(\mathbf{X} + t\mathbf{\Phi}) - f(\mathbf{X})}{t}$$

if the limit exists.

That is, $\partial f(\mathbf{X})/\partial\mathbf{\Phi}$ is the ordinary derivative of the function

$$h(t) = f(\mathbf{X} + t\mathbf{\Phi})$$

at $t = 0$, if $h'(0)$ exists.

Example: Let $\mathbf{\Phi} = (\phi_1, \phi_2, \phi_3)$ and

$$f(x, y, z) = 3xyz + 2x^2 + z^2.$$

Then

$$\begin{aligned} h(t) &= f(x + t\phi_1, y + t\phi_2, z + t\phi_3), \\ &= 3(x + t\phi_1)(y + t\phi_2)(z + t\phi_3) + 2(x + t\phi_1)^2 \\ &\quad +(z + t\phi_3)^2. \end{aligned}$$

$$h(t) = 3(x + t\phi_1)(y + t\phi_2)(z + t\phi_3) + 2(x + t\phi_1)^2 + (z + t\phi_3)^2$$

Then we have

$$\begin{aligned} h'(t) &= 3\phi_1(y + t\phi_2)(z + t\phi_3) + 3\phi_2(x + t\phi_1)(z + t\phi_3) \\ &\quad + 3\phi_3(x + t\phi_1)(y + t\phi_2) + 4\phi_1(x + t\phi_1) \\ &\quad +2\phi_3(z + t\phi_3). \end{aligned}$$

Therefore,

$$\frac{\partial f(\mathbf{X})}{\partial\mathbf{\Phi}} = h'(0) = (3yz + 4x)\phi_1 + 3xz\phi_2 + (3xy + 2z)\phi_3.$$

2.16.1 Partial Derivative

Consider the unit vectors

$$\mathbf{E}_1 = (1, 0, \ldots, 0), \quad \mathbf{E}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{E}_n = (0, \ldots, 0, 1).$$

Since $\mathbf{X}$ and $\mathbf{X} + t\mathbf{E}_i$ di er only in the ith coordinate, $\partial f(\mathbf{X})/\partial\mathbf{E}_i$ is called the partial derivative of f with respect to x_i at $\mathbf{X}$.

It is also denoted by $\partial f(\mathbf{X})/\partial x_i$ or $f_{x_i}(\mathbf{X})$; thus,

$$\frac{\partial f(\mathbf{X})}{\partial x_1} = f_{x1}(\mathbf{X}) = \lim_{t\to 0}\frac{f(x_1 + t, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{t},$$

$$f_{x_i}(\mathbf{X}) = \lim_{t\to 0}\frac{f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{t}$$

if $2 \le i \le n$, and

$$\frac{\partial f(\mathbf{X})}{\partial x_n} = f_{xn}(\mathbf{X}) = \lim_{t\to 0}\frac{f(x_1, \ldots, x_{n-1}, x_n + t) - f(x_1, \ldots, x_{n-1}, x_n)}{t},$$

if the limits exist. If we write $\mathbf{X} = (x, y)$, then we denote the partial derivatives accordingly; thus,

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= f_x(x,y) = \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}\\ \frac{\partial f(x,y)}{\partial y} &= f_y(x,y) = \lim_{h\to 0}\frac{f(x,y+h)-f(x,y)}{h}.\end{aligned}$$

It can be seen from these de nitions that to compute $f_{x_i}(\mathbf{X})$ we simply di erentiate f with respect to x_i according to the rules for ordinary di erentiation, while treating the other variables as constants.

Example: Let

$$f(x,y,z) = 3xyz + 2x^2 + z^2.$$

Taking $\mathbf{\Phi} = \mathbf{E}_1$ (that is, setting $\phi_1 = 1$ and $\phi_2 = \phi_3 = 0$), we nd that

$$\frac{\partial f(\mathbf{X})}{\partial x} = \frac{\partial f(\mathbf{X})}{\partial \mathbf{E}_1} = 3yz + 4x,$$

which is the result obtained by regarding y and z as constants in and taking the ordinary derivative with respect to x. Similarly,

$$\begin{aligned}\frac{\partial f(\mathbf{X})}{\partial y} &= \frac{\partial f(\mathbf{X})}{\partial \mathbf{E}_2} = 3xz\\ \frac{\partial f(\mathbf{X})}{\partial z} &= \frac{\partial f(\mathbf{X})}{\partial \mathbf{E}_3} = 3xy + 2z.\end{aligned}$$

Theorem: If $f_{x_i}(\mathbf{X})$ and $g_{x_i}(\mathbf{X})$ exist, then

$$\begin{aligned}\frac{\partial(f+g)(\mathbf{X})}{\partial x_i} &= f_{x_i}(\mathbf{X}) + g_{x_i}(\mathbf{X}),\\ \frac{\partial(fg)(\mathbf{X})}{\partial x_i} &= f_{x_i}(\mathbf{X})g(\mathbf{X}) + f(\mathbf{X})g_{x_i}(\mathbf{X}),\end{aligned}$$

and, if $g(\mathbf{X}) \neq 0$,

$$\frac{\partial(f/g)(\mathbf{X})}{\partial x_i} = \frac{g(\mathbf{X})f_{x_i}(\mathbf{X}) - f(\mathbf{X})g_{x_i}(\mathbf{X})}{[g(\mathbf{X})]^2}.$$

If $f_{x_i}(\mathbf{X})$ exists at every point of a set D, then it de nes a function f_{x_i} on D.

If this function has a partial derivative with respect to x_j on a subset of D, we denote the partial derivative by

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i} = f_{x_i x_j}.$$

Similarly,

$$\frac{\partial}{\partial x_k}\left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right) = \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i} = f_{x_i x_j x_k}.$$

The function obtained by differentiating f successively with respect to $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ is denoted by

$$\frac{\partial^r f}{\partial x_{i_r} \partial x_{i_{r-1}} \cdots \partial x_{i_1}} = f_{x_{i_1} \cdots x_{i_{r-1}} x_{i_r}};$$

it is an rth-order partial derivative of f. The function

$$f(x, y) = 3x^2y^3 + xy$$

has partial derivatives everywhere. Its first-order partial derivatives are

$$f_x(x, y) = 6xy^3 + y, \quad f_y(x, y) = 9x^2y^2 + x.$$

Its second-order partial derivatives are

$$\begin{aligned} f_{xx}(x, y) &= 6y^3, & f_{yy}(x, y) &= 18x^2y, \\ f_{xy}(x, y) &= 18xy^2 + 1, & f_{yx}(x, y) &= 18xy^2 + 1. \end{aligned}$$

There are eight third-order partial derivatives. Some examples are

$$f_{xxy}(x, y) = 18y^2, \quad f_{xyx}(x, y) = 18y^2, \quad f_{yxx}(x, y) = 18y^2.$$

Compute $f_{xx}(0, 0)$, $f_{yy}(0, 0)$, $f_{xy}(0, 0)$, and $f_{yx}(0, 0)$ if

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

If $(x, y) \neq (0, 0)$, the ordinary rules for differentiation, applied separately to x and y, yield

$$\begin{aligned} f_x(x, y) &= \frac{(2xy+y^2)\sin(x-y)+(x^2y+xy^2)\cos(x-y)}{x^2+y^2} \\ &\quad - \frac{2x(x^2y+xy^2)\sin(x-y)}{(x^2+y^2)^2}, \quad (x, y) \neq (0, 0), \end{aligned} \tag{2.32}$$

and

$$\begin{aligned} f_y(x, y) &= \frac{(x^2+2xy)\sin(x-y)-(x^2y+xy^2)\cos(x-y)}{x^2+y^2} \\ &\quad - \frac{2y(x^2y+xy^2)\sin(x-y)}{(x^2+y^2)^2}, \quad (x, y) \neq (0, 0). \end{aligned} \tag{2.33}$$

These formulas do not apply if $(x, y) = (0, 0)$, so we find $f_x(0, 0)$ and $f_y(0, 0)$ from their definitions as difference quotients:

$$\begin{aligned} f_x(0, 0) &= \lim_{x\to 0} \frac{f(x, 0) - f(0, 0)}{x} = \lim_{x\to 0} \frac{0 - 0}{x} = 0, \\ f_y(0, 0) &= \lim_{y\to 0} \frac{f(0, y) - f(0, 0)}{y} = \lim_{y\to 0} \frac{0 - 0}{y} = 0. \end{aligned}$$

Setting $y = 0$ in (2.32) and (2.33) yields

$$f_x(x, 0) = 0, \quad f_y(x, 0) = \sin x, \quad x \neq 0,$$

so

$$f_{xx}(0,0) = \lim_{x\to 0} \frac{f_x(x,0) - f_x(0,0)}{x} = \lim_{x\to 0} \frac{0-0}{x} = 0.$$

$$f_{yx}(0,0) = \lim_{x\to 0} \frac{f_y(x,0) - f_y(0,0)}{x} = \lim_{x\to 0} \frac{\sin x - 0}{x} = 1.$$

Setting $x = 0$ in (2.32) and (2.33) yields

$$f_x(0,y) = -\sin y, \quad f_y(0,y) = 0, \quad y \neq 0,$$

so

$$f_{xy}(0,0) = \lim_{y\to 0} \frac{f_x(0,y) - f_x(0,0)}{y} = \lim_{y\to 0} \frac{-\sin y - 0}{y} = -1,$$

$$f_{yy}(0,0) = \lim_{y\to 0} \frac{f_y(0,y) - f_y(0,0)}{y} = \lim_{y\to 0} \frac{0-0}{y} = 0.$$

2.16.2 Equality of Mixed Partial Derivatives

Theorem: Suppose that f, f_x, f_y, and f_{xy} exist on a neighborhood N of (x_0, y_0), and f_{xy} is continuous at (x_0, y_0).
Then $f_{yx}(x_0, y_0)$ exists, and

$$f_{yx}(x_0, y_0) = f_{xy}(x_0, y_0). \tag{2.34}$$

Proof: Suppose that $\varepsilon > 0$. Choose $\delta > 0$ so that the open square

$$S_\delta = \{(x,y) : |x - x_0| < \delta, |y - y_0| < \delta\}$$

is in N.

$$|f_{xy}(\hat{x}, \hat{y}) - f_{xy}(x_0, y_0)| < \varepsilon \quad \text{if} \quad (\hat{x}, \hat{y}) \in S_\delta. \tag{2.35}$$

This is possible because of the continuity of f_{xy} at (x_0, y_0). The function

$$A(h,k) = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0) \tag{2.36}$$

is de ned if $-\delta < h, k < \delta$.
Moreover,

$$A(h,k) = \phi(x_0 + h) - \phi(x_0), \tag{2.37}$$

where

$$\phi(x) = f(x, y_0 + k) - f(x, y_0).$$

Since

$$\phi'(x) = f_x(x, y_0 + k) - f_x(x, y_0), \quad |x - x_0| < \delta,$$

(2.37) and the mean value theorem imply that

$$A(h, k) = [f_x(\hat{x}, y_0 + k) - f_x(\hat{x}, y_0)]\, h. \tag{2.38}$$

where $\hat{x}$ is between x_0 and $x_0 + h$.

The mean value theorem, applied to $f_x(\hat{x}, y)$ (where $\hat{x}$ is regarded as constant), also implies that

$$f_x(\hat{x}, y_0 + k) - f_x(\hat{x}, y_0) = f_{xy}(\hat{x}, \hat{y})k,$$

where $\hat{y}$ is between y_0 and $y_0 + k$.

From this and (2.38),

$$A(h, k) = f_{xy}(\hat{x}, \hat{y})hk.$$

Now (2.35) implies that

$$\left| \frac{A(h, k)}{hk} - f_{xy}(x_0, y_0) \right| = |f_{xy}(\hat{x}, \hat{y}) - f_{xy}(x_0, y_0)| < \varepsilon$$
$$\text{if} \quad 0 < |h|, |k| < \delta.$$

Since (2.36) implies that

$$\begin{aligned} \lim_{k\to 0} \frac{A(h, k)}{hk} &= \lim_{k\to 0} \frac{f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)}{hk} \\ &\quad - \lim_{k\to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{hk} \\ &= \frac{f_y(x_0 + h, y_0) - f_y(x_0, y_0)}{h}. \end{aligned}$$

It follows from (2.39) that

$$\left| \frac{f_y(x_0 + h, y_0) - f_y(x_0, y_0)}{h} - f_{xy}(x_0, y_0) \right| \le \varepsilon \quad \text{if} \quad 0 < |h| < \delta.$$

Taking the limit as $h \to 0$ yields

$$|f_{yx}(x_0, y_0) - f_{xy}(x_0, y_0)| \le \varepsilon.$$

Since ε is an arbitrary positive number, this proves (2.34).

2.16.3 Generalization of Equality of Mixed Partial Derivative

Theorem: Suppose that f and all its partial derivatives of order $\le r$ are continuous on an open subset S of $\mathbb{R}^n$.

Then

$$f_{x_{i_1} x_{i_2},\ldots,x_{i_r}}(\mathbf{X}) = f_{x_{j_1} x_{j_2},\ldots,x_{j_r}}(\mathbf{X}), \qquad \mathbf{X} \in S. \tag{2.39}$$

If each of the variables $x_1, x_2, \ldots, x_n$ appears the same number of times in

$$\{x_{i_1}, x_{i_2}, \ldots, x_{i_r}\} \quad \text{and} \quad \{x_{j_1}, x_{j_2}, \ldots, x_{j_r}\}.$$

If this number is r_k, we denote the common value of the two sides of (2.39) by

$$\frac{\partial^r f(\mathbf{X})}{\partial x_1^{r_1}\partial x_2^{r_2}\cdots\partial x_n^{r_n}}. \tag{2.40}$$

It being understood that

$$0 \le r_k \le r, \quad 1 \le k \le n, \tag{2.41}$$

$$r_1 + r_2 + \cdots + r_n = r, \tag{2.42}$$

and, if $r_k = 0$, we omit the symbol ∂x_k^0 from the denominator of (2.40).

Remark: A function of several variables may have rst-order partial derivatives at a point $\mathbf{X}_0$ but fail to be continuous at $\mathbf{X}_0$.

Example: Consider the function

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \quad (0,0), \\ 0, & (x,y) = (0,0). \end{cases} \tag{2.43}$$

Then

$$\begin{aligned} f_x(0,0) &= \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0 \\ f_y(0,0) &= \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{0-0}{k} = 0, \end{aligned}$$

but f is not continuous at $(0,0)$.

Remark: If di erentiability of a function of several variables is to be a stronger property than continuity, as it is for functions of one variable, the de nition of di erentiability must require more than the existence of rst partial derivatives.

A function f is di erentiable at x_0 if and only if

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-m(x-x_0)}{x-x_0} = 0$$

for some constant m, in which case $m = f'(x_0)$.

2.17 Di erentiability of Functions of Several Variables

A function f is di erentiable at

$$\mathbf{X}_0 = (x_{10}, x_{20}, \ldots, x_{n0})$$

if $\mathbf{X}_0 \in D_f^0$ and there are constants $m_1, m_2, \ldots, m_n$ such that

$$\lim_{\mathbf{X}\to\mathbf{X}_0}\frac{f(\mathbf{X})-f(\mathbf{X}_0)-\sum_{i=1}^n m_i(x_i-x_{i0})}{|\mathbf{X}-\mathbf{X}_0|} = 0. \tag{2.44}$$

Example: Show that the following function $f(x, y) = x^2 + 2xy$, is di erentiable at any point (x_0, y_0).

$$\begin{aligned}
f(x, y) - f(x_0, y_0) &= x^2 + 2xy - x_0^2 - 2x_0y_0 \\
&= x^2 - x_0^2 + 2(xy - x_0y_0) \\
&= (x - x_0)(x + x_0) + 2(xy - x_0y) \\
&\quad +2(x_0y - x_0y_0) \\
&= (x + x_0 + 2y)(x - x_0) + 2x_0(y - y_0) \\
&= 2(x_0 + y_0)(x - x_0) + 2x_0(y - y_0) \\
&\quad + (x - x_0)(x - x_0 + 2y - 2y_0) \\
&= m_1(x - x_0) + m_2(y - y_0) + (x - x_0)(x - x_0 + 2y - 2y_0),
\end{aligned}$$

where

$$m_1 = 2(x_0 + y_0) = f_x(x_0, y_0) \quad \text{and} \quad m_2 = 2x_0 = f_y(x_0, y_0). \tag{2.45}$$

Therefore,

$$\begin{aligned}
&\frac{|f(x, y) - f(x_0, y_0) - m_1(x - x_0) - m_2(y - y_0)|}{|\mathbf{X} - \mathbf{X}_0|} \\
&= \frac{|x - x_0|\,|(x - x_0) + 2(y - y_0)|}{|\mathbf{X} - \mathbf{X}_0|} \\
&\le \sqrt{5}|\mathbf{X} - \mathbf{X}_0|,
\end{aligned}$$

by Schwarz's inequality. This implies that

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \frac{f(x, y) - f(x_0, y_0) - m_1(x - x_0) - m_2(y - y_0)}{|\mathbf{X} - \mathbf{X}_0|} = 0,$$

so f is di erentiable at (x_0, y_0).

Theorem: If f is di erentiable at $\mathbf{X}_0 = (x_{10}, x_{20}, \ldots, x_{n0})$, then $f_{x_1}(\mathbf{X}_0)$, $f_{x_2}(\mathbf{X}_0)$, $\ldots$, $f_{x_n}(\mathbf{X}_0)$ exist and the constants $m_1, m_2, \ldots, m_n$ in

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \frac{f(\mathbf{X}) - f(\mathbf{X}_0) - \sum_{i=1}^{n} m_i(x_i - x_{i0})}{|\mathbf{X} - \mathbf{X}_0|} = 0,$$

are given by

$$m_i = f_{x_i}(\mathbf{X}_0), \quad 1 \le i \le n; \tag{2.46}$$

that is,

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \frac{f(\mathbf{X}) - f(\mathbf{X}_0) - \sum_{i=1}^{n} f_{x_i}(\mathbf{X}_0)(x_i - x_{i0})}{|\mathbf{X} - \mathbf{X}_0|} = 0.$$

Proof: Let i be a given integer in $\{1, 2, \ldots, n\}$. Let $\mathbf{X} = \mathbf{X}_0 + t\mathbf{E}_i$, so that $x_i = x_{i0} + t$, $x_j = x_{j0}$ if $j \neq i$, and $|\mathbf{X} - \mathbf{X}_0| = |t|$. Then

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \frac{f(\mathbf{X}) - f(\mathbf{X}_0) - \sum_{i=1}^{n} m_i(x_i - x_{i0})}{|\mathbf{X} - \mathbf{X}_0|} = 0.$$

and the di erentiability of f at $\mathbf{X}_0$ imply that

$$\lim_{t\to 0} \frac{f(\mathbf{X}_0 + t\mathbf{E}_i) - f(\mathbf{X}_0) - m_i t}{t} = 0.$$

Hence,

$$\lim_{t\to 0} \frac{f(\mathbf{X}_0 + t\mathbf{E}_i) - f(\mathbf{X}_0)}{t} = m_i.$$

This proves (2.46), since the limit on the left is $f_{x_i}(\mathbf{X}_0)$, by de nition.

2.17.1 Linear Function

A linear function is a function of the form

$$L(\mathbf{X}) = m_1x_1 + m_2x_2 + \cdots + m_nx_n, \tag{2.47}$$

where $m_1, m_2, \ldots, m_n$ are constants. From de nition of di erentiability, f is di erentiable at $\mathbf{X}_0$ if and only if there is a linear function L such that $f(\mathbf{X}) - f(\mathbf{X}_0)$ can be approximated so well near $\mathbf{X}_0$ by

$$L(\mathbf{X}) - L(\mathbf{X}_0) = L(\mathbf{X} - \mathbf{X}_0)$$

that

$$f(\mathbf{X}) - f(\mathbf{X}_0) = L(\mathbf{X} - \mathbf{X}_0) + E(\mathbf{X})(|\mathbf{X} - \mathbf{X}_0|), \tag{2.48}$$

where

$$\lim_{\mathbf{X}\to\mathbf{X}_0} E(\mathbf{X}) = 0. \tag{2.49}$$

Theorem: If f is di erentiable at $\mathbf{X}_0$, then f is continuous at $\mathbf{X}_0$.

Proof: From $L(\mathbf{X}) = m_1x_1 + m_2x_2 + \cdots + m_nx_n$, and Schwarz's inequality,

$$|L(\mathbf{X} - \mathbf{X}_0)| \le M|\mathbf{X} - \mathbf{X}_0|,$$

where

$$M = (m_1^2 + m_2^2 + \cdots + m_n^2)^{1/2}.$$

This and $f(\mathbf{X}) - f(\mathbf{X}_0) = L(\mathbf{X} - \mathbf{X}_0) + E(\mathbf{X})(|\mathbf{X} - \mathbf{X}_0|)$, imply that

$$|f(\mathbf{X}) - f(\mathbf{X}_0)| \le (M + |E(\mathbf{X})|)|\mathbf{X} - \mathbf{X}_0|.$$

which, with (2.49), implies that f is continuous at $\mathbf{X}_0$.

2.17.2 Di erential

The linear function

$$L(\mathbf{X}) = f_{x_1}(\mathbf{X}_0)x_1 + f_{x_2}(\mathbf{X}_0)x_2 + \cdots + f_{x_n}(\mathbf{X}_0)x_n.$$

This function is called the di erential of f at $\mathbf{X}_0$. We will denote it by $d_{\mathbf{X}_0}f$ and its value by $(d_{\mathbf{X}_0}f)(\mathbf{X})$.

Thus,

$$(d_{\mathbf{X}_0}f)(\mathbf{X}) = f_{x_1}(\mathbf{X}_0)x_1 + f_{x_2}(\mathbf{X}_0)x_2 + \cdots + f_{x_n}(\mathbf{X}_0)x_n. \tag{2.50}$$

In terms of the di erential, di erentiability can be rewritten as

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \frac{f(\mathbf{X}) - f(\mathbf{X}_0) - (d_{\mathbf{X}_0}f)(\mathbf{X} - \mathbf{X}_0)}{|\mathbf{X} - \mathbf{X}_0|} = 0.$$

For convenience in writing $d_{\mathbf{X}_0}f$, and to conform with standard notation, we introduce the function dx_i, de ned by

$$dx_i(\mathbf{X}) = x_i;$$

that is, dx_i is the function whose value at a point in $\mathbb{R}^n$ is the ith coordinate of the point.

It is the di erential of the function $g_i(\mathbf{X}) = x_i$. From (2.50),

$$d_{\mathbf{X}_0}f = f_{x_1}(\mathbf{X}_0)\,dx_1 + f_{x_2}(\mathbf{X}_0\,dx_2 + \cdots + f_{x_n}(\mathbf{X}_0)\,dx_n. \tag{2.51}$$

If we write $\mathbf{X} = (x, y, \ldots,)$, then we write

$$d_{\mathbf{X}_0}f = f_x(\mathbf{X}_0)\,dx + f_y(\mathbf{X}_0)\,dy + \cdots,$$

where dx, dy, ... are the functions de ned by

$$dx(\mathbf{X}) = x, \quad dy(\mathbf{X}) = y, \ldots$$

When it is not necessary to emphasize the speci c point $\mathbf{X}_0$, (2.51) can be written more simply as

$$df = f_{x_1}\,dx_1 + f_{x_2}\,dx_2 + \cdots + f_{x_n}\,dx_n.$$

When dealing with a speci c function at an arbitrary point of its domain, we may use the hybrid notation

$$df = f_{x_1}(\mathbf{X})\,dx_1 + f_{x_2}(\mathbf{X})\,dx_2 + \cdots + f_{x_n}(\mathbf{X})\,dx_n.$$

Example: The function

$$f(x, y) = x^2 + 2xy$$

is di erentiable at every $\mathbf{X}$ in $\mathbb{R}^n$.

The di erential of the functions is

$$df = (2x + 2y)\,dx + 2xdy.$$

To nd $d_{\mathbf{X}_0}f$ with $\mathbf{X}_0 = (1, 2)$, we set $x_0 = 1$ and $y_0 = 2$; thus,

$$\begin{aligned} d_{\mathbf{X}_0}f &= 6\,dx + 2\,dy \\ (d_{\mathbf{X}_0}f)(\mathbf{X} - \mathbf{X}_0) &= 6(x - 1) + 2(y - 2). \end{aligned}$$

Since $f(1, 2) = 5$, the di erentiability of f at $(1, 2)$ implies that

$$\lim_{(x,y)\to(1,2)} \frac{f(x, y) - 5 - 6(x - 1) - 2(y - 2)}{\sqrt{(x - 1)^2 + (y - 2)^2}} = 0.$$

Example: The di erential of a function $f = f(x)$ of one variable is given by

$$d_{x_0}f = f'(x_0)\,dx,$$

where dx is the identity function; that is,

$$dx(t) = t.$$

For example, if

$$f(x) = 3x^2 + 5x^3,$$

then

$$df = (6x + 15x^2)\,dx.$$

If $x_0 = -1$, then

$$d_{x_0}f = 9\,dx, \quad (d_{x_0}f)(x - x_0) = 9(x + 1),$$

and, since $f(-1) = -2$,

$$\lim_{x\to-1} \frac{f(x) + 2 - 9(x + 1)}{x + 1} = 0.$$

Remark: Unfortunately, the notation for the di erential is so complicated that it obscures the simplicity of the concept. The peculiar symbols df, dx, dy, etc., were introduced in the early stages of the development of calculus to represent very small (in nitesimal) increments in the variables. However, in modern usage they are not quantities at all, but linear functions. This meaning of the symbol dx di ers from its meaning in $\int_a^b f(x)\,dx$, where it serves merely to identify the variable of integration; indeed, some authors omit it in the latter context and write simply $\int_a^b f$.

Lemma: If f is di erentiable at $\mathbf{X}_0$, then

$$f(\mathbf{X}) - f(\mathbf{X}_0) = (d_{\mathbf{X}_0}f)(\mathbf{X} - \mathbf{X}_0) + E(\mathbf{X})|\mathbf{X} - \mathbf{X}_0|,$$

where E is de ned in a neighborhood of $\mathbf{X}_0$ and

$$\lim_{\mathbf{X}\to\mathbf{X}_0} E(\mathbf{X}) = E(\mathbf{X}_0) = 0.$$

Theorem: If f and g are di erentiable at $\mathbf{X}_0$, then so are $f+g$ and fg. The same is true of f/g if $g(\mathbf{X}_0)\neq 0$. The di erentials are given by

$$\begin{aligned} d_{\mathbf{X}_0}(f+g) &= d_{\mathbf{X}_0}f + d_{\mathbf{X}_0}g, \\ d_{\mathbf{X}_0}(fg) &= f(\mathbf{X}_0)d_{\mathbf{X}_0}g + g(\mathbf{X}_0)d_{\mathbf{X}_0}f, \end{aligned}$$

and

$$d_{\mathbf{X}_0}\left(\frac{f}{g}\right) = \frac{g(\mathbf{X}_0)d_{\mathbf{X}_0}f - f(\mathbf{X}_0)d_{\mathbf{X}_0}g}{[g(\mathbf{X}_0)]^2}.$$

2.17.3 A su cient Condition for Di erentiability

Theorem: If $f_{x_1}, f_{x_2}, \dots, f_{x_n}$ exist on a neighborhood of $\mathbf{X}_0$ and are continuous at $\mathbf{X}_0$, then f is di erentiable at $\mathbf{X}_0$.

If f is dierentiable at X0, then f is continuous at X0.

Proof: Let $\mathbf{X}_0 = (x_{10}, x_{20}, \dots, x_{n0})$ and suppose that $\varepsilon > 0$. Our assumptions imply that there is a $\delta > 0$ such that $f_{x_1}, f_{x_2}, \dots, f_{x_n}$ are de ned in the n-ball

$$S_\delta(\mathbf{X}_0) = \{\mathbf{X} : |\mathbf{X} - \mathbf{X}_0| < \delta\}$$

and

$$|f_{x_j}(\mathbf{X}) - f_{x_j}(\mathbf{X}_0)| < \varepsilon \quad \text{if} \quad |\mathbf{X} - \mathbf{X}_0| < \delta, \quad 1 \le j \le n. \tag{2.52}$$

Let $\mathbf{X} = (x_1, x_, \dots, x_n)$ be in $S_\delta(\mathbf{X}_0)$. De ne

$$\mathbf{X}_j = (x_1, \dots, x_j, x_{j+1,0}, \dots, x_{n0}), \quad 1 \le j \le n-1,$$

and $\mathbf{X}_n = \mathbf{X}$. Thus, for $1 \le j \le n$, $\mathbf{X}_j$ di ers from $\mathbf{X}_{j-1}$ in the jth component only, and the line segment from $\mathbf{X}_{j-1}$ to $\mathbf{X}_j$ is in $S_\delta(\mathbf{X}_0)$. Now write

$$f(\mathbf{X}) - f(\mathbf{X}_0) = f(\mathbf{X}_n) - f(\mathbf{X}_0) = \sum_{j=1} [f(\mathbf{X}_j) - f(\mathbf{X}_{j-1})], \tag{2.53}$$

and consider the auxiliary functions

$$\begin{aligned} g_1(t) &= f(t, x_{20}, \dots, x_{n0}), \\ g_j(t) &= f(x_1, \dots, x_{j-1}, t, x_{j+1,0}, \dots, x_{n0}), \quad 2 \le j \le n-1, \\ g_n(t) &= f(x_1, \dots, x_{n-1}, t), \end{aligned} \tag{2.54}$$

where, in each case, all variables except t are temporarily regarded as constants. Since

$$f(\mathbf{X}_j) - f(\mathbf{X}_{j-1}) = g_j(x_j) - g_j(x_{j0}),$$

the mean value theorem implies that

$$f(\mathbf{X}_j) - f(\mathbf{X}_{j-1}) = g_j'(\tau_j)(x_j - x_{j0}),$$

where τ_j is between x_j and x_{j0}. From (2.54),

$$g_j'(\tau_j) = f_{x_j}(\hat{\mathbf{X}}_j),$$

where $\hat{\mathbf{X}}_j$ is on the line segment from $\mathbf{X}_{j-1}$ to $\mathbf{X}_j$. Therefore,

$$f(\mathbf{X}_j) - f(\mathbf{X}_{j-1}) = f_{x_j}(\hat{\mathbf{X}}_j)(x_j - x_{j0}),$$

and (2.53) implies that

$$\begin{aligned} f(\mathbf{X}) - f(\mathbf{X}_0) &= \sum_{j=1}^{n} f_{x_j}(\hat{\mathbf{X}}_j)(x_j - x_{j0}) \\ &= \sum_{j=1}^{n} f_{x_j}(\mathbf{X}_0)(x_j - x_{j0}) + \sum_{j=1}^{n} [f_{x_j}(\hat{\mathbf{X}}_j) - f_{x_j}(\mathbf{X}_0)](x_j - x_{j0}). \end{aligned}$$

From this and (2.52),

$$\left| f(\mathbf{X}) - f(\mathbf{X}_0) - \sum_{j=1}^{n} f_{x_j}(\mathbf{X}_0)(x_j - x_{j0}) \right| \le \varepsilon \sum_{j=1}^{n} |x_j - x_{j0}| \le n\varepsilon|\mathbf{X} - \mathbf{X}_0|,$$

which implies that f is di erentiable at $\mathbf{X}_0$.

2.17.4 Continuously Di erentiable Function

We say that f is continuously di erentiable on a subset S of $\mathbb{R}^n$ if S is contained in an open set on which $f_{x_1}, f_{x_2}, \ldots, f_{x_n}$ are continuous.

The above theorem implies that such a function is di erentiable at each $\mathbf{X}_0$ in S.

Example: If

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

then

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

$$f_y(x, y) = \frac{2y}{x - y} + \frac{x^2 + y^2}{(x - y)^2}.$$

Since f_x and f_y are continuous on

$$S = \{(x, y) : x \neq y\},$$

f is continuously di erentiable on S.

Remark: If $f_{x_1}, f_{x_2}, \ldots, f_{x_n}$ exist on a neighborhood of $\mathbf{X}_0$ and are continuous at $\mathbf{X}_0$, then f is di erentiable at $\mathbf{X}_0$. These conditions are not necessary for di erentiability; that is, a function may be di erentiable at a point $\mathbf{X}_0$ even if its rst partial derivatives are not continuous at $\mathbf{X}_0$.

Example: let

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

Then

$$f_x(x,y) = 2(x-y)\sin\frac{1}{x-y} - \cos\frac{1}{x-y}, \quad x \neq y,$$

and

$$f_x(x,x) = \lim_{h\to 0} \frac{f(x+h,x) - f(x,x)}{h} = \lim_{h\to 0} \frac{h^2 \sin(1/h) - 0}{h} = 0,$$

so f_x exists for all (x, y), but is not continuous on the line $y = x$.

Example: Let

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

The same is true of f_y, since

$$f_y(x,y) = -2(x-y)\sin\frac{1}{x-y} + \cos\frac{1}{x-y}, \quad x \neq y,$$

and

$$f_y(x,x) = \lim_{k\to 0} \frac{f(x,x+k) - f(x,x)}{k} = \lim_{k\to 0} \frac{k^2 \sin(-1/k) - 0}{k} = 0.$$

Now,

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

and Schwarz's inequality implies that

$$\left| \frac{(x-y)^2}{\sqrt{x^2+y^2}} \sin \frac{1}{x-y} \right| \le \frac{2(x^2+y^2)}{\sqrt{x^2+y^2}} = 2\sqrt{x^2+y^2}, \quad x \neq y.$$

Therefore,

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - f_x(0,0)x - f_y(0,0)y}{\sqrt{x^2+y^2}} = 0,$$

so f is di erentiable at $(0,0)$, but f_x and f_y are not continuous at $(0,0)$.

2.17.5 Geometric Interpretation of Di erentiability

If a function f of one variable is di erentiable at x_0, then the curve $y = f(x)$ has a tangent line

$$y = T(x) = f(x_0) + f'(x_0)(x - x_0).$$

The tangent line approximates it so well near x_0 that

$$\lim_{x \to x_0} \frac{f(x) - T(x)}{x - x_0} = 0.$$

Moreover, the tangent line is the limit of the secant line through the points $(x_1, f(x_0))$ and $(x_0, f(x_0))$ as x_1 approaches x_0. Di erentiability of a function of n variables has an analogous geometric interpretation. We will illustrate it for $n = 2$. If f is de ned in a region D in $\mathbb{R}^2$, then the set of points (x, y, z) such that

$$z = f(x, y), \qquad (x, y) \in D, \tag{2.55}$$

is a surface in $\mathbb{R}^3$ Geometric interpretation of di erentiability:

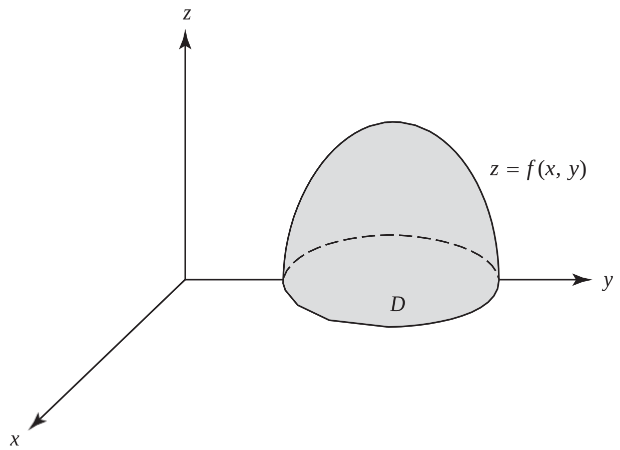

Figure 2.8: Domain of the function

If f is di erentiable at $\mathbf{X}_0 = (x_0, y_0)$, then the plane

$$z = T(x, y) = f(\mathbf{X}_0) + f_x(\mathbf{X}_0)(x - x_0) + f_y(\mathbf{X}_0)(y - y_0) \tag{2.56}$$

intersects the surface $z = f(x, y)$ at $(x_0, y_0, f(x_0, y_0))$ and approximates the surface so well near (x_0, y_0) that

$$\lim_{(x,y) \to (x_0,y_0)} \frac{f(x, y) - T(x, y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0.$$

Moreover, (2.56) is the only plane in $\mathbb{R}^3$ with these properties.

We say that this plane is tangent to the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$.

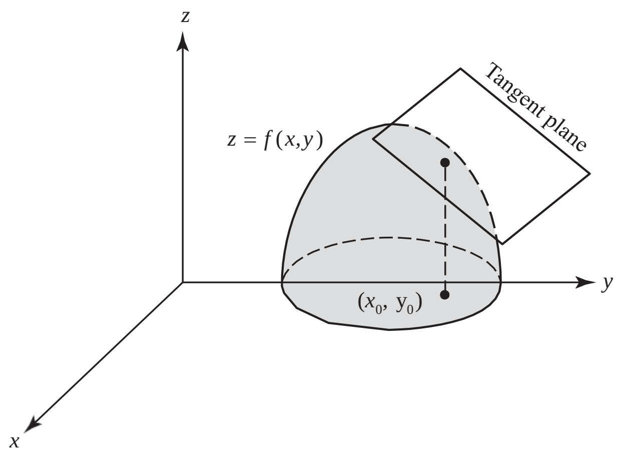

Figure 2.9: Geometric interpretation of di erentiability

Show that the tangent plane to the surface $z = f(x, y)$ is the limit of the secant planes.

Let $\mathbf{X}_i = (x_i, y_i)$ $(i = 1, 2, 3)$. The equation of the secant plane through the points $(x_i, y_i, f(x_i, y_i))$ $(i = 1, 2, 3)$ on the surface $z = f(x, y)$ is of the form

$$z = f(\mathbf{X}_0) + A(x - x_0) + B(y - y_0), \tag{2.57}$$

where A and B satisfy the system

$$\begin{aligned} f(\mathbf{X}_1) &= f(\mathbf{X}_0) + A(x_1 - x_0) + B(y_1 - y_0), \\ f(\mathbf{X}_2) &= f(\mathbf{X}_0) + A(x_2 - x_0) + B(y_2 - y_0). \end{aligned}$$

Solving for A and B yields

$$A = \frac{(f(\mathbf{X}_1) - f(\mathbf{X}_0))(y_2 - y_0) - (f(\mathbf{X}_2) - f(\mathbf{X}_0))(y_1 - y_0)}{(x_1 - x_0)(y_2 - y_0) - (x_2 - x_0)(y_1 - y_0)} \tag{2.58}$$

$$B = \frac{(f(\mathbf{X}_2) - f(\mathbf{X}_0))(x_1 - x_0) - (f(\mathbf{X}_1) - f(\mathbf{X}_0))(x_2 - x_0)}{(x_1 - x_0)(y_2 - y_0) - (x_2 - x_0)(y_1 - y_0)} \tag{2.59}$$

if

$$(x_1 - x_0)(y_2 - y_0) - (x_2 - x_0)(y_1 - y_0) \neq 0, \tag{2.60}$$

which is equivalent to the requirement that $\mathbf{X}_0$, $\mathbf{X}_1$, and $\mathbf{X}_2$ do not lie on a line. If we write

$$\mathbf{X}_1 = \mathbf{X}_0 + t\mathbf{U} \quad \text{and} \quad \mathbf{X}_2 = \mathbf{X}_0 + t\mathbf{V},$$

where $\mathbf{U} = (u_1, u_2)$ and $\mathbf{V} = (v_1, v_2)$ are xed nonzero vectors, then (2.58), (2.59), and (2.60) take the more convenient forms

$$A = \frac{\frac{f(\mathbf{X}_0+t\mathbf{U})-f(\mathbf{X}_0)}{t} v_2 - \frac{f(\mathbf{X}_0+t\mathbf{V})-f(\mathbf{X}_0)}{t} u_2}{u_1 v_2 - u_2 v_1}, \tag{2.61}$$

$$B = \frac{\frac{f(\mathbf{X}_0+t\mathbf{V})-f(\mathbf{X}_0)}{t} u_1 - \frac{f(\mathbf{X}_0+t\mathbf{U})-f(\mathbf{X}_0)}{t} v_1}{u_1 v_2 - u_2 v_1}, \tag{2.62}$$

and

$$u_1v_2 - u_2v_1 \neq 0.$$

If f is differentiable at $\mathbf{X}_0$, then

$$f(\mathbf{X}) - f(\mathbf{X}_0) = f_x(\mathbf{X}_0)(x - x_0) + f_y(\mathbf{X}_0)(y - y_0) + \epsilon(\mathbf{X})|\mathbf{X} - \mathbf{X}_0|, \tag{2.63}$$

where

$$\lim_{\mathbf{X}\to\mathbf{X}_0} \epsilon(\mathbf{X}) = 0. \tag{2.64}$$

Substituting first $\mathbf{X} = \mathbf{X}_0 + t\mathbf{U}$ and then $\mathbf{X} = \mathbf{X}_0 + t\mathbf{V}$ in (2.63) and dividing by t yields

$$\frac{f(\mathbf{X}_0 + t\mathbf{U}) - f(\mathbf{X}_0)}{t} = f_x(\mathbf{X}_0)u_1 + f_y(\mathbf{X}_0)u_2 + E_1(t)|\mathbf{U}| \tag{2.65}$$

and

$$\frac{f(\mathbf{X}_0 + t\mathbf{V}) - f(\mathbf{X}_0)}{t} = f_x(\mathbf{X}_0)v_1 + f_y(\mathbf{X}_0)v_2 + E_2(t)|\mathbf{V}|, \tag{2.66}$$

where

$$E_1(t) = \epsilon(\mathbf{X}_0 + t\mathbf{U})|t|/t \quad \text{and} \quad E_2(t) = \epsilon(\mathbf{X}_0 + t\mathbf{V})|t|/t,$$

so

$$\lim_{t\to 0} E_i(t) = 0, \quad i = 1, 2, \tag{2.67}$$

because of (2.64). Substituting (2.65) and (2.66) into (2.61) and (2.62) yields

$$A = f_x(\mathbf{X}_0) + \Delta_1(t), \quad B = f_y(\mathbf{X}_0) + \Delta_2(t), \tag{2.68}$$

where

$$\Delta_1(t) = \frac{v_2|\mathbf{U}|E_1(t) - u_2|\mathbf{V}|E_2(t)}{u_1v_2 - u_2v_1}$$

and

$$\Delta_2(t) = \frac{u_1|\mathbf{V}|E_2(t) - v_1|\mathbf{U}|E_1(t)}{u_1v_2 - u_2v_1},$$

so

$$\lim_{t\to 0} \Delta_i(t) = 0, \quad i = 1, 2, \tag{2.69}$$

because of (2.67).

From (2.57) and (2.68), the equation of the secant plane is

$$z = f(\mathbf{X}_0) + [f_x(\mathbf{X}_0) + \Delta_1(t)](x - x_0) + [f_y(\mathbf{X}_0) + \Delta_2(t)](y - y_0).$$

Therefore, because of (2.69), the secant plane approaches the tangent plane (2.56) as t approaches zero.

2.18 Maxima and Minima

We say that $\mathbf{X}_0$ is a local extreme point of f if there is a $\delta > 0$ such that

$$f(\mathbf{X}) - f(\mathbf{X}_0)$$

does not change sign in $S_\delta(\mathbf{X}_0) \cap D_f$.

More speci cally, $\mathbf{X}_0$ is a local maximum point if

$$f(\mathbf{X}) \leq f(\mathbf{X}_0)$$

or a local minimum point if

$$f(\mathbf{X}) \geq f(\mathbf{X}_0)$$

for all $\mathbf{X}$ in $S_\delta(\mathbf{X}_0) \cap D_f$.

Theorem: Suppose that f is de ned in a neighborhood of $\mathbf{X}_0$ in $\mathbb{R}^n$ and $f_{x_1}(\mathbf{X}_0)$, $f_{x_2}(\mathbf{X}_0), \dots, f_{x_n}(\mathbf{X}_0)$ exist.

Let $\mathbf{X}_0$ be a local extreme point of f. Then

$$f_{x_i}(\mathbf{X}_0) = 0, \quad 1 \leq i \leq n. \tag{2.70}$$

Proof: Let $\mathbf{E}_1 = (1, 0, \dots, 0)$, $\mathbf{E}_2 = (0, 1, 0, \dots, 0), \dots,$ $\mathbf{E}_n = (0, 0, \dots, 1)$, and

$$g_i(t) = f(\mathbf{X}_0 + t\mathbf{E}_i), \quad 1 \leq i \leq n.$$

Then g_i is di erentiable at $t = 0$, with

$$g_i'(0) = f_{x_i}(\mathbf{X}_0).$$

Since $\mathbf{X}_0$ is a local extreme point of f, $t_0 = 0$ is a local extreme point of g_i.

Remark: The converse of theorem is false, since (2.70) $f_{x_i}(\mathbf{X}_0) = 0, \quad 1 \leq i \leq n.$ may hold at a point $\mathbf{X}_0$ that is not a local extreme point of f.

For example, let $\mathbf{X}_0 = (0, 0)$ and

$$f(x, y) = x^3 + y^3.$$

We say that a point $\mathbf{X}_0$ where (2.70) holds is a critical point of f. Thus, if f is de ned in a neighborhood of a local extreme point $\mathbf{X}_0$, then $\mathbf{X}_0$ is a critical point of f; however, a critical point need not be a local extreme point of f.

2.19 Di erentiable Vector Valued Function

A vector-valued function $\mathbf{G} = (g_1, g_2, \ldots, g_n)$ is di erentiable at

$$\mathbf{U}_0 = (u_{10}, u_{20}, \ldots, u_{m0})$$

if its component functions $g_1, g_2, \ldots, g_n$ are di erentiable at $\mathbf{U}_0$.

Lemma: Suppose that $\mathbf{G} = (g_1, g_2, \ldots, g_n)$ is di erentiable at

$$\mathbf{U}_0 = (u_{10}, u_{20}, \ldots, u_{m0}),$$

and de ne

$$M = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{\partial g_i(\mathbf{U}_0)}{\partial u_j} \right)^2 \right)^{1/2}.$$

Then, if $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\frac{|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{U}_0)|}{|\mathbf{U} - \mathbf{U}_0|} < M + \varepsilon \quad \text{if} \quad 0 < |\mathbf{U} - \mathbf{U}_0| < \delta.$$

Proof: Since $g_1, g_2, \ldots, g_n$ are di erentiable at $\mathbf{U}_0$ to g_i shows that

$$\begin{aligned} g_i(\mathbf{U}) - g_i(\mathbf{U}_0) &= (d_{\mathbf{U}_0} g_i)(\mathbf{U} - \mathbf{U}_0) + E_i(\mathbf{U})|(\mathbf{U} - \mathbf{U}_0| \\ &= \sum_{j=1}^{m} \frac{\partial g_i(\mathbf{U}_0)}{\partial u_j}(u_j - u_{j0}) + E_i(\mathbf{U})|(\mathbf{U} - \mathbf{U}_0|, \end{aligned} \tag{2.71}$$

where

$$\lim_{\mathbf{U} \to \mathbf{U}_0} E_i(\mathbf{U}) = 0, \quad 1 \le i \le n. \tag{2.72}$$

From Schwarz's inequality,

$$|g_i(\mathbf{U}) - g_i(\mathbf{U}_0)| \le (M_i + |E_i(\mathbf{U})|)|\mathbf{U} - \mathbf{U}_0|,$$

where

$$M_i = \left(\sum_{j=1}^{m} \left(\frac{\partial g_i(\mathbf{U}_0)}{\partial u_j} \right)^2 \right)^{1/2}.$$

Therefore,

$$\frac{|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{U}_0)|}{|\mathbf{U} - \mathbf{U}_0|} \le \left(\sum_{i=1}^{n} (M_i + |E_i(\mathbf{U})|)^2 \right)^{1/2}.$$

From (2.72),

$$\lim_{\mathbf{U} \to \mathbf{U}_0} \left(\sum_{i=1}^{n} (M_i + |E_i(\mathbf{U})|)^2 \right)^{1/2} = \left(\sum_{i=1}^{n} M_i^2 \right)^{1/2} = M,$$

which implies the conclusion.

2.20 The Chain Rule

Theorem: Suppose that the real-valued function f is differentiable at $\mathbf{X}_0$ in $\mathbb{R}^n$. The vector-valued function $\mathbf{G} = (g_1, g_2, \ldots, g_n)$ is differentiable at $\mathbf{U}_0$ in $\mathbb{R}^m$, and $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$.

Then the real-valued composite function $h = f \circ \mathbf{G}$ defined by

$$h(\mathbf{U}) = f(\mathbf{G}(\mathbf{U})) \tag{2.73}$$

is differentiable at $\mathbf{U}_0$, and

$$d_{\mathbf{U}_0}h = f_{x_1}(\mathbf{X}_0)d_{\mathbf{U}_0}g_1 + f_{x_2}(\mathbf{X}_0)d_{\mathbf{U}_0}g_2 + \cdots + f_{x_n}(\mathbf{X}_0)d_{\mathbf{U}_0}g_n. \tag{2.74}$$

Proof: First we will show that $\mathbf{U}_0$ is an interior point of the domain of h. It is legitimate to ask if h is differentiable at $\mathbf{U}_0$. Let $\mathbf{X}_0 = (x_{10}, x_{20}, \ldots, x_{n0})$. Note that

$$x_{i0} = g_i(\mathbf{U}_0), \quad 1 \le i \le n,$$

by assumption.

Since f is differentiable at $\mathbf{X}_0$, which implies that

$$f(\mathbf{X}) - f(\mathbf{X}_0) = \sum_{i=1} f_{x_i}(\mathbf{X}_0)(x_i - x_{i0}) + E(\mathbf{X})|\mathbf{X} - \mathbf{X}_0|, \tag{2.75}$$

where

$$\lim_{\mathbf{X}\to\mathbf{X}_0} E(\mathbf{X}) = 0.$$

Substituting $\mathbf{X} = \mathbf{G}(\mathbf{U})$ and $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$ in (2.75) and recalling (2.73) yields

$$\begin{aligned} h(\mathbf{U}) - h(\mathbf{U}_0) &= \sum_{i=1} f_{x_i}(\mathbf{X}_0)(g_i(\mathbf{U}) - g_i(\mathbf{U}_0)) \\ &\quad + E(\mathbf{G}(\mathbf{U}))|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{U}_0)|. \end{aligned} \tag{2.76}$$

Substituting $g_i(\mathbf{U}) - g_i(\mathbf{U}_0) = d_{\mathbf{U}_0}g_i)(\mathbf{U} - \mathbf{U}_0) + E_i(\mathbf{U})|\mathbf{U} - \mathbf{U}_0|$ into (2.76) yields

$$\begin{aligned} h(\mathbf{U}) - h(\mathbf{U}_0) &= \sum_{i=1}^{n} f_{x_i}(\mathbf{X}_0)(d_{\mathbf{U}_0} g_i)(\mathbf{U} - \mathbf{U}_0) \\ &\quad + \left(\sum_{i=1}^{n} f_{x_i}(\mathbf{X}_0)E_i(\mathbf{U})\right)|\mathbf{U} - \mathbf{U}_0| \\ &\quad + E(\mathbf{G}(\mathbf{U}))|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{U}_0|. \end{aligned}$$

Since

$$\lim_{\mathbf{U}\to\mathbf{U}_0} E(\mathbf{G}(\mathbf{U})) = \lim_{\mathbf{X}\to\mathbf{X}_0} E(\mathbf{X}) = 0.$$

Due to Lemma we proved in previous module, imply that

$$\frac{h(\mathbf{U}) - h(\mathbf{U}_0) - \sum_{i=1}^{n} f_{x_i}(\mathbf{X}_0 d_{\mathbf{U}_0} g_i(\mathbf{U} - \mathbf{U}_0)}{|\mathbf{U} - \mathbf{U}_0|} = 0.$$

Therefore, h is di erentiable at $\mathbf{U}_0$, and $d_{\mathbf{U}_0} h$ is given by (2.74).

Example: Let

$$f(x, y, z) = 2x^2 + 4xy + 3yz,$$

$$g_1(u, v) = u^2 + v^2, \quad g_2(u, v) = u^2 - 2v^2, \quad g_3(u, v) = uv,$$

and

$$h(u, v) = f(g_1(u, v), g_2(u, v), g_3(u, v)).$$

Let $\mathbf{U}_0 = (1, -1)$ and

$$\mathbf{X}_0 = (g_1(\mathbf{U}_0), g_2(\mathbf{U}_0), g_3(\mathbf{U}_0)) = (2, -1, -1).$$

Then

$$f_x(\mathbf{X}_0) = 4, \quad f_y(\mathbf{X}_0) = 5, \quad f_z(\mathbf{X}_0) = -3,$$

Since

$$g_1(u, v) = u^2 + v^2, \quad g_2(u, v) = u^2 - 2v^2, \quad g_3(u, v) = uv,$$

$$\begin{aligned}
\frac{\partial g_1(\mathbf{U}_0)}{\partial u} &= 2, & \frac{\partial g_1(\mathbf{U}_0)}{\partial v} &= -2, \\
\frac{\partial g_2(\mathbf{U}_0)}{\partial u} &= 2, & \frac{\partial g_2(\mathbf{U}_0)}{\partial v} &= 4, \\
\frac{\partial g_3(\mathbf{U}_0)}{\partial u} &= -1, & \frac{\partial g_3(\mathbf{U}_0)}{\partial v} &= 1.
\end{aligned}$$

Therefore,

$$d_{\mathbf{U}_0} g_1 = 2\,du - 2\,dv, \quad d_{\mathbf{U}_0} g_2 = 2\,du + 4\,dv, \quad d_{\mathbf{U}_0} g_3 = -du + dv.$$

According to chain rule we have

$$d_{\mathbf{U}_0} h = f_{x_1}(\mathbf{X}_0) d_{\mathbf{U}_0} g_1 + f_{x_2}(\mathbf{X}_0) d_{\mathbf{U}_0} g_2 + \cdots + f_{x_n}(\mathbf{X}_0) d_{\mathbf{U}_0} g_n.$$

$$\begin{aligned}
d_{\mathbf{U}_0} h &= f_x(\mathbf{X}_0)\, d_{\mathbf{U}_0} g_1 + f_y(\mathbf{X}_0)\, d_{\mathbf{U}_0} g_2 + f_z(\mathbf{X}_0)\, d_{\mathbf{U}_0} g_3 \\
&= 4(2\,du - 2\,dv) + 5(2\,du + 4\,dv) - 3(-du + dv) \\
&= 21\,du + 9\,dv.
\end{aligned}$$

Since

$$d_{\mathbf{U}_0} h = h_u(\mathbf{U}_0)\, du + h_v(\mathbf{U}_0)\, dv$$

we conclude that

$$h_u(\mathbf{U}_0) = 21 \quad \text{and} \quad h_v(\mathbf{U}_0) = 9. \tag{2.77}$$

Alternatively: This can also be obtained by writing h explicitly in terms of (u, v) and di erentiating; thus,

$$\begin{aligned} h(u,v) &= 2[g_1(u,v)]^2 + 4g_1(u,v)g_2(u,v) + 3g_2(u,v)g_3(u,v) \\ &= 2(u^2+v^2)^2 + 4(u^2+v^2)(u^2-2v^2) + 3(u^2-2v^2)uv \\ &= 6u^4 + 3u^3v - 6uv^3 - 6v^4. \end{aligned}$$

Hence,

$$h_u(u,v) = 24u^3 + 9u^2v - 6v^3 \quad \text{and} \quad h_v(u,v) = 3u^3 - 18uv^2 - 24v^3,$$

so $h_u(1, -1) = 21$ and $h_v(1, -1) = 9$, consistent with (2.77).

Corollary: Under the assumptions of the chain rule theorem

$$\frac{\partial h(\mathbf{U}_0)}{\partial u_i} = \sum_{j=1} \frac{\partial f(\mathbf{X}_0)}{\partial x_j}\frac{\partial g_j(\mathbf{U}_0)}{\partial u_i}, \quad 1 \le i \le m. \tag{2.78}$$

Proof:Substituting

$$d_{\mathbf{U}_0} g_i = \frac{\partial g_i(\mathbf{U}_0)}{\partial u_1} du_1 + \frac{\partial g_i(\mathbf{U}_0)}{\partial u_2} du_2 + \cdots + \frac{\partial g_i(\mathbf{U}_0)}{\partial u_m} du_m, \quad 1 \le i \le n,$$

into (2.74) and collecting multipliers of $du_1, du_2, \ldots, du_m$ yields

$$d_{\mathbf{U}_0} h = \sum_{i=1}^{m} \left(\sum_{j=1} \frac{\partial f(\mathbf{X}_0)}{\partial x_j}\frac{\partial g_j(\mathbf{U}_0)}{\partial u_i} \right) du_i.$$

However, from Theorem ??,

$$d_{\mathbf{U}_0} h = \sum_{i=1}^{m} \frac{\partial h(\mathbf{U}_0)}{\partial u_i} du_i.$$

Comparing the last two equations yields (2.78).

Remark: When it is not important to emphasize the particular point $\mathbf{X}_0$, we write

$$\frac{\partial h}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial g_j}{\partial u_i}, \quad 1 \le i \le m, \tag{2.79}$$

with the understanding that in calculating $\partial h(\mathbf{U}_0)/\partial u_i$, $\partial g_j/\partial u_i$ is evaluated at $\mathbf{U}_0$ and $\partial f/\partial x_j$ at $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$.

$$\frac{\partial h}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial g_j}{\partial u_i}, \quad 1 \le i \le m, \tag{2.80}$$

with the understanding that in calculating $\partial h(\mathbf{U}_0)/\partial u_i$, $\partial g_j/\partial u_i$ is evaluated at $\mathbf{U}_0$ and $\partial f/\partial x_j$ at $\mathbf{X}_0 = \mathbf{G}(\mathbf{U}_0)$. By replacing the symbol $\mathbf{G}$ with $\mathbf{X} = \mathbf{X}(\mathbf{U})$; then we write

$$h(\mathbf{U}) = f(\mathbf{X}(\mathbf{U}))$$

and

$$\frac{\partial h(\mathbf{U}_0)}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial f(\mathbf{X}_0)}{\partial x_j}\frac{\partial x_j(\mathbf{U}_0)}{\partial u_i},$$

$$\text{or simply } \frac{\partial h}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial u_i}. \tag{2.81}$$

2.21 Higher derivatives of composite functions

Higher derivatives of composite functions can be computed by repeatedly applying the chain rule.

For example, di erentiating (2.81) with respect to u_k yields

$$\begin{aligned}\frac{\partial^2 h}{\partial u_k \partial u_i} &= \sum_{j=1}^{n} \frac{\partial}{\partial u_k}\left(\frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial u_i}\right)\\ &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial^2 x_j}{\partial u_k \partial u_i} + \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_i}\frac{\partial}{\partial u_k}\left(\frac{\partial f}{\partial x_j}\right).\end{aligned} \tag{2.82}$$

We must be careful nding

$$\frac{\partial}{\partial u_k}\left(\frac{\partial f}{\partial x_j}\right),$$

which really stands here for

$$\frac{\partial}{\partial u_k}\left(\frac{\partial f(\mathbf{X}(\mathbf{U}))}{\partial x_j}\right). \tag{2.83}$$

The safest procedure is to write temporarily

$$g(\mathbf{X}) = \frac{\partial f(\mathbf{X})}{\partial x_j};$$

then (2.83) becomes

$$\frac{\partial g(\mathbf{X}(\mathbf{U}))}{\partial u_k} = \sum_{s=1}^{n} \frac{\partial g(\mathbf{X}(\mathbf{U}))}{\partial x_s}\frac{\partial x_s(\mathbf{U})}{\partial u_k}.$$

Since

$$\frac{\partial g}{\partial x_s} = \frac{\partial^2 f}{\partial x_s \partial x_j},$$

this yields

$$\frac{\partial}{\partial u_k}\left(\frac{\partial f}{\partial x_k}\right) = \sum_{s=1}^{n} \frac{\partial^2 f}{\partial x_s \partial x_j}\frac{\partial x_s}{\partial u_k}.$$

Substituting this into (2.82) yields

$$\frac{\partial^2 h}{\partial u_k\,\partial u_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial^2 x_j}{\partial u_k\,\partial u_i} + \sum_{j=1}^{n} \frac{\partial x_j}{\partial u_i} \sum_{s=1}^{n} \frac{\partial^2 f}{\partial x_s\,\partial x_j}\frac{\partial x_s}{\partial u_k}. \tag{2.84}$$

To compute $h_{u_iu_k}(\mathbf{U}_0)$ from this formula, we evaluate the partial derivatives of x_1, x_2, ..., x_n at $\mathbf{U}_0$ and those of f at $\mathbf{X}_0 = \mathbf{X}(\mathbf{U}_0)$. The formula is valid if x_1, x_2, ..., x_n and their first partial derivatives are differentiable at $\mathbf{U}_0$ and f, f_{x_i}, f_{x_2}, ..., f_{x_n} and their first partial derivatives are differentiable at $\mathbf{X}_0$.

Example: Let (r, ϑ) be polar coordinates in the xy-plane; that is,

$$x = r\cos\vartheta, \quad y = r\sin\vartheta.$$

Suppose that $f = f(x, y)$ is differentiable on a set S, and let

$$h(r, \vartheta) = f(r\cos\vartheta, r\sin\vartheta).$$

We have

$$\begin{aligned} \frac{\partial h}{\partial r} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \cos\vartheta\frac{\partial f}{\partial x} + \sin\vartheta\frac{\partial f}{\partial y} \\ \frac{\partial h}{\partial \vartheta} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \vartheta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \vartheta} = -r\sin\vartheta\frac{\partial f}{\partial x} + r\cos\vartheta\frac{\partial f}{\partial y}, \end{aligned} \tag{2.85}$$

where f_x and f_y are evaluated at $(x, y) = (r\cos\vartheta, r\sin\vartheta)$.

Example: Suppose that f_x and f_y just calculated are differentiable on an open set S in $\mathbb{R}^2$. Differentiating (2.85) with respect to r yields

$$\begin{aligned} \frac{\partial^2 h}{\partial r^2} &= \cos\vartheta\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial x}\right) + \sin\vartheta\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial y}\right) \\ &= \cos\vartheta\left(\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y\,\partial x}\frac{\partial y}{\partial r}\right) + \sin\vartheta\left(\frac{\partial^2 f}{\partial x\,\partial y}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial r}\right). \end{aligned} \tag{2.86}$$

if $(x, y) \in S$. Since

$$\frac{\partial x}{\partial r} = \cos\vartheta, \quad \frac{\partial y}{\partial r} = \sin\vartheta, \quad \text{and} \quad \frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\,\partial x}$$

if $(x, y) \in S$. The equation (2.86) yields

$$\frac{\partial^2 h}{\partial r^2} = \cos^2\vartheta\frac{\partial^2 f}{\partial x^2} + 2\sin\vartheta\cos\vartheta\frac{\partial^2 f}{\partial x\partial y} + \sin^2\vartheta\frac{\partial^2 f}{\partial y^2}.$$

Differentiating (2.85) with respect to ϑ yields

$$\begin{aligned} \frac{\partial^2 h}{\partial \vartheta\,\partial r} &= -\sin\vartheta\frac{\partial f}{\partial x} + \cos\vartheta\frac{\partial f}{\partial y} + \cos\vartheta\frac{\partial}{\partial \vartheta}\left(\frac{\partial f}{\partial x}\right) + \sin\vartheta\frac{\partial}{\partial \vartheta}\left(\frac{\partial f}{\partial y}\right) \\ &= -\sin\vartheta\frac{\partial f}{\partial x} + \cos\vartheta\frac{\partial f}{\partial y} + \cos\vartheta\left(\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial \vartheta} + \frac{\partial^2 f}{\partial y\,\partial x}\frac{\partial y}{\partial \vartheta}\right) \\ &\quad + \sin\vartheta\left(\frac{\partial^2 f}{\partial x\partial y}\frac{\partial x}{\partial \vartheta} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial \vartheta}\right). \end{aligned}$$

Since

$$\frac{\partial x}{\partial \vartheta} = -r\sin\vartheta \quad \text{and} \quad \frac{\partial y}{\partial \vartheta} = r\cos\vartheta,$$

it follows that

$$\begin{aligned}\frac{\partial^2 h}{\partial\vartheta\,\partial r} &= -\sin\vartheta\frac{\partial f}{\partial x} + \cos\vartheta\frac{\partial f}{\partial y} - r\sin\vartheta\cos\vartheta\left(\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial y^2}\right)\\ &\quad + r(\cos^2\vartheta - \sin^2\vartheta)\frac{\partial^2 f}{\partial x\partial y}.\end{aligned}$$

Remark: For a composite function of the form

$$h(t) = f(x_1(t), x_2(t), \dots, x_n(t))$$

where t is a real variable, $x_1, x_2, \dots, x_n$ are differentiable at t_0, and f is differentiable at $\mathbf{X}_0 = \mathbf{X}(t_0)$. We have

$$h'(t_0) = \sum_{j=1} f_{x_j}(\mathbf{X}(t_0))x_j'(t_0). \tag{2.87}$$

Theorem: Let f be continuous at $\mathbf{X}_1 = (x_{11}, x_{21}, \dots, x_{n1})$ and $\mathbf{X}_2 = (x_{12}, x_{22}, \dots, x_{n2})$ and differentiable on the line segment L from $\mathbf{X}_1$ to $\mathbf{X}_2$.
Then

$$f(\mathbf{X}_2) - f(\mathbf{X}_1) = \sum_{i=1}^{n} f_{x_i}(\mathbf{X}_0)(x_{i2} - x_{i1}) = (d_{\mathbf{X}_0}f)(\mathbf{X}_2 - \mathbf{X}_1) \tag{2.88}$$

for some $\mathbf{X}_0$ on L distinct from $\mathbf{X}_1$ and $\mathbf{X}_2$.

Proof: An equation of L is

$$\mathbf{X} = \mathbf{X}(t) = t\mathbf{X}_2 + (1-t)\mathbf{X}_1, \quad 0 \le t \le 1.$$

Our hypotheses imply that the function

$$h(t) = f(\mathbf{X}(t))$$

is continuous on $[0, 1]$ and differentiable on $(0, 1)$.
Since

$$x_i(t) = tx_{i2} + (1-t)x_{i1},$$

We have

$$h'(t) = \sum_{i=1} f_{x_i}(\mathbf{X}(t))(x_{i2} - x_{i1}), \quad 0 < t < 1.$$

From the mean value theorem for functions of one variable

$$h(1) - h(0) = h'(t_0)$$

for some $t_0 \in (0, 1)$. Since $h(1) = f(\mathbf{X}_2)$ and $h(0) = f(\mathbf{X}_1)$, this implies (2.88) with $\mathbf{X}_0 = \mathbf{X}(t_0)$, i.e.,

$$f(\mathbf{X}_2) - f(\mathbf{X}_1) = \sum_{i=1} f_{x_i}(\mathbf{X}_0)(x_{i2} - x_{i1}) = (d_{\mathbf{X}_0} f)(\mathbf{X}_2 - \mathbf{X}_1).$$

Theorem: If $f_{x_1}, f_{x_2}, \ldots, f_{x_n}$ are identically zero in an open region S of $\mathbb{R}^n$, then f is constant in S.

Proof: We will show that if $\mathbf{X}_0$ and $\mathbf{X}$ are in S, then $f(\mathbf{X}) = f(\mathbf{X}_0)$.

Since S is an open region, S is polygonally connected.

Therefore, there are points

$$\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_n = \mathbf{X}$$

such that the line segment L_i from $\mathbf{X}_{i-1}$ to $\mathbf{X}_i$ is in S, $1 \leq i \leq n$. From mean value theorem

$$f(\mathbf{X}_i) - f(\mathbf{X}_{i-1}) = \sum_{i=1}^{n} (d_{\tilde{\mathbf{X}}_i} f)(\mathbf{X}_i - \mathbf{X}_{i-1}),$$

where $\tilde{\mathbf{X}}$ is on L_i and therefore in S.

Therefore,

$$f_{x_i}(\tilde{\mathbf{X}}_i) = f_{x_2}(\tilde{\mathbf{X}}_i) = \cdots = f_{x_n}(\tilde{\mathbf{X}}_i) = 0,$$

which means that $d_{\tilde{\mathbf{X}}_i} f \equiv 0$. Hence,

$$f(\mathbf{X}_0) = f(\mathbf{X}_1) = \cdots = f(\mathbf{X}_n);$$

that is, $f(\mathbf{X}) = f(\mathbf{X}_0)$ for every $\mathbf{X}$ in S.

Motivation: Suppose that f is de ned in an n-ball $B_\rho(\mathbf{X}_0)$, with $\rho > 0$.

If $\mathbf{X} \in B_\rho(\mathbf{X}_0)$, then

$$\mathbf{X}(t) = \mathbf{X}_0 + t(\mathbf{X} - \mathbf{X}_0) \in B_\rho(\mathbf{X}), \quad 0 \leq t \leq 1,$$

so the function

$$h(t) = f(\mathbf{X}(t))$$

is de ned for $0 \leq t \leq 1$.

We know that

$$h'(t) = \sum_{i=1} f_{x_i}(\mathbf{X}(t)(x_i - x_{i0}).$$

If f is di erentiable in $B_\rho(\mathbf{X}_0)$, and

$$h''(t) = \sum_{j=1} \frac{\partial}{\partial x_j}\left(\sum_{i=1}^{n} \frac{\partial f(\mathbf{X}(t))}{\partial x_i}(x_i - x_{i0})\right)(x_j - x_{j0})$$
$$= \sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbf{X}(t))}{\partial x_j \partial x_i}(x_i - x_{i0})(x_j - x_{j0})$$

If $f_{x_1}, f_{x_2}, \ldots, f_{x_n}$ are di erentiable in $B_\rho(\mathbf{X}_0)$. Continuing in this way, we see that

$$h^{(r)}(t) = \sum_{i_1,i_2,\ldots,i_r=1} \frac{\partial^r f(\mathbf{X}(t))}{\partial x_{i_r} \partial x_{i_{r-1}} \cdots \partial x_{i_1}} (x_{i_1} - x_{i_1,0})(x_{i_2} - x_{i_2,0}) \cdots (x_{i_r} - x_{i_r,0})$$

if all partial derivatives of f of order $\le r - 1$ are di erentiable in $B_\rho(\mathbf{X}_0)$.

2.22 rth Di erential

Suppose that $r \ge 1$ and all partial derivatives of f of order $\le r - 1$ are di erentiable in a neighborhood of $\mathbf{X}_0$.

Then the rth di erential of f at $\mathbf{X}_0$, denoted by $d^{(r)}_{\mathbf{X}_0} f$, is de ned by

$$d^{(r)}_{\mathbf{X}_0} f = \sum_{i_1,i_2,\ldots,i_r=1} \frac{\partial^r f(\mathbf{X}_0)}{\partial x_{i_r} \partial x_{i_{r-1}} \cdots \partial x_{i_1}} dx_{i_1} dx_{i_2} \cdots dx_{i_r}, \qquad (2.89)$$

where $dx_1, dx_2, \ldots, dx_n$ are the di erentials, that is, dx_i is the function whose value at a point in $\mathbb{R}^n$ is the ith coordinate of the point. For convenience, we de ne

$$(d^{(0)}_{\mathbf{X}_0} f) = f(\mathbf{X}_0).$$

Notice that $d^{(1)}_{\mathbf{X}_0} f = d_{\mathbf{X}_0} f$.

Remark: Suppose that $r \ge 1$ and all partial derivatives of f of order $\le r - 1$ are di erentiable in a neighborhood of $\mathbf{X}_0$, the value of

$$\frac{\partial^r f(\mathbf{X}_0)}{\partial x_{i_r} \partial x_{i_{r-1}} \cdots \partial x_{i_1}}$$

depends only on the number of times f is di erentiated with respect to each variable, and not on the order in which the di erentiations are performed.

Remark: The di erential can be rewritten as

$$d^{(r)}_{\mathbf{X}_0} f = \sum_{r} \frac{r!}{r_1! r_2! \cdots r_n!} \frac{\partial^r f(\mathbf{X}_0)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} (dx_1)^{r_1} (dx_2)^{r_2} \cdots (dx_n)^{r_n}, \qquad (2.90)$$

where $\sum_r$ indicates summation over all ordered n-tuples $(r_1, r_2, \ldots, r_n)$ of nonnegative integers such that

$$r_1 + r_2 + \cdots + r_n = r$$

and $\partial x_i^{r_i}$ is omitted from the denominators of all terms in (2.90) for which $r_i = 0$. In particular, if $n = 2$,

$$d^{(r)}_{\mathbf{X}_0} f = \sum_{j=0}^{r} \binom{r}{j} \frac{\partial^r f(x_0, y_0)}{\partial x^j\, \partial y^{r-j}} (dx)^j (dy)^{r-j}.$$

Example: Let

$$f(x, y) = \frac{1}{1 + ax + by},$$

where a and b are constants.

Then

$$\frac{\partial^r f(x, y)}{\partial x^j\, \partial y^{r-j}} = (-1)^r r! \frac{a^j b^{r-j}}{(1 + ax + by)^{r+1}},$$

so

$$\begin{aligned} d^{(r)}_{\mathbf{X}_0} f &= \frac{(-1)^r r!}{(1 + ax_0 + by_0)^{r+1}} \sum_{j=0}^{r} \binom{r}{j} a^j b^{r-j} (dx)^j (dy)^{r-j} \\ &= \frac{(-1)^r r!}{(1 + ax_0 + by_0)^{r+1}} (a\, dx + b\, dy)^r \end{aligned}$$

if $1 + ax_0 + by_0 \neq 0$.

Let

$$f(\mathbf{X}) = \exp\left(-\sum_{j=1}^{n} a_j x_j\right),$$

where $a_1, a_2, \ldots, a_n$ are constants. Then

$$\frac{\partial^r f(\mathbf{X})}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} = (-1)^r a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n} \exp\left(-\sum_{j=1}^{n} a_j x_j\right).$$

Therefore,

$$\begin{aligned} (d^{(r)}_{\mathbf{X}_0} f)(\mathbf{\Phi}) &= (-1)^r \left(\sum_r \frac{r!}{r_1! r_2! \cdots r_n!} a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n} (dx_1)^{r_1} (dx_2)^{r_2} \cdots (dx_n)^{r_n}\right) \\ &\quad \times \exp\left(-\sum_{j=1}^{n} a_j x_{j0}\right) \\ &= (-1)^r (a_1\, dx_1 + a_2\, dx_2 + \cdots + a_n\, dx_n)^r \exp\left(-\sum_{j=1}^{n} a_j x_{j0}\right). \end{aligned}$$

2.23 Taylor's Theorem for Functions of n Variables

Theorem: Suppose that f and its partial derivatives of order $\le k$ are di erentiable at $\mathbf{X}_0$ and $\mathbf{X}$ in $\mathbb{R}^n$ and on the line segment L connecting them.

Then

$$f(\mathbf{X}) = \sum_{r=0}^{k} \frac{1}{r!}(d^{(r)}_{\mathbf{X}_0} f)(\mathbf{X} - \mathbf{X}_0) + \frac{1}{(k+1)!}(d^{(k+1)}_{\tilde{\mathbf{X}}} f)(\mathbf{X} - \mathbf{X}_0) \tag{2.91}$$

for some $\tilde{\mathbf{X}}$ on L distinct from $\mathbf{X}_0$ and $\mathbf{X}$.

Proof: De ne

$$h(t) = f(\mathbf{X}_0 + t(\mathbf{X} - \mathbf{X}_0)). \tag{2.92}$$

With $\mathbf{\Phi} = \mathbf{X} - \mathbf{X}_0$, our assumptions and the discussion preceding De nition of di erentials imply that $h, h', \dots, h^{(k+1)}$ exist on $[0, 1]$.

From Taylor's theorem for functions of one variable,

$$h(1) = \sum_{r=0}^{k} \frac{h^{(r)}(0)}{r!} + \frac{h^{(k+1)}(\tau)}{(k+1)!}, \tag{2.93}$$

for some $\tau \in (0, 1)$.

From (2.92),

$$h(0) = f(\mathbf{X}_0) \quad \text{and} \quad h(1) = f(\mathbf{X}). \tag{2.94}$$

We have $\mathbf{\Phi} = \mathbf{X} - \mathbf{X}_0$,

$$h^{(r)}(0) = (d^{(r)}_{\mathbf{X}_0} f)(\mathbf{X} - \mathbf{X}_0), \quad 1 \le r \le k, \tag{2.95}$$

$$h^{(k+1)}(\tau) = \left(d^{(k+1)}_{\tilde{\mathbf{X}}} f\right)(\mathbf{X} - \mathbf{X}_0) \tag{2.96}$$

where

$$\tilde{\mathbf{X}} = \mathbf{X}_0 + \tau(\mathbf{X} - \mathbf{X}_0)$$

is on L and distinct from $\mathbf{X}_0$ and $\mathbf{X}$.

Substituting (2.94), (2.95), and (2.96) into (2.93) yields (2.91).

Let

$$f(x, y) = \frac{1}{1 + ax + by},$$

where a and b are constants.

Then

$$\frac{\partial^r f(x, y)}{\partial x^j \partial y^{r-j}} = (-1)^r r! \frac{a^j b^{r-j}}{(1 + ax + by)^{r+1}},$$

so

$$\begin{aligned} d^{(r)}_{\mathbf{X}_0} f &= \frac{(-1)^r r!}{(1 + ax_0 + by_0)^{r+1}} \sum_{j=0}^{r} \binom{r}{j} a^j b^{r-j} (dx)^j (dy)^{r-j} \\ &= \frac{(-1)^r r!}{(1 + ax_0 + by_0)^{r+1}} (a\,dx + b\,dy)^r \end{aligned}$$

if $1 + ax_0 + by_0 \neq 0$.

Example: The Taylor series with $\mathbf{X}_0 = (0, 0)$ and $\mathbf{\Phi} = (x, y)$ imply that if $1 + ax + by > 0$, then

$$\frac{1}{1 + ax + by} = \sum_{r=0}^{k} (-1)^r (ax + by)^r + (-1)^{k+1} \frac{(ax + by)^{k+1}}{(1 + a\tau x + b\tau y)^{k+2}}$$

for some $\tau \in (0, 1)$. (Note that τ depends on k as well as (x, y).)

Remark: By analogy with the situation for functions of one variable, we de ne the kth Taylor polynomial of f about $\mathbf{X}_0$ by

$$T_k(\mathbf{X}) = \sum_{r=0}^{k} \frac{1}{r!} (d^{(r)}_{\mathbf{X}_0} f)(\mathbf{X} - \mathbf{X}_0).$$

If the di erentials exist; then we have

$$f(\mathbf{X}) = T_k(\mathbf{X}) + \frac{1}{(k+1)!} (d^{(k+1)}_{\tilde{\mathbf{X}}} f)(\mathbf{X} - \mathbf{X}_0).$$

Theorem: Suppose that f and its partial derivatives of order $\leq k - 1$ are di erentiable in a neighborhood N of a point $\mathbf{X}_0$ in $\mathbb{R}^n$ and all kth-order partial derivatives of f are continuous at $\mathbf{X}_0$. Then

$$\lim_{\mathbf{X} \to \mathbf{X}_0} \frac{f(\mathbf{X}) - T_k(\mathbf{X})}{|\mathbf{X} - \mathbf{X}_0|^k} = 0. \tag{2.97}$$

Proof: If $\varepsilon > 0$, there is a $\delta > 0$ such that $B_\delta(\mathbf{X}_0) \subset N$ and all kth-order partial derivatives of f satisfy the inequality

$$\left| \frac{\partial^k f(\tilde{\mathbf{X}})}{\partial x_{i_k} \partial x_{i_{k-1}} \cdots \partial x_{i_1}} - \frac{\partial^k f(\mathbf{X}_0)}{\partial x_{i_k} \partial x_{i_{k-1}} \cdots \partial x_{i_1}} \right| < \varepsilon, \quad \tilde{\mathbf{X}} \in B_\delta(\mathbf{X}_0). \tag{2.98}$$

Now suppose that $\mathbf{X} \in B_\delta(\mathbf{X}_0)$. From Taylor series expansion, with k replaced by $k - 1$,

$$f(\mathbf{X}) = T_{k-1}(\mathbf{X}) + \frac{1}{k!} (d^{(k)}_{\tilde{\mathbf{X}}} f)(\mathbf{X} - \mathbf{X}_0), \tag{2.99}$$

where $\tilde{\mathbf{X}}$ is some point on the line segment from $\mathbf{X}_0$ to $\mathbf{X}$ and is therefore in $B_\delta(\mathbf{X}_0)$.

We can rewrite (2.99) as

$$f(\mathbf{X}) = T_k(\mathbf{X}) + \frac{1}{k!} \left[(d^{(k)}_{\tilde{\mathbf{X}}} f)(\mathbf{X} - \mathbf{X}_0) - (d^{(k)}_{\mathbf{X}_0} f)(\mathbf{X} - \mathbf{X}_0) \right]. \tag{2.100}$$

But de nition of di erential and (2.98) imply that

$$\left| (d^{(k)}_{\tilde{\mathbf{X}}} f)(\mathbf{X} - \mathbf{X}_0) - (d^{(k)}_{\mathbf{X}_0} f)(\mathbf{X} - \mathbf{X}_0) \right| < n^k \varepsilon |\mathbf{X} - \mathbf{X}_0|^k \tag{2.101}$$

which implies that

$$\frac{|f(\mathbf{X}) - T_k(\mathbf{X})|}{|\mathbf{X} - \mathbf{X}_0|^k} < \frac{n^k \epsilon}{k!}, \quad \mathbf{X} \in B_\delta(\mathbf{X}_0),$$

from (2.100). This implies (2.97).

2.23.1 Positive Definite

Let r be a positive integer and $\mathbf{X}_0 = (x_{10}, x_{20}, \dots, x_{n0})$. A function of the form

$$p(\mathbf{X}) = \sum_r a_{r_1 r_2 \dots r_n}(x_1 - x_{10})^{r_1}(x_2 - x_{20})^{r_2} \cdots (x_n - x_{n0})^{r_n}, \tag{2.102}$$

where the coefficients $\{a_{r_1 r_2 \dots r_n}\}$ are constants and the summation is over all n-tuples of nonnegative integers $(r_1, r_2, \dots, r_n)$ such that

$$r_1 + r_2 + \cdots + r_n = r,$$

is a homogeneous polynomial of degree r in $\mathbf{X} - \mathbf{X}_0$, provided that at least one of the coefficients is nonzero. For example, if f satisfies the conditions of rth differential, then the function

$$p(\mathbf{X}) = (d^{(r)}_{\mathbf{X}_0} f)(\mathbf{X} - \mathbf{X}_0)$$

is such a polynomial if at least one of the rth-order mixed partial derivatives of f at $\mathbf{X}_0$ is nonzero. Clearly, $p(\mathbf{X}_0) = 0$ if p is a homogeneous polynomial of degree $r \geq 1$ in $\mathbf{X} - \mathbf{X}_0$.

If $p(\mathbf{X}) \geq 0$ for all $\mathbf{X}$, we say that p is positive semidefinite; if $p(\mathbf{X}) > 0$ except when $\mathbf{X} = \mathbf{X}_0$, p is positive definite. Similarly, p is negative semidefinite if $p(\mathbf{X}) \leq 0$ or negative definite if $p(\mathbf{X}) < 0$ for all $\mathbf{X} \neq \mathbf{X}_0$. In all these cases, p is semidefinite. With p as in (2.102),

$$p(-\mathbf{X} + 2\mathbf{X}_0) = (-1)^r p(\mathbf{X}),$$

so p cannot be semidefinite if r is odd.

Example: The polynomial

$$p(x, y, z) = x^2 + y^2 + z^2 + xy + xz + yz$$

is homogeneous of degree 2 in $\mathbf{X} = (x, y, z)$. We can rewrite p as

$$p(x, y, z) = \frac{1}{2}\left[(x + y)^2 + (y + z)^2 + (z + x)^2\right].$$

so p is nonnegative, and $p(\overline{x}, \overline{y}, \overline{z}) = 0$ if and only if

$$\overline{x} + \overline{y} = \overline{y} + \overline{z} = \overline{z} + \overline{x} = 0,$$

which is equivalent to $(\overline{x}, \overline{y}, \overline{z}) = (0, 0, 0)$. Therefore, p is positive definite and $-p$ is negative definite.

Example: The polynomial

$$p_1(x, y, z) = x^2 + y^2 + z^2 + 2xy$$
$$p_1(x, y, z) = (x + y)^2 + z^2,$$

so p_1 is nonnegative. Since $p_1(1, -1, 0) = 0$, p_1 is positive semide nite and $-p_1$ is negative semide nite.

The polynomial

$$p_2(x, y, z) = x^2 - y^2 + z^2$$

is not semide nite, since, for example,

$$p_2(1, 0, 0) = 1 \quad \text{and} \quad p_2(0, 1, 0) = 1.$$

Theorem: Suppose that f and its partial derivatives of order $\leq k - 1$ are di erentiable in a neighborhood N of a point $\mathbf{X}_0$ in $\mathbb{R}^n$ and all kth-order partial derivatives of f are continuous at $\mathbf{X}_0$. with $k \geq 2$, and

$$d^{(r)}_{\mathbf{X}_0} f \equiv 0 \quad (1 \leq r \leq k - 1), \qquad d^{(k)}_{\mathbf{X}_0} f \not\equiv 0. \tag{2.103}$$

Then

- $\mathbf{X}_0$ is not a local extreme point of f unless $d^{(k)}_{\mathbf{X}_0} f$ is semi-de nite as a polynomial in $\mathbf{X} - \mathbf{X}_0$.
  In particular, $\mathbf{X}_0$ is not a local extreme point of f if k is odd.
- $\mathbf{X}_0$ is a local minimum point of f if $d^{(k)}_{\mathbf{X}_0} f$ is positive de nite, or a local maximum point if $d^{(k)}_{\mathbf{X}_0} f$ is negative de nite.
- If $d^{(k)}_{\mathbf{X}_0} f$ is semide nite, then $\mathbf{X}_0$ may be a local extreme point of f, but it need not be.

Corollary: Suppose that f, f_x, and f_y are di erentiable in a neighborhood of a critical point $\mathbf{X}_0 = (x_0, y_0)$ of f and f_{xx}, f_{yy}, and f_{xy} are continuous at (x_0, y_0).

Let

$$D = f_{xx}(x_0, y_0) f_{xy}(x_0, y_0) - f^2_{xy}(x_0, y_0).$$

Then

- (x_0, y_0) is a local extreme point of f if $D > 0$; (x_0, y_0) is a local minimum point if $f_{xx}(x_0, y_0) > 0$, or a local maximum point if $f_{xx}(x_0, y_0) < 0$.
- (x_0, y_0) is not a local extreme point of f if $D < 0$.

Proof: Write $(x - x_0, y - y_0) = (u, v)$ and

$$p(u, v) = (d^{(2)}_{\mathbf{X}_0} f)(u, v) = Au^2 + 2Buv + Cv^2,$$

where $A = f_{xx}(x_0, y_0)$, $B = f_{xy}(x_0, y_0)$, and $C = f_{yy}(x_0, y_0)$, so

$$D = AC - B^2.$$

If $D > 0$, then $A \neq 0$, and we can write

$$\begin{aligned} p(u, v) &= A\left(u^2 + \frac{2B}{A}uv + \frac{B^2}{A^2}v^2\right) + \left(C - \frac{B^2}{A}\right)v^2 \\ &= A\left(u + \frac{B}{A}v\right)^2 + \frac{D}{A}v^2. \end{aligned}$$

This cannot vanish unless $u = v = 0$. Hence, $d^{(2)}_{\mathbf{X}_0} f$ is positive de nite if $A > 0$ or negative de nite if $A < 0$, and Theorem implies the rst part of the corollary.

If $D < 0$, there are three possibilities:

1. $A \neq 0$; then $p(1, 0) = A$ and $p\left(-\frac{B}{A}, 1\right) = \frac{D}{A}$.
2. $C \neq 0$; then $p(0, 1) = C$ and $p\left(1, -\frac{B}{C}\right) = \frac{D}{C}$.
3. $A = C = 0$; then $B \neq 0$ and $p(1, 1) = 2B$ and $p(1, -1) = -2B$.

In each case the two given values of p di er in sign, so $\mathbf{X}_0$ is not a local extreme point of f, from Theorem part I.

Example: If

$$f(x, y) = e^{ax^2+by^2}.$$

We have

$$f_x(x, y) = 2axf(x, y), \quad f_y(x, y) = 2byf(x, y),$$

so

$$f_x(0, 0) = f_y(0, 0) = 0,$$

and $(0, 0)$ is a critical point of f.

To apply Corollary, we calculate

$$\begin{aligned} f_{xx}(x, y) &= (2a + 4a^2x^2)f(x, y), \\ f_{yy}(x, y) &= (2b + 4b^2y^2)f(x, y), \\ f_{xy}(x, y) &= 4abxyf(x, y). \end{aligned}$$

Therefore,

$$D = f_{xx}(0, 0)f_{yy}(0, 0) - f^2_{xy}(0, 0) = (2a)(2b) - (0)(0) = 4ab.$$

Corollary implies that $(0, 0)$ is a local minimum point if a and b are positive. A local maximum if a and b are negative. Neither if one is positive and the other is negative. Corollary does not apply if a or b is zero.

Chapter 3

Integral Calculus

Attempting to formulate de nition of Riemann integral for a function de ned on an in nite or semi-in nite interval would introduce questions concerning convergence of the resulting Riemann sums, which would be in nite series.

3.1 Locally Integrable Functions

We say f is locally integrable on an interval I if f is integrable on every nite closed subinterval of I.

For example,

$$f(x) = \sin x$$

is locally integrable on $(-\infty, \infty)$.

$$g(x) = \frac{1}{x(x-1)}$$

is locally integrable on $(-\infty, 0)$, $(0, 1)$, and $(1, \infty)$.

The function

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

is locally integrable on $[0, \infty)$.

If f is locally integrable on $[a, b)$, we de ne

$$\int_a^b f(x)\, dx = \lim_{c \to b-} \int_a^c f(x)\, dx \tag{3.1}$$

if the limit exists (nite). To include the case where $b = \infty$, we adopt the convention that $\infty- = \infty$.

Remarks:

- The limit in (3.1) always exists if $[a, b)$ is nite and f is locally integrable and bounded on $[a, b)$.
- In this case, the de nition of Riemann integral and locally integrable function assign the same value to $\int_a^b f(x)\, dx$ no matter how $f(b)$ is de ned. However, the limit may also exist in cases where $b = \infty$ or $b < \infty$ and f is unbounded as x approaches b from the left.

- In these cases, the de nition of locally integrable assigns a value to an integral that does not exist in the sense of Riemann integral, and $\int_a^b f(x)\,dx$ is said to be an improper integral that converges to the limit in (3.1).

Remarks:

- We also say in this case that f is integrable on $[a, b)$ and that $\int_a^b f(x)\,dx$ exists. If the limit in (3.1) does not exist (nite), we say that the improper integral $\int_a^b f(x)\,dx$ diverges, and f is nonintegrable on $[a, b)$.

- In particular, if $\lim_{c\to b-} \int_a^c f(x)\,dx = \pm\infty$, we say that $\int_a^b f(x)\,dx$ diverges to $\pm\infty$, and we write

$$\int_a^b f(x)\,dx = \infty \quad \text{or} \quad \int_a^b f(x)\,dx = -\infty,$$

whichever the case may be.

If f is locally integrable on $(a, b]$, we de ne

$$\int_a^b f(x)\,dx = \lim_{c\to a+} \int_c^b f(x)\,dx$$

provided that the limit exists (nite).

To include the case where $a = -\infty$, we adopt the convention that $-\infty + = -\infty$. If f is locally integrable on (a, b), we de ne

$$\int_a^b f(x)\,dx = \int_a^\alpha f(x)\,dx + \int_\alpha^b f(x)\,dx,$$

where $a < \alpha < b$, provided that both improper integrals on the right exist (nite).

Remarks: The existence and value of $\int_a^b f(x)\,dx$ according to the above de nition do not depend on the particular choice of α in (a, b).

When we wish to distinguish between improper integrals and integrals in the sense of de nition of Riemann integral, we will call the latter proper integrals.

Example: The function

$$f(x) = 2x \sin\frac{1}{x} - \cos\frac{1}{x}$$

is locally integrable and the derivative of

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

on $[-2/\pi, 0)$.

Hence,

$$\int_{-2/\pi}^{c} f(x)\,dx = \left. x^2 \sin\frac{1}{x}\right|_{-2/\pi}^{c} = c^2\sin\frac{1}{c} + \frac{4}{\pi^2}$$

$$\int_{-2/\pi}^{0} f(x)\,dx = \lim_{c\to 0-}\left(c^2\sin\frac{1}{c} + \frac{4}{\pi^2}\right) = \frac{4}{\pi^2}.$$

However, this is not an improper integral, even though $f(0)$ is not defined and cannot be defined so as to make f continuous at 0. If we define $f(0)$ arbitrarily (say $f(0) = 10$), then f is bounded on the closed interval $[-2/\pi, 0]$ and continuous except at 0. Therefore, $\int_{-2/\pi}^{0} f(x)\,dx$ exists and equals $4/\pi^2$ as a proper integral, in the sense of definition of improper integral.

Example: The function

$$f(x) = (1-x)^{-p}$$

is locally integrable on $[0, 1)$.

If $p \neq 1$ and $0 < c < 1$,

$$\int_0^c (1-x)^{-p}\,dx = \left.\frac{(1-x)^{-p+1}}{p-1}\right|_0^c = \frac{(1-c)^{-p+1}-1}{p-1}.$$

Hence,

$$\lim_{c\to 1-}\int_0^c (1-x)^{-p}\,dx = \begin{cases} (1-p)^{-1}, & p<1,\\ \infty, & p>1.\end{cases}$$

For $p = 1$,

$$\lim_{c\to 1-}\int_0^c (1-x)^{-1}\,dx = -\lim_{c\to 1-}\log(1-c) = \infty.$$

Hence,

$$\int_0^1 (1-x)^{-p}\,dx = \begin{cases} (1-p)^{-1}, & p<1,\\ \infty, & p\geq 1.\end{cases}$$

Example: The function

$$f(x) = x^{-p}$$

is locally integrable on $[1, \infty)$.

If $p \neq 1$ and $c > 1$,

$$\int_1^c x^{-p}\,dx = \left.\frac{x^{-p+1}}{-p+1}\right|_1^c = \frac{c^{-p+1}-1}{-p+1}.$$

Hence,

$$\lim_{c\to\infty}\int_1^c x^{-p}\,dx = \begin{cases} (p-1)^{-1}, & p>1,\\ \infty, & p<1.\end{cases}$$

For $p = 1$,

$$\lim_{c\to\infty}\int_1^c x^{-1}\,dx = \lim_{c\to\infty}\log c = \infty.$$

Hence,

$$\int_1^\infty x^{-p}\,dx = \begin{cases} (p-1)^{-1}, & p>1,\\ \infty, & p\le 1.\end{cases}$$

Example: If $1 < c < \infty$, then

$$\int_1^c \frac{1}{x}\log\frac{1}{x}\,dx = -\int_1^c \frac{1}{x}\log x\,dx = -\frac{1}{2}(\log x)^2\Big|_1^c = -\frac{1}{2}(\log c)^2.$$

Hence,

$$\lim_{c\to\infty}\int_1^c \frac{1}{x}\log\frac{1}{x}\,dx = -\infty,$$

so

$$\int_1^\infty \frac{1}{x}\log\frac{1}{x}\,dx = -\infty.$$

The function $f(x) = \log x$ is locally integrable on $(0,1]$, but unbounded as $x\to 0+$. Since

$$\lim_{c\to 0+}\int_c^1 \log x\,dx = \lim_{c\to 0+}(x\log x - x)\Big|_c^1 = -1 - \lim_{c\to 0+}(c\log c - c) = -1,$$

De nition ?? yields

$$\int_0^1 \log x\,dx = -1.$$

The function $f(x) = \cos x$ is locally integrable on $[0,\infty)$ and

$$\lim_{c\to\infty}\int_0^c \cos x\,dx = \lim_{c\to\infty}\sin c$$

does not exist; thus, $\int_0^\infty \cos x\,dx$ diverges, but not to $\pm\infty$.

In connection with De nition ??, it is important to recognize that the improper integrals $\int_a^\alpha f(x)\,dx$ and $\int_\alpha^b f(x)\,dx$ must converge separately for $\int_a^b f(x)\,dx$ to converge. For example, the existence of the symmetric limit

$$\lim_{R\to\infty}\int_{-R}^R f(x)\,dx,$$

which is called the principal value of $\int_{-\infty}^\infty f(x)\,dx$, does not imply that $\int_{-\infty}^\infty f(x)\,dx$ converges; thus,

$$\lim_{R\to\infty}\int_{-R}^R x\,dx = \lim_{R\to\infty} 0 = 0,$$

but $\int_0^\infty x\,dx$ and $\int_{-\infty}^0 x\,dx$ diverge and therefore so does $\int_{-\infty}^\infty x\,dx$.

Theorem: Suppose that $f_1, f_2, \dots, f_n$ are locally integrable on $[a,b)$.

The integrals $\int_a^b f_1(x)\,dx, \int_a^b f_2(x)\,dx, \dots, \int_a^b f_n(x)\,dx$ converge.

Let $c_1, c_2, \dots, c_n$ be constants. Then $\int_a^b (c_1 f + c_2 f_1 + \cdots + c_n f_n)(x)\,dx$ converges.

Furthermore,

$$\begin{aligned}\int_a^b (c_1f_1 + c_2f_2 + \cdots + c_nf_n)(x)\,dx &= c_1\int_a^b f_1(x)\,dx \\ &+c_2\int_a^b f_2(x)\,dx \\ &+\cdots + c_n\int_a^b f_n(x)\,dx.\end{aligned}$$

Proof: If $a < c < b$, then

$$\begin{aligned}\int_a^c (c_1f_1 + c_2f_2 + \cdots + c_nf_n)(x)\,dx &= c_1\int_a^c f_1(x)\,dx \\ &+c_2\int_a^c f_2(x)\,dx \\ &+\cdots + c_n\int_a^c f_n(x)\,dx.\end{aligned}$$

Letting $c \to b-$ yields the stated result.

Theorem: If f is nonnegative and locally integrable on $[a, b)$, then $\int_a^b f(x)\,dx$ converges if the function

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

is bounded on $[a, b)$, and $\int_a^b f(x)\,dx = \infty$ if it is not.

These are the only possibilities, and

$$\int_a^b f(t)\,dt = \sup_{a\le x<b} F(x)$$

in either case.

Proof: The function

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

is nondecreasing on $[a, b)$.

Recall: Suppose that f is monotonic on (a, b) and de ne

$$\alpha = \inf_{a<x<b} f(x), \qquad \beta = \sup_{a<x<b} f(x).$$

If f is nondecreasing, then $f(a+) = \alpha$ and $f(b-) = \beta$.

Remarks: We often write

$$\int_a^b f(x)\,dx < \infty$$

to indicate that an improper integral of a nonnegative function converges.

Similarly, if f is nonpositive and $\int_a^b f(x)\,dx$ converges, we write

$$\int_a^b f(x)\,dx > -\infty$$

because a divergent integral of this kind can only diverge to $-\infty$.

- These conventions do not apply to improper integrals of functions that assume both positive and negative values in $(a,\ b)$, since they may diverge without diverging to $\pm\infty$.

3.1.1 The Comparison Test

Theorem: If f and g are locally integrable on $[a, b)$ and

$$0 \leq f(x) \leq g(x), \quad a \leq x < b, \tag{3.2}$$

then

1. $\int_a^b f(x)\,dx < \infty$ if $\int_a^b g(x)\,dx < \infty$

2. $\int_a^b g(x)\,dx = \infty$ if $\int_a^b f(x)\,dx = \infty$.

Proof: Since

$$0 \leq f(x) \leq g(x), \quad a \leq x < b,$$

we have

$$\int_a^x f(t)\,dt \leq \int_a^x g(t)\,dt, \quad a \leq x < b.$$

So

$$\sup_{a \leq x < b} \int_a^x f(t)\,dt \leq \sup_{a \leq x \leq b} \int_a^x g(t)\,dt.$$

If $\int_a^b g(x)\,dx < \infty$, the right side of this inequality is nite by the previous Theorem, so the left side is also.

This implies that $\int_a^b f(x)\,dx < \infty$.

The proof is by contradiction. If $\int_a^b g(x)\,dx < \infty$, then (1) implies that $\int_a^b f(x)\,dx < \infty$, contradicting the assumption that $\int_a^b f(x)\,dx = \infty$.

Example: Determine the convergence of the improper integral

$$I = \int_0^1 \frac{2 + \sin \pi x}{(1-x)^p}\,dx.$$

Solution: We are going to show that the improper integral converges if $p < 1$. Since

$$0 < \frac{2 + \sin \pi x}{(1-x)^p} \leq \frac{3}{(1-x)^p}, \quad 0 \leq x < 1.$$

We have

$$\int_0^1 \frac{3\,dx}{(1-x)^p} < \infty, \quad p < 1.$$

Example: Determine the convergence of the improper integral

$$I = \int_0^1 \frac{2 + \sin \pi x}{(1-x)^p}\,dx.$$

Solution: However, I diverges if $p \geq 1$, since

$$0 < \frac{1}{(1-x)^p} \leq \frac{2 + \sin \pi x}{(1-x)^p}, \quad 0 \leq x < 1,$$

and

$$\int_0^1 \frac{dx}{(1-x)^p} = \infty, \quad p \geq 1.$$

Remark: If f is any function (not necessarily nonnegative) locally integrable on $[a, b)$. If a_1 and c are in $[a, b)$, then

$$\int_a^c f(x)\,dx = \int_a^{a_1} f(x)\,dx + \int_{a_1}^c f(x)\,dx.$$

Since $\int_a^{a_1} f(x)\,dx$ is a proper integral, on letting $c \to b-$ we conclude that if either of the improper integrals $\int_a^b f(x)\,dx$ and $\int_{a_1}^b f(x)\,dx$ converges then so does the other, and in this case

$$\int_a^b f(x)\,dx = \int_a^{a_1} f(x)\,dx + \int_{a_1}^b f(x)\,dx.$$

Remark: This means that any theorem implying convergence or divergence of an improper integral $\int_a^b f(x)\,dx$ remains valid if its hypotheses are satis ed on a subinterval $[a_1, b)$ of $[a, b)$ rather than on all of $[a, b)$.

For example, the comparison test remains valid if we have

$$0 \leq f(x) \leq g(x), \quad a_1 \leq x < b,$$

where a_1 is any point in $[a, b)$.

From this, you can see that if $f(x) \geq 0$ on some subinterval $[a_1, b)$ of $[a, b)$, but not necessarily for all x in $[a, b)$, we can still use the convention introduced earlier for positive functions; that is, we can write $\int_a^b f(x)\,dx < \infty$ if the improper integral converges or $\int_a^b f(x)\,dx = \infty$ if it diverges.

Theorem: Suppose that f and g are locally integrable on $[a, b)$, $g(x) > 0$ and $f(x) \geq 0$ on some subinterval $[a_1, b)$ of $[a, b)$, and

$$\lim_{x\to b-} \frac{f(x)}{g(x)} = M. \tag{3.3}$$

- If $0 < M < \infty$, then $\int_a^b f(x)\,dx$ and $\int_a^b g(x)\,dx$ converge or diverge together.
- If $M = \infty$ and $\int_a^b g(x)\,dx = \infty$, then $\int_a^b f(x)\,dx = \infty$.
- If $M = 0$ and $\int_a^b g(x)\,dx < \infty$, then $\int_a^b f(x)\,dx < \infty$.

Proof: From (3.3), there is a point a_2 in $[a_1, b)$ such that

$$0 < \frac{M}{2} < \frac{f(x)}{g(x)} < \frac{3M}{2}, \quad a_2 \leq x < b,$$

and therefore

$$\frac{M}{2} g(x) < f(x) < \frac{3M}{2} g(x), \quad a_2 \leq x < b. \tag{3.4}$$

The rst inequality in (3.4) imply that

$$\int_{a_2}^b g(x)\,dx < \infty \quad \text{if} \quad \int_{a_2}^b f(x)\,dx < \infty.$$

The second inequality in (3.4) imply that

$$\int_{a_2}^b f(x)\,dx < \infty \quad \text{if} \quad \int_{a_2}^b g(x)\,dx < \infty.$$

Therefore, $\int_{a_2}^b f(x)\,dx$ and $\int_{a_2}^b g(x)\,dx$ converge or diverge together, and in the latter case they must diverge to ∞, since their integrands are nonnegative. If $M = \infty$, there is a point a_2 in $[a_1, b)$ such that

$$f(x) \geq g(x), \quad a_2 \leq x \leq b,$$

We have $\int_a^b f(x)\,dx = \infty$. If $M = 0$, there is a point a_2 in $[a_1, b)$ such that

$$f(x) \leq g(x), \quad a_2 \leq x \leq b,$$

so we have $\int_a^b f(x)\,dx < \infty$.

3.2 Absolute integrability

We say that f is absolutely integrable on $[a, b)$ if f is locally integrable on $[a, b)$ and $\int_a^b |f(x)|\,dx < \infty$. In this case we also say that $\int_a^b f(x)\,dx$ converges absolutely or is absolutely convergent.

Remark: If f is nonnegative and integrable on $[a, b)$, then f is absolutely integrable on $[a, b)$, since $|f| = f$.

Example: Since

$$\left|\frac{\sin x}{x^p}\right| \le \frac{1}{x^p}$$

and $\int_1^\infty x^{-p}\,dx < \infty$ if $p > 1$.

The comparison theorem implies that

$$\int_1^\infty \frac{|\sin x|}{x^p}\,dx < \infty, \quad p > 1.$$

The function

$$f(x) = \frac{\sin x}{x^p}$$

is absolutely integrable on $[1, \infty)$ if $p > 1$.

Example: It is not absolutely integrable on $[1, \infty)$ if $p \le 1$.

To see this, we rst consider the case where $p = 1$.

Let k be an integer greater than 3. Then

$$\begin{aligned}\int_1^{k\pi} \frac{|\sin x|}{x}\,dx &> \int_\pi^{k\pi} \frac{|\sin x|}{x}\,dx \\ &= \sum_{j=1}^{k-1} \int_{j\pi}^{(j+1)\pi} \frac{|\sin x|}{x}\,dx \\ &> \sum_{j=1}^{k-1} \frac{1}{(j+1)\pi} \int_{j\pi}^{(j+1)\pi} |\sin x|\,dx.\end{aligned} \tag{3.5}$$

But

$$\int_{j\pi}^{(j+1)\pi} |\sin x|\,dx = \int_0^\pi \sin x\,dx = 2,$$

so (3.5) implies that

$$\int_1^{k\pi} \frac{|\sin x|}{x}\,dx > \frac{2}{\pi} \sum_{j=1}^{k-1} \frac{1}{j+1}. \tag{3.6}$$

However,

$$\frac{1}{j+1} \ge \int_{j+1}^{j+2} \frac{dx}{x}, \quad j = 1, 2, \ldots,$$

so (3.6) implies that

$$\begin{aligned}\int_1^{k\pi} \frac{|\sin x|}{x} &> \frac{2}{\pi} \sum_{j=1}^{k-1} \int_{j+1}^{j+2} \frac{dx}{x} \\ &= \frac{2}{\pi} \int_2^{k+1} \frac{dx}{x} = \frac{2}{\pi} \log \frac{k+1}{2}.\end{aligned}$$

Since $\lim_{k\to\infty} \log[(k+1)/2] = \infty$, implies that

$$\int_1^\infty \frac{|\sin x|}{x}\, dx = \infty.$$

Now implies that

$$\int_1^\infty \frac{|\sin x|}{x^p}\, dx = \infty, \qquad p \le 1. \tag{3.7}$$

Theorem: If f is locally integrable on $[a, b)$ and $\int_a^b |f(x)|\, dx < \infty$, then $\int_a^b f(x)\, dx$ converges; that is, an absolutely convergent integral is convergent.

Proof: If

$$g(x) = |f(x)| - f(x).$$

Then

$$0 \le g(x) \le 2|f(x)|$$

and $\int_a^b g(x)\, dx < \infty$, because of comparison theorem and the absolute integrability of f. Since

$$f = |f| - g,$$

Due to comparison test, we can conclude that $\int_a^b f(x)\, dx$ converges.

3.3 Nonoscillatory and Oscillatory Functions

A function f is nonoscillatory at $b-$ $(= \infty$ if $b = \infty)$ if f is de ned on $[a, b)$ and does not change sign on some subinterval $[a_1, b)$ of $[a, b)$.

If f changes sign on every such subinterval, f is oscillatory at $b-$.

Remark: For a function that is locally integrable on $[a, b)$ and nonoscillatory at $b-$, convergence and absolute convergence of $\int_a^b f(x)\, dx$ amount to the same thing, so absolute convergence is not an interesting concept in connection with such functions.

However, an oscillatory function may be integrable, but not absolutely integrable, on $[a, b)$, as the next example shows. We then say that f is conditionally integrable on $[a, b)$, and that $\int_a^b f(x)\, dx$ converges conditionally.

3.4 Conditional convergence

An oscillatory function may be integrable, but not absolutely integrable, on $[a, b)$, as the next example shows. We then say that f is conditionally integrable on $[a, b)$, and that $\int_a^b f(x)\, dx$ converges conditionally.

Example: The integral

$$I(p) = \int_1^\infty \frac{\sin x}{x^p}\, dx$$

is not absolutely convergent if $0 < p \le 1$.

We will show that it converges conditionally for these values of p.

Integration by parts yields

$$\int_1^c \frac{\sin x}{x^p}\,dx = -\frac{\cos c}{c^p} + \cos 1 - p\int_1^c \frac{\cos x}{x^{p+1}}\,dx. \tag{3.8}$$

Since

$$\left|\frac{\cos x}{x^{p+1}}\right| \le \frac{1}{x^{p+1}}$$

and $\int_1^\infty x^{-p-1}\,dx < \infty$ if $p > 0$, the comparison theorem implies that $x^{-p-1}\cos x$ is absolutely integrable $[1, \infty)$ if $p > 0$.

Therefore, we have an absolutely convergent integral, this implies that $x^{-p-1}\cos x$ is integrable $[1, \infty)$ if $p > 0$.

Letting $c \to \infty$ in (3.8), we nd that $I(p)$ converges, and

$$I(p) = \cos 1 - p\int_1^\infty \frac{\cos x}{x^{p+1}}\,dx \quad \text{if} \quad p > 0.$$

This and $\int_1^\infty \frac{|\sin x|}{x^p}\,dx = \infty, \quad p \le 1$, imply that $I(p)$ converges conditionally if $0 < p \le 1$.

3.5 Dirichlet's Test

Theorem: Suppose that f is continuous and its antiderivative $F(x) = \int_a^x f(t)\,dt$ is bounded on $[a, b)$.

Let g' be absolutely integrable on $[a, b)$, and suppose that

$$\lim_{x\to b-} g(x) = 0. \tag{3.9}$$

Then $\int_a^b f(x)g(x)\,dx$ converges.

Proof: The continuous function fg is locally integrable on $[a, b)$. Integration by parts yields

$$\int_a^c f(x)g(x)\,dx = F(c)g(c) - \int_a^c F(x)g'(x)\,dx,\ a \le c < b. \tag{3.10}$$

The comparison test implies that the integral on the right converges absolutely as $c \to b-$, since $\int_a^b |g'(x)|\,dx < \infty$ by assumption, and

$$|F(x)g'(x)| \le M\,|g'(x)|,$$

where M is an upper bound for $|F|$ on $[a, b)$. Moreover, (3.9) and the boundedness of F imply that $\lim_{c\to b-} F(c)g(c) = 0$.

Letting $c \to b-$ in (3.10) yields

$$\int_a^b f(x)g(x)\,dx = -\int_a^b F(x)g'(x)\,dx,$$

where the integral on the right converges absolutely.

Remark: Dirichlet's test is useful only if f is oscillatory at $b-$, since it can be shown that if f is nonoscillatory at $b-$ and F is bounded on $[a, b)$, then $\int_a^b |f(x)g(x)|\,dx < \infty$ if only g is locally integrable and bounded on $[a, b)$.

Remark: Dirichlet's test can also be used to show that certain integrals diverge.

Example: For example,

$$\int_1^\infty x^q \sin x\,dx$$

diverges if $q > 0$, but none of the other tests that we have studied so far implies this. It is not enough to argue that the integrand does not approach zero as $x \to \infty$ (a common mistake), since this does not imply divergence. To see that the integral diverges, we observe that if it converged for some $q > 0$, then $F(x) = \int_1^x x^q \sin x\,dx$ would be bounded on $[1, \infty)$.

We could let

$$f(x) = x^q \sin x \quad \text{and} \quad g(x) = x^{-q}$$

in Dirichlet's test and conclude that

$$\int_1^\infty \sin x\,dx$$

also converges. This is false.

3.6 Rectangles in $\mathbb{R}^n$

The

$$S_1 \times S_2 \times \cdots \times S_n$$

of subsets $S_1, S_2, \ldots, S_n$ of $\mathbb{R}$ is the set of points $(x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$ such that $x_1 \in S_1, x_2 \in S_2, \ldots, x_n \in S_n$. For example, the Cartesian product of the two closed intervals

$$[a_1, b_1] \times [a_2, b_2] = \{(x, y) : a_1 \le x \le b_1,\ a_2 \le y \le b_2\}$$

is a rectangle in $\mathbb{R}^2$ with sides parallel to the x- and y-axes.

The Cartesian product of three closed intervals

$$\begin{aligned} [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] &= \{(x, y, z) : a_1 \le x \le b_1, \\ &\qquad a_2 \le y \le b_2,\ a_3 \le z \le b_3\} \end{aligned}$$

is a rectangular parallelepiped in $\mathbb{R}^3$ with faces parallel to the coordinate axes. A coordinate rectangle R in $\mathbb{R}^n$ is the Cartesian product of n closed intervals; that is,

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

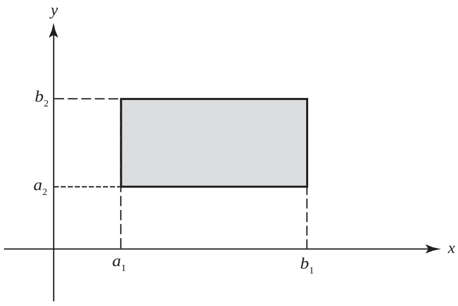

Figure 3.1: Rectangle in $\mathbb{R}^2$

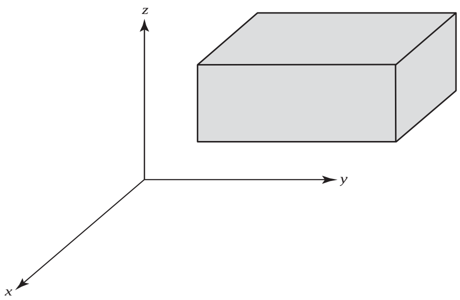

Figure 3.2: Rectangular parallelepiped in $\mathbb{R}^3$

The content of R is

$$V(R) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n).$$

The numbers $b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n$ are the edge lengths of R. If they are equal, then R is a coordinate cube. If $a_r = b_r$ for some r, then $V(R) = 0$ and we say that R is degenerate; otherwise, R is nondegenerate.

If $n = 1, 2$, or 3, then $V(R)$ is, respectively, the length of an interval, the area of a rectangle, or the volume of a rectangular parallelepiped. Henceforth, rectangle or cube will always mean coordinate rectangle or coordinate cube unless it is stated otherwise. If

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

and

$$P_r: \ a_r = a_{r0} < a_{r1} < \cdots < a_{rm_r} = b_r$$

is a partition of $[a_r, b_r]$, $1 \le r \le n$, then the set of all rectangles in $\mathbb{R}^n$ that can be written as

$$[a_{1,j_1-1}, a_{1j_1}] \times [a_{2,j_2-1}, a_{2j_2}] \times \cdots \times [a_{n,j_n-1}, a_{nj_n}], \ 1 \le j_r \le m_r, \qquad 1 \le r \le n,$$

is a partition of R. We denote this partition by

$$\mathbf{P} = P_1 \times P_2 \times \cdots \times P_n. \tag{3.11}$$

We de ne its norm to be the maximum of the norms of $P_1, P_2, \ldots, P_n$, thus,

$$\|\mathbf{P}\| = \max\{\|P_1\|, \|P_2\|, \ldots, \|P_n\|\}.$$

Put another way, $\|\mathbf{P}\|$ is the largest of the edge lengths of all the subrectangles in $\mathbf{P}$. Geometrically, a rectangle in $\mathbb{R}^2$ is partitioned by drawing horizontal and vertical lines through it; in $\mathbb{R}^3$, by drawing planes through it parallel to the coordinate axes. Partitioning divides a rectangle R into nitely many subrectangles that we can number in arbitrary order as $R_1, R_2, \ldots, R_k$. Sometimes it is convenient to write

$$\mathbf{P} = \{R_1, R_2, \ldots, R_k\}$$

rather than (3.11).

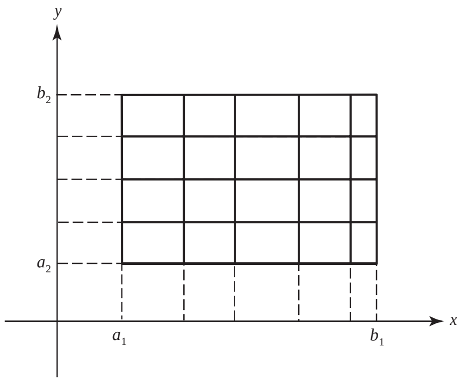

Figure 3.3: Partitioning of a rectangle in $\mathbb{R}^2$

- If $\mathbf{P} = P_1 \times P_2 \times \cdots \times P_n$ and $\mathbf{P}' = P_1' \times P_2' \times \cdots \times P_n'$ are partitions of the same rectangle, then $\mathbf{P}'$ is a re nement of $\mathbf{P}$ if P_i' is a re nement of P_i, $1 \le i \le n$.

3.7 Riemann Sum in $\mathbb{R}^n$

Suppose that f is a real-valued function de ned on a rectangle R in $\mathbb{R}^n$, $\mathbf{P} = \{R_1, R_2, \ldots, R_k\}$ is a partition of R.

Let $\mathbf{X}_j$ is an arbitrary point in R_j, $1 \le j \le k$.
Then

$$\sigma = \sum_{j=1}^{k} f(\mathbf{X}_j) V(R_j)$$

is a Riemann sum of f over $\mathbf{P}$.

Since $\mathbf{X}_j$ can be chosen arbitrarily in R_j, there are in nitely many Riemann sums for a given function f over any partition $\mathbf{P}$ of R.

3.8 Riemann Integral in $\mathbb{R}^n$

: Let f be a real-valued function de ned on a rectangle R in $\mathbb{R}^n$.

We say that f is Riemann integrable on R if there is a number L with the following property:

For every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|\sigma - L| < \varepsilon.$$

If σ is any Riemann sum of f over a partition $\mathbf{P}$ of R such that $\|\mathbf{P}\| < \delta$.

In this case, we say that L is the Riemann integral of f over R, and write

$$\int_R f(\mathbf{X})\, d\mathbf{X} = L.$$

Remarks: The integral $\int_R f(\mathbf{X}) d\mathbf{X}$ is also written as

$$\int_R f(x, y)\, d(x, y) \quad (n = 2), \qquad \int_R f(x, y, z)\, d(x, y, z) \quad (n = 3),$$

or

$$\int_R f(x_1, x_2, \dots, x_n)\, d(x_1, x_2, \dots, x_n) \quad (n \text{ arbitrary}).$$

Here $d\mathbf{X}$ does not stand for the di erential of $\mathbf{X}$.

It merely identi es $x_1, x_2, \dots, x_n$, the components of $\mathbf{X}$, as the variables of integration. $\int$ To avoid this minor inconsistency, some authors write simply $\int_R f$ rather than $\int_R f(\mathbf{X})\, d\mathbf{X}$.

As in the case where $n = 1$, we will say simply integrable or integral when we mean Riemann integrable or Riemann integral. If $n \geq 2$, we call the integral of above de nition a multiple integral ; for $n = 2$ and $n = 3$ we also call them double and triple integrals, respectively. When we wish to distinguish between multiple integrals and the integral we studied in Chapter ($n = 1$), we will call the latter an ordinary integral.

Example: Find $\int_R f(x, y)\, d(x, y)$, where

$$R = [a, b] \times [c, d]$$

and

$$f(x, y) = x + y.$$

Solution: Let P_1 and P_2 be partitions of $[a, b]$ and $[c, d]$; thus,

$$P_1 : a = x_0 < x_1 < \cdots < x_r = b$$

and

$$P_2 : c = y_0 < y_1 < \cdots < y_s = d.$$

A typical Riemann sum of f over $\mathbf{P} = P_1 \times P_2$ is given by

$$\sigma = \sum_{i=1} \sum_{j=1} (\xi_{ij} + \eta_{ij})(x_i - x_{i-1})(y_j - y_{j-1}), \tag{3.12}$$

$$\text{where} \quad x_{i-1} \le \xi_{ij} \le x_i \quad \text{and} \quad y_{j-1} \le \eta_{ij} \le y_j. \tag{3.13}$$

The midpoints of $[x_{i-1}, x_i]$ and $[y_{j-1}, y_j]$ are

$$\overline{x}_i = \frac{x_i + x_{i-1}}{2} \quad \text{and} \quad \overline{y}_j = \frac{y_j + y_{j-1}}{2}, \tag{3.14}$$

and (3.13) implies that

$$|\xi_{ij} - \overline{x}_i| \le \frac{x_i - x_{i-1}}{2} \le \frac{\|P_1\|}{2} \le \frac{\|\mathbf{P}\|}{2} \tag{3.15}$$

$$|\eta_{ij} - \overline{y}_j| \le \frac{y_j - y_{j-1}}{2} \le \frac{\|P_2\|}{2} \le \frac{\|\mathbf{P}\|}{2}. \tag{3.16}$$

Now we rewrite (3.12) as

$$\begin{aligned} \sigma &= \sum_{i=1}^{r} \sum_{j=1}^{s} (\overline{x}_i + \overline{y}_j)(x_i - x_{i-1})(y_j - y_{j-1}) \\ &+ \sum_{i=1}^{r} \sum_{j=1}^{s} \left[(\xi_{ij} - \overline{x}_i) + (\eta_{ij} - \overline{y}_j) \right] (x_i - x_{i-1})(y_j - y_{j-1}). \end{aligned} \tag{3.17}$$

To find $\int_R f(x, y)\, d(x, y)$ from (3.17), we recall that

$$\sum_{i=1} (x_i - x_{i-1}) = b - a, \qquad \sum_{j=1} (y_j - y_{j-1}) = d - c \tag{3.18}$$

and

$$\sum_{i=1} (x_i^2 - x_{i-1}^2) = b^2 - a^2, \qquad \sum_{j=1} (y_j^2 - y_{j-1}^2) = d^2 - c^2. \tag{3.19}$$

Because of (3.15) and (3.16) the absolute value of the second sum in (3.17) does not exceed

$$\begin{aligned} \|\mathbf{P}\| \sum_{j=1} \sum_{j=1} (x_i - x_{i-1})(y_j - y_{j-1}) &= \|\mathbf{P}\| \left[\sum_{i=1} (x_i - x_{i-1}) \right] \left[\sum_{j=1} (y_j - y_{j-1}) \right] \\ &= \|\mathbf{P}\|(b - a)(d - c) \end{aligned}$$

(see (3.18)), so (3.17) implies that

$$\left|\sigma - \sum_{i=1}^{r}\sum_{j=1}^{s}(\bar{x}_i + \bar{y}_j)(x_i - x_{i-1})(y_j - y_{j-1})\right| \le \|\mathbf{P}\|(b-a)(d-c). \tag{3.20}$$

It now follows that

$$\begin{aligned}
\sum_{i=1}^{r}\sum_{j=1}^{s} \bar{x}_i(x_i - x_{i-1})(y_j - y_{j-1}) &= \left[\sum_{i=1}^{r} \bar{x}_i(x_i - x_{i-1})\right]\left[\sum_{j=1}^{s}(y_j - y_{j-1})\right] \\
&= (d-c)\sum_{i=1}^{r} \bar{x}_i(x_i - x_{i-1}) \qquad \text{(from (3.18))} \\
&= \frac{d-c}{2}\sum_{i=1}^{r}(x_i^2 - x_{i-1}^2) \qquad \text{(from (3.14))} \\
&= \frac{d-c}{2}(b^2 - a^2) \qquad \text{(from (3.19))}.
\end{aligned}$$

Similarly,

$$\sum_{i=1}^{r}\sum_{j=1}^{s} \bar{y}_j(x_i - x_{i-1})(y_j - y_{j-1}) = \frac{b-a}{2}(d^2 - c^2).$$

Therefore, (3.20) can be written as

$$\left|\sigma - \frac{d-c}{2}(b^2 - a^2) - \frac{b-a}{2}(d^2 - c^2)\right| \le \|\mathbf{P}\|(b-a)(d-c).$$

Since the right side can be made as small as we wish by choosing $\|\mathbf{P}\|$ sufficiently small,

$$\int_R (x+y)\, d(x,y) = \frac{1}{2}\left[(d-c)(b^2 - a^2) + (b-a)(d^2 - c^2)\right].$$

Theorem: If f is unbounded on the nondegenerate rectangle R in $\mathbb{R}^n$, then f is not integrable on R.

Proof: We will show that if f is unbounded on R, $\mathbf{P} = \{R_1, R_2, \ldots, R_k\}$ is any partition of R, and $M > 0$, then there are Riemann sums σ and σ' of f over $\mathbf{P}$ such that

$$|\sigma - \sigma'| \ge M. \tag{3.21}$$

This implies that f cannot satisfy definition of Riemann integral. (Why?)

Let

$$\sigma = \sum_{j=1}^{k} f(\mathbf{X}_j)V(R_j)$$

be a Riemann sum of f over $\mathbf{P}$. Let

$$\sigma = \sum_{j=1}^{k} f(\mathbf{X}_j)V(R_j)$$

be a Riemann sum of f over $\mathbf{P}$.

There must be an integer i in $\{1, 2, \ldots, k\}$ such that

$$|f(\mathbf{X}) - f(\mathbf{X}_i)| \geq \frac{M}{V(R_i)} \tag{3.22}$$

for some $\mathbf{X}$ in R_i, because if this were not so, we would have

$$|f(\mathbf{X}) - f(\mathbf{X}_j)| < \frac{M}{V(R_j)}, \quad \mathbf{X} \in R_j, \quad 1 \leq j \leq k.$$

If this is so, then

$$\begin{aligned} |f(\mathbf{X})| &= |f(\mathbf{X}_j) + f(\mathbf{X}) - f(\mathbf{X}_j)| \leq |f(\mathbf{X}_j)| + |f(\mathbf{X}) - f(\mathbf{X}_j)| \\ &\leq |f(\mathbf{X}_j)| + \frac{M}{V(R_j)}, \quad \mathbf{X} \in R_j, \quad 1 \leq j \leq k. \end{aligned}$$

However, this implies that

$$|f(\mathbf{X})| \leq \max |f(\mathbf{X}_j)| + \frac{M}{V(R_j)} 1 \leq j \leq k, \quad \mathbf{X} \in R,$$

which contradicts the assumption that f is unbounded on R.

Now suppose that $\mathbf{X}$ satis es (3.22).

Consider the Riemann sum

$$\sigma' = \sum_{j=1}^{k} f(\mathbf{X}'_j) V(R_j)$$

over the same partition $\mathbf{P}$, where

$$\mathbf{X}'_j = \begin{cases} \mathbf{X}_j, & j \neq i, \\ \mathbf{X}, & j = i. \end{cases}$$

Since

$$|\sigma - \sigma'| = |f(\mathbf{X}) - f(\mathbf{X}_i)| V(R_i),$$

(3.22) implies (3.21).

3.9 Upper and Lower Integrals

If f is bounded on a rectangle R in $\mathbb{R}^n$ and $\mathbf{P} = \{R_1, R_2, \ldots, R_k\}$ is a partition of R.

Let

$$M_j = \sup_{\mathbf{X} \in R_j} f(\mathbf{X}), \quad m_j = \inf_{\mathbf{X} \in R_j} f(\mathbf{X}).$$

The upper sum of f over $\mathbf{P}$ is

$$S(\mathbf{P}) = \sum_{j=1}^{k} M_j V(R_j).$$

The upper integral of f over R, denoted by

$$\overline{\int_R} f(\mathbf{X})\, d\mathbf{X},$$

is the in mum of all upper sums.

Upper and Lower Integrals: The lower sum of f over $\mathbf{P}$ is

$$s(\mathbf{P}) = \sum_{j=1}^{k} m_j V(R_j).$$

The lower integral of f over R, denoted by

$$\underline{\int_R} f(\mathbf{X})\, d\mathbf{X},$$

is the supremum of all lower sums.

Theorem: Let f be bounded on a rectangle R and let $\mathbf{P}$ be a partition of R.

Then

1. The upper sum $S(\mathbf{P})$ of f over $\mathbf{P}$ is the supremum of the set of all Riemann sums of f over $\mathbf{P}$.
2. The lower sum $s(\mathbf{P})$ of f over $\mathbf{P}$ is the in mum of the set of all Riemann sums of f over $\mathbf{P}$.

Remarks: If

$$m \le f(\mathbf{X}) \le M \quad \text{for } \mathbf{X} \text{ in } R,$$

then

$$mV(R) \le s(\mathbf{P}) \le S(\mathbf{P}) \le MV(R);$$

therefore, $\overline{\int_R} f(\mathbf{X})\, d\mathbf{X}$ and $\underline{\int_R} f(\mathbf{X})\, d\mathbf{X}$ exist, are unique, and satisfy the inequalities

$$mV(R) \le \overline{\int_R} f(\mathbf{X})\, d\mathbf{X} \le MV(R)$$

and

$$mV(R) \le \underline{\int_R} f(\mathbf{X})\, d\mathbf{X} \le MV(R).$$

Remarks: The upper and lower integrals are also written as

$$\overline{\int_R} f(x, y)\, d(x, y) \quad \text{and} \quad \underline{\int_R} f(x, y)\, d(x, y) \quad (n = 2),$$

$$\overline{\int}_R f(x, y, z)\, d(x, y, z) \quad \text{and} \quad \underline{\int}_R f(x, y, z)\, d(x, y, z) \quad (n = 3),$$

or

$$\overline{\int}_R f(x_1, x_2, \dots, x_n)\, d(x_1, x_2, \dots, x_n)$$

and

$$\underline{\int}_R f(x_1, x_2, \dots, x_n)\, d(x_1, x_2, \dots, x_n) \qquad (n \text{ arbitrary}).$$

Example: Find $\underline{\int}_R f(x, y)\, d(x, y)$ and $\overline{\int}_R f(x, y)\, d(x, y)$, with $R = [a, b] \times [c, d]$ and

$$f(x, y) = x + y.$$

Solution: Let P_1 and P_2 be partitions of $[a, b]$ and $[c, d]$; thus,

$$P_1 : a = x_0 < x_1 < \cdots < x_r = b$$

$$P_2 : c = y_0 < y_1 < \cdots < y_s = d.$$

The maximum and minimum values of f on the rectangle $[x_{i-1}, x_i] \times [y_{j-1}, y_j]$ are $x_i + y_j$ and $x_{i-1} + y_{j-1}$, respectively.

Therefore,

$$S(\mathbf{P}) = \sum_{i=1} \sum_{j=1} (x_i + y_j)(x_i - x_{i-1})(y_j - y_{j-1}) \tag{3.23}$$

$$s(\mathbf{P}) = \sum_{i=1} \sum_{j=1} (x_{i-1} + y_{j-1})(x_i - x_{i-1})(y_j - y_{j-1}). \tag{3.24}$$

By substituting

$$x_i + y_j = \frac{1}{2}[(x_i + x_{i-1}) + (y_j + y_{j-1}) + (x_i - x_{i-1}) + (y_j - y_{j-1})]$$

into (3.23). We nd that

$$S(\mathbf{P}) = \frac{1}{2}(\Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4), \tag{3.25}$$

where

$$\begin{aligned}
\Sigma_1 &= \sum_{i=1}^{r} (x_i^2 - x_{i-1}^2) \sum_{j=1}^{s} (y_j - y_{j-1}) &&= (b^2 - a^2)(d - c),\\
\Sigma_2 &= \sum_{i=1}^{r} (x_i - x_{i-1}) \sum_{j=1}^{s} (y_j^2 - y_{j-1}^2) &&= (b - a)(d^2 - c^2),\\
\Sigma_3 &= \sum_{i=1}^{r} (x_i - x_{i-1})^2 \sum_{j=1}^{s} (y_j - y_{j-1}) &&\le \|\mathbf{P}\|(b - a)(d - c),\\
\Sigma_4 &= \sum_{i=1}^{r} (x_i - x_{i-1}) \sum_{j=1}^{s} (y_j - y_{j-1})^2 &&\le \|\mathbf{P}\|(b - a)(d - c).
\end{aligned}$$

Substituting these four results into (3.25) shows that

$$I < S(\mathbf{P}) < I + \|\mathbf{P}\|(b-a)(d-c),$$

where

$$I = \frac{(d-c)(b^2-a^2) + (b-a)(d^2-c^2)}{2}.$$

From this, we see that

$$\overline{\int_R}(x+y)\,d(x,y) = I.$$

After substituting

$$x_{i-1} + y_{j-1} = \frac{1}{2}[(x_i + x_{i-1}) + (y_j + y_{j-1}) - (x_i - x_{i-1}) - (y_j - y_{j-1})]$$

into (3.24), a similar argument shows that

$$I - \|\mathbf{P}\|(b-a)(d-c) < s(\mathbf{P}) < I.$$

So

$$\underline{\int_R}(x+y)\,d(x,y) = I.$$

Theorem: Suppose that $|f(\mathbf{X})| \le M$ if $\mathbf{X}$ is in the rectangle

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

Let $\mathbf{P} = P_1 \times P_2 \times \cdots \times P_n$ and $\mathbf{P}' = P_1' \times P_2' \times \cdots \times P_n'$ be partitions of R, where P_j' is obtained by adding r_j partition points to P_j, $1 \le j \le n$. Then

$$S(\mathbf{P}) \ge S(\mathbf{P}') \ge S(\mathbf{P}) - 2MV(R)\left(\sum_{j=1}^{n} \frac{r_j}{b_j - a_j}\right)\|\mathbf{P}\| \tag{3.26}$$

and

$$s(\mathbf{P}) \le s(\mathbf{P}') \le s(\mathbf{P}) + 2MV(R)\left(\sum_{j=1}^{n} \frac{r_j}{b_j - a_j}\right)\|\mathbf{P}\|. \tag{3.27}$$

Theorem: If f is bounded on a rectangle R, then

$$\underline{\int_R} f(\mathbf{X})\,d\mathbf{X} \le \overline{\int_R} f(\mathbf{X})\,d\mathbf{X}.$$

Theorem: If f is integrable on a rectangle R, then

$$\underline{\int_R} f(\mathbf{X})\,d\mathbf{X} = \overline{\int_R} f(\mathbf{X})\,d\mathbf{X} = \int_R f(\mathbf{X})\,d\mathbf{X}.$$

Theorem: If f is bounded on a rectangle R and $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\overline{\int_R} f(\mathbf{X})\,d\mathbf{X} \le S(\mathbf{P}) < \overline{\int_R} f(\mathbf{X})\,d\mathbf{X} + \varepsilon$$

and

$$\underline{\int_R} f(\mathbf{X})\,d\mathbf{X} \ge s(\mathbf{P}) > \underline{\int_R} f(\mathbf{X})\,d\mathbf{X} - \varepsilon$$

if $\|\mathbf{P}\| < \delta$.

Theorem: A bounded function f is integrable on a rectangle R if and only if

$$\underline{\int_R} f(\mathbf{X})\,d\mathbf{X} = \overline{\int_R} f(\mathbf{X})\,d\mathbf{X}.$$

Theorem: If f is bounded on a rectangle R, then f is integrable on R if and only if for every $\varepsilon > 0$ there is a partition $\mathbf{P}$ of R such that

$$S(\mathbf{P}) - s(\mathbf{P}) < \varepsilon.$$

Theorem: If f is bounded on a rectangle R and

$$\underline{\int_R} f(\mathbf{X})\,d\mathbf{X} = \overline{\int_R} f(\mathbf{X})\,d\mathbf{X} = L,$$

then f is integrable on R, and

$$\int_R f(\mathbf{X})\,d\mathbf{X} = L.$$

Theorem: If f is continuous on a rectangle R in $\mathbb{R}^n$, then f is integrable on R.

3.10 Sets with Zero Content

A subset E of $\mathbb{R}^n$ has zero content if for each $\varepsilon > 0$ there is a nite set of rectangles $T_1, T_2, \ldots, T_m$ such that

$$E \subset \bigcup_{j=1}^{m} T_j \tag{3.28}$$

and

$$\sum_{j=1}^{m} V(T_j) < \varepsilon. \tag{3.29}$$

Example: Since the empty set is contained in every rectangle, the empty set has zero content.

If E consists of nitely many points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m$, then $\mathbf{X}_j$ can be enclosed in a rectangle T_j such that

$$V(T_j) < \frac{\varepsilon}{m}, \quad 1 \le j \le m.$$

Then $E \subset \bigcup_{j=1}^{m} T_j$ and $\sum_{j=1}^{m} V(T_j) < \varepsilon$ hold, so E has zero content.

Example: Any bounded set E with only nitely many limit points has zero content.

To see this, we rst observe that if E has no limit points, then it must be nite, by the Bolzano Weierstrass theorem, and therefore must have zero content.

Now suppose that the limit points of E are $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m$. Let $R_1, R_2, \ldots, R_m$ be rectangles such that $\mathbf{X}_i \in R_i^0$ and

$$V(R_i) < \frac{\varepsilon}{2m}, \quad 1 \le i \le m. \tag{3.30}$$

The set of points of E that are not in $\cup_{j=1}^{m} R_j$ has no limit points (why?) and, being bounded, must be nite (again by the Bolzano Weierstrass theorem).

If this set contains p points, then it can be covered by rectangles $R'_1, R'_2, \ldots, R'_p$ with

$$V(R'_j) < \frac{\varepsilon}{2p}, \quad 1 \le j \le p. \tag{3.31}$$

Now,

$$E \subset \left(\bigcup_{i=1}^{m} R_i \right) \cup \left(\bigcup_{j=1}^{p} R'_j \right).$$

From (3.30) and (3.31),

$$\sum_{i=1}^{m} V(R_i) + \sum_{j=1}^{p} V(R'_j) < \varepsilon.$$

Example: If f is continuous on $[a, b]$, then the curve

$$y = f(x), \quad a \le x \le b \tag{3.32}$$

(that is, the set $\{(x, y) : y = f(x), a \le x \le b\}$), has zero content in $\mathbb{R}^2$.

Lemma: The union of finitely many sets with zero content has zero content.

Theorem: Suppose that f is bounded on a rectangle

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \tag{3.33}$$

and continuous except on a subset E of R with zero content. Then f is integrable on R.

Example: The function

$$f(x, y) = \begin{cases} x + y, & 0 \le x < y \le 1, \\ 5, & 0 \le y \le x \le 1, \end{cases}$$

is continuous on $R = [0, 1] \times [0, 1]$ except on the line segment

$$y = x, \quad 0 \le x \le 1$$

Since the line segment has zero content, f is integrable on R.

3.11 Integral Over Bounded Set

Suppose that f is bounded on a bounded subset of S of $\mathbb{R}^n$. Let

$$f_S(\mathbf{X}) = \begin{cases} f(\mathbf{X}), & \mathbf{X} \in S, \\ 0, & \mathbf{X} \notin S. \end{cases} \tag{3.34}$$

Let R be a rectangle containing S. Then the integral of f over S is defined to be

$$\int_S f(\mathbf{X})\, d\mathbf{X} = \int_R f_S(\mathbf{X})\, d\mathbf{X}$$

if $\int_R f_S(\mathbf{X})\, d\mathbf{X}$ exists.

Area and volume as integrals: If S is a bounded subset of $\mathbb{R}^n$ and the integral $\int_S d\mathbf{X}$ (with integrand $f \equiv 1$) exists.

We call $\int_S d\mathbf{X}$ the content (also, area if $n = 2$ or volume if $n = 3$) of S, and denote it by $V(S)$.

Thus,

$$V(S) = \int_S d\mathbf{X}.$$

Theorem: Suppose that f is bounded on a bounded set S and continuous except on a subset E of S with zero content.

Suppose also that ∂S has zero content. Then f is integrable on S.

3.12 Di erentiable Surfaces

A di erentiable surface S in $\mathbb{R}^n$ $(n > 1)$ is the image of a compact subset D of $\mathbb{R}^m$, where $m < n$, under a continuously di erentiable transformation $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^n$. If $m = 1$, S is also called a di erentiable curve.

Example: The circle

$$\{(x, y) : x^2 + y^2 = 9\}$$

is a di erentiable curve in $\mathbb{R}^2$.

Since it is the image of $D = [0, 2\pi]$ under the continuously di erentiable transformation $\mathbf{G} : \mathbb{R} \to \mathbb{R}^2$ de ned by

$$\mathbf{X} = \mathbf{G}(\vartheta) = \begin{bmatrix} 3\cos\vartheta \\ 3\sin\vartheta \end{bmatrix}.$$

Example: The sphere

$$\{(x, y, z) : x^2 + y^2 + z^2 = 4\}$$

is a di erentiable surface in $\mathbb{R}^3$.

Since it is the image of

$$D = \{(\vartheta, \phi) : 0 \le \vartheta \le 2\pi, -\pi/2 \le \phi \le \pi/2\}$$

under the continuously di erentiable transformation $\mathbf{G} : \mathbb{R}^2 \to \mathbb{R}^3$ de ned by

$$\mathbf{X} = \mathbf{G}(\vartheta, \phi) = \begin{bmatrix} 2\cos\vartheta\cos\phi \\ 2\sin\vartheta\cos\phi \\ 2\sin\phi \end{bmatrix}.$$

Theorem: A di erentiable surface in $\mathbb{R}^n$ has zero content.

Let S, D, and $\mathbf{G}$ be as in De nition ??. From Lemma ??, there is a constant M such that

$$|\mathbf{G}(\mathbf{X}) - \mathbf{G}(\mathbf{Y})| \le M|\mathbf{X} - \mathbf{Y}| \quad \text{if} \quad \mathbf{X}, \mathbf{Y} \in D. \tag{3.35}$$

Since D is bounded, D is contained in a cube

$$C = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_m, b_m],$$

where

$$b_i - a_i = L, \quad 1 \le i \le m.$$

Suppose that we partition C into N^m smaller cubes by partitioning each of the intervals $[a_i, b_i]$ into N equal subintervals. Let $R_1, R_2, \dots, R_k$ be the smaller cubes so produced that contain points of D, and select points $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$ such that $\mathbf{X}_i \in D \cap R_i$, $1 \le i \le k$. If $\mathbf{Y} \in D \cap R_i$, then (3.35) implies that

$$|\mathbf{G}(\mathbf{X}_i) - \mathbf{G}(\mathbf{Y})| \le M|\mathbf{X}_i - \mathbf{Y}|. \tag{3.36}$$

Since $\mathbf{X}_i$ and $\mathbf{Y}$ are both in the cube R_i with edge length L/N,

$$|\mathbf{X}_i - \mathbf{Y}| \le \frac{L\sqrt{m}}{N}.$$

This and (3.36) imply that

$$|\mathbf{G}(\mathbf{X}_i) - \mathbf{G}(\mathbf{Y})| \le \frac{ML\sqrt{m}}{N},$$

which in turn implies that $\mathbf{G}(\mathbf{Y})$ lies in a cube $\tilde{R}_i$ in $\mathbb{R}^n$ centered at $\mathbf{G}(\mathbf{X}_i)$, with sides of length $2ML\sqrt{m}/N$. Now

$$\sum_{i=1}^{k} V(\tilde{R}_i) = k\left(\frac{2ML\sqrt{m}}{N}\right)^n \le N^m\left(\frac{2ML\sqrt{m}}{N}\right)^n = (2ML\sqrt{m})^n N^{m-n}.$$

Since $n > m$, we can make the sum on the left arbitrarily small by taking N sufficiently large. Therefore, S has zero content.

Theorem: Suppose that S is a bounded set in $\mathbb{R}^n$, with boundary consisting of a finite number of differentiable surfaces.

Let f be bounded on S and continuous except on a set of zero content. Then f is integrable on S.

Example: Let

$$S = \{(x, y) : x^2 + y^2 = 1,\ x \ge 0\}.$$

The set S is bounded by a semicircle and a line segment, both differentiable curves in $\mathbb{R}^2$.

Let

$$f(x, y) = \begin{cases} (1 - x^2 - y^2)^{1/2}, & (x, y) \in S,\ y \ge 0, \\ -(1 - x^2 - y^2)^{1/2}, & (x, y) \in S,\ y < 0. \end{cases}$$

Then f is continuous on S except on the line segment

$$y = 0, \quad 0 \le x < 1,$$

which has zero content.

Hence, from the theorem we just stated implies that f is integrable on S.

Theorem: If f and g are integrable on S, then so is $f + g$, and

$$\int_S (f + g)(\mathbf{X})\, d\mathbf{X} = \int_S f(\mathbf{X})\, d\mathbf{X} + \int_S g(\mathbf{X})\, d\mathbf{X}.$$

Theorem: If f is integrable on S and c is a constant, then cf is integrable on S, and

$$\int_S (cf)(\mathbf{X})\, d\mathbf{X} = c\int_S f(\mathbf{X})\, d\mathbf{X}.$$

Theorem: If f and g are integrable on S and $f(\mathbf{X}) \le g(\mathbf{X})$ for $\mathbf{X}$ in S, then

$$\int_S f(\mathbf{X})\, d\mathbf{X} \le \int_S g(\mathbf{X})\, d\mathbf{X}.$$

Theorem: If f is integrable on S, then so is $|f|$, and

$$\left|\int_S f(\mathbf{X})\, d\mathbf{X}\right| \le \int_S |f(\mathbf{X})|\, d\mathbf{X}.$$

Theorem: If f and g are integrable on S, then so is the product fg.

Theorem: Suppose that u is continuous and v is integrable and nonnegative on a rectangle R. Then

$$\int_R u(\mathbf{X})v(\mathbf{X})\, d\mathbf{X} = u(\mathbf{X}_0)\int_R v(\mathbf{X})\, d\mathbf{X}$$

for some $\mathbf{X}_0$ in R.

Theorem: Suppose that S is contained in a bounded set T and f is integrable on S.

Then f_S is integrable on T, and

$$\int_T f_S(\mathbf{X})\, d\mathbf{X} = \int_S f(\mathbf{X})\, d\mathbf{X}.$$

Theorem: If f is integrable on disjoint sets S_1 and S_2, then f is integrable on $S_1 \cup S_2$, and

$$\int_{S_1\cup S_2} f(\mathbf{X})\, d\mathbf{X} = \int_{S_1} f(\mathbf{X})\, d\mathbf{X} + \int_{S_2} f(\mathbf{X})\, d\mathbf{X}. \tag{3.37}$$

Theorem: Suppose that f is integrable on sets S_1 and S_2 such that $S_1 \cap S_2$ has zero content. Then f is integrable on $S_1 \cup S_2$, and

$$\int_{S_1\cup S_2} f(\mathbf{X})\, d\mathbf{X} = \int_{S_1} f(\mathbf{X})\, d\mathbf{X} + \int_{S_2} f(\mathbf{X})\, d\mathbf{X}.$$

Example: Let

$$\begin{aligned} S_1 &= \{(x, y) : 0 \le x \le 1,\ 0 \le y \le 1 + x\} \\ S_2 &= \{(x, y) : -1 \le x \le 0,\ 0 \le y \le 1 - x\} \end{aligned}$$

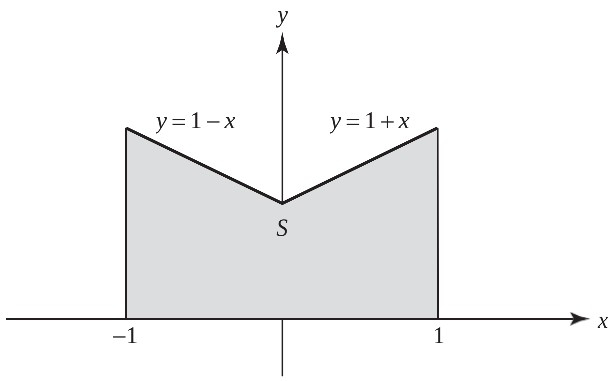

Figure 3.4: S_1 and S_2

Then

$$S_1 \cap S_2 = \{(0, y) : 0 \le y \le 1\}$$

has zero content.

Hence, by using corollary implies that if f is integrable on S_1 and S_2, then f is also integrable over

$$S = S_1 \cup S_2 = \{(x, y) : -1 \le x \le 1,\ 0 \le y \le 1 + |x|\}$$

and

$$\int_{S_1 \cup S_2} f(\mathbf{X})\, d\mathbf{X} = \int_{S_1} f(\mathbf{X})\, d\mathbf{X} + \int_{S_2} f(\mathbf{X})\, d\mathbf{X}.$$

3.13 Iterated Integrals

Let us rst assume that f is continuous on $R = [a, b] \times [c, d]$.

Then, for each y in $[c, d]$, $f(x, y)$ is continuous with respect to x on $[a, b]$, so the integral

$$F(y) = \int_a^b f(x, y)\, dx$$

exists.

Moreover, the uniform continuity of f on R implies that F is continuous and therefore integrable on $[c, d]$.

We say that

$$I_1 = \int_c^d F(y)\, dy = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy$$

is an iterated integral of f over R.

Iterated integrals: We will usually write it as

$$I_1 = \int_c^d dy \int_a^b f(x, y)\, dx.$$

Another iterated integral can be de ned by writing

$$G(x) = \int_c^d f(x, y)\, dy, \quad a \le x \le b,$$

De ning

$$I_2 = \int_a^b G(x)\, dx = \int_a^b \left(\int_c^d f(x, y)\, dy \right) dx,$$

which we usually write as

$$I_2 = \int_a^b dx \int_c^d f(x, y)\, dy.$$

Example: Let

$$f(x, y) = x + y$$

and $R = [0, 1] \times [1, 2]$. Then

$$F(y) = \int_0^1 f(x, y)\, dx = \int_0^1 (x + y)\, dx = \frac{1}{2} + y$$

$$I_1 = \int_1^2 F(y)\, dy = \int_1^2 \left(\frac{1}{2} + y \right) dy = \left(\frac{y}{2} + \frac{y^2}{2} \right) \Big|_1^2 = 2.$$

Also,

$$G(x) = \int_1^2 (x + y)\, dy = \left(xy + \frac{y^2}{2} \right) \Big|_{y=1}^2 = x + \frac{3}{2},$$

$$I_2 = \int_0^1 G(x)\, dx = \int_0^1 \left(x + \frac{3}{2} \right) dx = \left(\frac{x^2}{2} + \frac{3x}{2} \right) \Big|_0^1 = 2.$$

Theorem: Suppose that f is integrable on $R = [a, b] \times [c, d]$ and

$$F(y) = \int_a^b f(x, y)\, dx$$

exists for each y in $[c, d]$.

Then F is integrable on $[c, d]$, and

$$\int_c^d F(y)\, dy = \int_R f(x, y)\, d(x, y); \tag{3.38}$$

that is,

$$\int_c^d dy \int_a^b f(x, y)\, dx = \int_R f(x, y)\, d(x, y). \tag{3.39}$$

Theorem: If f is integrable on $[a, b] \times [c, d]$, then

$$\int_a^b dx \int_c^d f(x, y)\,dy = \int_c^d dy \int_a^b f(x, y)\,dx,$$

provided that $\int_c^d f(x, y)\,dy$ exists for $a \le x \le b$ and $\int_a^b f(x, y)\,dx$ exists for $c \le y \le d$.

In particular, these hypotheses hold if f is continuous on $[a, b] \times [c, d]$.

Example: The function

$$f(x, y) = x + y$$

is continuous everywhere.

For example, let $R = [0, 1] \times [1, 2]$.

Then we have

$$\begin{aligned}
\int_R (x + y)\,d(x, y) &= \int_1^2 dy \int_0^1 (x + y)\,dx \\
&= \int_1^2 \left[\left(\frac{x^2}{2} + xy\right)\Big|_{x=0}^{1}\right] dy \\
&= \int_1^2 \left(\frac{1}{2} + y\right) dy = \left(\frac{y}{2} + \frac{y^2}{2}\right)\Big|_1^2 = 2.
\end{aligned}$$

Since f also satisfies the hypotheses of Fubini's Theorem with x and y interchanged, we can calculate the double integral from the iterated integral in which the integrations are performed in the opposite order.

Thus,

$$\begin{aligned}
\int_R (x + y)\,d(x, y) &= \int_0^1 dx \int_1^2 (x + y)\,dy \\
&= \int_0^1 \left[\left(xy + \frac{y^2}{2}\right)\Big|_{y=1}^{2}\right] dx \\
&= \int_0^1 \left(x + \frac{3}{2}\right) dx = \left(\frac{x^2}{2} + \frac{3x}{2}\right)\Big|_0^1 = 2.
\end{aligned}$$

Remark: If $\int_c^d dy \int_a^b f(x, y)\,dx$ exists then so does $\int_R f(x, y)\,d(x, y)$. However, this need not to be true.

Example: If f is defined on $R = [0, 1] \times [0, 1]$ by

$$f(x, y) = \begin{cases} 2xy & \text{if } y \text{ is rational,} \\ y & \text{if } y \text{ is irrational,} \end{cases}$$

then

$$\int_0^1 f(x, y)\,dx = y, \quad 0 \le y \le 1,$$

and
$$\int_0^1 dy \int_0^1 f(x,y)\,dx = \int_0^1 y\,dy = \frac{1}{2}.$$
However, f is not integrable on R.

Theorem: Let $I_1, I_2, \ldots, I_n$ be closed intervals and suppose that f is integrable on $R = I_1 \times I_2 \times \cdots \times I_n$.

Suppose that there is an integer p in $\{1, 2, \ldots, n-1\}$ such that
$$F_p(x_{p+1}, x_{p+2}, \ldots, x_n) = \int_{I_1 \times I_2 \times \cdots \times I_p} f(x_1, x_2, \ldots, x_n)\, d(x_1, x_2, \ldots, x_p)$$
exists for each $(x_{p+1}, x_{p+2}, \ldots, x_n)$ in $I_{p+1} \times I_{p+2} \times \cdots \times I_n$.

Then
$$\int_{I_{p+1} \times I_{p+2} \times \cdots \times I_n} F_p(x_{p+1}, x_{p+2}, \ldots, x_n)\, d(x_{p+1}, x_{p+2}, \ldots, x_n)$$
exists and equals $\int_R f(\mathbf{X})\, d\mathbf{X}$.

Theorem: Let $I_j = [a_j, b_j]$, $1 \le j \le n$, and suppose that f is integrable on $R = I_1 \times I_2 \times \cdots \times I_n$.

Suppose also that the integrals
$$F_p(x_{p+1}, \ldots, x_n) = \int_{I_1 \times I_2 \cdots \times I_p} f(\mathbf{X})\, d(x_1, x_2, \ldots, x_p), \quad 1 \le p \le n-1,$$
exist for all
$$(x_{p+1}, \ldots, x_n) \quad \text{in} \quad I_{p+1} \times \cdots \times I_n.$$
Then the iterated integral
$$\int_{a_n}^{b_n} dx_n \int_{a_{n-1}}^{b_{n-1}} dx_{n-1} \cdots \int_{a_2}^{b_2} dx_2 \int_{a_1}^{b_1} f(\mathbf{X})\, dx_1$$
exists and equals $\int_R f(\mathbf{X})\, d\mathbf{X}$.