Real Analysis Definitions, Theorems, and Proofs

A Glossary for Walter Rudin's *Principles of Mathematical Analysis*

Erin Griffin

August 3, 2017

Contents

Chapter 1 The Real and Complex Number Systems

Definitions

Empty Set/ Nonempty If A is any set (whose elements may be numbers or any other objects), we write $x \in A$ to indicate that x is a member (or an element) of A. If x is not a member of A, we write: $x \notin A$. The set which contains no elements will be called the *empty set.* If a set has at least one element, it is called *nonempty*

Proper If A and B are sets, and if every element of A is an element of B, we say that A is a subset of B, and we write $A \subset B$, or $B \supset A$. If, in addition, there is an element of B which is not in A, then A is said to be a *proper* subset of B. Note that $A \subset A$ for every set A. If $A \subset B$ and $B \subset A$, we write $A = B$.

Order Let S be a set. An *order* on S is a relation, denoted by $<$, with the following two properties:

a) If $x \in S$ and $y \in S$ then one and only one of the statements:

$$x < y \qquad x = y \qquad x > y$$

is true.

b) If $x, y, z \in S$, if $x < y$, $y < z$, then $x < z$.

Ordered Set An *ordered set* is a set S in which an order is defined.

Bounded Suppose S is an ordered set, and $E \subset S$. IF there exists a $\beta \in S$ such that $x \leqslant \beta$ for every $x \in E$, we say that E is *bounded above*, and call β an *upper bound* of E.

Lower bounds are defined in the same way.

Least Upper Bound/ Greatest Lower Bound Suppose S is an ordered set $E \subset S$, and E is bounded above. Suppose there exists an $\alpha \in S$ with the following properties:

a) α is an upper bound of E

b) If $\gamma < \alpha$, then γ is not an upper bound of E

Then α is called the *least upper bound of* E or the *supremum of* E, and we write

$$\alpha = \sup E.$$

The *greatest lower bound*, or *infimum*, of a set E which is bounded below is defined in the same manner: The statement

$$\alpha = \inf E$$

means that α is a lower bound of E and that no β with $\beta > \alpha$ is a lower bound of E.

Least-Upper-Bound Property An ordered set S is said to have the *least-upper-bound property* if the following is true:

If $E \subset S$, E is not empty, and E is bounded above, then $\sup E$ exists in S.

Field A *field* is a set F with two operations, called *addition* and *multiplication*, which satisfy the following so-called "field axioms":

1. Axioms for Addition
   (a) If $x \in F$ and $y \in F$, then $x + y \in F$.
   (b) Addition is commutative: $x + y = y + x$ for all $x, y \in F$.
   (c) Addition is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$
   (d) F contains an element 0 such that $0 + x = x$ for every $x \in F$.
   (e) To every $x \in F$ corresponds an element $-x \in F$ such that $x + (-x) = 0$
2. Axioms for Multiplication
   (a) If $x \in F$ and $y \in F$, then $xy \in F$.
   (b) Multiplication is commutative: $xy = yx$ for all $x, y \in F$.
   (c) Multiplication is associative: $(xy)z = x(yz)$ for all $x, y, z \in F$
   (d) F contains an element $1 \neq 0$ such that $1x = x$ for every $x \in F$.
   (e) If $x \in F$ and $x \neq 0$ then there exists and element $1/x \in F$ such that $x(1/x) = 1$.
3. The Distributive Law: $x(y + z) = xy + xz$ for all $x, y, z \in F$.

Ordered Field An *ordered field* is a field F which is also an ordered set, such that

a) $x + y < x + z$ if $x, y, z \in F$ and $y < z$

b) $xy > 0$ if $x, y \in F$ and $x, y > 0$.

If $x > 0$ we call x positive; if $x < 0$ we call x negative.

Extended Real Numbers The *extended real number system* consists of the real field $\mathbb{R}$ and two symbols $+\infty$, $-\infty$. We preserve the original order in $\mathbb{R}$ and define

$$-\infty < x < +\infty$$

for every $x \in \mathbb{R}$.

Complex Number A *complex number* is an ordered pair (a, b) of real numbers. "Ordered" means that (a, b) and (b, a) are regarded as distinct if $a \neq b$. Let $x = (a, b)$ and $y = (c, d)$ be two complex numbers. We write $x = y$ if and only if $a = c, b = d$. We define:

$$x + y = (a + c, b + d) \qquad xy = (ac - bd, ad + bc)$$

i $i = (0, 1) \in \mathbb{C}$

Conjugate If $a, b \in \mathbb{R}$ and $z = a + bi$, the the complex number $\overline{z} = a - bi$ is called the *conjugate* of z. The numbers a and b are the real part and imaginary part of z respectively. Note these as

$$a = \Re(z) \qquad b = \Im(z)$$

Absolute Value If $z \in \mathbb{C}$, its absolute value $|z|$ is the non-negative square root of $z\bar{z}$; that is $|z| = (z\bar{z})^{1/2}$.

Coordinates For each positive integer k, let $\mathbb{R}^k$ be the set of all ordered k-tuples

$$\mathbf{x} = (x_1, x_2, \ldots, x_k)$$

where $x_1, x_2, \ldots, x_k \in \mathbb{R}$, called the *coordinates* of $\mathbf{x}$. The elements of $\mathbb{R}^k$ are called points, or vectors, especially when $k > 1$. We shall denote vectors by boldfaced letters. If $\mathbf{y} = (y_1, y_2, \ldots, y_k)$, and if $\alpha \in R$, then addition and multiplication are defined:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k) \in \mathbb{R}^k \qquad \alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_k) \in \mathbb{R}^k$$

These operations make $\mathbb{R}^k$ into a *vector space over the real field.* The inner product is defined by:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{k} x_i y_i$$

and the *norm* of $\mathbf{x}$ by:

$$|\mathbf{x}| = (\mathbf{x} \cdot \mathbf{x})^{1/2} = \left(\sum_{i=1}^{k} x_i^2 \right)^{1/2}$$

Theorems

Theorem 1.11 Suppose S is an ordered set with the least-upper-bound property, $B \subset S$, B is not empty, and B is bounded below. Let L be the set of all lower bound of B. Then

$$\alpha = \sup L$$

exists in S and $\alpha = \inf B$. In particular, $\inf B$ exists in S.

Proposition 1.14 The axioms for addition imply the following statements.

a) If $x + y = x + z$ then $y = z$

b) If $x + y = x$ then $y = 0$

c) If $x + y = 0$, then $y = -x$

d) $-(-x) = x$

Proposition 1.15 The axioms for multiplication imply the following statements.

a) If $x \neq 0$ $xy = xz$ then $y = z$

b) If $x \neq 0$ $xy = x$ then $y = 1$

c) If $x \neq 0$ $xy = 1$, then $y = 1/x$

d) If $x \neq 0$ $1/(1/x) = x$

Proposition 1.16 The field axioms imply the following statements, for any $x, y, z \in F$

a) $0x = 0$

b) If $x \neq 0$ and $y \neq 0$ then $xy \neq 0$

c) $(-x)y = -(xy) = x(-y)$

d) $(-x)(-y) = xy$

Proposition 1.18 The following statements are true in every ordered field.

a) If $x > 0$ then $-x < 0$ and vice versa

b) If $x > 0$ and $y < z$ then $xy < xz$

c) If $x < 0$ and $y < z$ then $xy > xz$

d) If $x \neq 0$ then $x^2 > 0$. In particular, $1 > 0$

e) If $0 < x < y$ then $0 < 1/y < 1/x$.

Theorem 1.19 There exists an ordered field $\mathbb{R}$ which has the least-upper-bound property. Moreover $\mathbb{R}$ contains $\mathbb{Q}$ as a subfield.

Theorem 1.20 a (Archimedean Property) If $x \in \mathbb{R}$, $y \in \mathbb{R}$, and $x > 0$, then there is a positive integer n such that $nx > y$.

b ($\mathbb{Q}$ is dense in $\mathbb{R}$) If $x, y \in \mathbb{R}$ and $x < y$, then there exists a $p \in \mathbb{Q}$ such that $x < p < y$. In other words, between and two real numbers there is a rational one.

Theorem 1.21 For every real $x > 0$ and every integer $n > 0$ there is one and only one positive real y such that $y^n = x$. This number is written $\sqrt[n]{x}$.

Corollary If a and b are positive real numbers and n is a positive integer, then $(ab)^{1/n} = a^{1/n}b^{1/n}$.

Theorem 1.25 These definitions of addition and multiplication turn the set of all complex numbers into a field with $(0,0)$ and $(1,0$ in the role of 0 and 1.

Theorem 1.26 For any real numbers $a, b \in \mathbb{R}$ we have

$$(a,0) + (b,0) = (a+b,0) \qquad (a,0)(b,0) = (ab,0)$$

Theorem 1.28 $i^2 = -1$

Theorem 1.29 If a and b are real, then $(a,b) = a + bi$

Theorem 1.31 If z and w are complex, then

a $\overline{z+w} = \overline{z} + \overline{w}$

b $\overline{zw} = \overline{z} \cdot \overline{w}$

c $z + \overline{z} = 2\Re(z)$, $z - \overline{z} = 2i\Im(z)$

d $z\bar{z}$ is real and positive (except when $z = 0$.)

Theorem 1.33 Let z and w are complex. then

a $|z| > 0$

b $|\bar{z}| = |z|$

c $|zw| = |z||w|$

d $|\Re(z)| \leqslant |z|$

e $|z + w| \leqslant |z| + |w|$.

Theorem 1.35 (Schwarz Inequality) If $a_1, \ldots, a_n$ and $b_1, \ldots b_n$ are complex numbers, then

$$\left|\sum_{j=1}^{n} a_j \bar{b}_j\right|^2 \leqslant \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$

Theorem 1.37 Suppose $\mathbf{x}, \mathbf{y}, \mathbf{z} \in R^k$ and $\alpha \in \mathbb{R}$. Then

a $|\mathbf{x}| \geqslant 0$;

b $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = \mathbf{0}$

c $|\alpha \mathbf{x}| = |\alpha||\mathbf{x}|$

d $|\mathbf{x} \cdot \mathbf{y}| \leqslant |\mathbf{x}||\mathbf{y}|$

e $|\mathbf{x} + \mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|$

f $|\mathbf{x} - \mathbf{z}| \leqslant |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}|$

Chapter 2 Basic Topology

Definitions

Function Consider two sets A and B, whose elements may be any objects whatsoever, and suppose that with each element x of A there is associated, in some manner, an element of B, which we denote by $f(x)$. Then f is said to be a *function* from A to B (or a *mapping* of A into B). The set A is called the *domain* of f (we also say f is defined on A), and the elements $f(x)$ are called the *values* of f. The set of all values of f is called the *range* of f.

One-to-One, Onto Let A and B be two sets and let f be a mapping of A into B. If $E \subset A$, $f(E)$ is defined to be the set of all elements $f(x)$ for $x \in E$. We call $f(E)$ the *image* of E under f. In this notation, $f(A)$ is the range of f. It is clear that $f(A) \subset B$. If $f(A) = B$, we say that f maps *onto* B. If $E \subset B$, $f^{-1}(E)$ denotes the set of all $x \in A$ such that $f(x) \in E$. We call $f^{-1}(E)$ the *inverse image* of E under f. If $y \in B$, $f^{-1}(y)$ is the set of all $x \in A$ such that $f(x) = y$. If, for each $y \in B$, $f^{-1}(y)$ consists of at most one element of A, then f is said to be a one-to-one mapping of A into B. This may also be expressed as follows: f is a 1-1 mapping of A into B provided that $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$, $x_1, x_2 \in A$.

Correspondence/ Equivalent If there exists a 1-1 mapping of A *onto* B, we say that A and B can be put in 1-1 *correspondence*, or that A and B have the same *cardinal number*, or, briefly, that A and B are *equivalent*, and we write $A \sim B$. This relation has the following properties:

a) It is reflexive: $A \sim A$

b) It is symmetric: If $A \sim B$, then $B \sim A$.

c) It is transitive: If $A \sim B$ and $B \sim C$, then $A \sim C$.

Any relation with these three properties is called an *equivalence relation*.

Finite For any positive integer n, let J_n be the set whose elements are the integers $1, 2, \ldots, n$; let J be the set consisting of all positive integers. For any set A,

a) we say A is *finite* if $A \sim J_n$ for some n (the empty set is also considered to be finite).

b) A is infinite if A is not finite

c) A is *countable* if $A \sim J$

d) A is *uncountable* if A is neither finite nor countable

e) A is *at most countable* if A is finite or countable

Sequence A *sequence* is a function f defined on the set J of all positive integers. If $f(n) = x_n$, for $n \in J$, it is customary to denote the sequence f by the symbol $\{x_n\}$. The values of f are called the *terms* of the sequence. If A is a set and if $x_n \in A$ for all $n \in J$, then $\{x_n\}$ is said to be a *sequence in* A, or a *sequence of elements of* A.

Subsets/ Family of Sets Let A and Ω be sets, and suppose that with each element α of A there is associated a subset of Ω which we denote by E_α. The set whose elements are set E_α will be denoted by $\{E_\alpha\}$. We shall call these a collection of sets or *family of sets.*

Union (From above) The *union* of the sets E_α is defined to be the set S such that $x \in S$ if and only if $x \in E_\alpha$ for at least one $\alpha \in A$. We use notation

$$S = \bigcup_{\alpha \in A} E_\alpha$$

If A consists of integers, we write one of the two following:

$$S = \bigcup_{i=1}^{n} E_i \qquad \text{or} \qquad S = E_1 \cup E_2 \cup \cdots \cup E_n$$

If A is the set of all positive integers, the usual notation is:

$$S = \bigcup_{i=1}^{\infty} E_i$$

Intersection The *intersection* of the sets E_α is defined to be the set P such that $x \in P$ if and only if $x \in E_\alpha$ for every $\alpha \in A$. We use notation:

$$P = \bigcap_{\alpha \in A} E_\alpha \quad \text{or} \quad P = \bigcap_{i=1}^{n} E_i = E_1 \cap E_2 \cap \ldots E_n \quad \text{or} \quad P = \bigcap_{i=1}^{\infty} E_i$$

If $A \cap B \neq \varnothing$ the we say that A and B *intersect*; otherwise they are *disjoint.*

Metric Space A set X, whose elements we shall call *points*, is said to be a *metric space* if with any two points p and q of X there is associated a real number $d(p,q)$ called the *distance* from p to q, such that:

a) $d(p,q) > 0$ if $p \neq q$; $d(p,p) = 0$.

b) $d(p,q) = d(q,p)$

c) $d(p,q) \leqslant d(p,r) + d(r,q)$ for any $r \in X$.

Segment By the *segment* (a,b) we mean the set of all real numbers x such that $a < x < b$.

Interval By the *interval* $[a,b]$ we mean the set of all real numbers x such that $a \leqslant x \leqslant b$.

K-Cell If $a_i < b_i$ for $i = 1, \ldots, k$, the set of all points $\mathbf{x} = (x_1, \ldots, x_k)$ in $\mathbb{R}^k$ whose coordinates satisfy the inequalities $a_i \leqslant x_i \leqslant b_i$, $(1 \leqslant i \leqslant k)$ is called a *k-cell.*

Ball If $\mathbf{x} \in \mathbb{R}^k$ and $r > 0$, the open (or closed) *ball* B with center at $\mathbf{x}$ and radius r is defined to be the set of all $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y} - \mathbf{x}| < r$ (or $|\mathbf{y} - \mathbf{x}| \leqslant r$).

Convex We call a set $E \subset \mathbb{R}^k$ *convex* if $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$ whenever $\mathbf{x} \in E$, $y \in E$, and $0 < \lambda < 1$.

Neighborhood Let X be a metric space. A *neighborhood* of p is a set $N_r(p)$ consisting of all q such that $d(p,q) < r$, for some $r > 0$. The number r is called the *radius* of $N_r(p)$.

Limit Point A point p is a *limit point* of the set E if *every* neighborhood of p contains a point $q \neq p$ such that $q \in E$.

Isolated Point If $p \in E$ and p is not a limit point of E, then p is called an *isolated point* of E.

Closed E is *closed* if every limit point of E is a point of E.

Interior A point p is an *interior* point of E if there is a neighborhood N of p such that $N \subset E$.

Open E is *open* if every point of E is an interior point of E.

Complement The *complement* of E (denoted E^c) is the set of all points $p \in X$ such that $p \notin E$

Perfect E is *perfect* if E is closed and if every point of E is a limit point of E.

Bounded E is *bounded* if there is a real number M and a point $q \in X$ such that $d(p,q) < M$ for all $p \in E$.

Dense E is *dense* in X if every point of X is a limit point of E, or a point of E (or both).

Closure If X is a metric space, if $E \subset X$, and if E' denotes the set of all limit points of E in X, then the *closure* of E is the set $\overline{E} = E \cup E'$.

Open Cover By an *open cover* of a set E in a metric space X we mean a collection $\{G_\alpha\}$ of open subsets of X such that $E \subset \bigcup_\alpha G_\alpha$.

Compact A subset K of a metric space X is said to be *compact* if every open cover of K contains a *finite* subcover. More explicitly, the requirement is that if $\{G_\alpha\}$is an open cover of K, then there are finitely many indices $\alpha_1, \ldots, \alpha_n$ such that $K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$.

Separated Two subsets A and B of a metric space X are said to be *separated* if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty, i.e. if no point of A lies in the closure of B and no point of B lies in the closure of A

Connected A set $E \subset X$ is said to be *connected* if E is not a union of two nonempty separated sets.

Theorems

Theorem 2.8 Every infinite subset of a countable set A is countable.

Proof. Suppose $E \subset A$, and E is infinite. Arrange the elements x of A in a sequence $\{x_n\}$ of distinct elements. Construct a sequence $\{n_k\}$ as follows: Let n_1 be the smallest positive integer such that $x_{n_1} \in E$. Having chosen $n_1, \dots, n_{k-1}$ $(k = 2, 3, 4, \dots)$, let n_k be the smallest integer greater than n_{k-1} such that $x_{n_k} \in E$.

Putting $f(k) = x_{n_k}$ $(k = 1, 2, 3, \dots)$, we obtain a 1-1 correspondence between E and J, the set of all positive integers. □

Theorem 2.12 Let $\{E_n\}$, $n = 1, 2, 3, \dots$, be a sequence of countable sets, and put

$$S = \bigcup_{n=1}^{\infty} E_n$$

Then S is countable.

Proof. Let every set E_n be arranged in a sequence $\{x_{nk}\}$, $k = 1, 2, 3, \dots$, and consider the infinite array

$$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{array}$$

in which the elements of E_n form the nth row. The array contains all the elements of S. These elements can be arranged in a sequence going diagonal up and to the right:

$$x_{11};\ x_{21},\ x_{12};\ x_{31},\ x_{22},\ x_{13};\ x_{41},\ x_{32},\ x_{23},\ x_{14};\ \dots$$

If any two of the sets E_n have elements in common, these will appear more than once in the sequence above. Hence there is a subset R of the set of all positive integers such that $S \sim T$, which show that S is at most countable. Since $E_1 \subset S$, and E_1 is infinite, S is infinite, and thus countable. □

Corollary Suppose A is at most countable, and, for every $\alpha \in A$, B_α is at most countable. Put

$$T = \bigcup_{\alpha \in A} B_\alpha$$

Then T is at most countable.

Theorem 2.13 Let A be a countable set, and let B_n be the set of all n-tuples $(a_1, \dots, a_n)$, where $a_k \in A$, $(k = 1, 2, \dots, n)$, and the elements $a_1, \dots, a_n$ need not be distinct. Then B_n is countable.

Proof. That B_1 is countable is evident, since $B_1 = A$. Suppose B_{n-1} is countable . The elements of B_n are of the form (b, a) where $b \in B_{n-1}$ and $a \in A$. For every fixed b, the set of pairs (b, a) is equivalent to A and hence countable. Thus B_n is the union of a countable set of countable sets. By Theorem 2.12, B_n is countable. The theorem follows by induction. □

Corollary The set of all rational numbers is countable.

Theorem 2.14 Let A be the set of all sequences whose elements are the digits 0 and 1. This set A is uncountable.

Proof. Let E be a countable subset of A, and let E consist of the sequences $s_1, s_2, s_3, \ldots$. We construct a sequence s as follows. If the nth digit in s_n is 1, we let the nth digit of s be 0, and vice versa. Then the sequence s differs from every member of E in at least one place; hence $s \notin E$. But clearly $s \in A$, so E is a proper subset of A. Thus, every countable subset of A is a proper subset of A. It follows that A is uncountable (otherwise A would be a proper subset of itself). □

Theorem 2.19 Every neighborhood is an open set.

Proof. Consider a neighborhood $E = N_r(p)$, and let q be any point of E. Then there is a positive real number h such that $d(p, q) = r - h$. For all points s such that $d(q, s) < h$, we have then

$$d(p, s) \leqslant d(p, q) + d(q, s) < r - h + h = r.$$

so that $s \in N_r(p)$. Thus q is an interior point of E. □

Theorem 2.20 If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E.

Proof. Suppose, to the contrary, there is a neighborhood N of p which contains only a finite number of points of E. Let $q_1, \ldots, q_n$ be those points of $N \cap E$, which are distinct from p, and put

$$r = \min_{1 \leqslant m \leqslant n} d(p, q_m)$$

The minimum of a finite set of positive numbers is clearly positive, so that $r > 0$. The neighborhood $N_r(p)$ contains no points q of E such that $q \neq p$, so that p is not a limit point of E. Thus, by contradiction, every neighborhood of a limit point p contains infinitely many points of E. □

Corollary A finite point set has no limit points.

Theorem 2.22 Let $\{E_\alpha\}$ be a (finite or infinite) collection of sets E_α. Then

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c)$$

Proof. Let $A = (\bigcup_\alpha E_\alpha)^c$ and $B = \bigcap_\alpha (E_\alpha^c)$. If $x \in A$, then $x \notin \bigcup_\alpha E_\alpha$, hence $x \notin E_\alpha$ for any α, hence $x \in E_\alpha^C$ for every α, so that $x \in \bigcap_\alpha E_\alpha^c$. Thus $A \subset B$.

Conversely, if $x \in B$, then $x \in E_\alpha c$ for every α, hence $x \notin E_\alpha$ for any α, hence $x \notin \bigcup_\alpha E_\alpha$, so that $x \in (\bigcup_\alpha E_\alpha)^c$. Thus $B \subset A$.

It follows that $A = B$. □

Theorem 2.23 A set E is open if and only if its complement is closed.

Proof. First, suppose E^c is closed. Choose $x \in E$. Then $x \notin E^c$, and x is not a limit point of E^c. Hence there exists a neighborhood N of x such that $E^c \cap N$ is empty, that is $N \subset E$. Thus x is an interior point of E, and E is open.

Next suppose E is open. Let x be a limit point of E^c, so that x is not an interior point of E. Since E is open, this means that $x \in E^c$. It follows that E^c is closed. □

Corollary A set F is closed if and only if its complement is open.

Theorem 2.24 a) For any collection $\{G_\alpha\}$ of open sets, $\bigcup_\alpha G_\alpha$ is open.

b) For any collection $\{F_\alpha\}$ of closed sets, $\bigcap_\alpha F_\alpha$ is closed.

c) For any finite collection $G_1, \ldots, G_n$ of open sets, $\bigcap_{i=1}^n G_i$ is open.

d) For any finite collection $F_1, \ldots, F_n$ of closed sets, $\bigcup_{i=1}^n F_i$ is closed.

Proof. Put $G = \bigcup_\alpha G_\alpha$. If $x \in G$, then $x \in G_\alpha$ for some α. Since x is an interior point of G_α, x is also an interior point of G, and G is open, proving (a).

By Theorem 2.22 $(\bigcup_\alpha F_\alpha)^c = \bigcap_\alpha (F_\alpha^c)$, and F_α^c is open, by Theorem 2.23. Hence (a) implies that $(\bigcup_\alpha F_\alpha)^c$ is open, so $\bigcup_\alpha F_\alpha$ is closed.

Next put $H = \bigcap_{i=1}^n G_i$. For any $x \in H$, there exist neighborhoods N_i of x, with radii r_i, such that $N_i \subset G_i$ $(i = 1, \ldots, n)$. Put $r = \min(r_1, \ldots, r_n)$, and let N be the neighborhood of x of radius r. Then $N \subset G_i$ for $i = 1, \ldots, n$, so that $N \subset H$, and H is open. By taking complements, (d) follows from (c). $(\bigcup_{i=1}^n F_\alpha)^c = \bigcap_{i=1}^n (F_\alpha^c)$ □

Theorem 2.27 If X is a metric space and $E \subset X$, then

a) $\overline{E}$ is closed,

b) $E = \overline{E}$ if an only if E is closed

c) $\overline{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$.

Proof. a) If $p \in X$ and $p \notin \overline{E}$ then p is neither a point of E nor a limit point of E. Hence p has a neighborhood which does not intersect E. The complement of $\overline{E}$ is therefore open. Hence $\overline{E}$ is closed.

b) If $E = \overline{E}$, (a) implies that E is closed. If E is closed, then $E' \subset E$. Hence $\overline{E} = E$.

c) If F is closed and $E \subset F$, then $F' \subset F$, hence $E' \subset F$. Thus $\overline{E} \subset F$

□

Theorem 2.28 Let E be a nonempty set of real numbers which is bounded above. Let $y = \sup E$. Then $y \in \overline{E}$. Hence $y \in E$ if E is closed.

Proof. If $y \in E$ then $y \in \overline{E}$. Assume $y \notin E$. For every $h > 0$ there exists then a point $x \in E$ such that $y - h < x < y$, for otherwise $y - h$ would be an upper bound of E. Thus y is a limit point of E. Hence $y \in \overline{E}$ □

Theorem 2.30 Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.

Proof. Suppose E is open relative to Y. To each $p \in E$ there is a positive number r_p such that the conditions $d(p,q) < r_p$, $q \in Y$ imply that $q \in E$. Let V_p be the set of all $q \in X$ such that $d(p,q) < r_p$, and define $G = \bigcup_{p \in E} V_p$. Then G is an open subset of X, by Theorems 2.19 and 2.24. Since $p \in V_p$ for all $p \in E$, it is clear that $E \subset G \cap Y$. By our choice of V_p, we have $V_p \cap Y \subset E$ for every $p \in E$, so that $G \cap Y \subset E$. Thus $E = G \cap Y$, and one half of the theorem is proved. Conversely, if G is open in X and $E = G \cap Y$, every $p \in E$ has a neighborhood $V_p \subset G$. Then $V_p \cap Y \subset E$, so that E is open relative to Y. □

Theorem 2.33 Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.

Proof. Suppose K is compact relative to X, and let $\{V_\alpha\}$ be a collection of sets, open relative to Y, such that $K \subset \bigcup_\alpha V_\alpha$. By Theorem 2.30, there are sets G_α, open relative to X, such that $V_\alpha = Y \cap G_\alpha$, for all α; and since K is compact relative to X, we have

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$$

for some choice of finitely many indices $\alpha_1, \ldots, \alpha_n$. Since $K \subset Y$, implies

$$K \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}$$

This proves that K is compact relative to Y.

Conversely, suppose K is compact relative to Y, let $\{G_\alpha\}$ be a collection of open subsets of X which covers K, and put $V_\alpha = Y \cap G_\alpha$. Then $K \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}$ holds for some choice of $\alpha_1, \ldots, \alpha_n$; and since $V_\alpha \subset G_\alpha$, this implies $K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$. □

Theorem 2.34 Compact subsets of metric spaces are closed.

Proof. Let K be a compact subset of a metric space X. We shall prove that the complement of K is an open subset of X.

Suppose $p \in X$, $p \notin K$. If $q \in K$, let V_q and W_q be neighborhoods of p and q respectively, of radius less that $\frac{1}{2}d(p,q)$. Since K is compact, there are finitely many points $q_1, \ldots, q_n$ in K such that $K \subset W_{q_1} \cup \cdots \cup W_{q_n} = W$. If $V = V_{q_1} \cap \cdots \cap V_{q_n}$, then V is a neighborhood of p which does not intersect W. Hence $V \subset K^c$, so that P is an interior point of K^c. The theorem follows. □

Theorem 2.35 Closed subsets of compact sets are compact.

Proof. Suppose $F \subset K \subset X$, F is closed (relative to X), and K is compact. Let $\{V_\alpha\}$ be an open cover of F. If F^c is adjoined to $\{V_\alpha\}$, we obtain an open cover Ω of K. Since K is compact, there is a finite subcollection Φ of Ω which covers K, and hence F. If F^c is a member of Φ, we may remove it from Φ and still retain an open cover of F. We have thus shown that a finite subcollection of $\{V_\alpha\}$ covers F. □

Corollary If F is closed and K is compact, then $F \cap K$ is compact.

Theorem 2.36 If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X such that the intersection of every finite sub-collection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is nonempty.

Proof. Fix a member K_1 of $\{K_\alpha\}$ and put $G_\alpha = K_\alpha^c$. Assume that no point of K_1 belongs to every K_α. Then the sets G_α form an open cover of K_1; and since K_1 is compact, there are finitely many indices $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that $K_1 \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$. But this means that $K_1 \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_n}$ is empty, in contradiction to our hypothesis that the intersection of every finite sub-collection of $\{K_\alpha\}$ is nonempty. □

Corollary If $\{K_n\}$ is a sequence of nonempty compact sets such that $K_n \supset K_{n+1}$ $(n = 1, 2, 3, \ldots)$, then $\bigcap_1^\infty K_n$ is not empty.

Theorem 2.37 If E is an infinite subset of a compact set K, then E has a limit point in K.

Proof. If no point of K were a limit point of E, then each $q \in K$ would have a neighborhood V_q which contains at most one point of E (namely q, if $q \in E$). It is clear that no finite subcollection of $\{V_q\}$ can cover E; and the same is true of K since $E \subset K$. This contradicts the compactness of K. □

Theorem 2.38 If $\{I_n\}$ is a sequence of intervals in $\mathbb{R}$ such that $I_n \supset I_{n+1}$ $(n = 1, 2, 3, \ldots)$, then $\bigcap_1^\infty I_n$ is not empty.

Proof. If $I_n = [a_n, b_n]$, let E be the set of all a_n. Then E is nonempty and bounded above (by b_1). Let x be the sup of E. If m and n are positive integers, then

$$a_n \leqslant a_{m+n} \leqslant b_{m+n} \leqslant b_m$$

so that $x \leqslant b_m$ for each m. Since it is obvious that $a_m \leqslant x$, we see that $x \in I_m$ for $m = 1, 2, 3, \ldots$. □

Theorem 2.39 Let k be a positive integer. If $\{I_n\}$ is a sequence of k-cells such that $I_n \subset I_{n+1}$ $(n = 1, 2, 3, \dots)$, then $\bigcap_1^\infty I_n$ is not empty.

Proof. Let I_n consist of all points $\mathbf{x} = (x_1, \dots, x_n)$ such that

$$a_{n,j} \leqslant x_j \leqslant b_{n,j} \qquad (1 \leqslant j \leqslant k;\ n = 1, 2, 3, \dots)$$

and put $I_{n,j} = [a_{n,j}, b_{n,j}]$. For each j the sequence $\{I_{n,j}\}$ satisfies the hypotheses of Theorem 2.38. Hence there are real numbers $x_j^*(1 \leqslant j \leqslant k)$ such that

$$a_{n,j} \leqslant x_j^* \leqslant b_{n,j} \qquad (1 \leqslant j \leqslant k;\ n = 1, 2, 3, \dots)$$

Setting $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_k^*)$, we see that $\mathbf{x}^* \in I_n$ for $n = 1, 2, 3, \dots$. The theorem follows. ☐

Theorem 2.40 Every k-cell is compact.

Proof. Let I be a k-cell, consisting of all points $\mathbf{x} = (x_1, \dots, x_k)$ such that $a_j \leqslant x_j \leqslant b_j$ $(1 \leqslant j \leqslant k)$. Put

$$\delta = \left\{ \sum_1^k (b_j - a_j)^2 \right\}^{1/2}.$$

Then $|\mathbf{x} - \mathbf{y}| \leqslant \delta$ if $x, y \in I$. Seeking a contradiction, suppose that there exists an open cover $\{G_\alpha\}$ of I which contains no finite subcover of I. Put $c_j = (a_j + b_j)/2$. The intervals $[a_j, c_j]$ and $[c_j, b_j]$ then determine 2^k k-cells Q_i whose union is I. At least one of these sets of Q_i, call it I_1 cannot be covered by any finite subcollection of $\{G_\alpha\}$ (otherwise I could be covered). We next subdivide I_1 and continue the process. We obtain a sequence $\{I_n\}$ with the following properties:

a) $I \supset I_1 \supset I_2 \supset \dots$;

b) I_n is not covered by any finite subcollection of $\{G_\alpha\}$;

c) if $\mathbf{x}, \mathbf{y} \in I_n$, then $|\mathbf{x} - \mathbf{y}| \leqslant 2^{-n}\delta$.

By (a) and Theorem 2.39, there is a point $\mathbf{x}^*$ which lies in every I_n. For some α, $\mathbf{x}^* \in G_\alpha$. Since G_α is open, there exists $r > 0$ such that $|\mathbf{y} - \mathbf{x}^*| < r$ implies that $\mathbf{y} \in G_\alpha$. If n is so large that $2^{-n}\delta < r$ (there is such an n, for otherwise $2^n \leqslant \delta/r$ for all positive integers n, which is absurd since $\mathbb{R}$ is archimedean), then (c) implies that $I_n \subset G_\alpha$, which contradicts (b), completing the proof. ☐

Theorem 2.41 If a set E in $\mathbb{R}^k$ has one of the following three properties, then it has the other two.

a) E is closed and bounded

b) E is compact

c) Every infinite subset of E has a limit point in E.

Proof. If (a) holds, then $E \subset I$ for some k-cell I, and (b) follows from Theorem 2.40 and 2.35. Theorem 2.37 shows that (b) implies (c). If remains to be shown that (c) implies (a).

If E is not bounded, then E contains points $\mathbf{x}_n$ with $|\mathbf{x}_n| > n$ $(n = 1, 2, 3, \dots)$. The set S consisting of these points $\mathbf{x}_n$ is infinite and clearly has no limit points in $\mathbb{R}^k$, hence has none in E. Thus, (c) implies that E is bounded.

IF E is not closed, then there is a point $\mathbf{x}_0 \in \mathbb{R}^k$ which is a limit point of E but not a point of E. For $n = 1, 2, 3, \dots$, there are points $\mathbf{x}_n \in E$ such that $|\mathbf{x}_n - \mathbf{x}_0| < \frac{1}{n}$. Let S be the set of these points $\mathbf{x}_n$. Then S is infinite (otherwise $|\mathbf{x}_n - \mathbf{x}_0|$ would have a constant positive value, for infinitely many n), S has $\mathbf{x}_0$ as a limit point, and S has no other limit point in $\mathbb{R}^k$. For if $\mathbf{y} \in \mathbb{R}^k$, $\mathbf{y} \neq \mathbf{x}_0$, then $|\mathbf{x}_n - \mathbf{y}| \geqslant |\mathbf{x}_0 - \mathbf{y}| - |\mathbf{x}_n - \mathbf{x}_0| \geqslant |\mathbf{x}_0 - \mathbf{y}| - \frac{1}{n} \geqslant \frac{1}{2}|\mathbf{x}_0 - \mathbf{y}|$ for all but finitely many n; this shows that $\mathbf{y}$ is not a limit point of S. Thus s has no limit point in E; hence E must be closed if (c) holds. □

Theorem 2.42 (Weierstrass) Every bounded infinite subset of $|R^k$ has a limit point in $\mathbb{R}^k$.

Proof. Being bounded, the set E in question is a subset of a k-cell $I \subset \mathbb{R}^k$. By Theorem 2.40 I is compact, and so E has a limit point in I, by Theorem 2.37. □

Theorem 2.43 Let P be a nonempty perfect set in $\mathbb{R}^k$. Then P is uncountable.

Proof. Since P has limit points, P must be infinite. Suppose P is countable, and denote the points of P by $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots$. We shall construct a sequence $\{V_n\}$ of neighborhoods, as follows.

Let V_1 be any neighborhood of $\mathbf{x}_1$. If V_1 consists of all $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y} - \mathbf{x}_1| < r$, the closure $\overline{V}_1$ of V_1 is the set of all $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y} - \mathbf{x}_1| \leqslant r$.

Suppose V_n has been constructed, so that $V_n \cap P$ is not empty. Since every point of P is a limit point of P, there is a neighborhood V_{n+1} such that (i) $\overline{V}_{n+1} \subset V_n$, (ii) $\mathbf{x}_n \notin \overline{V}_{n+1}$, (iii) $V_{n+1} \cap P$ is not empty. By (iii), V_{n+1} satisfies out induction hypothesis, and the construction can proceed.

Put $K_n = \overline{V}_n \cap P$. Since $\overline{V}_n$ is closed and bounded, $\overline{V}_n$ is compact. Since $\mathbf{x}_n \notin K_{n+1}$, no point of P lies in $\bigcap_1^\infty K_n$. Since $K_n \subset P$, this implies that $\bigcap_1^\infty K_n$ is empty. But each K_n is nonempty, by (iii), and $K_n \supset K_{n+1}$ by (i); this contradicts the Corollary to Theorem 2.36. □

Corollary Every interval $[a, b]$ is uncountable. In particular, the set of all real numbers is uncountable.

Theorem 2.47 A subset E of the real line $\mathbb{R}$ is connected if and only if it has the following property: if $x \in E$ and $y \in E$, and $x < z < y$, then $z \in E$.

Proof. If there exists $x \in E$, $y \in E$, and some $z \in (x, y)$ such that $z \notin E$, then $E = A_z \cup B_z$ where $A_z = E \cap (-\infty, z)$, $B_z = E \cap (z, \infty)$. Since $x \in A_z$ and $y \in B_z$, A and B are nonempty. Since $A_z \subset (-\infty, z)$ and $B_z \subset (z, \infty)$, they are separated. Hence E is not connected.

To prove the converse, suppose E is not connected. Then there are nonempty separated sets A and B such that $A \cup B = E$. Pick $x \in A$, $y \in B$, and assume (WLOG) that $x < y$. Define $z = \sup(A \cap [x, y])$. By Theorem 2.28, $z \in \overline{A}$; hence $z \notin \overline{B}$. In particular, $x \leqslant z < y$. If $z \notin A$, it follows that $x < z < y$ and $z \notin E$. If $z \in A$, then $z \notin \overline{B}$, hence there exists z_1 such that $z < z_1 < y$ and $z_1 \notin B$. Then $x < z_1 < y$ and $z_1 \notin E$. □

Chapter 3 Numerical Sequences

Definitions

Converge/ Diverge A sequence $\{p_n\}$ in a metric space X is said to *converge* if there is a point $p \in X$ with the following property: for every $\varepsilon > 0$ there is an integer N such that $n \geqslant N$ implies that $d(p_n, p) < \varepsilon$. In this case we also say that $\{p_n\}$ converges to p, or that p is the limit of $\{p_n\}$, and we write $p_n \to p$, or

$$\lim_{n\to\infty} p_n = p$$

If $\{p_n\}$ does no converge, it is said to *diverge*.

Range/ Bounded The set of all points p_n is the *range* of $\{p_n\}$. The range of a sequence may be a finite set, or it may be infinite. The sequence $\{p_n\}$ is said to be *bounded* if it's range is bounded.

Subsequence Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3 < \ldots$. Then the sequence $\{p_{n_i}\}$ is called a *subsequence* of $\{p_n\}$. If $\{p_{n_i}\}$ converges, its limit is called a *subsequential limit* of $\{p_n\}$.

It is clear that $\{p_n\}$ converges to p is and only if every subsequence of $\{p_n\}$ converges to P.

Cauchy Sequence A sequence $\{p_n\}$ in a metric space X is said to be a *Cauchy sequence* if for every $\varepsilon > 0$ there is an integer N such that $d(p_n, p_m) < \varepsilon$ if $n \geqslant N$ and $m \geqslant N$.

Diameter Let E be a nonempty subset of a metric space X, and let S be the set of all real numbers of the form $d(p, q)$, with $p \in E$ and $q \in E$. The sup of S is called the *diameter of E*.

Complete A metric space in which every Cauchy sequence converges is said to be *complete*.

Monotonic A sequence $\{s_n\}$ of real numbers is said to be

a) *monotonically increasing* if $s_n \leqslant s_{n+1}$ $(n = 1, 2, 3, \ldots)$

b) *monotonically decreasing* if $s_n \geqslant s_{n+1}$ $(n = 1, 2, 3, \ldots)$

The class of monotonic sequences consists of the increasing and the decreasing sequences.

Convergence to Infinity Let $\{s_n\}$ be a sequence of real numbers with the following property: For every real M there is an integer N such that $n \geqslant N$ implies $s_n \geqslant M$. We then write: $s_n \to +\infty$.

Similarly, if for every real M there is an integer N such that $n \geqslant N$ implies $s_n \leqslant M$. We then write: $s_n \to -\infty$.

Upper/ Lower Limits Let $\{s_n\}$ be a sequence of real numbers. Let E be the set of numbers X such that $s_{n_k} \to x$ for some subsequence $\{s_{n_k}\}$. This set E contains all subsequential limits as defined above. Let:

$$s^* = \sup E \qquad s_* = \inf E$$

The numbers s^* and s_* are called the *upper* and *lower limits* of $\{s_n\}$; we use the notation:

$$\limsup_{n\to\infty} s_n = s^* \qquad \liminf_{n\to\infty} s_n = s_*$$

Series Given a sequence $\{a_n\}$, we use the notation $\sum_{n=p}^{q} a_n$ to denote the sum $a_p + a_{p+1} + \dots + a_q$. With $\{a_n\}$ we associate a sequence $\{s_n\}$ where $s_n = \sum_{k=p}^{n} a_k$. For $\{s_n\}$ we also use the symbolic expression $a_1 + a_2 + a_3 + \dots$ or, more concisely,

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

We call this an *infinite series*, or just a *series*. The numbers s_n are called the *partial sums* of the series. If $\{s_n\}$ converges to s, we say that the series *converges*, and write

$$\sum_{n=1}^{\infty} a_n = s.$$

The number s is called the sum of the series; but it should be clearly understood that s is the *limit of a sequence of sums*, and is not obtained by simple addition.

If $\{s_n\}$ diverges, the series is said to *diverge*.

e $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Power Series Given a sequence $\{c_n\}$ of complex numbers, the series

$$\sum_{n=0}^{\infty} c_n z^n$$

is called a *power series*. The numbers c_n are called the *coefficients* of the series; z is a complex number.

Absolute Convergence The series $\sum a_n$ is said to *converge absolutely* if the series $\sum |a_n|$ converges.

Product Given $\sum a_n$ and $\sum b_n$, we put

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

and call $\sum c_n$ the *product* of the two given series. Equivalently,

$$\text{product} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k}$$

Rearrangement Let $\{k_n\}$ $n = 1, 2, 3, \ldots$ be a sequence in which every positive integer appears once and only once (that is, $\{k_n\}$ is a one-to-one function from J onto J, Putting $a'_n = a_{k_n}$ $(n = 1, 2, 3, \ldots)$, we say that $\sum a'_n$ is a *rearrangement* of $\sum a_n$.

Theorems

Theorem 3.2 Let $\{p_n\}$ be a sequence in a metric space X.

a) $\{p_n\}$ converges to $p \in X$ if and only if every neighborhood of p contains p_n for all but finitely many n.

b) If $p \in X$, $p' \in X$, and if $\{p_n\}$ converges to p and to p', then $p' = p$.

c) If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.

d) If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$in E such that $p = \lim_{n\to\infty} p_n$

Proof. a) Suppose $p_n \to p$ and let V be a neighborhood of P. For some $\varepsilon > 0$, the conditions $d(q,p) < \varepsilon$, $q \in X$ imply $q \in V$. Corresponding to this ε, there exists N such that $n \geqslant N$ implies $d(p_n, p) < \varepsilon$. Thus $m \geqslant N$ implies $p_n \in V$.

Conversely suppose every neighborhood of p contains all but finitely many of the p_n. Fix $\varepsilon > 0$, and let V be the set of all $q \in X$ such that $d(p,q) < \varepsilon$. By assumption, there exists N (corresponding to this V) such that $p_n \in V$ if $n \geqslant N$. Thus $d(p_n, p) < \varepsilon$ if $n \geqslant N$; hence $p_n \to p$.

b) Let $\varepsilon > 0$ be given. There exist integers N, N' such that

$$n \geqslant N \text{ implies } d(p_n, p) < \frac{\varepsilon}{2} \qquad n \geqslant N' \text{ implies } d(p_n, p') < \frac{\varepsilon}{2}$$

Hence if $n \geqslant max(N, N')$, we have: $d(p,p') \leqslant d(p,p_n) + d(p_n,p')M\varepsilon$. Since ε was arbitrary, we conclude that $d(p,p') = 0$.

c) Suppose $p_n \to p$. There is an integer N such that $n > N$ implies $d(p_n, p) < 1$. Put $r = \max\{1, d(p_1,p), \ldots, d(p_N,p)\}$. Then $d(p_n,p) <\leqslant r$ for $n = 1, 2, 3, \ldots$.

d) For each positive integer n, there is a point $p_n \in E$ such that $d(p_n, p) < \frac{1}{n}$. Given $\varepsilon > 0$, choose N so that $N\varepsilon > 1$. If $n > N$, if follows that $d(p_n, p) < \varepsilon$. Hence $p_n \to p$.

□

Theorem 3.3 Suppose $\{s_n\}$, $\{t_n\}$ are complex sequences, and $\lim_{n\to\infty} s_n = s$, $\lim_{n\to\infty} t_n = t$. Then,

a) $\lim_{n\to\infty}(s_n + t_n) = s + t$;

b) $\lim_{n\to\infty} cs_n = cs$, $lim_{n\to\infty}(c + s_n) = c + s$;

c) $\lim_{n\to\infty} s_n t_n = st$

d) $\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s}$, provided $s_n \neq 0$ and $s \neq 0$.

Proof. a) Given $\varepsilon > 0$, there exists integers N_1, N_2 such that

$$n \geqslant N_1 \text{ implies } |s_n - s| < \frac{\varepsilon}{2} \qquad n \geqslant N_2 \text{ implies } |t_n - t| < \frac{\varepsilon}{2}$$

If $N = \max(N_1, N_2)$, then $n \geqslant N$ implies

$$|(s_n + t_n) - (s + t)| \leqslant |s_n - s| + |t_n - t| < \varepsilon$$

This proves (a).

b) Trivial.

c) We use the identity:

$$s_n t_n - st = (s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s)$$

Given $\varepsilon > 0$, there are integers N_1, N_2 such that

$$n \geqslant N_1 \text{ implies } |s_n - s| < \sqrt{\varepsilon}, \qquad n \geqslant N_2 \text{ implies } |t_n - t| < \sqrt{\varepsilon}.$$

If we take $N = max(N_1, N_2)$, $n \geqslant N$ implies $|(s_n - s)(t_n - t)| < \varepsilon$, so that $\lim_{n\to\infty}(s_n - s)(t_n - t) = 0$. We now apply (a) and (b) to our original equivalence, and conclude that $\lim_{n\to\infty}(s_n t_n - st) = 0$.

d) Choose m such that $|s_n - s| < \frac{1}{2}|s|$ if $n \geqslant m$, we see that $|s_n| > \frac{1}{2}|s|$. Given $\varepsilon > 0$, there is an integer $N > m$ such that $n \geqslant N$ implies $|s_n - s| < \frac{1}{2}|s|^2\varepsilon$. Hence, for $n \geqslant N$,

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| = \left|\frac{s_n - s}{s_n s}\right| < \frac{2}{|s|^2}|s_n - s| < \varepsilon.$$

□

Theorem 3.4 Suppose $\mathbf{x}_n \in \mathbb{R}^k$ and $\mathbf{x}_n = (\alpha_{1,n}, \alpha_{2,n}, \ldots, \alpha_{k,n})$. Then $\{\mathbf{x}_n\}$ converges to $\mathbf{x} = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ if and only if

$$\lim_{n\to\infty} \alpha_{j,n} = \alpha_j \qquad (1 \leqslant j \leqslant k)$$

Suppose $\{\mathbf{x}_n\}$, $\{y_n\}$ are sequences in $\mathbb{R}^k$, $\{\beta_n\}$ is a sequence of real numbers, and $\mathbf{x}_n \to \mathbf{x}$, $\mathbf{y}_n \to \mathbf{y}$, $\beta_n \to \beta$. Then

$$\lim_{n\to\infty}(\mathbf{x}_n + \mathbf{y}_n) = \mathbf{x} + \mathbf{y} \qquad \lim_{n\to\infty}(\mathbf{x}_n \cdot \mathbf{y}_n) = \mathbf{x} \cdot \mathbf{y} \qquad \lim_{n\to\infty}(\beta_n \mathbf{x}_n) = \beta\mathbf{x}$$

Proof. a) If $\mathbf{x}_n \to \mathbf{x}$, the inequalities $|\alpha_{j,n} - \alpha_j| \leqslant |\mathbf{x}_n - \mathbf{x}|$, which follow immediately from the definition of the norm in $\mathbb{R}^k$, show that $\lim_{n\to\infty} \alpha_{j,n} = \alpha_j$ for $(1 \leqslant j \leqslant k)$. Conversely, if $\lim_{n\to\infty} \alpha_{j,n} = \alpha_j$ for $(1 \leqslant j \leqslant k)$, then to each $\varepsilon > 0$ there corresponds an integer N such that $n \geqslant N$ implies $|\alpha_{j,n} - \alpha_j| < \frac{\varepsilon}{\sqrt{k}}$ for $(1 \leqslant j \leqslant k)$. Hence $n \geqslant N$ implies

$$|\mathbf{x}_n - x| = \left\{\sum_{j=1}^{k} |\alpha_{j,n} - \alpha_j|^2\right\}^{1/2} < \varepsilon,$$

so that $\mathbf{x}_n \to \mathbf{x}$.

b) Follows from (a) and Theorem 3.3.

□

Theorem 3.6 a) If $\{p_n\}$ is a sequence in a compact metric space X, then some subsequence of $\{p_n\}$ converges to a point of X.

b) Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.

Proof. a) Let E be the range of $\{p_n\}$. If E is finite then there is a $p \in E$ and a sequence $\{n_i\}$ with $n_1 < n_2 < \dots$, such that $p_{n_1} = p_{n_2} = \cdots = p$. The subsequence $\{p_{n_i}\}$ so obtained converges evidently to p

If E is infinite, Theorem 2.37 shows that E has a limit point $P \in X$. Choose n_1 so that $d(p, p_{n_1}) < 1$. Having chosen $n - 1, \dots, n_{i-1}$, we see from Theorem 2.20 that there is an integer $n_i > n_{i-1}$ such that $d(p, p_{n_i}) < 1/i$. Then $\{p_{n_i}\}$ converges to p.

b) This follows from (a), since Theorem 2.41 implies that every bounded subset of $\mathbb{R}^k$ lies in a compact subset of $\mathbb{R}^k$.

□

Theorem 3.7 The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subset of X.

Proof. Let E^* be the set of all subsequential limits of $\{p_n\}$ and let q be a limit point of E^* . We have to show that $q \in E^*$.

Choose n_1 so that $p_{n_1} \neq q$. (If no such n_1 exists, then E^* has only one point and there is nothing to prove.) Put $\delta = d(q, p_{n_1})$. Suppose $n_1, \dots, n_{i-1}$ are chosen. Since q is a limit point of E^*, there is a $x \in E^*$ with $d(x, q) < 2^{-i}\delta$. Since $x \in E^*$ there is an $n_i > n_{i-1}$ such that $d(x, p_{n_i}) < 2^{-i}\delta$. Thus $d(q, p_{n_i}) \leqslant 2^{1-i}\delta$ for $i = 1, 2, 3, \dots$. This says that $\{p_{n_i}\}$ converges to q. Hence $q \in E^*$. □

Theorem 3.10 a) If $\overline{E}$ is the closure of a set E in a metric space X, then

$$\operatorname{diam} \overline{E} = \operatorname{diam} E$$

b) If K_n is a sequence of compact sets in X such that $K_n \supset K_{n+1}$ and if

$$\lim_{n \to \infty} \operatorname{diam} K_n = 0,$$

then $\bigcap_1^\infty K_n$ consists of exactly one point.

Proof. a) Since $E \subset \overline{E}$, it is clear that $\operatorname{diam} E \leqslant \operatorname{diam} \overline{E}$. Fix $\varepsilon > 0$, and choose $p \in \overline{E}$, $q \in \overline{E}$. By definition of $\overline{E}$, there are points p', q', in E such that $d(p, p') < \varepsilon$, $d(q, q') < \varepsilon$. Hence

$$d(p, q) \leqslant d(p, p') + d(p', q') + d(q', q) < 2\varepsilon + d(p', q') \leqslant 2\varepsilon + \operatorname{diam} E$$

Thus, $\operatorname{diam} \overline{E} \leqslant diam E$. Therefore $\operatorname{diam} \overline{E} = \operatorname{diam} E$.

b) Put $K = \bigcap +1^{\infty} K_n$. By Theorem 2.36, K is not empty. If K contains more than one point, then diam $K > 0$. But for each n, $K_n \supset K$, so that diam $K_n \geqslant$ diam K. This contradicts the assumption that diam $K_n \to 0$.

□

Theorem 3.11 a) In any metric space X, every convergent sequence is a Cauchy sequence.

b) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some point of X.

c) In $\mathbb{R}^k$, every Cauchy sequence converges

Proof. a) If $p_n \to p$ and if $\varepsilon > 0$, there is an integer N such that $d(p, p_n) < \varepsilon$ for all $n \geqslant N$. Hence, for $n, m \geqslant N$:

$$d(p_n, p_m) \leqslant d(p_n, p) + d(p_m, p) < 2\varepsilon$$

Thus $\{p_n\}$ is Cauchy.

b) Let $\{p_n\}$ be a Cauchy sequence in the compact space X. For $N = 1, 2, 3, \ldots$, let E_N be the set consisting of $p_N, p_{N+1}, \ldots$. Then $\lim_{N\to\infty} \operatorname{diam} \overline{E}_N = 0$, by Definition 3.9 and Theorem 3.10(a). Being a closed subset of the compact metric space X, each $\overline{E}_N$ is compact (Theorem 2.35). Also $E_N \supset E_{N+1}$, so that $\overline{E}_N \supset \overline{E}_{N+1}$.

Theorem 3.10(b) shows now that there is a unique $p \in X$ which lies in every $\overline{E}_N$. Let $\varepsilon > 0$ be given. Because $\lim_{N\to\infty} \operatorname{diam} \overline{E}_N = 0$ there is an integer N_0 such that $\operatorname{diam} \overline{E}_N < \varepsilon$ if $N \geqslant N_0$. Since $p \in \overline{E}_N$, it follows that $d(p, q) < \varepsilon$ for every $q \in \overline{E}_N$, hence for every $q \in E_n$. In other words, $d(p, p_n) < \varepsilon$ if $n \geqslant N_0$. This says precisely that $p_n \to p$.

c) Let $\{\mathbf{x}_n\}$ be a Cauchy sequence in $\mathbb{R}^k$. Define E_N as in (b) with $\mathbf{x}_i$ in place of p_i. For some N, diam $E_N < 1$. The range of $\{\mathbf{x}_n\}$ is the union of E_N and the finite set $\{\mathbf{x}_1, \ldots, \mathbf{x}_{N-1}\}$. Hence $\{\mathbf{x}_n\}$ is bounded. Since every bounded subset of $\mathbb{R}^k$ has compact closure in $\mathbb{R}^k$ (c) follows from (b).

□

Corollary All Compact metric spaces and all Euclidean spaces are complete.

Corollary Every Closed subset E of a complete metric space X is complete.

Theorem 3.14 Suppose $\{s_n\}$ is monotonic. Then $\{s_n\}$ converges if and only if it is bounded.

Proof. Suppose $s_n \leqslant s_{n+1}$ (the proof is analogous in the other case). Let E be the range of $\{s_n\}$. If $\{s_n\}$ is bounded, let s be the least upper bound of E. Then $s_n \leqslant s$ $(n = 1, 2, 3, \ldots)$. For every $\varepsilon > 0$, there is an integer N such that $s - \varepsilon < s_N \leqslant s$, for otherwise $s - \varepsilon$ would be an upper bound of E. Since $\{s_n\}$ increases, $n \geqslant N$ therefore implies $s - \varepsilon < s_n \leqslant s$, which shows that $\{s_n\}$ converges to s. The converse follows from Theorem 3.2 (c). □

Theorem 3.17 Let $\{s_n\}$ be a sequence of real numbers. Let E and s^* be as defined above. Then s^* has the following two properties:

a) $s^* \in E$

b) If $x > s^*$, there is an integer N such that $n \geqslant N$ implies $s_n < x$.

Moreover, s^* is the only number with these properties. Furthermore, the analogous result is true for s_*.

Proof. a) If $s^* = +\infty$, then E is not bounded above; hence $\{s_n\}$ is not bounded above, and there is a subsequence $\{s_{n_k}\}$ such that $s_{n_k} \to +\infty$.

If s^* is real, then E is bounded above, and at least one subsequential limit exists, so that (a) follows from Theorems 3.7 and 2.28.

If $s^* = -\infty$, the E contains only one element, namely $-\infty$, and there is no subsequential limit. Hence, for any real M, $s_n > M$ for at most finite number of values of n, so that $s_n \to -\infty$.

b) Suppose there is a number $x > s^*$ such that $s_n \geqslant x$ for infinitely many values of n. In that case, there is a number $y \in E$ such that $y \geqslant x \geqslant s^*$, contradicting the definition of s^*.

Thus s^* satisfies (a) and (b).

To show the uniqueness, suppose there are two numbers, p and q, which satisfy (a) and (b), and suppose $p < q$. Choose x such that $p < x < q$. Since p satisfies (b), we have $s_n < x$ for $n \geqslant N$. But then q cannot satisfy (a).

□

Theorem 3.19 If $s_n \leqslant t_n$ for $n \geqslant N$, where N is fixed, then:

$$\liminf_{n\to\infty} s_n \leqslant \liminf_{n\to\infty} t_n \qquad \limsup_{n\to\infty} s_n \leqslant \limsup_{n\to\infty} t_n$$

Theorem 3.20 a) If $p > 0$, the $\lim_{n\to\infty} \frac{1}{n^p} = 0$

b) If $p > 0$, the $\lim_{n\to\infty} \sqrt[n]{p} = 1$

c) $\lim_{n\to\infty} \sqrt[n]{n} = 1$

d) If $p > 0$ and $\alpha \in \mathbb{R}$, then $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$

e) If $|x| < 1$, then $\lim_{n\to\infty} x^n = 0$.

Proof. a) Take $n > \left(\frac{1}{\varepsilon}\right)^{1/p}$. (Note that the archimedean property of the real number system is used here.)

b) If $p > 1$, put $x_n = \sqrt[n]{p} - 1$. Then $x_n > 0$, and, by the binomial theorem, $1 + nx_n \leqslant (1 + x_n)^n = p$, so that $0 < x_n \leqslant \frac{p-1}{n}$. Hence $x_n \to 0$. If $p = 1$, (b) is trivial, and if $0 < p < 1$, the result is obtained by taking reciprocals.

c) Put $x_n = \sqrt[n]{n} - 1$. Then $x_n \geqslant 0$, and, by the binomial theorem, $n = (1 + x_n)^n \geqslant \frac{n(n-1)}{2} x_n^2$. Hence, $0 \leqslant x_n \leqslant \sqrt{\frac{2}{n-1}}$ for $n \geqslant 2$.

d) Let Let k be an integer such that $k > \alpha$, $k > 0$. For $n > 2k$:

$$(1+p)^n > \binom{n}{k}^k = \frac{n(n-1)\dots(n-k+1)}{k!} p^k > \frac{n^k p^k}{2^k k!}$$

Hence $0 < \frac{n^\alpha}{(a+p)^n} < \frac{2^k k!}{p^k} n^{\alpha - k}$ for $n > 2k$. Since $\alpha - k < 0$, $n^{\alpha-k} \to 0$, by (a).

e) Take $\alpha = 0$ in (d).

□

Theorem 3.22 $\sum a_n$ converges if and only if for every $\varepsilon > 0$ there is an integer N such that

$$\left| \sum_{k=n}^{m} a_k \right| \leqslant \varepsilon$$

if $m \geqslant n \geqslant N$. In particular, by taking $m = n$, $|a_n| \leqslant \varepsilon$.

Theorem 3.23 If $\sum a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Theorem 3.24 A series of nonnegative terms converges if and only if its partial sums form a bounded sequence.

Theorem 3.25 (Comparison Test) a) If $|a_n| \leqslant c_n$ for $n \geqslant N_0$, where N_0 is some fixed integer, and if $\sum c_n$ converges, then $\sum a_n$ converges.

b) If $a_n \geqslant d_n \geqslant 0$ for $n \geqslant N_0$, and if $\sum d_n$ diverges, then $\sum a_n$ diverges.

Proof. Given $\varepsilon > 0$, there exists $N \geqslant N_0$ such that $m \geqslant n \geqslant N$ implies $\sum_{k=n}^{m} c_k \leqslant \varepsilon$, by the Cauchy criterion. Hence

$$\left| \sum_{k=n}^{m} a_k \right| \leqslant \sum_{k=n}^{m} |a_k| \leqslant \sum_{k=n}^{m} c_k \leqslant \varepsilon,$$

and (a) follows.

Next, (b) follows from (a), for if $\sum a_n$ converges, so must $\sum d_n$. Note, (b) also follows from Theorem 3.24. □

Theorem 3.26 (Geometric Series) If $0 \leqslant x < 1$, then

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

If $x \geqslant 1$ this series diverges.

Proof. If $x \neq 1$, $s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$. The result follows if we let $n \to \infty$. For $x = 1$, we get $1 + 1 + 1 + \dots$, which evidently diverges. □

Theorem 3.27 (Cauchy Condensation Test) Suppose $a_1 \geqslant a_2 \geqslant a_3 \geqslant \cdots \geqslant 0$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the series

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \dots$$

converges.

Proof. By Theorem 3.24, if suffices to consider boundedness of the partial sums. Let $s_n = a_1 + a_2 + \cdots + a_n$ $t_k = a_1 + 2a_2 + \cdots + 2^k a_{2^k}$. For $n < 2^k$,

$$s_n \leqslant a_1 + (a + 2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \leqslant a_1 + 2a_2 + \cdots + 2^k a_{2^k} = t_k$$

so that $s_n \leqslant t_k$. On the other hand, if $n > 2^k$,

$$s_n \geqslant a_1 + a_2 + (a + 3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \leqslant \frac{1}{2} a_1 + a_2 + 2a_4 \cdots + 2^{k-1} a_{2^k} = \frac{1}{2} t_k$$

so that $2s_n \geqslant t_k$. So, $s_n \leqslant t_k \leqslant s_n$. Thus, the sequences $\{s_n\}$ and $\{t_k\}$ are either both bounded or both unbounded. This completes the proof. □

Theorem 3.28 (p-Test) $\sum \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \leqslant 1$.

Proof. If $p \leqslant -$, divergence follows from Theorem 3.23. If $p > 0$, Theorem 3.27 is applicable, and we are led to the series

$$\sum_{k=0}^{\infty} 2^k \frac{1}{2^{kp}} = \sum_{k=0}^{\infty} 2^{(1-p)k}.$$

Now, $2^{1-p} < 1$ if and only if $1 - p < 0$, and the result follows by comparison with the geometric series (take $x = 2^{1-p}$ in Theorem 3.26). □

Theorem 3.29 If $p > 1$

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$$

converges. If $p \leqslant 1$, the series diverges.

Proof. The monotonicity of the logarithmic function implies that $\{\log n\}$ increases. Hence $\{1/x \log n\}$decrease, and we can apply Theorem 3.27; this leads us to the series

$$\sum_{k=1}^{\infty} 2^k \frac{1}{2^k (\log 2^k)^p} = \sum_{k=1}^{\infty} \frac{1}{(k \log 2)^p} = \frac{1}{(\log 2)^p} \sum_{k=1}^{\infty} \frac{1}{k^p},$$

and Theorem 3.29 follows from Theorem 3.28. □

Thoerem 3.31 $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$.

Proof. Let $s_n = \sum_{k=0}^{n}\frac{1}{k!}$, $t_n = \left(1\frac{1}{n}\right)^n$. By the binomial theorem:

$$t_n = 1+1+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{2!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\dots+\frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{n-1}{n}\right)$$

Hence $t_n \leqslant s_n$, so that $\limsup_{n\to\infty} t_n \leqslant e$, by Theorem 3.19. Next if $n \geqslant m$,

$$t_n \geqslant 1+1+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\dots+\frac{1}{m!}\left(1-\frac{1}{n}\right)\dots\left(1-\frac{m-1}{n}\right).$$

Let $n \to \infty$, keeping m fixed. We get $\liminf t_n \geqslant 1+1+\frac{1}{2!}+\dots+\frac{1}{m!}$, so that $s_m \leqslant \liminf_{n\to\infty} t_n$. Letting $m\to\infty$, we finally get $e \leqslant \liminf_{n\to\infty} t_n$. The theorem follows. □

Theorem 3.32 e is irrational.

Proof. Suppose e is rational. Then $e = p/q$, where p and q are positive integers. Because $0 < e - s_n < \frac{1}{n!n}$, we know that $0 < q!(e-s_q) < \frac{1}{q}$. By our assumption $q!e$ is an integer. Since $q!s_q = q!(1+1+\frac{1}{2!}+\dots+\frac{1}{q!})$ is an integer, we see that $q!(e-s_n)$ is an integer. Since $q \geqslant 1$, this implies the existence of an integer between 0 and 1. We have thus reached a contradiction. □

Theorem 3.33 (Root Test) Given $\sum a_n$, put $\alpha - \limsup_{n\to\infty}\sqrt[n]{|a_n|}$. Then:

a) if $\alpha < 1$, $\sum a_n$ converges

b) if $\alpha > 1$, $\sum a_n$ diverges

c) if $\alpha = 1$, the test gives no information

Proof. If $\alpha < 1$, we can choose β so that $\alpha < \beta < 1$, and an integer N such that $\sqrt[n]{|a_n|} < \beta$ for $n \geqslant N$. That is, $n \geqslant N$ implies $|a_n| < \beta^n$. Since $0 < \beta < 1$, $\sum\beta^n$ converges. Convergence of σa_n follows now from the comparison test.

If $\alpha > 1$, then, again by Theorem 3.17, there is a sequence $\{n_k\}$ such that $\sqrt[n_k]{|a_{n_k}|} \to \alpha$. Hence $|a_n| > 1$ for infinitely many values of n, so that the condition $a_n \to 0$, necessary for convergence of σa_n, does not hold.

To prove (c), we consider the series $\sum\frac{1}{n}$, $\sum\frac{1}{n^2}$. For each of these series $\alpha = 1$, but the first diverges, the second converges. □

Theorem 3.34 (Ratio Test) The series $\sum a_n$

a) Converges if $\limsup_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1$

b) Diverges if $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \geqslant 1$ for all $n \geqslant n_0$, where n_0 is some fixed integer.

Proof. If condition (a) holds, we can find $\beta < 1$ and an integer N, such that $\left|\frac{a_{n+1}}{a_n}\right| < \beta$ for $n \geqslant N$. In particular, $|a_{N+1}| < \beta|a_N|$, $|a_{N+2}| < \beta|a_{N+1}| < \beta^2|a_N|$, ..., $|a_{N+p}| < \beta^p|a_N|$. That is, $|a_n| < |a_N|\beta^{-N} \cdot \beta^n$ for $n \geqslant N$, and (a) follows from the comparison test, since $\sigma\beta^n$ converges.

If $|a_{n+1}| \geqslant |a_n|$ for $n \geqslant n_0$, it is easily seen that the condition $a_n \to 0$ does not hold, and (b) follows. □

Theorem 3.37 For any sequence $\{c_n\}$ of positive numbers,

$$\liminf_{n\to\infty} \frac{c_{n+1}}{c_n} \leqslant \liminf_{n\to\infty} \sqrt[n]{c_n} \qquad \limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

Proof. We shall prove the second inequality; the proof of the first is quite similar. Put $\alpha = \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}$. If $\alpha = +\infty$, there is nothing to prove. If α is finite, choose $\beta > \alpha$. There is an integer N such that $\frac{c_{n+1}}{c_n} \leqslant \beta$ for $n \geqslant N$. In particular, for any $p > 0$, $c_{N+k+1} \leqslant \beta c_{N+k}$. Multiplying these inequalities, we obtain $c_{N+p} \leqslant \beta^p c_N$, or $c_n \leqslant c_N\beta^{-N}\beta^n$ for $n \geqslant N$. Hence $\sqrt[n]{c_n} \leqslant \sqrt[n]{c_n\beta^{-N}} \cdot \beta$, so that $\limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \beta$, by Theorem 3.20 (b). Since this is true for every $\beta > \alpha$, we have $\limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \alpha$. □

Theorem 3.39 Given the power series $\sum c_n z^n$, put

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|}, \qquad R = \frac{1}{\alpha}$$

If $\alpha = 0$, then $R = +\infty$; if $\alpha = +\infty$, $R = 0$. (Note, R is called the *radius of convergence*). Then $\sum c_n z^n$ converges if $|z| < R$, and diverges if $|z| > R$.

Proof. Put $a_n = c_n z^n$, and apply the root test:

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = |z| \qquad \limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{|z|}{R}.$$

□

Theorem 3.41 Given two sequence $\{a_n\}$, $\{b_n\}$, put $A_n = \sum_{k=0}^{n} a_k$ if $n \geqslant 0$; put $A_{-1} = 0$. Then, if $0 \leqslant p \leqslant q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

Proof.

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q}(A_n - A_{n-1})b_n = \sum_{n=p}^{q} A_n b_n - \sum_{n=p}^{q-1} A_n b_{n+1} = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

□

Theorem 3.42 Suppose

a) the partial sums A_n of $\sum a_n$ form a bounded sequence;

b) $b_0 \geqslant b_1 \geqslant b_2 \geqslant \dots$

c) $\lim_{n\to\infty} b_n = 0$

Then $\sum a_n b_n$ converges.

Proof. Choose M such that $|A_n| \leqslant M$ for all n. Given $\varepsilon > 0$, there is an integer N such that $b_N \leqslant (\varepsilon/2M)$. For $N \leqslant p \leqslant q$, we have

$$\left|\sum_{n=p}^{q} a_n b_n\right| = \left|\sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1}b_p\right| \leqslant M\left|\sum_{n=p}^{q-1}(b_n - b_{n+1}) + b_q - b_p\right| = 2Mb_{\leqslant} Mb_N \leqslant \varepsilon.$$

Convergence now follows from the Cauchy criterion. We note that the first inequality in the above chain depends of course on the fact that $b_n - b_{n+1} \geqslant 0$. □

Theorem 3.43 (Alternative Series Test) Suppose

a) $|c_1| \geqslant |c_2| \geqslant |c_3| \geqslant \dots;$

b) $c_{2m-1} \geqslant 0$, $c_{2m} \leqslant 0$ $(m = 1, 2, 3, \dots)$

c) $\lim_{n\to\infty} c_n = 0$

Then $\sum c_n$ converges

Proof. Apply Theorem 3.42, with $a_n = (-1)^{n+1}$, $b_n = |c_n|$ □

Theorem 3.44 Suppose the radius of convergence of $\sum c_n z^n$ is 1, and suppose $c_0 \geqslant c_1 \geqslant c_2 \geqslant \dots$, $\lim_{n\to\infty} c_n = 0$. Then $\sum c_n z^n$ converges at every point on the circle $|z| = 1$, except possibly at $z = 1$.

Proof. Put $a_n = z^n$, $b_n = c_n$. The hypotheses of Theorem 3.42 are then satisfied, since $|A_n| = |\sum_{m=0}^{n} z^m| = \left|\frac{1-z^{n+1}}{1-z}\right| \leqslant \frac{2}{|1-z|}$, if $|z| = 1$, $z \neq 1$ □

Theorem 3.45 If $\sum a_n$ converges absolutely, then $\sum a_n$ converges.

Proof. The assertion follows from the inequality $|\sum_{k=n}^{m} a_k| \leqslant \sum_{k=n}^{m} |a_k|$, plus the Cauchy criterion. □

Theorem 3.47 If $\sum a_n = A$ and $\sum b_n = B$, then $\sum(a_n + b_n) = A + B$ and $\sum ca_n = cA_n$ for any fixed c.

Proof. Let $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$. Then $A_n + B_n = \sum_{k=0}^{n}(a_k + b_k)$. Since $\lim_{n\to\infty} A_n = A$ and $\lim_{n\to\infty} B_n = B$, we see that $\lim_{n\to\infty}(A_n + B_n) = A + B$. The proof of the second assertion follows. ☐

Theorem 3.50 Suppose

a) $\sum_{n=0}^{\infty} a_n$ converges absolutely

b) $\sum_{n=0}^{\infty} a_n = A$

c) $\sum_{n=0}^{\infty} b_n = B$

d) $c_n = \sum_{k=0}^{n} a_k b_{n-k} \qquad (n = 0, 1, 2, \dots)$

Then $\sum_{n=0}^{\infty} c_n = AB$.

Proof. Put $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$, $C_n = \sum_{k=0}^{n} c_k$, $\beta_n = B_n - B$. Then:

$$\begin{aligned} C_n &= a_0 b_0) a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + \cdots + a_n B_0 \\ &= a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \cdots + a_n(B + \beta_0) \\ &= A_n B + a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0 \end{aligned}$$

Put $\gamma_n = a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0$. We wish to show that $C_n \to AB$. Since $A_n B \to AB$, it suffices to show that $\lim_{n\to\infty} \gamma_n = 0/$ Put $\alpha = \sum_{n=0}^{\infty} |a_n|$. It is here that we use (a). Let $\varepsilon > 0$ be given. By (c), $\beta_n \to 0$. Hence we can choose N such that $|\beta_n| \leqslant \varepsilon$ for $n \geqslant N$, in which case

$$|\gamma_n| \leqslant |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + |\beta_{N+1} a_{n-N-1} + \cdots + \beta_n a_0| \leqslant |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + \varepsilon\alpha.$$

Keeping N fixed, and letting $n \to \infty$, we get $\limsup_{n\to\infty} |\gamma_n| \leqslant \varepsilon\alpha$, since $a_k \to 0$ as $k \to \infty$. Since ε is arbitrary, $\lim_{n\to\infty} \gamma_n = 0$ follows. ☐

Theorem 3.51 If the series $\sum a_n = A$, $\sum b_n = B$, $\sum c_n = C$, and $c_n = a_0 b_n + \cdots + a_n b_0$, then $C = AB$.

Theorem 3.54 Let $\sum a_n$ be a series of real numbers which converges, but not absolutely. Suppose $-\infty \leqslant \alpha \leqslant \beta \leqslant \infty$. Then there exists a rearrangement $\sum a'_n$ with partial sums s'_n such that

$$\liminf_{n\to\infty} s'_n = \alpha \qquad \limsup_{n\to\infty} s'_n = \beta$$

Proof. Let $p_n = \frac{|a_n| + a_n}{2}$, $q_n = \frac{|a_n| - a_n}{2}$ for $(n = 1, 2, 3, \dots)$. Then $p_n - q_n = a_n$, $p_n + q_n = |a_n|$, $p_n, q_n \geqslant 0$. The series $\sum p_n$, $\sum q_n$ must both diverge. If, on the contrary, both were convergent, the $\sum(p_n + q_n) = \sum |a_n|$ would converge, contradicting the hypothesis. Since $\sum_{n=1}^{N} a_n = \sum_{n=1}^{N}(p_n - q_n) = \sum_{n=1}^{N} p_n - \sum_{n=1}^{N} q_n$, divergence of $\sum p_n$ and convergence of $\sum q_n$ or vice versa implies divergence of $\sum a_n$, again contradicting the hypothesis.

Now let $P_1, P_2, P_3, \dots$ denote the nonnegative terms of $\sum a_n$, in the order in which they occur, and let $Q_1, Q_2, Q_3, \dots$ be the absolute values of the negative terms of $\sum a_n$, also in their original order. The series $\sum P_n$, $\sum Q_n$ differ from $\sum p_n$, $\sum q_n$ differ only by zero terms, and are therefore divergent. We shall construct sequences $\{m_n\}, \{k_n\}$, such that the series

$$P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \dots P_{m_2} - Q_{k_1+1} - \cdots - Q_{k_2} + \dots,$$

which is clearly a rearrangement of $\sum a_n$, satisfies

$$\liminf_{n\to\infty} s'_n = \alpha \qquad \limsup_{n\to\infty} s'_n = \beta$$

Choose real-valued sequences $\{\alpha_n\}, \{\beta_n\}$ such that $\alpha_n \to \alpha$, $\beta_n \to \beta$, $\alpha_n < \beta_n$, $\beta_1 > 0$.

Let m_1, k_1 be the smallest integers such that

$$P_1 + \cdots + P_{m_1} > \beta_1 \quad \text{and} \quad P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} < \alpha_1$$

Let m_2, k_2 be the smallest integers such that

$$P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \dots P_{m_2} > \beta_2,$$

$$P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \dots P_{m_2} - Q_{k_1+1} - \cdots - Q_{k_2} < \alpha_2;$$

and continue in this way. This is possible since $\sum P_n$ and $\sum Q_n$ diverge.

If x_n, y_n denote the partial sums of the string whose last terms are P_{m_n} and $-Q_{k_n}$, then $|x_n - \beta_n| \leqslant P_{m_n}$, $|y_n - \alpha_n| \leqslant Q_{k_n}$. Since $P_n \to 0$ and $Q_n \to 0$ as $n \to \infty$, we see that $x_n \to \beta$, $y_n \to \alpha$.

Finally, it is clear that no number less than α or greater than β can be a subsequential limit of the partial sums. □

Theorem 3.55 If $\sum a_n$ is a series of complex numbers which converges absolutely, then every rearrangement of $\sum a_n$ converges, and they all converge to the same sum.

Proof. Let $\sum a'_n$ be a rearrangement, with partial sums s'_n. Given $\varepsilon > 0$, there exists an integer N such that $m \geqslant n \geqslant N$ implies $\sum_{i=n}^{m} |a_i| \leqslant \varepsilon$. Now choose p such that the integers $1, 2, 3, \dots, N$ are all contained in the set $k_1, k_2, \dots, k_p$. Then if $n > p$, the numbers $a_1, \dots, a_N$ will cancel in the difference $s_n - s'_n$, so that $|s_n - s'_n| \leqslant \varepsilon$. Hence $\{s'_n\}$ converges to the same sum as $\{s_n\}$. □

Chapter 4 Continuity

Definitions

Limit of a Function Let X and Y be metric spaces; suppose $E \subset X$, f maps E into Y and p is a limit point of E. We write $f(x) \to q$ as $x \to p$ or

$$\lim_{x \to p} f(x) = q$$

if there is a point $q \in Y$ with the following property: Fore very $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), q) < \varepsilon$ for all points $x \in E$ for which $0 < d_X(x, p) < \delta$.

Sum/ Difference/ Product/ Quotient of Function Suppose we have two complex functions, f and g, both defined on E. By $f + g$ we mean the function which assigns to each point x of E the number $f(x) + g(x)$. Similarly we define $f - g$, fg, f/g for $g(x) \neq 0$. If f assigns to each point of x of E the same number c, f is said to be a constant function, or constant, and we write $f = c$. If f and g are real functions then $f(x) \geqslant g(x)$ is the same as $f \geqslant g$. The same holds for $\mathbf{f}, \mathbf{g} : E \to \mathbb{R}^K$.

Continuous Suppose X an Y are metric spaces, $E \subset X$, $p \in E$, and $f : E \to Y$. Then f is said to be *continuous at* p if for every $\varepsilon > 0$ there exists a δ such that

$$d_Y(f(x), f(p)) < \varepsilon$$

for all points $x \in E$ for which $d_X(x, p) < \delta$.

If f is continuous at every point of E, then f is said to be *continuous on* E.

Bounded A mapping $\mathbf{f}$ of a set E into $\mathbb{R}^k$ is said to be *bounded* if there is a real number M such that $|\mathbf{f}(x)| \leqslant M$ for all $x \in E$.

Uniformly Continuous Let f be a mapping of a metric space X into a metric space Y. We say that f is *uniformly continuous* on X if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(p), f(q)) < \varepsilon$$

for all p and q in X for which $d_X(p, q) < \delta$.

One-Sided Limit of a Function Let f be defined on (a, b). Consider any point x such that $a \leqslant x < b$. We write:

$$f(x+) = q$$

if $f(t_n) \to q$ as $n \to \infty$, for all sequences $\{t_n\}$ in (x, b) such that $t_n \to x$. To obtain the definition of $f(x-)$, for $a < x \leqslant b$, we restrict ourselves to sequences $\{t_n\}$ in (a, x).

It is clear that at any point $x \in (a, b)$ the $\lim_{t \to x} f(t)$ exists if and only if

$$f(x+) = f(x-) = \lim_{t \to x} f(t).$$

Simple Discontinuity Let f be defined on (a,b). If f is discontinuous at a point x and if $f(x+)$ and $f(x-)$ exists, then f is said to have a discontinuity of the *first kind*, or a *simple discontinuity*, at x. Otherwise the discontinuity is said to be of the *second kind*.

There are two ways in which a function can have a simple discontinuity:

$$f(x+) \neq f(x-) \qquad \text{or} \qquad f(x+) = f(x-) \neq f(x).$$

Monotonic Let f be real on (a,b). Then f is said to be *monotonically increasing* on (a,b) if $a < x < y < b$ implies $f(x) \leqslant f(y)$. If the last inequality is reversed, we obtain the definition of a *monotonically decreasing* function. The class of monotonic functions consists of both the increasing and the decreasing functions.

Neighborhood of Infinity For any real c, the set of real numbers x such that $x > c$ is called a neighborhood of $+\infty$ and is written $(c,+\infty)$. Similarly, the set $(-\infty, c)$ is a neighborhood of $-\infty$.

Limit at Infinity Let f be a real function defined on $E \subset \mathbb{R}$. We say that $f(t) \to A$ as $t \to x$, where A and x are in the extended real number system, if for every neighborhood U of A there is a neighborhood V of x such that $V \cap E$ is not empty, and such that $f(t) \in U$ for all $t \in V \cap E$, $t \neq x$.

Theorems

Theorem 4.2 Let X and Y be metric spaces; suppose $E \subset X$, f maps E into Y and p is a limit point of E. Then $\lim_{x\to p} f(x) = q$ if and only if $\lim_{n\to\infty} f(p_n) = q$ for every sequence $\{p_n\}$ in E such that $p_n \neq p$ $\lim_{n\to\infty} p_n = p$.

Proof. Suppose that $\lim_{x\to p} f(x) = q$. Choose $\{p_n\}$ in E as stipulated. Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that $d_Y(f(x),q) < \varepsilon$ if $x \in E$. and $0 < d_X(x,p) < \delta$. Also, there exists N such that $n > N$ implies $0 < d_X(p_n, p) < \delta$. Thus for $n > N$ we have $d_Y(f(p_n), q) < \varepsilon$, which show s that holds.

Conversely, suppose $\lim_{x\to p} f(x) \neq q$. then there exists some $\varepsilon > 0$ such that for every $\delta > 0$ there exists a point $x \in E$ (depending on δ), for which $d_Y(f(x),q) \geqslant \varepsilon$ but $0 < d_X(x,p) < \delta$. Taking $\delta_n = \frac{1}{n}$ we thus find a sequence $\{p_n\}$ in E satisfying the aforementioned requirements, for which $\lim_{n\to\infty} f(p_n) = q$ is false. □

Corollary If f has a limit at p, this limit is unique.

Theorem 4.4 Suppose $E \subset X$, a metric space, p is a limit point of E, f and g are complex functions on E, and

$$\lim_{x\to p} f(x) = A \qquad \lim_{x\to p} g(x) = B$$

Then:

a) $\lim_{x\to p}(f+g)(x) = A + B$

b) $\lim_{x \to p}(fg) = AB$

c) $\lim_{x \to p}\left(\frac{f}{g}\right) = \frac{A}{B}$

Proof. In view of Theorem 4.2, these assertions follow immediately from the analogous properties of sequences in Theorem 3.3. □

Theorem 4.6 Suppose X an Y are metric spaces, $E \subset X$, $p \in E$ such that p is a limit point of E, and $f : E \to Y$. Then f is continuous at P if and only if $\lim_{x \to p} f(x) = f(p)$.

Proof. This is clear if we compare Definitions 4.1 and 4.5. □

Theorem 4.7 Suppose X, Y, Z are metric spaces, $E \subset X$, f maps E into Y. g maps the range of f, $f(E)$, into Z, and H is the mapping of E into Z defined by

$$h(x) = g(f(x)) \qquad (x \in E)$$

If f is continuous at a point $p \in E$ and if g is continuous at the point $f(p)$, then h is continuous at P. The function h is called the *composition* or the *composite* of f and g, most commonly noted: $h = g \circ f$.

Proof. Let $\varepsilon > 0$ be given. Since g is continuous at $f(p)$, there exists $\eta > 0$ such that $d_Z(g(y), g(f(p))) < \varepsilon$ if $d_Y(y, f(p)) < \eta$ and $y \in f(E)$. Since f is continuous at p, there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \eta$ if $d_X(x,p) < \delta$ and $x \in E$. It follows that $d_Z(h(x), h(p)) = d_Z(g(f(x)), g(f(p))) < \varepsilon$. if $d_X(x,p) < \delta$ and $x \in E$. Thus h is continuous at p. □

Theorem 4.8 A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y.

Proof. Suppose f is continuous on X and V is an open set in Y. We have to show that every point of $f^{-1}(V)$ is an interior point of $f^{-1}(V)$. So, suppose $p \in X$ and $f(p) \in V$. Since V is open, there exists $\varepsilon > 0$ such that $y \in V$ if $d_Y(f(p), y) < \varepsilon$; and since f is continuous at p, there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \varepsilon$ if $d_X(x,p) < \delta$. Thus $x \in f^{-1}(V)$ as soon as $d(x,p) < \delta$.

Conversely, suppose $f^{-1}(V)$ is open in X for every open set V in Y . Fix $p \in X$ and $\varepsilon > 0$, let V be the set of all $y \in Y$ such that $d_Y(y, f(p)) < \varepsilon$. Then V is open; hence $f^{-1}(V)$ is open; hence there exists $\delta > 0$ such that $x \in f^{-1}(V)$ as soon as $d_X(p,x) < \delta$. But if $x \in f^{-1}(V)$, then $f(x) \in V$, so that $d_Y(f(x), f(p)) < \varepsilon$. □

Corollary A mapping f of a metric space X into a metric space Y is continuous if and only if $f^{-1}(C)$ is closed in X for every closed set C in Y. (Note, $f^{-1}(E^c) = [f^{-1}(E)]^c$).

Theorem 4.9 Let f and g be complex continuous functions on a metric space X. Then $f + g$, fg, f/g are continuous on X.

Proof. At isolated points of X there is nothing to prove. At limit points, the statement follows from Theorems 4.4 and 4.6. □

Theorem 4.10 a) Let $f_1, \dots, f_k$ be real functions on a metric space X, and let $\mathbf{f}$ be the mapping of X into $\mathbb{R}^k$ defined by $\mathbf{f}(x) = (f_1(x), \dots, f_k(x))$, then $\mathbf{f}$ is continuous if and only if each of the functions f_i is continuous.

b) If $\mathbf{f}$ and $\mathbf{g}$ are continuous mappings of X into $\mathbb{R}^k$, then $\mathbf{f}+\mathbf{g}$ and $\mathbf{f}\cdot\mathbf{g}$ are continuous on X.

The functions f_i are called the *components* of $\mathbf{f}$.

Proof. Part (a) follows from the inequalities

$$|f_j(x) - f_j(y)| \leqslant |\mathbf{f}(x) - \mathbf{f}(y)| = \left(\sum_{i=1}^{k} |f_i(x) - f_i(y)|^2\right)^{\frac{1}{2}}$$

for $j = 1, \dots, k$. Part (b) follows from (a) and Theorem 4.9. □

Theorem 4.14 Suppose f is a continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact.

Proof. Let $\{V_\alpha\}$be an open cover of $f(X)$. Since f is continuous, Theorem 4.8 shows that each of the sets $f^{-1}(V_\alpha)$ is open. Since X is compact, there are finitely many indices, say $\alpha_1, \dots, \alpha_n$, such that $X \subset f^{-1}(V_{\alpha_1}) \cup \cdots \cup f^{-1}(V_{\alpha_n})$. Since $f(f^{-1}(E)) \subset E$ for every $E \subset Y$, this implies that $f(X) \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}$. □

Theorem 4.15 If $\mathbf{f}$ is a continuous mapping of a compact metric space X into $\mathbb{R}^k$, then $\mathbf{f}(X)$ is closed and bounded. Thus, $\mathbf{f}$ is bounded.

Proof. This follows from Theorem 2.41 □

Theorem 4.16 Suppose f is a continuous real function on a compact metric space X, and

$$M = \sup_{p \in X} f(p) \qquad m = \inf_{p \in X} f(p)$$

Then there exists points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$. That is, f attains it's maximum and minimum.

Proof. By Theorem 4.15 $f(X)$ is closed and bounded set of real numbers; hence $f(X)$ contains $M = \sup f(X)$ and $m = \inf f(X)$, by Theorem 2.28. □

Theorem 4.17 Suppose f is a continuous 1-1 mapping of a compact metric space X onto a metric space Y. Then the inverse mapping f^{-1} defined on Y by

$$f^{-1}(f(x)) = x \qquad (x \in X)$$

is a continuous mapping of Y onto X.

Proof. Applying Theorem 4.8 to f^{-1} in place of f we see that it suffices to prove that $f(V)$ is an open set in Y for every open set V in X. Fix such a set V.

The complement V^c of V is closed in X, hence compact by Theorem 2.35; hence $f(V^c)$ is a compact subset of Y by Theorem 4.14 and so is closed in Y by Theorem 2.34. Since f is one-to-one and onto, $f(V)$ is the complement of $f(V^c)$. Hence $f(V)$ is open. □

Theorem 4.19 Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is uniformly continuous on X.

Proof. Let $\varepsilon > 0$ be given. Since f is continuous, we can associate to each point $p \in X$ a positive number $\phi(p)$ such that $q \in X$, $d_X(p,q) < \phi(p)$ implies $d_Y(f(p), f(q)) < \frac{\varepsilon}{2}$. Let $J(p)$ be the set of all $q \in X$ for which $d_X(p,q) < \frac{1}{2}\phi(p)$. Since $p \in J(p)$, the collection of all sets $J(p)$ is an open cover of X; and since X is compact, there is a finite set of points $p_1, \ldots, p_n$ in X, such that $X \subset J(p_1) \cup \cdots \cup J(p_n)$. We put $\delta = \frac{1}{2}\min[\phi(p_1), \ldots, \phi(p_n)]$. Then $\delta > 0$. (That is, because the minimum of a finite set of positive numbers is positive. On the contrary the inf of an infinite set of positive numbers may be 0.)

Now let q and p be points of X, such that $d_X(p,q) < \delta$. By our finite open cover, there is an integer m, $1 \leqslant m \leqslant n$, such that $p \in J(p_m)$; hence $d_X(p, p_m) < \frac{1}{2}\phi(p_m)$, and we also have $d_X(q, p_m) \leqslant d_X(p,q) + d_X(p, p_m) < \delta + \frac{1}{2}\phi(p_m) \leqslant \phi(p_m)$. Finally, because $d_Y(f(p), f(q)) < \frac{\varepsilon}{2}$ when $d_X(p,q) < \phi(p)$, $d_Y(f(p), f(q)) \leqslant d_Y(f(p), f(p_m)) + d_Y(f(q), f(p_m)) < \varepsilon$. □

Theorem 4.20 Let E be a non-compact set in $\mathbb{R}$. Then

a) there exists a continuous function of E which is not bounded;

b) there exists a continuous and bounded function on E which has no maximum.

c) If, in addition, E is bounded then: there exists a continuous function on E which is not uniformly continuous.

Proof. Suppose first that E is bounded so that there exists a limit point x_0 of E which is not a point of E. Consider $f(x) = \frac{1}{x - x_0}$. This is continuous on E, but evidently unbounded. To see that it is not uniformly continuous, let $\varepsilon > 0$ and $\delta > 0$ be arbitrary, and choose a point $x \in E$ such that $|x - x_0| < \delta$. Taking t close enough to x_0, we can then make the difference $|f(t) - f(x)|$ greater than ε, although $|t - x| < \delta$. Since this is true for every $\delta > 0$, f is not uniformly continuous on E.

The function g given by $g(x) = \frac{1}{1 + (x - x_0)^2}$ is continuous on E, and is bounded, since $0 < g(x) < 1$. It is clear that $\sup_{x \in E} g(x) = 1$, whereas $g(x) < 1$ for all $x \in E$. Thus g has no maximum on E.

Having proved the theorem for bounded sets E, let us now suppose that E is unbounded. Then $f(x) = x$ establishes (a), whereas $h(x) = \frac{x^2}{1 + x^2}$ establishes (b), since $\sup_{x \in E} h(x) = 1$ and $h(x) < 1$ for all $x \in E$.

Assertion (c) would be false if boundedness were omitted from the hypotheses. For, let E be the set of all integers. Then every function defined on E is uniformly continuous on E. To see this we need merely take $\delta < 1$ in Def 4.18. □

Theorem 4.22 If f is a continuous mapping of a metric space X into a metric space Y, and if E is a connected subset of X, then $f(E)$ is connected.

Proof. Assume, to the contrary, that $f(E) = A \cup B$ where A and B are nonempty separated subsets of Y. Put $G = E \cap f^{-1}(A)$, $H = E \cap f^{-1}(B)$. then $E = G \cup H$, and neither G nor H is empty. Since $A \subset \overline{A}$, we have $G \subset f^{-1}(\overline{A})$; the latter set is closed, since f is continuous; hence $\overline{G} \subset f^{-1}(\overline{A})$. It follows that $f(\overline{G}) \subset \overline{A}$. Since $f(H) = B$ and $\overline{A} \cap B$ is empty, we conclude that $\overline{G} \cap H$ is empty.

The same argument shows that $G \cap \overline{H}$ is empty. Thus G and H are separated. This is impossible if E is connected. □

Theorem 4.23 (Intermediate Value Theorem) Let f be a continuous real function on the interval $[a, b]$. If $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.

Proof. By Theorem 2.47, $[a, b]$ is connected; hence Theorem 4.22 shows that $f([a, b])$ is a connected subset of $\mathbb{R}$, and the assertion follows if we appeal once more to Theorem 2.47. □

Theorem 4.29 Let f be monotonically increasing on (a, b). Then $f(x+)$ and $f(x-)$ exists at every point of x of (a, b). More precisely,

$$\sup_{a<t<x} f(t) = f(x-) \leqslant f(x) \leqslant f(x+) = \inf_{x<t<b} f(t)$$

Furthermore, if $a < x < y < b$, then

$$f(x+) \leqslant f(y-)$$

. Analogous results hold for monotonically decreasing functions.

Proof. By hypothesis, the set of numbers $f(t)$, where $a < t < x$, is bounded above by the number $f(x)$, and therefore has a least upper bound which we shall denote by A. Evidently $A \leqslant f(x)$. We have to show that $A = f(x-)$.

Let $\varepsilon > 0$ be given. It follows from the definition of A as a least upper bound that there exists $\delta > 0$ such that $a < x - \delta < x$. and $A - \varepsilon < f(x - d) \leqslant A$. Since f is monotonic, we have, for $x - \delta < t < x$, $f(x - \delta) \leqslant f(t) \leqslant A$. Combining these two, we see that $|f(t) - A| < \varepsilon$. Hence $f(x-) = A$.

Likewise, $f(x+) = \inf_{x<t<b} f(t)$.

If $a < x < y < b$, we see that $f(x+) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t)$. The last equality is obtained by applying the aforementioned equality to (a, y) in place of (a, b). Similarly $f(x-) = \sup_{a<t<y} f(t) = \sup_{x<t<y} f(t)$. Comparison of these two strings of equalities gives us that $f(x+) \leqslant f(y-)$. □

Corollary Monotonic functions have no discontinuities of the second kind.

Theorem 4.30 Let f be monotonic on (a, b). Then the set of points of (a, b) at which f is discontinuous is at most countable.

Proof. Suppose, for the sake of definiteness, that f is increasing, and let E be the set of points at which f is discontinuous. With every point x of E we associate a rational number $r(x)$ such that $f(x-) < r(x) < f(x+)$. Since $x_1 < x_2$ implies $f(x_1+) \leqslant f(x_2-)$, we see that $r(x_1) \neq r(x_2)$ if $x_1 \neq x_2$.

We have thus established a 1-1 correspondence between the set E and a subset of the set of rational numbers. The latter, as we know, is countable. □

Theorem 4.34 Let f and g be defined on $E \subset \mathbb{R}$ Suppose

$$f(t) \to A \qquad g(t) \to B \qquad \text{as } t \to x$$

Then

a) $f(t) \to A'$ implies $A' = A$

b) $(f + g)(x) \to A + B$

c) $(fg)(t) \to AB$

d) $(f/g)(t) \to A/B$

Note: $\infty - \infty$, $0 \cdot \infty$, ∞/∞, $A/0$ are not defined.

Chapter 5 Differentiation

Definitions

Deivative Let f be defined (and real-valued) on $[a, b]$. For any $x \in [a, b]$ form the quotient

$$\phi(t) = \frac{f(t) - f(x)}{t - x} \qquad (a < t < b, t \neq x)$$

and define

$$f'(x) = \lim_{t \to x} \phi(t)$$

provided this limit exists.

We thus associate with the function f and a function f' whose domain is the set of points x at which the above limit exists; f' is called the *derivative* of f.

Differentiable If f' is defined at a point x we say that f is *differentiable* at x. If f' is defined at every point of a set $E \subset [a, b]$, we say that f is *differentiable* on E.

Local Maximum Let f be a real function defined on a metric space X. We say that f has a *local maximum* at a point $p \in X$ if there exists $\delta > 0$ such that $f(q) \leqslant f(p)$ for all $q \in X$ with $d(p.q) < \delta$. (Local minima are defined likewise.)

Higher Order Derivatives If f has a derivative f' on an interval. and if f' is itself differentiable, we denote the derivative of f' by f'' and call f'' the *second derivative* of f. Continuing in this manner, we obtain functions

$$f, f', f'', f^{(3)}, f^{(4)}, \ldots, f^{(n)}$$

each of which is the derivative of the preceding one. $f^{(n)}$ is called the *nth derivative*, or the derivative of order n, of f.

Note, In order for $f^{(n)}(x)$ to exist at a point x, $f^{(n-1)}(t)$ must exist in a neighborhood of x (or in a one-sided neighborhood, if x is an endpoint of the interval on which f is defined), and $f^{(n-1)}(x)$ must be differentiable at x.

Theorems

Theorem 5.2 Let f be defined on $[a, b]$. If f is differentiable at a point $x \in [a, b]$, then f is continuous at x.

Proof. As $t \to x$, we have, by Theorem 4.4, $f(t) - f(x) = \frac{f(t)-f(x)}{t-x} \cdot (t - x) \to f'(x) \cdot 0 = 0$. □

Theorem 5.3 Suppose f and g are defined on $[a, b]$ and are differentiable at a point $x \in [a, b]$. Then $f + g$, fg, f/g are differentiable at x, and:

a) $(f+g)'(x) = f'(x) + g'(x)$

b) $(fg)'(x) = f'(x)g(x) + g'(x)f(x)$

c) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-g'(x)f(x)}{g(x)^2}$

Proof. (a) is clear, by Theorem 4.4. Let $h = fg$. Then $h(t) = h(x) = f(t)[g(t) - g(x)] + g(x)[f(t) - f(x)]$. If we divide this by $t - x$ and not that $f(t) \to f(x)$ as $t \to x$ (Theorem 5.2) then (b) follows. Next let $h = f/g$. Then

$$\frac{h(t)-h(x)}{t-x} = \frac{1}{g(t)g(x)}\left[g(x)\frac{f(t)-f(x)}{t-x} - f(x)\frac{g(t)-g(x)}{t-x}\right].$$

Letting $t \to x$, and applying Theorem 4.4 and 5.2, we obtain (c). □

Theorem 5.5 (Chain Rule) Suppose f is continuous on $[a,b]$, $f'(x)$ exists at some point $x \in [a,b]$, g is defined on an interval I which contains the range of f, and g is differentiable at the point $f(x)$. If

$$h(t) = g(f(t)) \qquad (a \leqslant t \leqslant b)$$

then h is differentiable at x, and

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

Proof. Let $y = f(x)$. By definition of the derivative, we have $f(t) - f(x) = (t - x)[f'(x) + u(t)]$ and $g(s) - g(y) = (s-y)[g'(y) + v(s)]$, where $t \in [a,b]$, $s \in I$, and $u(t) \to 0$ as $t \to x$ and $v(s) \to 0$ as $s \to y$. Let $s = f(t)$. Using this, we obtain $h(t) - h(x) = g(f(t)) - g(f(x)) = [f(t) - f(x)] \cdot [g'(y) + v(s)] = (t-x) \cdot [f'(x) + u(t)] \cdot [g'(y) + v(s)]$, or, if $t \neq x$,

$$\frac{h(t)-h(x)}{t-x} = [g'(y) + v(s)] \cdot [f'(x) + u(t)].$$

Letting $t \to x$, we see that $s \to y$, by the continuity of f, so that the right side of the previous equation tends to $g'(y)f'(x)$. □

Theorem 5.8 (Rolle's Theorem) Let f be defined on $[a,b]$; if f has a local maximum at a point $x \in (a,b)$, and if $f'(x)$ exists, then $f'(x) = 0$. (The analogous statement for local minima also holds.)

Proof. Choose δ in accordance with Definition 5.7, so that $a < x-\delta < x < x+\delta < b$. If $x - \delta < t < x$, then $\frac{f(t)-f(x)}{t-x} \geqslant 0$. Letting $t \to x$, we see that $f'(x) \geqslant 0$. If $x < t < x+\delta$, then $\frac{f(t)-f(x)}{t-x} \leqslant 0$, which shows that $f'(x) \leqslant 0$. Hence $f'(x) = 0$. □

Theorem 5.9 (Generalized Mean Value Theorem) If f and g are continuous real functions on $[a,b]$ which are differentiable in (a,b), then there is a point $x \in (a,b)$ at which:

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x)$$

Proof. Put $h(t) = [f(b)-f(a)]g(t)-[g(b)-g(a)]f(t)$. Then h is continuous on $[a, b]$, h is differentiable in (a, b), and $h(a) = f(b)g(a) - f(a)g(b) = h(b)$. To prove the theorem, we have to show that $h'(x) = 0$ for some $x \in (a, b)$.

If h is constant, this holds for every $x \in (a, b)$. If $h(t) > h(a)$ for some $t \in (a, b)$, let x be a point on $[a, b]$ at which h attains its maximum (Theorem 4.16). Because $h(a) = h(b)$ we know that $x \in (a, b)$, and Theorem 5.8 shows that $h'(x) = 0$. If $h(t) < h(a)$ for some $t \in (a, b)$, the same argument applies if we choose for x a point on $[a, b]$ where h attains its minimum. □

Theorem 5.10 (Mean Value Theorem) If f is a real continuous function on $[a, b]$ which is differentiable in (a, b), then there is a point $x \in (a, b)$ at which

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

Proof. Take $g(x) = x$ in Theorem 5.9. □

Theorem 5.11 Suppose f is differentiable in (a, b).

a) If $f'(x) \geqslant 0$ for all $x \in (a, b)$, then f is monotonically increasing.
b) If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.
c) If $f'(x) \leqslant 0$ for all $x \in (a, b)$, then f is monotonically decreasing.

Proof. All conclusions can be read off from the equation $f(x_2)-f(x_1) = (x_2-x_1)f'(x)$, which is valid, for each pair of numbers x_1, x_2 in (a, b) for some x between x_1 and x_2. □

Theorem 5.12 (Intermediate Value Theorem for Derivatives) Suppose f is a real differentiable function on $[a, b]$ and suppose $f'(a) < \lambda < f'(b)$. Then there is a point $x \in (a, b)$ such that $f'(x) = \lambda$.

Proof. Put $g(t) = f(t) - \lambda t$. Then $g'(a) < 0$, so that $g(t_1) < g(a)$ for some $t_1 \in (a, b)$ and $g'(b) > 0$, so that $g(t_2) < g(b)$ for some $t_2 \in (a, b)$. Hence g attains its minimum on $[a, b]$ (Theorem 4.16) at some point x such that $a < x < b$. By Theorem 5.8, $g'(x) = 0$. Hence $f'(x) = \lambda$. □

Corollary If f is differentiable on $[a, b]$, then f' cannot have any simple discontinuities on $[a, b]$.

Theorem 5.13 (L'Hospital's Rule) Suppose f and g are real and differentiable in (a, b), and $g'(x) \neq 0$ for all $x \in (a, b)$, where $-\infty \leqslant a < b \leqslant +\infty$. Suppose

$$\frac{f'(x)}{g'(x)} \to A \quad as \quad x \to a.$$

If

$$f(x) \to 0 \quad and \quad g(x) \to 0 \quad as \quad x \to a$$

or if

$$g(x) \to +\infty \qquad as \qquad x \to a,$$

then

$$\frac{f(x)}{g(x)} \to A \qquad as \qquad x \to a$$

The analogous statement is also true if $x \to b$ or if $g(x) \to -\infty$.

Proof. We first consider the case in which $-\infty \leqslant A < +\infty$. Choose a real number q such that $A < q$, and then choose r such that $A < r < q$. Because $\frac{f'(x)}{g'(x)} \to A$ as $x \to a$, there is a point $c \in (a, b)$ such that $a < x < c$ implies $\frac{f'(x)}{g'(x)} < r$. If $a < x < y < c$, then Theorem 5.9 shows that there is a point $t \in (x, y)$ such that $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(t)}{g'(t)} < r$. Suppose that $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$.. Letting $x \to a$ in the previous equation, we see that $\frac{f(y)}{g(y)} \leqslant r < q$ $(a < y < c)$.

Next, suppose that $g(x) \to +\infty$ as $x \to a$. Keeping y fixed in $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(t)}{g'(t)} < r$, we can choose a point $c_1 \in (a, y)$ such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < c_1$. Multiplying this by $[g(x) - g(y)]/g(x)$, we obtain $\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}$. If we let $x \to a$ in this equation, we know that there is a point $c_2 \in (a, c_1)$ such that $\frac{f(x)}{g(x)} < q$ for $(a < x < c_2)$.

Summing up $\frac{f(y)}{g(y)} \leqslant r < q$ and $\frac{f(x)}{g(x)} < q$ show that for any q, subject only to the condition $A < q$, there is a point c_2 such that $f(x)/g(x) < q$ if $a < x < c_2$.

In the same manner, if $-\infty < A \leqslant +\infty$, and p is chosen so that $p < A$, we can find a point c_3 such that $p < \frac{f(x)}{g(x)}$ for $a < x < c_3$. And thus the conclusion follows. □

Theorem 5.15 (Taylor's Theorem) Suppose f is a real function on $[a, b]$, n is a positive integer, $f^{(n-1)}$ is continuous on $[a, b]$, $f^{(n)}(t)$ exists for every $t \in (a, b)$. Let α, β be distinct points of $[a, b]$, and define:

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k$$

Then there exists a point x between α and β such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n.$$

Proof. Let M be the number defined by $f(\beta) = P(\beta) + M(\beta - \alpha)^n$ and put $g(t) = f(t) - P(t) - M(t - \alpha)^n$. We have to show that $n!M = f^{(n)}(x)$ for some x between α and β. By $P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k$ and $g(t) = f(t) - P(t) - M(t - \alpha)^n$,

$$g^{(n)}(t) = f^{(n)}(t) - n!M.$$

Hence the proof will be complete if we can show that $g^{(n)}(x) = 0$ for some x between α and β.

Since $P^k(\alpha) = f^{(k)}(\alpha)$ for $k = 0, \ldots, n-1$, we have $g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0$. Out choice of M shows that $g(\beta) = 0$, so that $g'(x_1) = 0$ for some x_1 between α and β, by the mean value theorem. Since $g'(\alpha) = 0$, we conclude similarly that $g''(x_2) = 0$ for some x_2 between α and x_1. After n steps we arrive at the conclusion that $g^{(n)}(x_n) = 0$ for some x_n between α and x_{n-1}, that is, between α and β.

□

Theorem 5.19 Suppose $\mathbf{f}$ is a continuous mapping of $[a, b]$ into $\mathbb{R}^k$ and $\mathbf{f}$ is differentiable in (a, b). Then there exists $x \in (a, b)$ such that

$$|\mathbf{f}(b) - \mathbf{f}(a)| \leqslant (b-a)|\mathbf{f}'(x)|.$$

Proof. Put $\overline{e} = \mathbf{f}(b) - \mathbf{f}(a)$, and define $\varphi(t) = \overline{e} \cdot \mathbf{f}(t)$ for $a \leqslant t \leqslant b$. The φ is a real-valued continuous function on $[a, b]$ which is differentiable in (a, b). The mean value theorem shows therefore that $\varphi(b) - \varphi(a) = (b-a)\phi'(x) = (b-a)\overline{e} \cdot \mathbf{f}'(x)$ for some $x \in (a, b)$. On the other hand,

$$\varphi(b) - \varphi(a) = \overline{e} \cdot \mathbf{f}(b) - \overline{e} \cdots \mathbf{f}(a) = \overline{e} \cdot \overline{e} = |\overline{e}|^2.$$

The Schwarz inequality now gives $|\overline{e}|^2 = (b-a)|\overline{e} \cdot \mathbf{f}'(x)| \leqslant (b-a)|\overline{e}||\mathbf{f}'(x)|$. Hence $|\overline{e}| \leqslant (b-a)|\mathbf{f}'(x)|$, which is the desired conclusion. □

Chapter 6 The Riemann-Stieltjes Integral

Definitions

Partition Let $[a, b]$ be a given interval. By a *partition* P of $[a, b]$ we mean a finite set of points $x_0, x_1, \ldots, x_n$, where

$$a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b.$$

We write

$$\Delta x_i = x_i - x_{i-1} \quad (i = 1, \ldots, n).$$

Integral Components Suppose f is a bounded real function on $[a, b]$. Corresponding to each partition P of $[a, b]$ we put:

$$M_i = \sup f(x) \quad (x_{i-1} \leqslant x \leqslant x_i), \qquad m_i = \inf f(x) \quad (x_{i-1} \leqslant x \leqslant x_i),$$

$$U(P, f) = \sum_{i=1}^{n} M_i \Delta x_i, \qquad L(P, f) = \sum_{i=1}^{n} m_i \Delta x_i,$$

$$\overline{\int_a^b} f\,dx = \inf_P U(P, f), \qquad \underline{\int_a^b} f\,dx = \sup_P L(P, f)$$

The last two are called the *upper* and *lower Riemann integrals* of f over $[a, b]$ respectively.

Riemann Integrable If the upper and lower integrals are equal we say that f is *Riemann-integrable* on $[a, b]$, we write $f \in \mathscr{R}$ (that is, $\mathscr{R}$ denotes the set of Riemann-integrable functions), and we denote the common value of the upper and lower integrals by:

$$\int_a^b f\,dx \quad \text{or by} \quad \int_a^b f(x)\,dx$$

Alpha Let α be a monotonically increasing function on $[a, b]$. Corresponding to each partition P of $[a, b]$ we write

$$\Delta \alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \quad (i = 1, \ldots, n).$$

Note, $\Delta \alpha_i \geqslant 0$.

Integral Components Suppose f is a bounded real function on $[a, b]$. Corresponding to each partition P of $[a, b]$ we put:

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta \alpha_i, \qquad L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta \alpha_i,$$

$$\overline{\int_a^b} f\,d\alpha = \inf_P U(P, f, \alpha), \qquad \underline{\int_a^b} f\,d\alpha = \sup_P L(P, f, \alpha)$$

Riemann- Stieltjes Integrable If the upper and lower integrals are equal we denote their common value by

$$\int_a^b f\,d\alpha \quad \text{or by} \quad \int_a^b f(x)\,d\alpha.$$

This is the *Riemann- Stieltjes integral* of f with respect to α over $[a, b]$. If this exists, we say that f is integrable with to α and we write $f \in \mathscr{R}(\alpha)$

Refinement We say that the partition P^* is a *refinement* of P is $P^* \supset P$.

Common Refinement Given two partitions, P_1 and P_2, we say that P_* is their *common refinement* if $P^* = P_1 \cup P_2$.

Unit Step Function The *unit step function* I is defined by:

$$I(x) = \begin{cases} 0 & (x \leqslant 0), \\ 1 & (x > 0). \end{cases}$$

Vector-valued Functions Let $f_1, \ldots, f_k$ be real functions on $[a, b]$ and let $\mathbf{f} = (f_1, \ldots, f_k)$ be the corresponding mapping of $[a, b]$ into $\mathbb{R}^k$. If α increases monotonically on $[a, b]$, to say that $\mathbf{f} \in \mathscr{R}(\alpha)$ means that $f_j \in \mathscr{R}(\alpha)$ for $j = 1, \ldots, k$. If this is the case, we define

$$\int_a^b \mathbf{f}\,d\alpha = \left(\int_a^b f_1\,d\alpha, \ldots, \int_a^b f_k\,d\alpha\right).$$

In other words, $\int \mathbf{f}\,d\alpha$ is the point in $\mathbb{R}^k$ whose jth coordinate is $\int f_j\,d\alpha$.

Curve/ Arc/ Closed Curve/ Length/ Rectifiable A continuous mapping γ of an interval $[a, b]$ into $\mathbb{R}^k$ is called a *curve* in $\mathbb{R}^k$. To emphasize the parameter interval $[a, b]$, we may also say that γ is a curve on $[a, b]$. If γ is one-to-one, γ is called an *arc*. If $\gamma(a) = \gamma(b)$, γ is said to be a *closed curve*. It should be noted that we define a curve to be a *mapping*, not a point set.

We associate to each partition $P = \{x_0, \ldots, \mathbf{x}_n\}$ of $[a, b]$ and to each curve γ on $[a, b]$ the number

$$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1}|.$$

The ith term in this sum is the distance between the points. Hence $\Lambda(P, \gamma)$ is the length of a polygonal path with vertices at $\gamma(x_0), \gamma(x_1), \ldots, \gamma(x_n)$, in this order. As our partition becomes finer and finer, this polygon approaches the range of γ more closely. Thus, *length* of γ is:

$$\Lambda(\gamma) = \sup \Lambda(P, \gamma),$$

If $\Lambda(\gamma) < \infty$ we say that γ is *rectifiable*.

Theorems

Theorem 6.4 If P^* is a refinement of P, then

$$L(P, f, \alpha) \leqslant L(P^*, f, \alpha) \quad \text{and} \quad U(P^*, f, \alpha) \leqslant U(P, f, \alpha).$$

Proof. To prove that $L(P, f, \alpha) \leqslant L(P^*, f, \alpha)$, suppose first that P^* contains just one point more than P. Let this extra point be x^*, and suppose $x_{i-1} < x^* < x_i$, where x_{i-1} and x_i are two consecutive points of P. Put $w_1 = \inf f(x)$ for $x_{i-1} < x < x^*$ and $w_2 = \inf f(x)$ for $x^* < x < x_i$. Clearly $w_1 \geqslant m_i$ and $w_2 \geqslant m_i$ (previously $m_i = \inf f(x)$ for $x_{i-1} \leqslant x \leqslant x_i$. Hence

$$L(P^*, f, \alpha) - L(P, f, \alpha) = w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_{i-1})] \quad = (w_1 -$$

If P^* contains k points more than P, we repeat this reasoning k times, and arrive at our conclusion. The proof of $U(P^*, f, \alpha) \leqslant U(P, f, \alpha)$ is analogous. □

Theorem 6.5 $\underline{\int}_a^b f \, d\alpha \leqslant \overline{\int}_a^b f \, d\alpha$

Proof. Let P^* be the common refinement of two partitions P_1 and P_2. By Theorem 6.4, $L(P_1, f, \alpha) \leqslant L(P^*, f, \alpha) \leqslant U(P^*, f, \alpha) \leqslant U(P_2, f, \alpha)$. Hence $L(P_1, f, \alpha) \leqslant U(P_2, f, \alpha)$. If P_2 is fixed and the sup is taken over all P_1, this gives $\underline{\int} f \, d\alpha \leqslant U(P_2, f, \alpha)$. The theorem follows by taking the inf over all P_2 in the last equation. □

Theorem 6.6 $f \in \mathscr{R}(\alpha)$ on $[a, b]$ if and only if for every $\varepsilon > 0$ there exists a partition P such that:

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Proof. For every P we have $L(P, f, \alpha) \leqslant \underline{\int} f \, d\alpha \leqslant \overline{\int} f \, d\alpha \leqslant U(P, f, \alpha)$. Thus, $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ implies that $0 \leqslant \overline{\int} f \, d\alpha - \underline{\int} f \, d\alpha < \varepsilon$. Hence, if $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ for every $\varepsilon > 0$, we have $\overline{\int} f \, d\alpha = \underline{\int} f \, d\alpha$, that is $f \in \mathscr{R}(\alpha)$.

Conversely, suppose $f \in \mathscr{R}(\alpha)$, and let $\varepsilon > 0$ be given. Then there exist partitions P_1 and P_2 such that $U(P_2, f, \alpha) - \int f \, d\alpha < \frac{\varepsilon}{2}$, and $\int f \, d\alpha - L(P_1, f, \alpha) < \frac{\varepsilon}{2}$. We choose P to be the common refinement of P_1 and P_2. Then Theorem 6.4 together with the aforementioned inequalities show that:

$$U(P, f, \alpha) \leqslant U(P_2, f, \alpha) < \int f \, d\alpha + \frac{\varepsilon}{2} < L(P_1, f, \alpha) + \varepsilon \leqslant L(P, f, \alpha) + \varepsilon$$

Thus, for partition P, $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ □

Theorem 6.7 a) If $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ holds for some P and some ε, then it holds for every refinement of P.

b) If $U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon$ holds for $P = \{x_0, \ldots, x_n\}$ and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \delta\alpha_i < \varepsilon$$

c) If $f \in \mathscr{R}(\alpha)$ and the hypotheses of (b) hold, then

$$\left| \sum_{i=1}^{n} f(t_i)\delta\alpha_i - \int_a^b f\, d\alpha \right| < \varepsilon$$

Proof. Theorem 6.4 implies (a). Under the assumptions made in (b), both $f(s_i)$ and $f(t_i)$ lie in $[m_i, M_i]$, so that $|f(s_i) - f(t_i)| \leqslant M_i - m_i$. Thus

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta\alpha_i \leqslant U(P,f,\alpha) - L(P,f,\alpha),$$

which proves (b). The obvious inequalities $L(P,f,\alpha) \leqslant \sum f(t_i)\Delta\alpha_i \leqslant U(P,f,\alpha)$ and $L(P,f,\alpha) \leqslant \int f\, d\alpha \leqslant U(P,f,\alpha)$ prove (c). □

Theorem 6.8 If f is continuous on $[a,b]$ then $f \in \mathscr{R}(\alpha)$ on $[a,b]$.

Proof. Let $\varepsilon > 0$ be given. Choose $\eta > 0$ so that $[\alpha(b) - \alpha(a)]\eta < \varepsilon$. Since f is uniformly continuous on $[a,b]$ (Theorem 4.19), there exists a $\delta > 0$ such that $|f(x) - f(t)| < \eta$ if $x, t \in [a,b]$ and $|x - t| < \delta$. If P is any partition of $[a,b]$ such that $\Delta x_i < \delta$ for all i, then the above equation implies that $M_i - m_i \leqslant \eta$ $(i - 1, \ldots, n)$ and therefore

$$U(P,f,\alpha) - L(P,f,\alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \leqslant \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \varepsilon.$$

By Theorem 6.6, $f \in \mathscr{R}(\alpha)$. □

Theroem 6.9 If f is monotonic on $[a,b]$, and if α is continuous on $[a,b]$, then $f \in \mathscr{R}(\alpha)$.

Proof. Let $\varepsilon > 0$ be given. For any positive integer n, choose a partition such that $\Delta\alpha_i = \frac{\alpha(b)-\alpha(a)}{n}$. This is possible since α is continuous (Theorem 4.23).

We suppose that f is monotonically increasing (the proof is analogous in the other case). Then $M_i = f(x_i)$, $m_i = f(x_{i-1})$ so that

$$U(P,f,\alpha) - L(P,f,\alpha) = \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] = \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)] < \varepsilon$$

if n is taken large enough. By Theorem 6.6, $f \in \mathscr{R}(\alpha)$. □

Theorem 6.10 Suppose f is bounded on $[a, b]$, f has only finitely many points of discontinuity on $[a, b]$ and α is continuous at every point at which f is discontinuous. Then $f \in R\alpha$.

Proof. Let $\varepsilon > 0$ be given. Put $M = \sup|f(x)|$, let E be the set f points at which f is discontinuous. Since E is finite and α is continuous at every point of E, we can cover E by finitely many disjoint intervals $[u_j, v_j] \subset [a, b]$ such that the sum of the corresponding differences $\alpha(v_j) - \alpha(u_j)$ is less that ε. Furthermore, we can place these intervals in such a way that every point of $E \cap (a, b)$ lies in the interior of some $[u_j, v_j]$. Remove the segments (u_j, v_j) from $[a, b]$. The remaining set K is compact. Hence f is uniformly continuous on K, and there exists $\delta > 0$ such that $|f(s) - f(t)| < \varepsilon$ if $s \in K$, $t \in K$, $|s - t| < \delta$.

Now form a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$, as follows: Each u_j occurs in P. Each v_j occurs in P. No point of any segment (u_j, v_j) occurs in P. If x_{i-1} is not one of the u_j, then $\Delta x_i < \delta$.

Note that $M_i - m_i \leqslant 2M$ for every i, and that $M_i - m_i \leqslant \varepsilon$ unless x_{i-1} is one of the u_j. Hence, as in the proof of Theorem 6.8

$$U(P, f, \alpha) - L(P, f, \alpha) \leqslant [\alpha(b) - \alpha(a)]\varepsilon + 2M\varepsilon.$$

Since ε is arbitrary, Theorem 6.6 shows that $f \in \mathscr{R}(\alpha)$. (Note: if f and α have common points of discontinuity, then f need not be in $\mathscr{R}(\alpha)$. □

Theorem 6.11 (Composition of Functions) Suppose $f \in \mathscr{R}(\alpha)$ on $[a, b]$, $m \leqslant f \leqslant M$, ϕ is continuous on $[m, M]$, and $h(x) = \phi(f(x))$ on $[a, b]$. Then $h \in \mathscr{R}(\alpha)$ on $[a, b]$.

Proof. Choose $\varepsilon > 0$. Since $|phi$ is uniformly continuous on $[m, M]$, there exists $\delta > 0$ such that $\delta < \varepsilon$ and $|\phi(s) - \phi(t)| < \varepsilon$ if $|s - t| \leqslant \delta$ and $s, t \in [m, M]$.

Since $f \in \mathscr{R}(\alpha)$, there is a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$. Let M_i, m_i have the same meaning as in Definition 6.1 and let M_i^*, m_i^* be the analogous numbers for h. Divide the numbers $1, \ldots, n$ into to classes: $i \in A$ if $M_i - m_i < \delta$ and $i \in B$ if $M_i - m_i \geqslant \delta$.

For $i \in A$, our choice of δ shows that $M_i^* - m_i^* \leqslant \varepsilon$.

For $i \in B$, $M_i^* - m_i^* \leqslant 2K$, where $K = \sup|\phi(t)|$, $m \leqslant t \leqslant M$. Because $U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$, we have $\delta \sum_{i \in B} \Delta\alpha_i \leqslant \sum_{i \in B}(M_i - m_i)\Delta\alpha_i < \delta^2$ so that $\sum_{i \in B} \Delta\alpha_i < \delta$. If follows that:

$$\begin{aligned} U(P, h, \alpha) - L(P, h, \alpha) &= \sum_{i \in A}(M_i^* - m_i^*)\Delta\alpha_i + \sum_{i \in B}(M_i^* - m_i^*)\Delta\alpha_i \\ &\leqslant \varepsilon[\alpha(b) - \alpha(a)] + 2K\delta < \varepsilon[\alpha(b) - \alpha(a) + 2K] \end{aligned}$$

Since ε was arbitrary, Theorem 6.6 implies $h \in \mathscr{R}(\alpha)$. □

Theorem 6.12 (Properties of Integrals) a) If $f_1, f_2 \in \mathscr{R}(\alpha)$ on $[a,b]$, then $f_1 + f_2 \in \mathscr{R}(\alpha)$, $cf \in \mathscr{R}(\alpha)$ for every constant c, and

$$\int_a^b (cf_1 + f_2)\, d\alpha = c\int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$$

b) If $f_1(x) \leqslant f_2(x)$ on $[a,b]$, then

$$\int_a^b f_1\, d\alpha \leqslant \int_a^b f_2\, d\alpha.$$

c) If $f \in \mathscr{R}(\alpha)$ on $[a,b]$ and $a < c < b$, then $f \in \mathscr{R}(\alpha)$ on $[a,c]$ and $f \in \mathscr{R}(\alpha)$ on $[c,b]$, and

$$\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.$$

d) If $f \in \mathscr{R}(\alpha)$ on $[a,b]$ and if $|f(x)| \leqslant M$ on $[a,b]$, then

$$\left|\int_a^b f\, d\alpha\right| \leqslant M[\alpha(a) - \alpha(b)].$$

e) If $f \in \mathscr{R}(\alpha_1)$ and $f \in \mathscr{R}(\alpha_2)$, then $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ and

$$\int_a^b f\, d(\alpha_1 + \alpha_2) = \int_a^b f\, d(\alpha_1) + \int_a^b f\, d(\alpha_2);$$

If $f \in \mathscr{R}(\alpha)$ and c is positive constant, then $f \in \mathscr{R}(c\alpha)$ and

$$\int_a^b f\, d(c\alpha) = c\int_a^b f\, d\alpha$$

Proof. If $f = f_1 + f_2$ and P is any partition of $[a,b]$, we have

$$L(P,f_1,\alpha) + L(P,f_2,\alpha) \leqslant L(P,f,\alpha) \leqslant U(P,f,\alpha) \leqslant U(P,f_1,\alpha) + U(P,f_2,\alpha).$$

If $f_1 \in \mathscr{R}(\alpha)$ and $f_2 \in \mathscr{R}(\alpha)$, let $\varepsilon > 0$ be given. There are partitions P_j $(j = 1,2)$ such that $U(P_j,f_j,\alpha) - L(P_j,f_j,\alpha) < \varepsilon$. These inequalities persist if P_1 and P_2 are replaced by their common refinement P. Then our original string of inequalities implies $U(P,f,\alpha) - L(P,f,\alpha) < 2\varepsilon$, which proves that $f \in \mathscr{R}(\alpha)$.

With this same P we have $U(P,f_j,\alpha) < \int f_j\, d\alpha + \varepsilon$. Hence our original string of inequalities implies $\int f\, d\alpha \leqslant U(P,f,\alpha) < \int f_1\, d\alpha + \int f_2\, d\alpha + 2\varepsilon$. Since ε was arbitrary, we conclude that $\int f\, d\alpha \leqslant \int f_1\, d\alpha + \int f_2\, d\alpha$. If we replace f_1 and f_2 with $-f_1$ and $-f_2$, the inequality is reversed and the equalities is proved. □

Theorem 6.13 If $f \in \mathscr{R}(\alpha)$ and $f \in \mathscr{R}(\alpha)$ on $[a,b]$, then:

a) $fg \in \mathscr{R}(\alpha)$;

b) $|f| \in \mathscr{R}(\alpha)$ and $\left|\int_a^b f\,d\alpha\right| \leqslant \int_a^b |f|\,d\alpha$.

Proof. If we take $|phi(t) = t^2$, Theorem 6.11 shows that $f^2 in\mathscr{R}(\alpha)$ if $f \in \mathscr{R}(\alpha)$. The identity $4fg = (f+g)^2 - (f-g)^2$ completes the proof of (a).

If we take $\phi(t) = |t|$, Theorem 6.11 shows similarly that $|f| \in \mathscr{R}(\alpha)$. Choose $c = \pm 1$, so that $c \in f\,d\alpha \geqslant 0$. Then $\left|\int f\,d\alpha\right| = c\int f\,d\alpha = \int cf\,d\alpha \leqslant \int |f|\,d\alpha$, since $cf \leqslant |f|$. □

Theorem 6.15 If $a < s < b$, f is bounded on $[a,b]$, f is continuous at s and $\alpha = I(x-s)$, then

$$\int_a^b f\,d\alpha = f(s)$$

Proof. Consider partitions $P + \{x_0, x_1, x_2, x_3\}$, where $x_0 = a$, and $x_1 = s < x_2 < x_3 = b$. Then $U(P,f,\alpha) = M_2$, $L(P,f,\alpha) = m_2$. Since f is continuous at s we see that M_2 and m_2 converge to $f(s)$ as $x_2 \to s$. □

Theorm 6.16 Suppose $c_n \geqslant 0$ for $1, 2, 3, \ldots, \sum c_n$ converges, $\{s_n\}$ is a sequence of distinct points in (a,b) and

$$\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n)$$

Let f be continuous on $[a,b]$. Then

$$\int_a^b f\,d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Proof. The comparison test shows that the series $\sum_{n=1}^{\infty} c_n I(x-s_n)$ converges for every x. Its sum $\alpha(x)$ is evidently monotonic, and $\alpha(a) = 0$, $\alpha(b) = \sum c_n$. Let $\varepsilon > 0$ be given, and choose N so that $\sum_{N+1}^{\infty} c_n < \varepsilon$. Put

$$\alpha_1(x) = \sum_{n=1}^{N} c_n I(x-s_n) \qquad \alpha_2(x) = \sum_{N+1}^{\infty} c_n I(x-s_n)$$

By theorem 6.12 and 6.15, $\int_a^b f\,d\alpha_1 = \sum_{i=1}^{N} c_n f(s_n)$. Since $a_2(b) - a_2(a) < \varepsilon$, we see that $\left|\int_a^b f\,d\alpha_2\right| \leqslant M\varepsilon$, where $M = \sup|f(x)|$. Since $\alpha = \alpha_1 + \alpha_2$, it follows from these equations that $\left|\int_a^b f\,d\alpha - \sum_{i=1}^{N} c_n f(s_n)\right| \leqslant M\varepsilon$. If we let $N \to \infty$, we obtain the desired equality. □

Theorem 6.17 Assume α increases monotonically and $\alpha' \in \mathscr{R}$ on $[a,b]$. Let f be a bounded real function on $[a,b]$. Then $f \in \mathscr{R}(\alpha)$ if and only if $f\alpha' \in \mathscr{R}$. In that case

$$\int_a^b f\,d\alpha = \int_a^b f(x)\alpha'(x)\,dx.$$

Proof. Let $\varepsilon > 0$ be given and apply Theorem 6.6 to α': There is a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that $U(P,\alpha') - L(P,\alpha') < \varepsilon$. The mean value theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that $\Delta\alpha_i = \alpha'(t_i)\Delta x_i$ for $i = 1, \ldots, n$. If $s_i \in [x_{i-1}, x_i]$, then $\sum_{i=1}^n |\alpha'(s_i) - \alpha'(t_i)\Delta x_i| < \varepsilon$, because $U(P,\alpha') - L(P,\alpha') < \varepsilon$ and by Theorem 6.7 (b). Put $M = \sup|f(x)|$. Since $\sum_{i=1}^n f(s_i)\Delta\alpha_i = \sum_{i=1}^n f(s_i)\alpha'(t_i)\Delta x_i$ it follows from the aformentioned sum that $|\sum_{i=1}^n f(s_i)\Delta\alpha_i - \sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i| \leqslant M\varepsilon$. In particular, $\sum_{i=1}^n f(s_i)\Delta\alpha_i \leqslant U(P, f\alpha') + M\varepsilon$, for all choices of $s_i \in [x_{i-1}, x_i]$ so that $U(P,f,\alpha) \leqslant U(P,f\alpha') + M\varepsilon$. The same argument leads to $U(P,f\alpha') \leqslant U(P,f,\alpha) + M\varepsilon$. Thus $|U(P,f,\alpha) - U(P,f\alpha')| \leqslant M\varepsilon$. Now we note that this remains true if P is replaced by any refinement. We conclude that $\left|\overline{\int_a^b} f\,d\alpha - \overline{\int_a^b} f(x)\alpha'(x)\,dx\right| \leqslant M\varepsilon$. But ε is arbitrary. Hence $\overline{\int_a^b} f\,d\alpha = \overline{\int_a^b} f(x)\alpha'(x)\,dx$, for any bounded f. The equality of the lower integrals follows in the exact same way. □

Theorem 6.19 (Change of Variables) Suppose φ is a strictly increasing continuous function that maps an interval $[A,B]$ onto $[a,b]$. Suppose α is a monotonically increasing function on $[a,b]$ and $f \in \mathscr{R}(\alpha)$ on $[a,b]$. Define β and g on $[A,B]$ by

$$\beta(y) = \alpha(\varphi(y)), \qquad g(y) = f(\varphi(y)).$$

Then $g \in \mathscr{R}(\beta)$ and

$$\int_A^B d\,d\beta = \int_a^b f\,d\alpha$$

Proof. To each partition $P = \{x_-, \ldots, x_n\}$ of $[a,b]$ corresponds a partition $Q = \{y_0, \ldots, y_n\}$ of $[A,B]$, so that $x_i = \varphi(y_i)$. All partitions of $[A,B]$ are obtained in this way. Since the values taken by f on $[x_{i-1}, x_i]$ are exactly the same as those taken by g in $[y_{i-1}, y_i]$, we see that $U(Q,g,\beta) = U(P,f,\alpha)$, $L(Q,g,\beta) = L(P,f,\alpha)$.

Since $f \in \mathscr{R}(\alpha)$, P can be chosen so that both $U(P,f,\alpha)$ and $L(P,f,\alpha)$ are close to $\int f\,d\alpha$. Hence the combinations of the equalities with Theorem 6.6 shows that $g \in \mathscr{R}(\beta)$ and that the desired equivalence is true. □

Theorem 6.20 Let $f \in \mathscr{R}$ on $[a,b]$, for $a \leqslant x \leqslant b$, put

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

Then F is continuous on $[a,b]$; furthermore, if f is continuous at a point x_0 of $[a,b]$, then F is differentiable at x_0, and $F'(x_0) = f(x_0)$.

Proof. Since $f \in \mathscr{R}$, f is bounded. Suppose $|f(t)| \leqslant M$ for $a \leqslant t \leqslant b$. If $a \leqslant x < y \leqslant b$, then $|F(y) - F(x)| = |\int_x^y f(t)\,dt| \leqslant M(y-x)$, by Theorem 6.12 (c) and (d). Given $\varepsilon > 0$, we see that $|F(y) - F(x)| < \varepsilon$ provided that $|y - x| < \varepsilon/M$. This proves continuity (and, in fact, uniform continuity) of F.

Now suppose f is continuous at x_0. Given $\varepsilon > 0$, choose $\delta > 0$ such that $|f(t)-f(x_0)| < \varepsilon$ if $|t-x_0| < \delta$, and $a \leqslant t \leqslant b$. Hence, if $x_0-\delta < s \leqslant x_0 \leqslant t \leqslant x_0+\delta$ and $a \leqslant s < t \leqslant b$, we have by Theorem 6.12(d),

$$\left|\frac{F(t)-F(s)}{t-s} - f(x_0)\right| = \left|\frac{1}{t-s}\int_s^t [f(u)-f(x_0)]\,du\right| < \varepsilon.$$

It follows that $F'(x_0) = f(x_0)$. □

The Fundamental Theorem of Calculus If $f \in \mathscr{R}$ on $[a,b]$ and if there is a differentiable function F on $[a,b]$ such that $F' = F$, then

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

Proof. Let $\varepsilon > 0$ be given. Choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ so that $U(P,f) - L(P,f) < \varepsilon$. The mean value theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that $F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$ for $i - 1, \ldots, n$. Thus $\sum_{i=1}^n f(t_i)\Delta x_i = F(b) - F(a)$. It now follows from Theorem 6.7(c) that $\left|F(b) - F(a) - \int_a^b f(x)\,dx\right| < \varepsilon$. Since this holds for every $\varepsilon > 0$, the proof is complete. □

Theorem 6.22 (Integration by Parts) Suppose F and G are differentiable functions on $[a,b]$, $F' = f \in \mathscr{R}$, and $G' = g \in \mathscr{R}$. Then

$$\int_a^b F(x)g(x)\,dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\,dx.$$

Proof. Put $H(x) = F(x)G(x)$ and apply Theorem 6.21 to H and its derivative. Note that $H' \in \mathscr{R}$, by Theorem 6.13. □

Theorem 6.24 If $\mathbf{f}$ and $\mathbf{F}$ map $[a,b]$ into $\mathbb{R}^k$, if $\mathbf{f} \in \mathscr{R}$ on $[a,b]$, and if $\mathbf{F}' = \mathbf{f}$, then

$$\int_a^b \mathbf{f}(t)\,dt = \mathbf{F}(b) - \mathbf{F}(a).$$

Theorem 6.25 If $\mathbf{f}$ maps $[a,b]$ into $\mathbb{R}^k$ and if $\mathbf{f} \in \mathscr{R}(\alpha)$ for some monotonically increasing function α on $[a,b]$, then $|\mathbf{f}| \in \mathscr{R}(\alpha)$ and

$$\left|\int_a^b \mathbf{f}\,d\alpha\right| \leqslant \int_a^b |\mathbf{f}|\,d\alpha.$$

Proof. If $f_1, \ldots, f_k$ are the components of $\mathbf{f}$, then $|\mathbf{f}| = (f_1^2 + \cdots + f_k^2)^{1/2}$. By Theorem 6.11, each of the functions f_i^2 belongs to $\mathscr{R}(\alpha)$; hence so does their sum. Since x^2 is a continuous functions of x, Theorem 4.17 shows that the square-roots function is continuous on $[0, M]$, for every real M. If we apply Theorem 6.11 once more, we see that $|\mathbf{f}| \in \mathscr{R}(\alpha)$.

To prove the desired inequality, put $\mathbf{y} = (y_1, \ldots, y_k)$ where $y_j = \int f_j\, d\alpha$. Then we have $\mathbf{y} = \int \mathbf{f}\, d\alpha$, and $|\mathbf{y}|^2 = \sum y_i^2 = \sum y_j \int f_j\, d\alpha = \int (\sum y_j f_j)\, d\alpha$. By the Schwarz inequality, $\sum y_j f_j(t) \leqslant |\mathbf{y}||\mathbf{f}(t)|$, hence Theorem 6.12 (b) implies $|\mathbf{y}|^2 \leqslant |\mathbf{y}| \int |\mathbf{f}|\, d\alpha$. If $\mathbf{y} = 0$, then the inequality is trivial. If $\mathbf{y} \neq 0$, division by $|\mathbf{y}|$ gives the inequality. □

Theorem 6.27 If γ' is continuous on $[a, b]$, then γ is rectifiable, and

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\, dt.$$

Proof. If $a \leqslant x_{i-1} < x_i \leqslant b$, then $|\gamma(x_i) - \gamma(x_{i-1})| = \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\, dt\right| \leqslant \int_{x_{i-1}}^{x_i} |\gamma'(t)|\, dt$. Hence $\Gamma(P, \gamma) \leqslant \int_a^b |\gamma'(t)|\, dt$ for every partition P of $[a, b]$. Consequently, $\Gamma(\gamma) \leqslant \int_a^b |\gamma'(t)|\, dt$.

To prove the opposite inequality, let $\varepsilon > 0$ be given. Since γ' is uniformly continuous on $[a, b]$, there exists $\delta > 0$ such that $|\gamma'(s) - \gamma'(t)| < \varepsilon$ if $|s - t| < \delta$. Let $P = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$, with $\Delta x_i < \delta$ for all i. If $x_{i-1} \leqslant t \leqslant x_i$, it follows that $|\gamma'(t)| \leqslant |\gamma'(x_i)| + \varepsilon$. Hence

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} |\gamma'(t)|\, dt &\leqslant |\gamma'(x_i)|\Delta x_i + \varepsilon \Delta x_i \\
&= \left|\int_{x_{i-1}}^{x_i} [\gamma'(t) + \gamma'(x_i) - \gamma'(t)]\, dt\right| + \varepsilon \Delta x_i \\
&= \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\, dt\right| + \left|\int_{x_{i-1}}^{x_i} \gamma'(x_i) - \gamma'(t)]\, dt\right| + \varepsilon \Delta x_i \\
&\leqslant |\gamma(x_i) - \gamma(x_{i-1})| + 2\varepsilon \Delta x_i.
\end{aligned}$$

If we add these inequalities, we obtain

$$\int_a^b |\gamma'(t)|\, dt \leqslant \Gamma(P, \gamma) + 2\varepsilon(b - a) \leqslant \Gamma(\gamma) + 2\varepsilon(b - a).$$

Since ε was arbitrary $\int_a^b |\gamma'(t)|\, dt \leqslant \Gamma(\gamma)$. This completes the proof. □

Chapter 7 Sequences and Series of Functions

Definitions

Limit/ pointwise convergence/ sum Suppose $\{f_n\}$, $n = 1, 2, 3, \ldots$, is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in E$. We an then define a function f by

$$f(x) = \lim_{n \to \infty} f_n(x) \qquad (x \in E)$$

Under these circumstances we say that $\{f_n\}$ converges on E, that f is the *limit*, or the *limit function*, of $\{f_n\}$, and that $\{f_n\}$ converges to f *pointwise* on E. Similarly, if $\sum f_n(x)$ converges for every $x \in E$, and if we define:

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (x \in E)$$

the function f is called the *sum* of the series σf_n.

Uniform Convergence We say that a sequence of functions $\{f_n\}$, $n = 1, 2, 3, \ldots$, converges *uniformly* on E to a function f is for every $\varepsilon > 0$ there is an integer N such that $n \geqslant N$ implies the following for all $x \in E$:

$$\mathrm{f}_n(x) - f(x) \leqslant \varepsilon$$

Supremum Norm If X is a metric space, $\mathscr{C}(X)$ will denote the set of all complex=values, continuous, bounded functions with domain X. We associate each $f \in \mathscr{C}(X)$ with its *supremum norm*

$$||f|| = \sup_{x \in X} |f(x)|.$$

Since f is assumed to be bounded, $||f|| \leqslant \infty$. It is obvious that $||f|| = 0$ if and only if $f(x) = 0$ for every $x \in X$. If $h = f + g$:

$$|h(x)| \leqslant |f(x)| + |g(x)| \leqslant ||f|| + ||g||$$

for all $x \in X$; hence

$$||f + g|| \leqslant ||f|| + ||g||.$$

$\mathscr{C}(X)$ as a Metric Space If we define the distance between $f \in \mathscr{C}(X)$ and $g \in CX$ to be $||f - g||$, it follows that the Axioms for a metric are satisfied. Therefore, A sequence $\{f_n\}$ converges to f with respect to the metric of $\mathbb{C}$ if and only if $f_n \rightrightarrows f$ on X.

Pointwise Bounded Let $\{f_n\}$ be a sequence of functions defined on a set E. We say that $\{f_n\}$ is *pointwise bounded* on E if the sequence $\{f_n(x)\}$ is bounded for every $x \in E$, that is, if there exists a finite-valued function ϕ defined on E such that

$$|f_n(x)| < \phi(x) \qquad (x \in E,\ n = 1, 2, 3, \ldots).$$

(If $\{f_n\}$ is pointwise bounded on E and E_1 is a countable subset of E it is always possible to find a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E_1$.)

Uniformly Bounded We say that $\{f_n\}$ is *uniformly bounded* on E if there exists a number M such that

$$|f_n(x)| < M \qquad (x \in E,\ n = 1, 2, 3, \dots).$$

(If $\{f_n\}$ is a uniformly bounded sequence of continuous functions on a compact set E, there need not exists a subsequence which converges pointwise on E.)

Equicontinuous A family $\mathscr{F}$ of complex functions f defined on a set E in a metric space X is said to be *equicontinuous* on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon$$

whenever $d(x, y) < \delta$, $x, y \in E$, $f \in \mathscr{F}$. (Note: Every member of an equicontinuous family is uniformly continuous.)

Algebra A family $\mathscr{A}$ of complex functions defined on a set E is said to be an *algebra* if: (i) $f + g \in \mathscr{A}$, (ii) $fg \in \mathscr{A}$ (iii) $cf \in \mathscr{A}$. for all $f, g \in \mathscr{A}$ and for all complex constants c, that is $\mathscr{A}$ is closed under addition, multiplication, and scalar multiplication.

Uniformly Closed If $\mathscr{A}$ has the property that $f \in \mathscr{A}$ whenever $f_n \in A$ $(n = 1, 2, 3, \dots)$ and $f_n \rightrightarrows f$ on E, then $\mathscr{A}$ is said to be *uniformly closed*.

Uniform Closure Let $\mathscr{B}$ be the set of all functions which are limits of uniformly convergent sequences of members of $\mathscr{A}$. Then $\mathscr{B}$ is called the *uniform closure* of $\mathscr{A}$.

Seperate Points Let $\mathscr{A}$ be a family of functions on a set E. Then $\mathscr{A}$ is said to *separate points* on E if every pair of distinct points $x_1, x_2 \in E$ there corresponds a function $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$.

Vanishes At No Point If to each $x \in E$ there corresponds a function $g \in \mathscr{A}$ such that $g(x) \neq 0$, we say that $|A$ vanishes at no point of E.

Theorems

Theorem 7.8 (Cauchy) The sequence of functions $\{f_n\}$, defined on E, converges uniformly on E if and only if for every $\varepsilon > 0$ there exists an integer N such that $m \geqslant N$, $n \geqslant N$, $x \in E$ implies:

$$|f_n(x) - f_m(x)| \leqslant \varepsilon$$

Proof. Suppose $\{f_n\}$ converges uniformly on E, and let f be the limit function. Then there is an integer N, such that $n \geqslant N$, $x \in E$ implies $|f_n(x) - f(x)| \leqslant \frac{\varepsilon}{2}$, so that $|f_n(x) - f_m(x)| \leqslant |f_n(x) - f(x)| + |f(x) - f_m(x)| \leqslant \varepsilon$ if $n, m \geqslant N$, $x \in E$.

Conversely, suppose the Cauchy condition holds. By Theorem 3.11, the sequence $\{f_n(x)\}$ converges, for every x, to a limit which we may call $f(x)$. Thus the sequence $\{f_n\}$ converges on E, to f. We have to prove that the convergence is uniform. Let $\varepsilon > 0$ be given, and choose N such that $|f_n(x) - f_m(x)| \leqslant \varepsilon$. Fix n and let $m \to \infty$. Since $f_m(x) \to f(x)$ as $m \to \infty$, this give $|f_n(x) - f(x)| \leqslant \varepsilon$ for every $n \geqslant N$ and every $\in E$, which completes the proof. □

Theorem 7.9 Suppose

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad (x \in E)$$

Put

$$M_n = \sup_{x\in E} |f_n(x) - f(x)|$$

Then $f_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n \to infty$.

Theorem 7.10 (M-test) Suppose $\{f_n\}$ is a sequence of functions defined on E, and suppose

$$|f_n(x)| \leqslant M_n \qquad (x \in E,\ n = 1, 2, 3, \dots).$$

Then $\sum f_n$ converges uniformly on E if $\sum M_n$ converges.

Proof. If $\sum M_n$ converges, then, for arbitrary $\varepsilon > 0$, $|\sum_{i=n}^{m} f_i(x)| \leqslant \sum_{i=n}^{m} M_i \leqslant \varepsilon$ provided m and n are sufficiently large. Uniform convergence now follows from Theorem 7.8. □

Theorem 7.11 Suppose $f_n \to f$ uniformly on a set E in a metric space. Let x be a limit point of E, and suppose that

$$\lim_{t\to x} f_n(t) = A_n \qquad (n = 1, 2, 3, \dots).$$

Then $\{A_n\}$ converges, and

$$\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n.$$

In other words, the conclusion is that:

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

Proof. Let $\varepsilon > 0$ be given. By the uniform convergence of $\{f_n\}$, here exists N such that $n \geqslant N$, $m \geqslant N$, $t \in E$ imply $|f_n(t) - f_m(t)| \leqslant \varepsilon$. Letting $t \to x$ we obtain $|A_n - A_m| \leqslant \varepsilon$ for $n, m \geqslant N$, so that $\{A_n\}$ is a Cauchy sequence and therefore converges, say to A.

Next $|f(t) - A| \leqslant |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|$. We first choose n such that $|f(t) - f_n(t)| \leqslant \frac{\varepsilon}{3}$ for all $t \in E$ (this is made possible by the uniform convergence), and such that $|A_n - A| \leqslant \frac{\varepsilon}{3}$. Then, for this n, we choose a neighborhood V of x such that $|f_n(t) - A_n| \leqslant \frac{\varepsilon}{3}$ if $t \in V \cap E$, $t \neq x$.

Substituting the inequalities, we see that $|f(t) - A| \leqslant \varepsilon$, provided $t \in V \cap E$, $t \neq x$. This is equivalent to our desired equivalence. □

Theorem 7.12 If $\{f_n\}$ is a sequence of continuous functions on E and if $f_n \rightrightarrows f$ on E, then f is continuous on E.

Theorem 7.13 Suppose K is compact, and

a) $\{f_n\}$ is a sequence of continuous functions on K

b) $\{f_n\}$ converges pointwise to a continuous function f on K,

c) $f_n(x) \geqslant f_{n+1}(x)$ for all $x \in K$, $n = 1, 2, 3 \ldots$ ($\{f_n\}$ is a decreasing sequence)

Then $f_n \rightrightarrows f$ on K.

Proof. Put $g_n = f_n - f$. Then g_n is continuous, $g_n \to 0$ pointwise, and $g_n \geqslant g_{n+1}$. We have to prove that $g_n \to 0$ uniformly on K.

Let $\varepsilon > 0$ be given. Let K_n be the set of all $x \in K$ with $g_n(x) \geqslant \varepsilon$. Since g_n is continuous, K_n is closed (Theorem 4.8), hence compact (Theorem 2.35). Since $g_n \geqslant g_{n+1}$, we have $K_n \supset K_{n+1}$. Fix $x \in K$. Since $g_n(x) \to 0$, we see that $x \notin K_n$ if N is sufficiently large. Thus $x \notin \bigcap K_n$. In other words $\bigcap K_n$ is empty. Hence K_N is empty for some N (Theorem 2.36). It follows that $0 \leqslant g_n(x) < \varepsilon$ for all $x \in K$ and for all $n \geqslant N$. This proves the theorem. □

Theorem 7.15 The aforementioned metric makes $\mathscr{C}(X)$ a complete metric space. (That is, a metric space in which every Cauchy Sequence converges

Proof. Let $\{f_n\}$ be a Cauchy sequence in $\mathscr{C}(X)$. This means that to each $\varepsilon > 0$ corresponds an N such that $\|f_n - f_m\| < \varepsilon$ if $n \geqslant N$ and $m \geqslant N$. If follows (by Theorem 7.8) that there is a function f with domain X to which $\{f_n\}$ converges uniformly. By Theorem 7.12, f is continuous. Moreover, f is bounded, since there is an N such that $|f(x) - f_n(x)| < 1$ for all $x \in X$, and f_n is bounded.

Thus $f \in \mathscr{C}(X)$, and since $f_n \rightrightarrows f$ on X, we have $\|f - f_n\| \to 0$ as $n \to \infty$. □

Theorem 7.16 Let α be monotonically increasing on $[a, b]$, for $n = 1, 2, 3, \ldots$, and suppose $f_n \rightrightarrows f$ on $[a, b]$. Then $f \in \mathscr{R}(\alpha)$ on $[a, b]$ and

$$\int_a^b f\, d\alpha = \lim_{n \to \infty} \int_a^b f_n\, d\alpha.$$

Proof. It suffices to prove this for real f_n. Put $\varepsilon_n = \sup |f_n(x) - f(x)|$, the supremum being taken over $a \leqslant x \leqslant b$. Then $f_n - \varepsilon_n \leqslant f \leqslant f_n + \varepsilon_n$, so that the upper and lower integrals of f satisfy $\int_a^b (f_n - \varepsilon_n)\, d\alpha \leqslant \underline{\int} f\, d\alpha \leqslant \overline{\int} f\, d\alpha \leqslant \int_a^b (f_n + \varepsilon_n)\, d\alpha$. Hence $0 \leqslant \overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha \leqslant 2\varepsilon_n[\alpha(b) - \alpha(a)]$. Since $\varepsilon_n \to 0$ as $n \to \infty$ (Theorem 7.9), the upper and lower integrals of f are equal.

Thus $f \in \mathscr{R}(\alpha)$. Another application of the previous formula now yields $\left|\int_a^b f\, d\alpha - \int_a^b f_n\, d\alpha\right| \leqslant \varepsilon_n[\alpha(b) - \alpha(a)]$. This implies the desired equivalence. □

Corollary If $f_n \in \mathscr{R}(\alpha)$ on $[a, b]$ and if

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (a \leqslant x \leqslant b)$$

the series converging uniformly on $[a, b]$, then

$$\int_a^b f\,d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n\,d\alpha$$

In other words, the series may be integrated term by term.

Theorem 7.17 Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a, b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a, b]$. In $\{f_n'\}$ converges uniformly on $[a, b]$, then $\{f_n\}$ converges uniformly on $[a, b]$, to a function f, and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \qquad (a \leqslant x \leqslant b).$$

Proof. Let $\varepsilon > 0$ be given. Choose N such that $n, m \geqslant N$, implies $|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$ and $|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)}$.

If we apply the mean value theorem 5.19 to the function $f_n - f_m$ shows that $|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leqslant \frac{|x-t|\varepsilon}{2(b-a)} \leqslant \frac{\varepsilon}{2}$for any x and t on $[a, b]$, if $n, m \geqslant N$. The inequality

$$|f_n(x) - f_m(x)| \leqslant |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|$$

implies that $|f_n(x) - f_m(x)| < \varepsilon$ so that $\{f_n\}$ converges uniformly on $[a, b]$. Let $f(x) = \lim_{n\to\infty} f_n(x)$. Let us now fix a point x on $[a, b]$ and define $\phi_n(t) = \frac{f_n(t)-f_n(x)}{t-x}$, $\phi(t) = \frac{f(t)-f(x)}{t-x}$ for $a \leqslant t \leqslant b$, $t \neq x$. Then $\lim_{t\to x} \phi_n(t) = f_n'(x)$. From $|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leqslant \frac{|x-t|\varepsilon}{2(b-a)}$ we know that $|\phi_n(t) - \phi_m(t)| \leqslant \frac{\varepsilon}{2(b-a)}$ so that $\{\phi_n\}$ converges uniformly, for $t \neq x$. Since $\{f_n\}$ converges to f, we conclude that $\lim_{n\to\infty} \phi_n(t) = \phi(t)$ uniformly for $a \leqslant t \leqslant b$, $t \neq x$.

If we now apply Theorem 7.11 to $\{\phi_n\}$, we see that $\lim_{t\to x} \phi(t) = \lim_{n\to\infty} f_n'(x)$; and this is the desired equality by the definition of $\phi(t)$. □

Theorem 7.18 There exists a real continuous function on the real line which is nowhere differentiable.

Proof. Define $\varphi(x) = |x|$ $(-1 \leqslant x \leqslant 1)$ and extend the definition of $\varphi(x)$ to all real x by requiring that $\varphi(x + 2) = \varphi(x)$. Then, for all s and t, $|\varphi(s) - \varphi(t)| \leqslant |s - t|$. In particular, φ is continuous on $\mathbb{R}$. Define $f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x)$. Since $0 \leqslant \varphi \leqslant 1$, Theorem 7.10 shows that the series converges uniformly on $\mathbb{R}$. By Theorem 7.12, f is continuous on $\mathbb{R}$.

Now fix a real number x and a positive integer m. Put $\delta_m = \pm\frac{1}{2} \cdot 4^{-m}$ where the sign is so chosen that no integer lies between $4^m x$ and $4^m(x + \delta_m)$. This can be dome, since $4^m|\delta_m| = \frac{1}{2}$. Define $\gamma_n = \frac{\varphi(4^n(x+\delta_m)) - \varphi(4^n x)}{\delta_m}$. When $n > m$, then $4^n\delta_m$ is an even integer, so that $\gamma_n = 0$. When $- \leqslant n \leqslant m$, $|\gamma_n| \leqslant 4^n$.

Since $|\gamma_m| = 4^m$, we conclude that

$$\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n\right| \geqslant 3^m - \sum_{n=0}^{m-1} 3^n = \frac{1}{2}(3^m + 1)$$

As $m \to \infty$, $\delta_m \to 0$. If follows that f is not differentiable at x. □

Theorem 7.23 If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.

Proof. Let $\{x_i\}$, $i = 1, 2, 3, \dots$, be the points of E, arranged in a sequence. Since $\{f_n(x_1)\}$ is bounded, there exists a subsequence, which we shall denote by $\{f_{i,k}\}$, such that $\{f_{i,k}(x_1)\}$ converges as $k \to \infty$.

Let us now consider sequences $S_1, S_2, S_3, \dots$ which we represent by the array:

$$S_1 : f_{1,1}, f_{1,2}, f_{1,3}, \dots$$

$$S_2 : f_{2,1}, f_{2,2}, f_{2,3}, \dots$$

$$S_3 : f_{3,1}, f_{3,2}, f_{3,3}, \dots$$

and which have the following properties:

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

b) $\{f_{n,k}(x_n)\}$ converges as $k \to \infty$ (the boundedness of $\{f_n(x_n)\}$ makes it possible to choose S_n in this way);

c) The order in which the functions appear is the same in each sequence; i.e.m 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 now go down the diagonal of the array; i.e., we consider the sequence $S : f_{1,1}, f_{2,2}, f_{3,3}, \dots$. By (c), the sequence S is a subsequence of S_n, for $n = 1, 2, 3, \dots$. Hence (b) implies that $f_{n,n}(x_i)$ converges, as $n \to \infty$, for every $x_i \in E$. □

Theorem 7.24 If K is a compact metric space, if $f_n \in \mathscr{C}(K)$ for $n = 1, 2, 3, dots$, and if $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.

Proof. Let $\varepsilon > 0$ be given. Since $\{f_n\}$ converges uniformly, there is an integer N such that $||f_n - f_N|| < \varepsilon$. Since continuous functions are uniformly continuous on compact sets, there is a $\delta > 0$ such that $|f_i(x) - f_i(y)| < \varepsilon$ if $1 \leqslant i \leqslant N$ and $d(x, y) < \delta$.

If $n > N$ and $d(x, y) < \delta$, it follows that

$$|f_n(x) - f_n(y)| \leqslant |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\varepsilon.$$

This proves the theorem. □

Theorem 7.25 If K is compact, if $\mathrm{f}_n \in \mathscr{C}(K)$ for $n = 1, 2, 3, dots$, and if $\{f_n\}$ is pointwise bounded and equicontinuous on K, then

a) $\{f_n\}$ is uniformly bounded on K,

b) $\{f_n\}$ contains a uniformly convergent subsequence.

Proof. a) Let $\varepsilon > 0$ be given and choose $\delta > 0$, in accordance with Definition 7.22, so that $|f_n(x) - f_n(y)| < \varepsilon$ for all n provided that $d(x, y) < \delta$.

Since K is compact, there are finitely many points $p_1, \ldots, p_r$ in K such that to every $x \in K$ corresponds at least one p_i with $d(x, p_i) < \delta$. Since $\{f_n\}$ is pointwise bounded, there exist $M_i < \infty$ such that $|f_n(p_i)| < M_i$ for all n. If $M = \max(M_1, \ldots, M_r)$, then $|f_n(x)| < M + \varepsilon$ for every $x \in K$. This proves (a).

b) Let E be a countable dense subset of K. Theorem 7.23 shows that $\{f_n\}$ has a subsequence $\{f_{n_i}\}$ such that $\{f_{n_i}(x)\}$ converges for every $x \in E$.

Put $f_{n_i} = g_i$, to simplify the notation. We shall prove that $\{g_i\}$converges uniformly on K.

Let $\varepsilon > 0$, and pick $\delta > 0$ as in the beginning of this proof. Let $V(x, \delta)$ be the set of all $y \in K$ with $d(x, y) < \delta$. Since E is dense in K, and K is compact, there are finitely many points $x_1, \ldots, x_m$ in E such that $K \subset V(x_1, \delta) \cup \cdots \cup V(x_m, \delta)$.

Since $\{g_i(x)\}$ converges for every $x \in E$, there is an integer N such that $|g_i(x_s) - g_j(x_s)| < \varepsilon$ whenever $i \geqslant N$, $j \geqslant N$, $1 \leqslant s \leqslant m$.

If $x \in K$, then $x \in V(x_s, \delta)$for some s, so that $|f_i(x) - g_i(x_s)| < \varepsilon$ for every i. If $i \geqslant N$, $j \geqslant N$, it follows that

$$|g_i(x) - g_j(x)| \leqslant |g_i(x) - g_i(x_s)| + |g_i(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)| < 3\varepsilon$$

This completes the proof.

□

Theorem 7.26 (Stone-Weierstrass Theorem) If f is a continuous complex 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 $[a, b]$. If f is real, the P_n may be taken real.

Proof. We may assume, without loss of generality, that $[a, b] = [0, 1]$. We may also assume that $f(0) = f(1) = 0$. For if the theorem is proved for this case, consider $g(x) = f(x) - f(0) - x[f(1) - f(0)]$. Here $g(0) = g(1) = 0$, and if g can be obtained as the limit of a uniformly convergent sequence of polynomials, it is clear that the same is true for f, since $f - g$ is a polynomial. Furthermore, we define $f(x)$ to be zero for x outside of $[0, 1]$. Then f is uniformly continuous on the whole line.

We put $Q_n(x) = c_n(1 - x^2)^n$ where c_n is chosen so that $\int_{-1}^{1} Q_n(x)\, dx = 1$. We need some information about the order of magnitude of c_n. Since

$$\int_{-1}^{1} (1-x^2)^n\, dx = 2\int_0^1 (1-x^2)^n\, dx \geqslant 2\int_0^{1/\sqrt{n}} (1-x^2)^n\, dx \geqslant 2\int_0^{1/\sqrt{n}} (1-nx^2)\, dx = \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}},$$

it follows that $c_n < \sqrt{n}$.

The inequality $(1-x^2)^n \geqslant 1-nx^2$ which is used above is easily shown to be true by considering the function $(1-x^2)^n - 1 + nx^2$ which is zero at $x = 0$ and whose derivative is positive in $(0,1)$. For any $\delta > 0$ the fact that $c_n < \sqrt{n}$ implies $Q_n(x) \leqslant \sqrt{n}(1-\delta^2)^n$ $(\delta \leqslant |x| \leqslant 1)$.

Now set $P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)\,dt$. Our assumptions about f show, by a simple change of variable, 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, which are real if f is real.

Given $\varepsilon > 0$, we choose $\delta > 0$ such that $|y-x| < \delta$ implies $|f(y) - f(x)| < \frac{\varepsilon}{2}$. Let $M = \sup |f(x)|$. Using the fact that $Q_n(x) \geqslant 0$ in contingency with our findings, we see that for $0 \leqslant x \leqslant 1$,

$$|P_n(x) - f(x)| = \left| \int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)\,dt \right| \leqslant \int_{-1}^{1} |f(x+t) - f(x)|Q_n(t)\,dt$$

$$\leqslant 2M \int_{-1}^{-\delta} Q_n(t)\,dt + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)\,dt + 2M \int_{\delta}^{1} Q_n(t)dt \leqslant 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} < \varepsilon$$

for all large enough N, which proves the theorem. □

Corollary 7.27 For every interval $[-a,a]$ there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and such that

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

uniformly on $[-a,a]$.

Proof. By Theorem 7.26, there exists a sequence $\{P_n^*\}$ of real polynomials which converge to $|x|$ uniformly on $[-a,a]$. In particular $P_n^*(0) \to 0$ as $n \to \infty$. The polynomials $P_n(x) = P_n^*(x) - P_n^*(0)$ have desired properties. □

Theorem 7.29 Let $\mathscr{B}$ be the uniform closure of an algebra $\mathscr{A}$ of bounded functions. Then $\mathscr{B}$ is a uniformly closed algebra.

Proof. If $f \in \mathscr{B}$ and $g \in \mathscr{B}$, there exist uniformly convergent sequences $\{f_n\}$, $\{g_n\}$ such that $f_n \to f$, $g_n \to g$ and $f_n \in \mathscr{A}$, $g_n \in \mathscr{A}$. Since we are dealing with bounded functions, it is easy to show that $f_n + g_n \to f + g$, $f_n g_n \to fg$, $cf_n \to cf$, where c is any constant, the convergence being uniform in each case.

Hence $f + g \in \mathscr{B}$, $fg \in \mathscr{B}$, and $cf \in \mathscr{B}$, so that $\mathscr{B}$ is an algebra.. By Theorem 2.27, $\mathscr{B}$ is (uniformly) closed. □

Theorem 7.31 Suppose $\mathscr{A}$ is an algebra of functions on a set E, $\mathscr{A}$ separates points on E, and $|A$ vanishes at no points of E. Suppose x_1, x_2 are distinct points of E, and c_1, c_2 are constants. Then $\mathscr{A}$ contains a function of f such that:

$$f(x_1) = c_1 \qquad f(x_2) = c_2$$

Proof. The assumptions show that $\mathscr{A}$ contains functions g, h, and k such that $g(x_1) \neq g(x_2)$, $h(x_1) \neq 0$, $k(x_2) \neq 0$. Put $u = gk - g(x_1)k$ and $v = gh - g(x_2)h$. Then $u, v \in \mathscr{A}$, $x(x_1) = v(x_2) = 0$, $u(x_2) \neq 0$, and $v(x_1) \neq 0$. Therefore $f = \frac{c_1 v}{v(x_1)} + \frac{c_2 u}{u(x_2)}$ has the desired properties. □

Theorem 7.32 Let $\mathscr{A}$ be an algebra of real continuous functions on a compact set K. IF $\mathscr{A}$ separates points on K and if $\mathscr{A}$ vanishes at no point of K, then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all real continuous functions on K.

Proof. Step 1: If $f \in \mathscr{B}$, then $|f| \in \mathscr{B}$.

Let $a = \sup|f(x)|$ and let $\varepsilon > 0$ be given. By Corollary 7.27 there exists real numbers $c_1, \ldots, c_n$ such that $|\sum_{i=1}^{n} c_i y^i - |y|| < \varepsilon$ for $-a \leqslant y \leqslant a$. Since $\mathscr{B}$ is an algebra, the function $g = \sum_{i=1}^{n} c_i f^i$ is a member of $\mathscr{B}$. Thus, $|g(x) - |f(x)|| < \varepsilon$ for $x \in K$. Since $\mathscr{B}$ is uniformly closed, this shows that $|f| \in \mathscr{B}$.

Step 2: If $f \in \mathscr{B}$ and $g \in \mathscr{B}$, then $\max(f, g) \in \mathscr{B}$ and $\min(f, g) \in \mathscr{B}$. (By $\max(f, g)$ we mean the function h defined by

$$h(x) = \begin{cases} f(x) & f(x) \geqslant g(x) \\ g(x) & f(x) < g(x) \end{cases}$$

and $\min(f, g)$ is defined likewise)

Step 2 follows from step 1 and the identities $\max(f, g) = \frac{f+g}{2} + \frac{|f-g|}{2}$, $\min(f, g) = \frac{f+g}{2} - \frac{|f-g|}{2}$. By iteration, the results can of course be extended to any finite set of functions: If $f_1, \ldots, f_n \in \mathscr{B}$, then $\max(f_1, \ldots, f_n) \in \mathscr{B}$, and $\min(f_1, \ldots, f_n \in \mathscr{B}$.

Step 3: Given a real function f, continuous on K, a point $x \in K$ and $\varepsilon > 0$, there exists a function $g_x \in \mathscr{B}$ such that $g_x(x) = f(x)$ and $g_x(t) > f(t) - \varepsilon$ for $t \in K$.

Since $\mathscr{A} \subset \mathscr{B}$ and $\mathscr{A}$ satisfies the hypotheses of Theorem 7.31 so does $\mathscr{B}$. Hence, for every $y \in K$, we can find a function $h_y \in \mathscr{B}$ such that $h_y(x) = f(x)$, $h_y(y) = f(y)$. By the continuity of h_y, there exists an open set J_y, containing y, such that $h_y(t) > f(t) - \varepsilon$. Since K is compact, there is a finite set of points $y_1, \ldots, y_n$ such that $K \subset J_{y_1} \cup \cdots \cup J_{y_n}$. Put $g_x = \max(h_{y_1}, \ldots, h_{y_n})$. By Step 2, $g_x \in \mathscr{B}$, and the aforementioned relations show that g_x has the other required properties.

Step 4: Given a real function f, continuous on K, and $\varepsilon > 0$ there exists a function $h \in \mathscr{B}$ such that $|h(x) - f(x)| < \varepsilon$. Since $\mathscr{B}$ is uniformly closed, this statement is equivalent to the conclusion of the theorem.

Let us consider the functions g_x, for each $x \in K$, constructed in step 3. By the continuity of g_x, there exist open sets V_x containing x, such that $g_x(t) < f(t) + \varepsilon$ ($t \in V_x$). Since K is compact, there exists a finite set of points $x_1, \ldots, x_m$ such that $K \subset V_{x_1} \cup V_{x_m}$. Put $h = \min(g_{x_1}, \ldots, g_{x_m})$. By step 2, $h \in \mathscr{B}$ and $h(t) > f(t) - \varepsilon$. However By Step 3, $h(t) < f(t) + \varepsilon$. Thus, $|h(x) - f(x)| < \varepsilon$.

□

Theorem 7.33 Suppose $\mathscr{A}$ is a self-adjoint algebra of complex continuous functions on a compact set K, $\mathscr{A}$ separates points on K, and $\mathscr{A}$ vanishes at no point of K. Then the uniform closure $\mathscr{B}$ of $|A$ consists of all complex continuous functions on K. In other words, $\mathscr{A}$ is dense in $\mathscr{C}(K)$.

Proof. Let $\mathscr{A}_R$ be the set of all real functions on K which belong to $\mathscr{A}$.

If $f \in \mathscr{A}$ and $f = u + iv$, with $u, v \in \mathbb{R}$, then $2u = f + \overline{f}$, and since $\mathscr{A}$ is self-adjoint, we see that $u \in \mathscr{A}_R$. If $x_1 \neq x_2$, there exists $f \in \mathscr{A}$ such that $f(x_1) = 1$, $f(x_2) = 0$; hence $0 = u(x_2) \neq u(x_1) = 1$, which shows that $\mathscr{A}_R$ separates points on K. If $x \in K$, then $g(x) \neq 0$ for some $g \in \mathscr{A}$, and there is a complex number λ such that $\lambda g(x) > 0$; if $f = \lambda g$, $f = u + iv$, it follows that $u(x) > 0$; hence $\mathscr{A}_R$ vanishes at no point of K.

Thus $\mathscr{A}_R$ satisfies the hypotheses of Theorem 7.32. It follows that every real continuous function on K lies in the uniform closure of $\mathscr{A}_R$, hence lies in $\mathscr{B}$. If f is a complex continuous function on K, $f = u+iv$, then $u \in \mathscr{B}$, $v \in \mathscr{B}$, hence $f \in \mathscr{B}$. This completes the proof. □

Chapter 8 Some Special Functions

Definitions

Analytic Functions Functions of the form

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

Exponential Function Define

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

The ratio test shows that this series converges for every complex z. Note:

$$E(z)E(w) = \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = E(z+w)$$

Thus, $E(z)E(-z) = 1$. Further,

$$E'(z) = \lim_{h\to 0} \frac{E(z+h)-E(z)}{h} = \lim_{h\to 0} \frac{E(z+h)-1}{h} E(z)$$

Let $E(1) = e$. So $E(n) = E(1+1+1+\cdots+1) = E(1)E(1)\ldots E(1) = e^n$. This holds for any $n \in \mathbb{Q}$. Furthermore, $E(x) = e^x = \sup e^p$ ($p < x$, p rational).

Triginometric Functions Define the following:

$$C(x) = \frac{1}{2}[E(ix) + E(-ix)] \qquad S(x) = \frac{1}{2i}[E(ix) - E(-ix)]$$

Note: $E(ix) = C(x) + iS(x)$. Further,

$$C'(x) = -S(x) \qquad S'(x) = C(X)$$

Ultimately equivalent to cos and sin.

Trigonometric Polynomial A trigonometric polynomial is a finite sum of the form

$$f(x) = a_0 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) \qquad (x real),$$

where $a_0, \ldots, a_N, b_1, \ldots, b_N$ are complex numbers. The above identities can also be written in the form

$$f(x) = \sum_{-N}^{N} c_n e^{inx}$$

It follows that:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\,dx = \begin{cases} 1 & (if n = 0) \\ 0 & (if n = \pm 1, \pm 2, \ldots) \end{cases}$$

Fourier Coefficients If f is an integrable function on $[-\pi, \pi]$, the numbers c_m defines by:

$$c_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}\,dx$$

for all integers m are called the *Fourier coefficients* of f,

Fourier Series The series:

$$\sum_{-\infty}^{\infty} c_n e^{inx}$$

formed with the Fourier coefficients is called the *Fourier series* of f.

Orthogonal System of Functions/ Orthonormal Let $\{\phi_n\}$ $(n = 1, 2, 3, \dots)$ be a sequence of complex functions on $[a, b]$ such that

$$\int_a^b \phi_n(x)\overline{\phi_m(x)}\,dx = 0 \qquad (n \neq m).$$

Then $\{\phi_n\}$ is said to be an *orthogonal system of functions* on $[a, b]$. If in addition:

$$\int_a^b |\phi_b(x)|^2\,dx = 1$$

for all n, $\{\phi_n\}$ is said to be *orthonormal*.

Gamma Function For $0 < x < \infty$

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

The integral converges for these x. (When $x < 1$, both 0 and ∞ have to be looked at.)

Theorems

Theorem 8.1 Suppose the series

$$\sum_{n=0}^{\infty} c_n x^n$$

converges for $|x| < R$, and define

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < R).$$

Then the series converges uniformly on $[-R+\varepsilon, R-\varepsilon]$ no matter which $\varepsilon > 0$ is chosen. The function f is continuous and differentiable in $(-R, R)$.

Corollary Under the hypotheses of Theorem 8.1, f has derivatives of all orders in $(-R, R)$, which are given by:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\ldots(n-k+1)c_n x^{n-k}.$$

In particular, $f^{(k)}(0) = k!c_k$.

Theorem 8.2 Suppose $\sum c_n$ converges. Put

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < 1).$$

Then

$$\lim_{x \to 1} f(x) = \sum_{n=0}^{\infty} c_n.$$

Theorem 8.3 Given a double sequence $\{a_{ij}\}$, $i = 1, 2, 3, \ldots, j = 1, 2, 3, \ldots$, suppose that:

$$\sum_{j=1}^{\infty} |a_{ij}| = b_i$$

and $\sum b_i$ converges. Then

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$$

Theorem 8.4 (Taylor's Theorem) Suppose

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < R).$$

If $-R < a < R$, then f can be expanded in a power series about the point $x - a$ which converges in $|x - a| < R - |a|$, and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \qquad (|x - a| < R - |a|)$$

Theorem 8.5 Suppose the series $\sum a_n x^n$ and $\sum b_n x^n$ converge in the segment $S = (-R, R)$. Let E be the set of all $x \in S$ at which

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

If E has a limit point in S, then $a_n = b_n$ for $n = 0, 1, 2, \ldots$. Hence the above equation holds for all $x \in S$.

Theorem 8.6 Let e^x be defined on R^1 as it is above. Then:

a) e^x is continuous and differentiable for all x;

b) $(e^x)' = e^x$;

c) e^x is a strictly increasing function of x and $e^x > 0$;

d) $e^{x+y} = e^x e^y$;

e) $e^x \to +\infty$ as $x \to +\infty$, $e^x \to 0$ as $x \to -\infty$;

f) $\lim_{x\to+\infty} x^n e^{-x} = 0$ for every n.

Theorem 8.7 a) The function E is periodic, with period $2\pi i$.

b) The functions C and S are periodic with period 2π.

c) If $0 < t < 2\pi$ then $E(it) \neq 1$.

d) If z is a complex number with $|z| = 1$, there is a unique $r \in [0, 2\pi)$ such that $E(it) = z$.

Theorem 8.8 Suppose $a_0, \ldots, a_n$ are complex number, $n \geqslant 1$, $a_n \neq 0$,

$$P(z) = \sum_0^n a_k z^k.$$

Then $P(z) = 0$ for some complex number z.

Theorem 8.11 Let $\{\phi_n\}$ be orthonormal on $[a, b]$. Let

$$s_n(x) = \sum_{m=1}^n c_m \phi_m(x)$$

be the nth partial sum of the Fourier series of f, and suppose

$$t_n(x) = \sum_{m=1}^n \gamma_m \phi_m(x).$$

Then

$$\int_a^b |f - s_n|^2\, dx \leqslant \int_a^b |f - t_n|^2\, dx,$$

and equality holds if and only if $\gamma_m = c_m$.

Corollary If $f(x) = 0$ for all x in some segment J, then $\lim_{s_N}(f; x) = 0$ for every $x \in J$.

Theorem 8.15 If f is continuous (with period 2π) and if $\varepsilon > 0$, then there is a trigonometric polynomial P such that $|P(x) - f(x)| < \varepsilon$ for all real x.

Theorem 8.16 (Parseval's Theorem) Suppose f and g are Riemann-integrable functions with period 2π, and

$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx} \qquad g(x) \sim \sum_{-\infty}^{\infty} \gamma_n e^{inx}.$$

Then

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 \, dx = 0$$

$$\frac{1}{2\pi} int_{-\pi}^{\pi} f(x)\overline{g(x)} \, dx = \sum_{-\infty}^{\infty} c_n \overline{\gamma}_n$$

$$\frac{1}{2\pi} int_{-\pi}^{\pi} |f(x)^2 \, dx = \sum_{-\infty}^{\infty} |c_n|^2$$

Theorem 8.18 a) The function equation:

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

holds if $0 < x < \infty$.

b) $\Gamma(n+1) = n!$ for $n = 1, 2, 3, \ldots$.

c) $\log \Gamma$ is convex on $(0, \infty)$.

Theorem 8.19 If f is a positive function on $(0, \infty)$ such that

a) $f(x+1) = xf(x)$

b) $f(1) = 1$

c) $\log f$ is convex

then $f(x) = \Gamma(x)$.

Theorem 8.20 If $x > 0$ and $y > 0$, then

$$\int_0^1 t^{x-1}(1-t)^{y-1} \, dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

This integral is the so called *beta function* $B(x, y)$.

Stirling's Formula This provides a simple approximate expression for $\Gamma(x+1)$ when x is large (hence for $n!$ when n is large). The formula is

$$\lim_{x\to\infty} \frac{\Gamma(x+1)}{(x/e)^x \sqrt{2\pi x}} = 1$$

Chapter 9 Functions of Several Variables

Definitions

Vector Space A nonempty set $X \subset \mathbb{R}^n$ is a *vector space* if $\mathbf{x} + \mathbf{y} \in X$ and $c\mathbf{x} \in X$ for all $\mathbf{x}, \mathbf{y} \in X$, $c \in \mathbb{R}$.

Linear Combination If $\mathbf{x}_1, \dots, \mathbf{x}_k \in R^n$ and $c_1, \dots, c_k$ are scalars, the vector

$$c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k$$

is called a *linear combination* of $\mathbf{x}_1, \dots, \mathbf{x}_k$.

Span If $S \subset R^n$ and if E is the set of all linear combinations of elements of S we say that S *spans* E, or that E is *the span* of S. Observe that every span is a vector space.

Independent/ Dependent A set consisting of vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ (we shall use the notation $\{x_1, \dots, \mathbf{x}_k\}$ for such a set) is said to be *independent* if the relation $c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = 0$ implies that $c_1 = \cdots = c_k = 0$. Otherwise $\{x_1, \dots, \mathbf{x}_k\}$ is said to be *dependent.*

Dimension If a vector space X contains an independent set of r vectors but contains no independent set of $r+1$ vectors, we say that X has *dimension* r, and write: $\dim X = r$.

Basis/ Coordinates/ Standard Basis An independent subset of a vector space X which spans X is called a *basis* of X. Observe that if $B = \{x_1, \dots, \mathbf{x}_r\}$ is the basis of X, then every $\mathbf{x} \in X$ has a unique representation of the for $\mathbf{x} = \sum c_j\mathbf{x}_j$. Such a representation exists since V spans X, and it is unique since B is independent. The numbers $c_1, \dots, c_r$ are called the *coordinates* of $\mathbf{x}$ with respect to the basis B. The most familiar example of a basis is the set $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, where $\mathbf{e}_j$ is the vector in $\mathbb{R}^n$ whose jth coordinate is 1 and whose other coordinates are all 0. If $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = (x_1, \dots, x_n)$, then $\mathbf{x} = \sum x_j\mathbf{e}_j$. We shall call $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ the *standard basis*

Linear Transformation A mapping A of a vector space X into a vector space Y is said to be a *linear transformation* if

$$A(c\mathbf{x}_1 + \mathbf{x}_2) = cA(\mathbf{x}_1) + A(x_2)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and all scalars c. Note that $A\mathbf{x} = A(\mathbf{x})$. Further, a linear transformation A of X into Y is completely determined by its action on any basis.

Linear Operators A linear transformations of X into X are often called *linear operators* on X.

Invertible If A is a linear operator on X which (i) is one-to-one and (ii) maps X onto X, we say that A is invertible. In this case we can define an operator A^{-1} on X by requiring that $A^{-1}(A\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in X$.

Set of Linear Transformation Let $L(X, Y)$ be the *set of all linear transformations* of the vector space X into the vector space Y. Instead of $L(X, X)$ we shall simply write $L(X)$. If $A_1, A_2 \in L(X, Y)$ and if c_1, c_2 are scalars, define $c_1A_1 + c_2A_2$ by

$$(c_1A_1 + c_2A_2)\mathbf{x} = c_1A_1\mathbf{x} + c_2A_2\mathbf{x}$$

Clearly $c_1A_1 + c_2A_2 \in L(X, Y)$

Product If X, Y, Z are vector spaces and if $A \in L(X, Y)$ and $B \in L(Y, Z)$, we define their *product* BA to be the composition of A and B:

$$(BA)\mathbf{x} = B(A\mathbf{x}) \qquad (\mathbf{x} \in X)$$

Then $BA \in L(X, Z)$. Note that BA need not be the same as AB, even if $X = Y = Z$.

Norm For $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, define the *norm* $||A||$ of A to be the sup of all numbers $|A\mathbf{x}|$, where $\mathbf{x}$ ranges over all vectors in $\mathbb{R}^n$ with $|\mathbf{x}| \leqslant 1$. Observe that the inequality

$$|A\mathbf{x}| \leqslant ||A|||\mathbf{x}|$$

holds for all $\mathbf{x} \in \mathbb{R}^n$. Also, if λ is such the $|A\mathbf{x}| \leqslant \lambda|\mathbf{x}|$ for all $\mathbf{x} \in \mathbb{R}^n$, then $||A|| \leqslant \lambda$

Matrices Omitted, trivial.

Differentiable Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, and $\mathbf{x} \in E$. If there exists a linear transformation A of $\mathbb{R}^n$ into $\mathbb{R}^m$ such that

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

then we say that $\mathbf{f}$ is *differentiable* at $\mathbf{x}$, and we write:

$$\mathbf{f}'(\mathbf{x}) = A$$

If $\mathbf{f}$ is differentiable at every $x \in E$ we say that $\mathbf{f}$ is differentiable in E.

If $|\mathbf{h}|$ is small enough then $\mathbf{x} + \mathbf{h} \in E$, since E is open. Thus $\mathbf{f}(\mathbf{x} + \mathbf{h})$ is defined, $\mathbf{f}(\mathbf{x} + \mathbf{h}) \in \mathbb{R}^m$, and since $\mathscr{A} \in L(\mathbb{R}^n, \mathbb{R}^m)$, $A\mathbf{h} \in \mathbb{R}^m$. Thus

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(x) - A\mathbf{h} \in \mathbb{R}^m.$$

Notes a)

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

can be rewritten in the form:

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

where the remainder $r(\mathbf{h})$ satisfies: $\lim_{\mathbf{h}\to 0} \frac{|\mathbf{r}(\mathbf{h})|}{|h|} = 0$. That is, for fixed $\mathbf{x}$ and small $\mathbf{h}$ the left side is approximately equal to $\mathbf{f}'(\mathbf{x})\mathbf{h}$, that is, to the value of the linear transformation applied to $\mathbf{h}$.

b) If $\mathbf{f}$ is differentiable in E then $\mathbf{f}'(\mathbf{x})$ is a function that maps E into $L(\mathbb{R}^n, \mathbb{R}^m)$.

c) $\mathbf{f}$ is continuous at any point at which $\mathbf{f}$ is differentiable.

d) The aforementioned derivative in part (a) is called the differential of $\mathbf{f}$ at $\mathbf{x}$, or the total derivative, to distinguish it from the partial derivatives.

Components Consider $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. The *components* of $\mathbf{f}$ are the real functions $f_1, \ldots, f_m$ defined by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u}_i \qquad (\mathbf{x} \in E)$$

Partial Derivative For $\mathbf{x} \in E$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, we define:

$$(D_j f_i)(\mathbf{x}) = \lim_{t \to 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t},$$

provided the limit exists. Writing $f_i(x_1, \ldots, x_n)$ in place of $f_i(\mathbf{x})$ we see that $D_j f_i$ is the derivative of f_i with respect to x_j, keeping the other variables fixed. The notation $\frac{\partial f_i}{\partial x_j}$ is therefore often used in place of $D_j f_i$, and $D_j f_i$ is called a *partial derivative.*

Continuously Differentiable A differentiable mapping $\mathbf{f}$ of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is said to be *continuously differentiable* in E if $\mathbf{f}'$ is a continuous mapping of E into $L(\mathbb{R}^n, \mathbb{R}^m)$. More explicitly, it is required that to every $\mathbf{x} \in E$ and to every $\varepsilon > 0$ corresponds a $\delta > 0$ such that

$$||\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})|| < \varepsilon$$

if $\mathbf{y} \in E$ and $|\mathbf{x} - \mathbf{y}| < \delta$.

Contraction Let X be a metric space, with metric d. If φ maps X into x and if there is a number $c < 1$ such that

$$d(\varphi(x), \varphi(y)) \leqslant cd(x, y)$$

for all $x, y \in X$, then φ is said to be a *contraction* of X into X.

Fixed Point For $\varphi : X \to X$ a point $x \in X$ such that $\phi(x) = x$ is called a *fixed point.*

Notation for Implicit Function Theorem If $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$, let us write $(\mathbf{x}, \mathbf{y})$ for the point (or vector)

$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}$$

In what follows, the first entry in $(\mathbf{x}, \mathbf{y})$ or in a similar symbol will always be a vector in $\mathbb{R}^n$ and the second a vector in $\mathbb{R}^m$.

Every $A \in L(R^{n+m}, R^n)$ can be split into two linear transformations A_x and A_y defined by

$$A)x\mathbf{h} = A(\mathbf{h}, 0), \qquad A_y\mathbf{k} = A(0, \mathbf{k})$$

for any $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{k} \in \mathbb{R}^m$. Then $A_x \in L(\mathbb{R}^n)$ and $A_y \in L(R^m, R^n)$, and

$$A(\mathbf{h}, \mathbf{k}) = A_x\mathbf{h} + a_y\mathbf{k}.$$

Null Space The *null space* of A, $\mathscr{N}(A)$, is the set of all $\mathbf{x} \in X$ at which $A\mathbf{x} = \mathbf{0}$. It is clear that $\mathscr{N}(A)$ is a vector space in X.

Range The *range* of A, $\mathscr{R}(A)$, is a vector space in Y.

Rank The *rank* of A is defined to be the dimension of $\mathscr{R}(A)$.

Projection Let X be a vector space. An Operator $P \in L(X)$ is said to be a *projection* in X if $P^2 = P$.

More explicitly, the requirement is that $P(P\mathbf{x}) = P\mathbf{x}$ for every $\mathbf{x} \in X$. IN other words, p fixes every vector in its range $\mathscr{R}(P)$. Some elementary properties:

a) If P is a projection in X, then every $\mathbf{x} \in X$ has a unique representation of the form $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where $\mathbf{x}_1 \in \mathscr{R}(P)$, $\mathbf{x}_2 \in \mathscr{N}(P)$.

b) If X is a finite-dimensional vector space and if X_1 is a vector space in X, then there is a projection P in X with $\mathscr{R}(P) = X_1$.

Determinants If $(j_1, \ldots, j_n)$ is an ordered n-tuples, define

$$s(j_1, \ldots, j_n) = \prod_{p<q} sgn(j_q - j_p)$$

where sgn is the sign. Let $[A]$ be the matrix of a linear operator A on $\mathbb{R}^n$, relative to the standard basis $\{e_1, \ldots, e_n\}$, with entries a_{ij} in the ith row and jth column.

$$\det[A] = \sum s(j_1, \ldots, j_n) a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

The sum extends over n-tuples of integers. Let $\mathbf{x}_i$ be the ith column vector of A.

$$\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det[A].$$

Jacobians If $\mathbf{f}$ maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and if $\mathbf{f}$ is differentiable at a point $\mathbf{x} \in E$, the determinant of the linear operator $\mathbf{f}'(\mathbf{x})$ is called the *Jacobian* of $\mathbf{f}$ at $\mathbf{x}$:

$$J_{\mathbf{f}}(\mathbf{x}) = \det \mathbf{f}'(\mathbf{x})$$

For $(y_1, \ldots, y_n) = \mathbf{f}(x_1, \ldots, x_n)$, we shall also use the notation:

$$\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}$$

Second-order Partial Derivatives Suppose f is a real function defined in an open set $E \subset \mathbb{R}^n$ with partial derivatives $D_1 f, \ldots, D_n f$. If the functions $D_j f$ are themselves differentiable, then the *second-order partial derivatives* of f are defined by

$$D_{ij} f = D_i D_j f \qquad (i, j = 1, \ldots, n)$$

Theorems

Theorem 9.2 Let r be a positive integer. If a vector space X is spanned by a set of r vectors, then $\dim X \leqslant r$.

Proof. If this is false, there is a vector space X which contains an independent set $Q = \{\mathbf{y}_1, \ldots, \mathbf{y}_{r+1}\}$ and which is spanned by a set S_0 consisting of r vectors.

Suppose $0 \leqslant i < r$, and suppose a set S_i has been constructed which spans X and which consists of all $\mathbf{y}_j$ with $1 \leqslant j \leqslant i$ plus a certain collection of $r - i$ members of S_0, say $\mathbf{x}_1, \ldots, \mathbf{x}_{r-i}$. (In other words, S_i is obtained from S_0 by replacing i of its elements by members of $\mathbb{Q}$, without altering the span.) Since S_i spans X, $\mathbf{y}_{i+1}$ is in the span of S_i; hence there are scalars $a_1, \ldots, a_{i+1}$, $b_1, \ldots, b_{r-i}$, with $a_{i+1} = 1$, such that $\sum_{j=1}^{i+1} a_j \mathbf{y}_j + \sum_{k=1}^{r-i} b_k \mathbf{x}_k = 0$. If all b_k's were 0, the independence of Q would force all a_j's to be 0, a contradiction. If follows that some $\mathbf{x}_k \in S_i$ is a linear combination of the other members of $T_i = S_i \cup \{\mathbf{y}_{i+1}\}$. Remove this $\mathbf{x}_k$ from T_i and call the remaining set S_{i+1}. Then S_{i+1} spans the same set as T_i, namely X, so that S_{i+1} has the properties postulated for S_i with $i + 1$ in place of i.

Starting with S_0, we thus construct sets $S_1, \ldots, S_r$. The last of these consists of $\mathbf{y}_1, \ldots, \mathbf{y}_r$, and our construction shows that it spans X. But $\mathbb{Q}$ is independent; hence $\mathbf{y}_{r+1}$ is not in the span of S_r. This contradiction establishes the Theorem. □

Corollary $\dim \mathbb{R}^n = n$

Proof. Sine $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ spans $\mathbb{R}^n$, the theorem shows that $\dim \mathbb{R}^n \leqslant n$. Since $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is independent, $\dim \mathbb{R}^n \geqslant n$. □

Theorem 9.3 Suppose X is a vector space, and $\dim X = n$.

a) A set E of n vectors in X spans X if and only if E is independent.
b) X has a basis, and every basis consists of n vectors.
c) If $1 \leqslant r \leqslant n$ and $\{\mathbf{y}_1, \ldots \mathbf{y}_r\}$ is an independent set in X, then X has a basis containing $\{\mathbf{y}_1, \ldots \mathbf{y}_r\}$.

Proof. Suppose $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. Since $\dim X = n$, the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}\}$ is dependent, for every $\mathbf{y} \in X$. If E is independent, it follows that y is in the span of E; hence E spans X. Conversely, if E is dependent, one of its members can be removed without changing the span of E. Hence E cannot span X, by Theorem 9.2. This proves (a).

Since $\dim X = n$, X contains an independent set of n vectors, and (a) shows that every such set is a basis of X; (b) now follows from 9.1(d) and 9.2.

To prove (c), let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be a basis of X. The set $S = \{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$ spans X and is dependent, since it contains more than n vectors. The argument used in the proof of Theorem 9.2 shows that one of the $\mathbf{x}_i$'s is a linear combination of the other members of S. If we remove this $\mathbf{x}_i$ from S, the remaining set still spans X. This process can be repeated r times and leads to a basis of X which contains $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$, by (a). □

Theorem 9.5 A linear operator A on a finite-dimensional vector space X is one-to-one if and only if the range of A is all of X.

Proof. Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be a basis of X. The linearity of A shows that its range $\mathscr{R}(A)$ is the span of the set $Q = \{A\mathbf{x}_1, \ldots, A\mathbf{x}_n\}$. We therefore infer from Theorem 9.3(a) that $\mathscr{R}(A) = X$ if and only if Q is independent. We have to prove that this happens if and only if A is one-to-one.

Suppose A is one-to -one and $\sum c_i A\mathbf{x}_i = 0$. Then $A(\sum c_i \mathbf{x}_i) = 0$, hence $\sum c_i x_i = 0$, hence $c_1 = \cdots = c_n = 0$, and we conclude that Q is independent.

Conversely, suppose Q is independent and $A(\sum c_i x_i) = 0$. Then $\sum c_i A\mathbf{x}_i = 0$, hence $c_1 = \cdots = c_n = 0$, and we conclude: $A\mathbf{x} = 0$ only if $\mathbf{x} = 0$. If now $A\mathbf{x} = A\mathbf{y}$, then $A(x - y) = Ax - Ay = 0$, so that $x - y = 0$, and this says that A is one-to-one. □

Theorem 9.7 a) If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $||A|| < \infty$ and A is a uniformly continuous mapping of $\mathbb{R}^n$ into $\mathbb{R}^m$.

b) If $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ and c is a scalar, then

$$||A + B|| \leqslant ||A|| + ||B||, \qquad ||cA|| = |c|||A||$$

With the distance between A and B defined as $||A - B||$, $L(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space.

c) If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^n, \mathbb{R}^k)$ then

$$||BA|| \leqslant ||B||||A||$$

Proof. a) Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be the standard basis in $\mathbb{R}^n$and suppose $\mathbf{x} = \sum c_i \mathbf{e}_i$, $|\mathbf{x}| \leqslant 1$, so that $|c_i| \leqslant 1$ for $i = 1, \ldots, n$. Then

$$|A\mathbf{x}| = \left|\sum c_i A\mathbf{e}_i\right| \leqslant \sum |c_i||A\mathbf{e}_i| \leqslant \sum |A\mathbf{e}_i|$$

so that $||A|| \leqslant \sum_{i=1}^{n} |A\mathbf{e}_i < \infty$. Since $|A\mathbf{x} - A\mathbf{y}| \leqslant ||A|||x - y|$ if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we see that A is uniformly continuous.

b) The inequality in (b) follows from

$$|(A + B)\mathbf{x}| = |A\mathbf{x} + B\mathbf{x}| \leqslant |A\mathbf{x}| + |B\mathbf{x}| \leqslant (||A|| + ||B||)|\mathbf{x}|.$$

The second part of (b) is proved in the same manner. If $A, B, C \in L(\mathbb{R}^n, \mathbb{R}^m)$, we have the triangle inequality

$$||A - c|| = ||(A - B) + (B - C)|| \leqslant ||A - B|| + ||B - C||,$$

and it is easily verified that $||A - B||$ has the other properties of a metric.

c) Finally, (c) follows from

$$|(BA)\mathbf{x}| = |B(A\mathbf{x})| \leqslant ||B|| |A\mathbf{x}| \leqslant ||B|| ||A|| |\mathbf{x}|.$$

Since we now have metrics in the space $L(\mathbb{R}^n, \mathbb{R}^m)$, the concepts of open set, continuity, etc., make sense for these spaces. Our next theorem utilizes these concepts.

□

Theorem 9.8 Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.

a) If $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and

$$||B - A|| \cdot ||A^{-1}|| < 1$$

then $B \in \Omega$

b) Ω is an open subset of $L(\mathbb{R}^n)$, and the mapping $A \to A^{-1}$ is continuous on Ω.

Proof. a) Put $||A^{-1}|| = 1/\alpha$, put $||B - A|| = \beta$. Then $\beta < \alpha$. For every $\mathbf{x} \in \mathbb{R}^n$,

$$\alpha|\mathbf{x}| = \alpha|A^{-1}A\mathbf{x}| \leqslant \alpha||A^{-1}|| \cdot |A\mathbf{x}| = |A\mathbf{x}| \leqslant |(A - B)\mathbf{x}| + |B\mathbf{x}| \leqslant \beta|\mathbf{x}| + |B\mathbf{x}|,$$

so that $(\alpha - \beta)|\mathbf{x}| \leqslant |B\mathbf{x}|$. Since $\alpha - \beta > 0$, this shows that $B\mathbf{x} \neq 0$ if $\mathbf{x} = 0$. Hence B is 1-1. By Theorem 9.5, $B \in \Omega$. This holds for all B with $||B - A|| < \alpha$. Thus we have (a) and the fact that Ω is open.

b) Next, replace $\mathbf{x}$ by $B^{-1}\mathbf{y}$ in $(\alpha - \beta)|\mathbf{x}| \leqslant |B\mathbf{x}|$. The resulting inequality $(\alpha - \beta)|B^{-1}\mathbf{y}| \leqslant |BB^{-1}\mathbf{y}| = |\mathbf{y}|$ shows that $||B^{-1}|| \leqslant (\alpha - \beta)^{-1}$. The identity $B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}$, combined with Theorem 9.7 (c), implies therefore that $||B^{-1} - A^{-1}|| \leqslant ||B^{-1}|| ||A - B|| ||A^{-1}|| \leqslant \frac{\beta}{\alpha(\alpha - \beta)}$. This establishes the continuity assertion made in (b), since $\beta \to 0$ as $B \to A$.

□

Theorem 9.12 Suppose E and $\mathbf{f}$ are as in the definition of differentiable, $\mathbf{x} \in E$, and the following holds with $A = A_1$ and $A = A_2$:

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

Then $A_1 = A_2$.

Proof. If $B = A_1 - A_2$, the inequality $|B\mathbf{h}| \leqslant |\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_1\mathbf{h}| + |\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_2\mathbf{h}|$ shows that $|B\mathbf{h}|/|\mathbf{h}| \to 0$ as $\mathbf{h} \to 0$. For fixed $\mathbf{h} \neq 0$, it follows that $\frac{|B(t\mathbf{h})|}{|t\mathbf{h}|} \to 0$ as $t \to 0$. The linearity of B shows that the left side is independent of r. Thus $B\mathbf{h} = 0$ for every $\mathbf{h} \in \mathbb{R}^n$. Hence $B = 0$. □

Theorem 9.15 (Chain Rule) Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, $\mathbf{f}$ is differentiable at $\mathbf{x}_0 \in E$, $\mathbf{g}$ maps an open set containing $\mathbf{f}(E)$ into $\mathbb{R}^k$, and $\mathbf{g}$ is differentiable at $\mathbf{f}(\mathbf{x}_0)$. Then the mapping $\mathbf{F}$ of E into $\mathbb{R}^k$ defined by

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

is differentiable at $\mathbf{x}_0$, and

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0).$$

Proof. Put $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$, $A = \mathbf{f}'(\mathbf{x}_0)$, $B = \mathbf{g}'(\mathbf{y}_0)$, and define $\mathbf{u}(\mathbf{h}) = \mathbf{f}(\mathbf{x}_0+\mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - A\mathbf{h}$ and $\mathbf{v}(\mathbf{k}) = \mathbf{g}(\mathbf{y}_0 + \mathbf{k}) - \mathbf{g}(\mathbf{y}_0) - B\mathbf{k}$, for all $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{k} \in \mathbb{R}^m$ for which $\mathbf{f}(\mathbf{x}_0 + \mathbf{h})$ and $\mathbf{g}(\mathbf{y}_0 + \mathbf{k})$ are defined. Then $|\mathbf{u}(\mathbf{h})| = \varepsilon(\mathbf{h})|\mathbf{h}|$, $|\mathbf{v}(\mathbf{k})| = \eta(\mathbf{k})|\mathbf{k}|$, where $\varepsilon(\mathbf{h}) \to 0$ as $\mathbf{h} \to 0$ and $\eta(\mathbf{k}) \to 0$ as $\mathbf{k} \to 0$.

Given $\mathbf{h}$, put $\mathbf{k} = \mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0)$. The $|\mathbf{k}| = |A\mathbf{h} + \mathbf{u}(\mathbf{h})| \leqslant [||A|| + \varepsilon(h)]|h|$, and

$$\mathbf{F}(\mathbf{x}_0+\mathbf{h}) - \mathbf{F}(\mathbf{x}_0) - BA\mathbf{h} = \mathbf{g}(\mathbf{y}_0+\mathbf{k}) - \mathbf{g}(\mathbf{y}_0) - BA\mathbf{h} = B(\mathbf{k} - A\mathbf{h}) + \mathbf{v}(\mathbf{k}) = B\mathbf{u}(\mathbf{h}) + \mathbf{v}(\mathbf{k}).$$

Hence, for $\mathbf{h} \neq 0$,

$$\frac{\mathbf{F}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{F}(\mathbf{x}_0) - BA\mathbf{h}|}{|h|} \leqslant ||B||\varepsilon(\mathbf{h}) + [||A|| + \varepsilon(\mathbf{h})]\eta(\mathbf{k}).$$

Let $\mathbf{h} \to 0$. Then $\varepsilon(\mathbf{h}) \to 0$. Also, $\mathbf{k} \to 0$, so that $\eta(\mathbf{k}) \to 0$. It follows that $\mathbf{F}'(\mathbf{x}_0) = BA = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0)$. □

Theorem 9.17 Suppose $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$, and $\mathbf{f}$ is differentiable at a point $\mathbf{x} \in E$. Then the partial derivatives $(D_j f_i)(\mathbf{x})$ exists, and

$$\mathbf{f}'(\mathbf{x})\mathbf{e}_j = \sum_{i=1}^{m} (D_j f_i)(\mathbf{x})\mathbf{u}_i \qquad (1 \leqslant j \leqslant n).$$

Proof. Fix j. Since $\mathbf{f}$ is differentiable at $\mathbf{x}$, $\mathbf{f}(\mathbf{x} + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})(t\mathbf{e}_j) + \mathbf{r}(t\mathbf{e}_j)$ where $|\mathbf{r}(t\mathbf{e}_j)| / t \to 0$ as $t \to 0$. The linearity of $\mathbf{f}'(\mathbf{x})$ shows therefore that

$$\lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x})}{t} = \mathbf{f}'(\mathbf{x})\mathbf{e}_j.$$

If we now represent $\mathbf{f}$ in terms of its components, the limit becomes

$$\lim_{t\to 0} \sum_{i=1}^{m} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t}\mathbf{u}_i = \mathbf{f}'(\mathbf{x})\mathbf{e}_j.$$

It follows that each quotient in this sum has a limit, as $t \to 0$, so that each $(D_j f_i)(\mathbf{x})$ occupies the spot in the ith row and jth column of $[\mathbf{f}'(\mathbf{x})]$. Thus

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} (D_1 f_1)(\mathbf{x}) & \dots & (D_n f_1)(\mathbf{x}) \\ \dots & \dots & \dots \\ (D_1 f_m)(\mathbf{x}) & \dots & (D_n f_m)(\mathbf{x}) \end{bmatrix}$$

If $\mathbf{h} = \sum h_j \mathbf{e}_j$ is any vector in $\mathbb{R}^n$, then we see that $\mathbf{f}'(\mathbf{x})\mathbf{h} = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n} (D_j f_i)(\mathbf{x}) h_j \right\} \mathbf{u}_i$. □

Theorem 9.19 Suppose $\mathbf{f}$ maps a convex open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, $\mathbf{f}$ is differentiable in E, and there is a real number M such that $\|\mathbf{f}'(x)\| \leqslant M$ for every $\mathbf{x} \in E$. Then

$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| \leqslant M|\mathbf{b} - \mathbf{a}|$$

Proof. Fix $\mathbf{a} \in E$, $\mathbf{b} \in E$. Define $\gamma(t) = (1-t)\mathbf{a} + t\mathbf{b}$ for all $t \in \mathbb{R}$ such that $\gamma(t) \in E$. Since E is convex, $\gamma(t) \in E$ if $0 \leqslant t \leqslant 1$. Put $\mathbf{g}(t) = \mathbf{f}(\gamma(t))$. Then $\mathbf{g}'(t) = \mathbf{f}'(\gamma(t))\gamma'(t) = \mathbf{f}'(\gamma(t))(\mathbf{b} - \mathbf{a})$, so that $|\mathbf{g}'(t)| \leqslant \|\mathbf{f}'(\gamma(t))\||\mathbf{b} - \mathbf{a}| \leqslant M|\mathbf{b} - \mathbf{a}|$ for all $t \in [0,1]$. By Theorem 5.19, $|\mathbf{g}(1) - \mathbf{g}(0)| \leqslant M|\mathbf{b} - \mathbf{a}|$. But $\mathbf{g}(0) = \mathbf{f}(\mathbf{a})$ and $\mathbf{g}(1) = \mathbf{f}(\mathbf{b})$. This completes the proof. □

Corollary If, in addition $\mathbf{f}'(\mathbf{x}) = 0$ for all $\mathbf{x} \in E$, then $\mathbf{f}$ is constant.

Proof. To prove this, note that the hypotheses of the theorem hold now with $M = 0$. □

Theorem 9.21 Suppose $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$. Then $\mathbf{f} \in \mathscr{C}'(E)$ if and only if the partial derivatives $D_j f_i$ exists and are continuous on E for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$.

Proof. Assume first that $\mathbf{f} \in \mathscr{C}'(E)$. We know from Theorem 9.17 that $(D_j f_i)(\mathbf{x}) = (\mathbf{f}'(\mathbf{x})\mathbf{e}_j) \cdot \mathbf{u}_i$ for all i, j and for all $\mathbf{x} \in E$. Hence

$$(D_j f_i)(\mathbf{y}) - (D_j f_i)(\mathbf{x}) = \{[\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})]\mathbf{e}_j\} \cdot \mathbf{u}_i$$

and since $|\mathbf{u}_i| = |\mathbf{e}_j| = 1$, it follows that

$$|(D_j f_i)(\mathbf{y}) - (D_j f_i)(\mathbf{x})| \leqslant |[\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})]\mathbf{e}_j| \leqslant \|\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})\|.$$

Hence $D_j f_i$ is continuous.

For the converse, it suffices to consider the case $m = 1$. (Why?) Fix $\mathbf{x} \in E$ and $\varepsilon > 0$. Since E is open, there is an open ball $S \subset E$, with center at $\mathbf{x}$ and radius f, and the continuity of the functions $D_j f$ shows that r can be chosen so that $|(D_j f)(\mathbf{y}) - (D_j f)(\mathbf{x})| < \frac{\varepsilon}{n}$ for $\mathbf{y} \in S$.

Suppose $\mathbf{h} = \sum h_j \mathbf{e}_j$, $|\mathbf{h}| < r$, put $\mathbf{v}_0 = 0$, and $\mathbf{v}_k = h_1 e_1 + \cdots + h_k e_k$, for $1 \leqslant k \leqslant n$. Then

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} [f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + v_{j-1})].$$

Since $|\mathbf{v}_k| < r$ for $1 \leqslant k \leqslant n$ and since S is convex, the segments with end points $\mathbf{x} + \mathbf{v}_{j-1}$ and $\mathbf{x} + \mathbf{v}_j$ lie in S. Since $\mathbf{v}_j = \mathbf{v}_{j-1} + h_j \mathbf{e}_j$, the mean value theorem shows that the jth summand is equal to $h_j(D_j f)(\mathbf{x} + \mathbf{v}_{j-1} + \theta_j h_j \mathbf{e}_j)$ for some $\theta_j \in (0,1)$ and this differs from $h_j(D_j f)(\mathbf{x})$ by less that $|h_j|\varepsilon/n$. It follows that

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - \sum_{j=1}^{n} h_j (D_j f)(\mathbf{x}) \leqslant \frac{1}{n} \sum_{j=1}^{n} |h_j| e[s \leqslant |\mathbf{h}|\varepsilon$$

for all $\mathbf{h}$ such that $|\mathbf{h}| < r$.

This says that f is differentiable at $\mathbf{x}$ and that $f'(\mathbf{x})$ is the linear function which assigns the number $\sum h_j(D_jf)(\mathbf{x})$ to the vector $\mathbf{h} = \sum h_j\mathbf{e}_j$. The matrix $[f'(\mathbf{x})]$ consists of the row $(D_1f)(\mathbf{x}), \ldots, (D_nf)(\mathbf{x})$; and since these are continuous functions on E, $f \in \mathscr{C}'(E)$. □

Theorem 9.23 (Contraction Mapping Principle If X is a complete metric space, and if φ is a contraction of X into X, then there exists one and only one $x \in X$ such that $\varphi(x) = x$.

Proof. Pick $x_0 in X$ arbitrarily, and define $\{x_n\}$ recursively, by setting $x_{n+1} = \varphi(x_n)$. Choose $c < 1$ so that $d(\varphi(x), \varphi(y)) \leqslant cd(x, y)$ holds. For $n \geqslant 1$ we then have $d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1}) \leqslant cd(x_n, x_{n-1})$. Hence induction gives $d(x_{n+1}, x_n) \leqslant c^n d(x_1, x_0)$.

If $n < m$ it follows that $d(x_n, x_n) \leqslant \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \leqslant (c^n + c^{n+1} + \cdots + c^{m-1})d(x_1, x_0) \leqslant [(1-c)^{-1}]d(x_1, x_0)]c^n$. Thus $\{x_n\}$ is a Cauchy sequence. Since X is complete, $\lim x_n = x$ for some $x \in X$.

Since φ is a contraction, φ is continuous (in fact, uniformly continuous) on X. Hence $\varphi(x) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = x$. □

Theorem 9.24 (Inverse Function Theorem) Suppose $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $\mathbf{f}'(\mathbf{a})$ is invertible for some $\mathbf{a} \in E$, and $\mathbf{b} = \mathbf{f}(\mathbf{a})$. Then

a) there exists open sets U and V in $\mathbb{R}^n$ such that $\mathbf{a} \in U$, $\mathbf{b} \in V$, $\mathbf{f}$ is one-to-one on U, and $\mathbf{f}(U) = V$;

b) if $\mathbf{g}$ is the inverse of $\mathbf{f}$, defined in V by

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x} \qquad (\mathbf{x} \in U),$$

then $\mathbf{g} \in \mathscr{C}'(V)$.

Proof. a) Put $\mathbf{f}'(\mathbf{a}) = A$, and choose λ so that $2\lambda||A^{-1}|| = 1$. Since $\mathbf{f}'$ is continuous at $\mathbf{a}$, there is an open ball $U \subset E$, with center at $\mathbf{a}$, such that $||\mathbf{f}'(\mathbf{x}) - A|| < \lambda$ for $x \in U$.

We associate to each $\mathbf{y} \in \mathbb{R}^n$ a function φ, defined by $\varphi(\mathbf{x}) = \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x}))$ for $\mathbf{x} \in E$. Note that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ if and only if $\mathbf{x}$ is a fixed point of φ.

Since $\varphi'(\mathbf{x}) = I - A^{-1}\mathbf{f}'(\mathbf{x}) = A^{-1}(A - \mathbf{f}'(\mathbf{x}))$, we see that $||\varphi'(\mathbf{x})|| < \frac{1}{2}$ for $\mathbf{x} \in U$. Hence $|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \leqslant \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$ by Theorem 9.19. It follows that φ has at most one fixed point in U, so that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ for at most one $\mathbf{x} \in U$. Thus $\mathbf{f}$ is 1-1 in U.

Next put $V = f(U)$, and pick $\mathbf{y}_0 \in V$. Then $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ for some $\mathbf{x}_0 \in U$. Let B be an open ball with center at $\mathbf{x}_0$ and radius $r > 0$, so small that its closure $\overline{B}$ lies in U. We will show that $\mathbf{y} \in V$ whenever $|\mathbf{y} - \mathbf{y}_0| < \lambda r$. This proves, of course, that V is open.

Fix $\mathbf{y}$, $|\mathbf{y} - \mathbf{y}_0| < \lambda r$. With φ as defined earlier, $|\varphi(\mathbf{x}_0) - \mathbf{x}_0| = |A^{-1}(\mathbf{y} - \mathbf{y}_0)| < ||A^{-1}||\lambda r = \frac{r}{2}$. If $x \in \overline{B}$, it therefore follows that

$$|\varphi(\mathbf{x}) - \mathbf{x}_0| \leqslant |\varphi(\mathbf{x}) - \varphi(\mathbf{x}_0)| + |\varphi(\mathbf{x}_0) - \mathbf{x}_0 < \frac{1}{2}|\mathbf{x} - \mathbf{x}_0| + \frac{r}{2} \leqslant r;$$

hence $\varphi(\mathbf{x}) \in B$. Note that $|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \leqslant \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$ holds if $\mathbf{x}_1, \mathbf{x}_2 \in \overline{B}$.

Thus, φ is a contraction of $\overline{B}$ into $\overline{B}$. Being a closed subset of $\mathbb{R}^n$, $\overline{B}$ is complete. Theorem 9.23 implies therefore that φ has a fixed point $\mathbf{x} \in \overline{B}$. For this $\mathbf{x}$, $f(\mathbf{x}) = \mathbf{y}$. Thus $\mathbf{y} \in \mathbf{f}(\overline{B}) \subset \mathbf{f}(U) = V$

This proves part (a) of the theorem.

b) Pick $\mathbf{y} \in V$, $\mathbf{y} + \mathbf{k} \in V$. Then there exist $\mathbf{x} \in U$, $\mathbf{x} + \mathbf{h} \in U$, so that $\mathbf{y} = \mathbf{f}(x)$, $\mathbf{y} + \mathbf{k} = \mathbf{f}(\mathbf{x} + \mathbf{h})$. With φ as defined in (a),

$$\varphi(\mathbf{x} + \mathbf{h}) - \varphi(\mathbf{x}) = \mathbf{h} + A^{-1}[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x} + h)] = \mathbf{h} - A^{-1}\mathbf{k}$$

Because $|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \leqslant \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$, $|\mathbf{h} - A^{-1}\mathbf{k} \leqslant \frac{1}{2}|\mathbf{h}|$. Hence $|A^{-1}\mathbf{k}| \geqslant \frac{1}{2}|\mathbf{h}|$, and $|\mathbf{h}| \leqslant 2||A^{-1}|||\mathbf{k}| = \lambda^{-1}|\mathbf{k}|$.

Recognize that $\mathbf{f}'(x)$ has an inverse (we conclude this from our original assertions and Theorem 9.8), say T. Since $\mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k} = \mathbf{h} - T\mathbf{k} = -T[\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}]$, we see that

$$\frac{|\mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k}|}{|\mathbf{k}|} \leqslant \frac{||T||}{\lambda} \cdot \frac{|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}|}{|\mathbf{h}|}.$$

As $\mathbf{k} \to 0$, we see that $\mathbf{h} \to 0$. The right side of the last inequality thus tends to 0. Hence the same is true of the left. We have thus proved that $\mathbf{g}'(\mathbf{y}) = T$. But T was chosen to be the inverse of $\mathbf{f}'(\mathbf{x}) = \mathbf{f}'(\mathbf{g}(\mathbf{y}))$. Thus $\mathbf{g}'(\mathbf{y}) = \{\mathbf{f}'(\mathbf{g}(\mathbf{y}))\}^{-1}$.

Finally, note that $\mathbf{g}$ is a continuous mapping of V onto U (since $\mathbf{g}$ is differentiable, that $\mathbf{f}'$ is a continuous mapping of U into the set Ω of all invertible elements of $L(\mathbb{R}^n)$, and that inversion is a continuous mapping of Ω onto Ω, by Theorem 9.8. If we combine these facts with the last equation, we see that $\mathbf{g} \in \mathscr{C}'(V)$.

This completes the proof.

□

Theorem 9.25 If $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and if $\mathbf{f}'(x)$ is invertible for every $\mathbf{x} \in E$, then $\mathbf{f}(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$. In other words, $\mathbf{f}$ is an *open mapping* of E into $\mathbb{R}^n$

Theorem 9.27 If $A \in L(R^{m+n}, R^n)$ and if A_x is invertible, there there corresponds to every $\mathbf{k} \in \mathbb{R}^m$ a unique $\mathbf{h} \in R^n$ such that $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$. This $\mathbf{h}$ can be computed from $\mathbf{k}$ by the formula:

$$\mathbf{h} = -(A_x)^{-1}A_y\mathbf{k}$$

Proof. Because $A(\mathbf{h},\mathbf{k}) = A_x\mathbf{h} + A_y\mathbf{k}$, $A(\mathbf{h},\mathbf{k}) - 0$ if and only if $A_x\mathbf{h} + A_y\mathbf{k} = 0$, which is the same as $\mathbf{h} = -(A_x)^{-1}A_y\mathbf{k}$ when A_x is invertible. □

Theorem 9.28 (Implicit Function Theorem) Let $\mathbf{f}$ be a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^{n+m}$ into $\mathbb{R}^n$, such that $\mathbf{f}(\mathbf{a},\mathbf{b}) = \mathbf{0}$ for some point $(\mathbf{a},\mathbf{b}) \in E$.

Put $A = \mathbf{f}'(\mathbf{a},\mathbf{b})$ and assume that A_x is invertible. (That is, the Jacobian, the determinant of the $n \times n$ matrix A_x, is nonzero.) Then there exists open sets $U \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$, with $(\mathbf{a},\mathbf{b}) \in U$ and $\mathbf{b} \in W$m having the following property:

To every $\mathbf{y} \in W$ corresponds a unique $\mathbf{x}$ such that

$$(\mathbf{x},\mathbf{y}) \in U \quad \text{and} \quad \mathbf{f}(\mathbf{x},\mathbf{y}) = 0.$$

If this $\mathbf{x}$ is defined to be $\mathbf{g}(\mathbf{y})$, the $\mathbf{g}$ is a $\mathscr{C}'$-mapping of W into $\mathbb{R}^n$ $\mathbf{g}(\mathbf{b}) = \mathbf{a}$,

$$\mathbf{f}(\mathbf{g}(\mathbf{y}),\mathbf{y}) = \mathbf{0} \qquad (y \in W),$$

and

$$\mathbf{g}'(\mathbf{b}) = -(A_x)^{-1}A_y$$

Proof. Define $\mathbf{F}$ be $\mathbf{F}(\mathbf{x},\mathbf{y}) = (\mathbf{f}(\mathbf{x},\mathbf{y}),\mathbf{y})$. Then $\mathbf{F}$ is a $\mathscr{C}'$-mapping of E into $\mathbb{R}^{n+m}$. We claim that $\mathbf{F}'(\mathbf{a},\mathbf{b})$ is an invertible element of $L(\mathbb{R}^{n+m})$:

Since $\mathbf{f}(\mathbf{a},\mathbf{b}) = 0$, we have $\mathbf{f}(\mathbf{a}+\mathbf{h},\mathbf{b}+\mathbf{k}) = A(\mathbf{h},\mathbf{k}) + \mathbf{r}(\mathbf{h},\mathbf{k})$. where $\mathbf{r}$ is the remainder that occurs in the definition of $\mathbf{f}'(\mathbf{a},\mathbf{b})$. Since

$$\mathbf{F}(\mathbf{a}+\mathbf{h},\mathbf{b}+\mathbf{k}) - \mathbf{F}(\mathbf{a},\mathbf{b}) = (\mathbf{f}(\mathbf{a}+\mathbf{h},\mathbf{b}+\mathbf{k}),\mathbf{k}) = (A(\mathbf{h},\mathbf{k}),\mathbf{k}) + (\mathbf{r}(\mathbf{h},\mathbf{k}),0)$$

It follows that $\mathbf{F}'(\mathbf{a},\mathbf{b})$ is the linear operator on $\mathbb{R}^{n+m}$ that maps $(\mathbf{h},\mathbf{k})$ to $(A(\mathbf{h},\mathbf{k}),\mathbf{k})$. If this image vector is 0, then $A(\mathbf{h},\mathbf{k}) = 0$ and $\mathbf{k} = 0$, hence $\mathscr{A}(\mathbf{h},0)) = 0$, and Theorem 9.27 implies that $\mathbf{h} = 0$. It follows that $\mathbf{F}'(\mathbf{a},\mathbf{b})$ is 1-1; hence it is invertible.

The inverse function theorem can therefore be applied to $\mathbf{F}$. It shows that there exist open set U and V in $\mathbb{R}^{m+n}$, with $(\mathbf{a},\mathbf{b}) \in U$, $(0,\mathbf{b}) \in V$ such that $\mathbf{F}$ is a 1-1 mapping of U onto V.

We let W be the set of all $\mathbf{y} \in \mathbb{R}^m$ such that $(0,\mathbf{y}) \in V$. Note that $\mathbf{b} \in W$. It is clear that W is open since V is open.

If $\mathbf{y} \in W$, then $(0,\mathbf{y}) = \mathbf{F}(\mathbf{x},\mathbf{y})$ for some $(\mathbf{x},\mathbf{y}) \in U$. By our definition of $\mathbf{F}(\mathbf{x},\mathbf{y})$, $\mathbf{f}(\mathbf{x},\mathbf{y}) = 0$ for this $\mathbf{x}$.

Suppose, with the same $\mathbf{y}$, that $(\mathbf{x}',\mathbf{y}) \in U$ and $\mathbf{f}(\mathbf{x}',\mathbf{y}) = 0$. Then $\mathbf{F}(\mathbf{x}',\mathbf{y}) = (\mathbf{f}(\mathbf{x}',\mathbf{y}),\mathbf{y}) = (\mathbf{f}(\mathbf{x},\mathbf{y}),\mathbf{y}) = \mathbf{F}(\mathbf{x},\mathbf{y})$. Since $\mathbf{F}$ is 1-1 in U, it follows that $\mathbf{x}' = \mathbf{x}$. Thus proving the first part of the theorem.

For the second part, define $\mathbf{g}(\mathbf{y})$, for $\mathbf{y} \in W$, so that $(\mathbf{g}(\mathbf{y}),\mathbf{y}) \in U$ and $\mathbf{f}(\mathbf{g}(\mathbf{y}),\mathbf{y}) = 0$. Then $\mathbf{F}(\mathbf{g}(\mathbf{y}),\mathbf{y}) = (0,\mathbf{y})$. If $\mathbf{G}$ is the mapping of V onto U that inverts $\mathbf{F}$, then $\mathbf{G} \in \mathscr{C}'$ by the inverse function theorem, and $(\mathbf{g}(\mathbf{y}),\mathbf{y}) = \mathbf{G}(0,\mathbf{y})$. Since $\mathbf{G} \in \mathscr{C}'$, this shows that $\mathbf{g} \in \mathscr{C}'$.

Finally, to compute $\mathbf{g}'(\mathbf{b})$, put $(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \Phi(\mathbf{y})$. Then $\Phi(\mathbf{y})\mathbf{k}) = (\mathbf{g}'(\mathbf{y})\mathbf{k}, \mathbf{k})$. Because $\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = 0$, $\mathbf{f}(\Phi(\mathbf{y})) = 0$ in W. The chain rule shows therefore that $\mathbf{f}'(\Phi(\mathbf{y}))\Phi'(y) = 0$. When $\mathbf{y} = \mathbf{b}$, then $\Phi(\mathbf{y}) = (\mathbf{a}, \mathbf{b})$, and $\mathbf{f}'(\Phi(\mathbf{y})) = A$. Thus $A\Phi'(\mathbf{b}) = 0$.

If now follows that $A_x\mathbf{g}'(\mathbf{b})\mathbf{k} + A_y\mathbf{k} = A(\mathbf{g}'(b)\mathbf{k}, \mathbf{k}) = A\Phi'(\mathbf{b})\mathbf{k}) = 0$ for every $\mathbf{k} \in \mathbb{R}^m$. Thus $A_x\mathbf{g}'(\mathbf{b}) + A_y = 0$. This is equivalent to $\mathbf{g}'(\mathbf{b}) = -(A_x)^{-1}A_y$. □

Theorem 9.32 Suppose m, n, r are nonnegative integers, $m \geqslant r$, $n \geqslant r$, $\mathbf{F}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and $\mathbf{F}'(\mathbf{x})$ has rank r for every $\mathbf{x} \in E$. Fix $\mathbf{a} \in E$, put $A = \mathbf{F}'(\mathbf{a})$, let Y_1 be the range of A, and let P be a projection in $\mathbb{R}^m$ whose range is Y_1. Let Y_2 be the null space of P. Then there are open sets U and V in $\mathbb{R}^n$ with $\mathbf{a} \in U$, $U \subset E$, and there is a 1-1 $\mathscr{C}'$-mapping $\mathbf{H}$ of V onto U (whose inverse is also of class $\mathscr{C}'$) such that

$$\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x} + \varphi(A\mathbf{x}) \qquad (x \in V)$$

where φ is a $\mathscr{C}'$-mapping of the open set $A(V) \subset Y_1$ into Y_2.

Proof. If $r = 0$, Theorem 9.19 shows that $\mathbf{F}(\mathbf{x})$ is constant in a neighborhood U of $\mathbf{a}$, and now $\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x} + \varphi(A\mathbf{x})$ holds trivially, with $V = U$, $\mathbf{H}(\mathbf{x}) = \mathbf{x}$, $\varphi(0) = \mathbf{F}(\mathbf{a})$.

From now on we assume $r > 0$. Since $\dim Y_1 = r$, Y_1 has a basis $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$. Choose $\overline{e}_i \in \mathbb{R}^n$ so that $A\overline{e}_i = \mathbf{y}_i$ ($1 \leqslant i \leqslant r$), and define a linear mapping S of Y_1 into $\mathbb{R}^n$ by setting $S(c_1\mathbf{y}_1 + \cdots + c_r\mathbf{y}_r) = c_1\overline{e}_1 + \cdots + c_r\overline{e}_r$ for all scalars $c_1, \ldots, c_r$.

The $AS\mathbf{y}_i = A\overline{e}_i = \mathbf{y}_i$ for $1 \leqslant i \leqslant r$. Thus $AS\mathbf{y} = \mathbf{y}$. Define a mapping $\mathbf{G}$ of E into $\mathbb{R}^n$ by setting $\mathbf{G}(\mathbf{x}) = \mathbf{x} + SP[\mathbf{F}(\mathbf{x}) - A\mathbf{x}]$. Since $\mathbf{F}'(\mathbf{a}) = A$, differentiation of this shows that $\mathbf{G}'(\mathbf{a}) = I$, the identity operator on $\mathbb{R}^n$. By the inverse function theorem, there are open sets U and V in $\mathbb{R}^n$, with $\mathbf{a} \in U$, such that $\mathbf{G}$ is a 1-1 mapping of U and V, if necessary, we can arrange it so that V is convex and $\mathbf{H}'(\mathbf{x})$ is invertible for every $\mathbf{x} \in V$.

Note that $ASPA = A$, since $PA = A$. Therefore $A\mathbf{G}(\mathbf{x}) = P\mathbf{F}(\mathbf{x})$. In particular, this holds for $\mathbf{x} \in U$. If we replace $\mathbf{x}$ by $\mathbf{H}(\mathbf{x})$, we obtain $P\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x}$.

Define $\psi(\mathbf{x}) = \mathbf{F}(\mathbf{H}(\mathbf{x})) - A\mathbf{x}$. Since $PA = A$, this implies that ¶$\psi(\mathbf{x}) = 0$ for all $\mathbf{x} \in V$. Thus ψ is a $\mathscr{C}'$-mapping of V into Y_2.

Since V is open, it is clear that $A(V)$ is an open subset of its range $\mathscr{R}(A) = Y_1$.

To complete the proof, we have to show that there is a $\mathscr{C}'$-mapping φ of $A(V)$ into Y_2 which satisfies $\varphi(A\mathbf{x}) = \psi(\mathbf{x})$.

We will first prove that $\psi(\mathbf{x}_1) = \psi(\mathbf{x}_2)$ if $\mathbf{x}_1, \mathbf{x}_2 \in V$, $A\mathbf{x}_1 = A\mathbf{x}_2$.

Put $\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{H}(\mathbf{x}))$ for $\mathbf{x} \in V$. Since $\mathbf{H}'(\mathbf{x})$ has rank n for every $\mathbf{x} \in V$, and $\mathbf{F}'(\mathbf{x})$ has rank r for every $\mathbf{x} \in U$, it follows that $\operatorname{rank}\Phi'(x) = \operatorname{rank}\mathbf{F}(\mathbf{H}(\mathbf{x}))\mathbf{H}'(\mathbf{x}) = r$

Fix $\mathbf{x} \in V$. Let M be the range of $\Phi(\mathbf{x})$. Then $M \subset \mathbb{R}^m$, $\dim M = r$. Because $P\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x}$, $P\Phi'(\mathbf{x}) = A$. Thus P maps M onto $\mathscr{R}(A) = Y_1$. Since M and Y_1 have the same dimension, it follows that P (restricted to M) is 1-1.

Suppose now that $A\mathbf{h} = 0$. The $P\Phi(\mathbf{x})\mathbf{h} = 0$. But $\Phi'(\mathbf{x})\mathbf{h} \in M$, and P is 1-1 on M. Hence $\Phi(\mathbf{x})\mathbf{h} = 0$. Thus, we've proved that if $\mathbf{x} \in V$ and $A\mathbf{h} = 0$, then $\psi'(\mathbf{x})\mathbf{h} = 0$.

We can now prove $\psi(\mathbf{x}_1) = \psi(\mathbf{x}_2)$. Suppose $\mathbf{x}_1 \in V$, $\mathbf{x}_2 \in V$, $A\mathbf{x}_1 = A\mathbf{x}_2$. Put $\mathbf{h} = \mathbf{x}_2 - \mathbf{x}_1$ and define $\mathbf{g}(t) = \psi(\mathbf{x}_1 + t\mathbf{h})$. The convexity of V shows that $\mathbf{x}_1 + t\mathbf{h} \in V$ for these $t \in [0,1]$. Hence, $\mathbf{g}'(t) = \psi'(\mathbf{x}_1 + t\mathbf{h})\mathbf{h} = 0$, so that $g(1) = g(0)$. But $\mathbf{g}(1) = \psi(\mathbf{x}_2)$ and $\mathbf{g}(0) = \psi(\mathbf{x}_1)$. This proves the desired equivalence.

By this equivalence, $\psi(\mathbf{x})$ depends only on $A\mathbf{x}$ for $\mathbf{x} \in V$. Hence our definition of φ unambiguously in $A(V)$. It only remains to be proved that $\varphi \in \mathscr{C}'$.

Fix $\mathbf{y}_0 \in A(V)$, fix $\mathbf{x}_0 \in V$ so that $A\mathbf{x}_0 = \mathbf{y}_0$. Since V is open, $\mathbf{y}_0$ has a neighborhood W in Y_1 such that the vector $\mathbf{x} = \mathbf{x}_0 + S(\mathbf{y} - \mathbf{y}_0)$ lies in V for all $\mathbf{y} \in W$. Thus $A\mathbf{x} = A\mathbf{x}_0 + \mathbf{y} - \mathbf{y}_0 = \mathbf{y}$. Thus $\varphi(\mathbf{y}) = \psi(\mathbf{x}_0 - S\mathbf{y}_0 + S\mathbf{y})$. This formula shows that $\varphi \in \mathscr{C}'$ in W, hence in $A(V)$, since $\mathbf{y}_0$ was chosen arbitrarily in $A(V)$. This completes the proof. □

Theorem 9.34 a) If I is the identity operator on $\mathbb{R}^n$, then

$$\det[I] = \det(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$$

b) det is a linear function of each of the column vectors $\mathbf{x}_j$, if the others are held fixed.

c) If $[A]_1$ is obtained from $[A]$ by interchanging two columns, then $\det[A_1] = -\det[A]$.

d) If $[A]$ has two equal columns, then $\det[A] = 0$.

Proof. If $A = I$, then $a(i,i) = 1$ and $a(i,j) = 0$ for $i \neq j$. Hence $\det[I] = s(1, 2, \ldots, n) = 1$, which proves (a). By the definition of sgn, $s(j_1, \ldots, j_n) = 0$ if any two of the j's are equal. Each of the remaining $n!$ products in the computation of the determinant contains exactly one factor from each column. This proves (b). Part (c) is an immediate consequence of the fact that $s(j_1, \ldots, j_n)$ changes sign if any two of the j's are interchanged, and (d) is a corollary of (c). □

Theorem 9.35 If $[A]$ and $[B]$ are $n \times n$ matrices then:

$$\det([B][A]) = \det([B]) \det[A].$$

Proof. If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are the columns of $[A]$ define $\Delta_B(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \Delta_B[A] = \det([B][A])$. The columns of $[B][A]$ are the vectors $B\mathbf{x}_1, \ldots, B\mathbf{x}_n$. Thus $\Delta_B(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det(B\mathbf{x}_1, \ldots, B\mathbf{x}_n)$. Using this and Theorem 9.34, Δ_B also has properties 9.34(b) to (d). By (b) and the definition of $\mathbf{x}_j$,

$$\Delta_B[A] = \Delta_B\left(\sum_i a(i,1)\mathbf{e}_i, x_2, \ldots, \mathbf{x}_n\right) = \sum_i a(i,1)\Delta_B(\mathbf{e}_i, \mathbf{x}_2, \ldots, \mathbf{x}_n).$$

Repeating this process with $\mathbf{x}_2, \ldots, \mathbf{x}_n$, we obtain $\Delta_B[A] = \sum a(i_1, 1)a(i_2, 2) \ldots a(i_n, n)\Delta_B(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ the sum being extended over all ordered n-tuples $(i_1, \ldots, i_n)$ with $1 \leqslant i_r \leqslant n$. By (c) and (d), $\Delta_B(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = t(i_1, \ldots, i_n)\Delta_B(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, where $t = 1-, 0, 1$, and since $[B][I] = [B]$, we see that$\Delta_B(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det[B]$. Substituting things, we obtain $\det([B][A]) = \{\sum a(i_1, 1) \ldots a(i_n, n)t(i_1, \ldots, i_n)\} \det[B]$, for all n by N matrices $[A]$ and $[B]$. Taking $B = I$, we see that the above sum in braces is $\det[A]$. This proves the theorem. □

Theorem 9.36 A linear operator A in $\mathbb{R}^n$ is invertible if and only if $\det[A] \neq 0$.

Proof. If A is invertible, Theorem 9.35 shows that $\det[A]\det[A^{-1}] = \det[AA^{-1}] = \det[I] = 1$ so that $\det[A] \neq 0$.

if A is not invertible, the columns $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of $[A]$ are dependent; hence there is one, say, $\mathbf{x}_k$, such that $\mathbf{x}_k + \sum_{j \neq k} c_j \mathbf{x}_j = 0$ for certain scalars c_j. By 9.34 (b) and (d), $\mathbf{x}_k$ can be replaced by $\mathbf{x}_k + c_j \mathbf{x}_j$ without altering the determinant, if $j \neq k$. Repeating, we see that $\mathbf{x}_k$ can be replaced by the left side of the previous sum, i.e., by 0, without altering the determinant. But a matrix which has 0 for one column has determinant 0. Hence $\det[A] = 0$. □

Theorem 9.40 Suppose f is defined in an open set $E \subset \mathbb{R}^2$, and $D_1 f$ and $D_{21} f$ exist at every point of E. Suppose $Q \subset E$ is a closed rectangle with sides parallel to the coordinate axes, having (a, b) and $(a + h, b + k)$ as opposite vertices. Put

$$\Delta(f, Q) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)$$

Then there is a point (x, y) in the interior of Q such that $\Delta(f, Q) = hk(D_21 f)(x, y)$.

Proof. Put $u(t) = f(t, b_k) - f(t, b)$. Two applications of Theorem 5.10 show that there is an x between a and $a + h$, and that there is a y between b and $b + k$, such that

$$\Delta(f, Q) = u(a+h) - u(a) = hu'(x) = h[D_1 f)(x, b+k) - (D_1 f)(x, b)] = hk(D_{21} f)(x, y).$$

□

Theorem 9.41 Suppose f is defined in an open set $E \subset \mathbb{R}^2$, and $D_1 f$. $D_{21} f$. and $D_2 f$ exist at every point of E, and $D_{21} f$ is continuous at some point $(a, b) \in E$. Then $D_{12} f$ exists at (a, b) and $(D_{12} f)(a, b) = (D_{21} f)(a, b)$

Corollary $D_{21} f = D_{12} f$ if $f \in \mathscr{C}''(E)$

Proof. put $A = (D_{21} f)(a, b)$. Choose $\varepsilon > 0$. If Q is a rectangle as in Theorem 9.40, and if h and k are sufficiently small, we have $|A - (D_{21} f)(x, y)| < \varepsilon$ for all $(x, y) \in Q$. Thus $\left|\frac{\Delta(f,Q)}{hk} - A\right| < \varepsilon$. Fix h and let $k \to 0$. Since $D_2 f$ exists in E, the last inequality implies that $\left|\frac{(D_2 f)(a+h,b) - (D_2 f)(a,b)}{h} - A\right| \leqslant \varepsilon$. Since ε was arbitrary, and since this holds for all sufficiently small $h \neq 0$, it follows that $D_{12} f)(a, b) = A$. This gives the desired equality. □

Theorem 9.42 Suppose

a) $\varphi(x,t)$ is defined for $a \leqslant x \leqslant b$, $c \leqslant t \leqslant d$;
b) α is an increasing function on $[a,b]$;
c) $\phi^t \in \mathscr{R}(\alpha)$ for every $t \in [c,d]$;
d) $c < s < d$, and to every $\varepsilon > 0$ corresponds a $\delta > 0$ such that $|(D_2\varphi)(x,t) - (D_2\varphi)(x,s)| < \varepsilon$ for all $x \in [a,b]$ and for all $t \in (s-\delta, s_\delta)$.

Define

$$f(t) = \int_a^b \varphi(x,t)\,d\alpha \qquad (c \leqslant t \leqslant d).$$

Then $(D_2\varphi)^s \in \mathscr{R}(\alpha)$, $f'(s)$ exists, and

$$f'(s) = \int_a^b (D_2\varphi)(x,s)\,d\alpha$$

Proof. Consider the difference quotients $\psi(x,t) = \frac{\varphi(x,t)-\varphi(x,s)}{t-s}$ for $0 < |t-s| < \delta$. By Theorem 5.10 there corresponds to each (x,t) a number u between s and t such that $\psi(x,t) = (D_2\varphi)(x,u)$. Hence (d) implies that $|\psi(x,t) - (D_2\varphi)(x,s)| < \varepsilon$. Note that $\frac{f(t)-f(s)}{t-s} = \int_a^b \psi(x,t)\,d\alpha(x)$, By the last inequality, $\psi^t \to (D_2\varphi)^s$, uniformly on $[a,b]$, as $t \to s$. Since each $\psi^t \in \mathscr{R}(\alpha)$, the desired conclusion follows from the last equality and Theorem 7.16. □

Real Analysis Definitions and Theorems

A Glossary for Walter Rudin's *Principles of Mathematical Analysis*

Erin Griffin

May 15, 2017

Contents

Chapter 1 The Real and Complex Number Systems

Definitions

Empty Set/ Nonempty If A is any set (whose elements may be numbers or any other objects), we write $x \in A$ to indicate that x is a member (or an element) of A. If x is not a member of A, we write: $x \notin A$. The set which contains no elements will be called the *empty set*. If a set has at least one element, it is called *nonempty*

Proper If A and B are sets, and if every element of A is an element of B, we say that A is a subset of B, and we write $A \subset B$, or $B \supset A$. If, in addition, there is an element of B which is not in A, then A is said to be a *proper* subset of B. Note that $A \subset A$ for every set A. If $A \subset B$ and $B \subset A$, we write $A = B$.

Order Let S be a set. An *order* on S is a relation, denoted by $<$, with the following two properties:

a) If $x \in S$ and $y \in S$ then one and only one of the statements:

$$x < y \qquad x = y \qquad x > y$$

is true.

b) If $x, y, z \in S$, if $x < y$, $y < z$, then $x < z$.

Ordered Set An *ordered set* is a set S in which an order is defined.

Bounded Suppose S is an ordered set, and $E \subset S$. IF there exists a $\beta \in S$ such that $x \leqslant \beta$ for every $x \in E$, we say that E is *bounded above*, and call β an *upper bound* of E.

Lower bounds are defined in the same way.

Least Upper Bound/ Greatest Lower Bound Suppose S is an ordered set $E \subset S$, and E is bounded above. Suppose there exists an $\alpha \in S$ with the following properties:

a) α is an upper bound of E

b) If $\gamma < \alpha$, then γ is not an upper bound of E

Then α is called the *least upper bound of* E or the *supremum of* E, and we write

$$\alpha = \sup E.$$

The *greatest lower bound*, or *infimum*, of a set E which is bounded below is defined in the same manner: The statement

$$\alpha = \inf E$$

means that α is a lower bound of E and that no β with $\beta > \alpha$ is a lower bound of E.

Least-Upper-Bound Property An ordered set S is said to have the *least-upper-bound property* if the following is true:

If $E \subset S$, E is not empty, and E is bounded above, then $\sup E$ exists in S.

Field A *field* is a set F with two operations, called *addition* and *multiplication*, which satisfy the following so-called "field axioms":

1. Axioms for Addition
   (a) If $x \in F$ and $y \in F$, then $x + y \in F$.
   (b) Addition is commutative: $x + y = y + x$ for all $x, y \in F$.
   (c) Addition is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$
   (d) F contains an element 0 such that $0 + x = x$ for every $x \in F$.
   (e) To every $x \in F$ corresponds an element $-x \in F$ such that $x + (-x) = 0$
2. Axioms for Multiplication
   (a) If $x \in F$ and $y \in F$, then $xy \in F$.
   (b) Multiplication is commutative: $xy = yx$ for all $x, y \in F$.
   (c) Multiplication is associative: $(xy)z = x(yz)$ for all $x, y, z \in F$
   (d) F contains an element $1 \neq 0$ such that $1x = x$ for every $x \in F$.
   (e) If $x \in F$ and $x \neq 0$ then there exists and element $1/x \in F$ such that $x(1/x) = 1$.
3. The Distributive Law: $x(y + z) = xy + xz$ for all $x, y, z \in F$.

Ordered Field An *ordered field* is a field F which is also an ordered set, such that

a) $x + y < x + z$ if $x, y, z \in F$ and $y < z$
b) $xy > 0$ if $x, y \in F$ and $x, y > 0$.

If $x > 0$ we call x positive; if $x < 0$ we call x negative.

Extended Real Numbers The *extended real number system* consists of the real field $\mathbb{R}$ and two symbols $+\infty$, $-\infty$. We preserve the original order in $\mathbb{R}$ and define

$$-\infty < x < +\infty$$

for every $x \in \mathbb{R}$.

Complex Number A *complex number* is an ordered pair (a, b) of real numbers. "Ordered" means that (a, b) and (b, a) are regarded as distinct if $a \neq b$. Let $x = (a, b)$ and $y = (c, d)$ be two complex numbers. We write $x = y$ if and only if $a = c, b = d$. We define:

$$x + y = (a + c, b + d) \qquad xy = (ac - bd, ad + bc)$$

i $i = (0, 1) \in \mathbb{C}$

Conjugate If $a, b \in \mathbb{R}$ and $z = a + bi$, the the complex number $\overline{z} = a - bi$ is called the *conjugate* of z. The numbers a and b are the real part and imaginary part of z respectively. Note these as

$$a = \Re(z) \qquad b = \Im(z)$$

Absolute Value If $z \in \mathbb{C}$, its absolute value $|z|$ is the non-negative square root of $z\bar{z}$; that is $|z| = (z\bar{z})^{1/2}$.

Coordinates For each positive integer k, let $\mathbb{R}^k$ be the set of all ordered k-tuples

$$\mathbf{x} = (x_1, x_2, \ldots, x_k)$$

where $x_1, x_2, \ldots, x_k \in \mathbb{R}$, called the *coordinates* of $\mathbf{x}$. The elements of $\mathbb{R}^k$ are called points, or vectors, especially when $k > 1$. We shall denote vectors by boldfaced letters. If $\mathbf{y} = (y_1, y_2, \ldots, y_k)$, and if $\alpha \in R$, then addition and multiplication are defined:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k) \in \mathbb{R}^k \qquad \alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_k) \in \mathbb{R}^k$$

These operations make $\mathbb{R}^k$ into a *vector space over the real field.* The inner product is defined by:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{k} x_i y_i$$

and the *norm* of $\mathbf{x}$ by:

$$|\mathbf{x}| = (\mathbf{x} \cdot \mathbf{x})^{1/2} = \left(\sum_{i=1}^{k} x_i^2 \right)^{1/2}$$

Theorems

Theorem 1.11 Suppose S is an ordered set with the least-upper-bound property, $B \subset S$, B is not empty, and B is bounded below. Let L be the set of all lower bound of B. Then

$$\alpha = \sup L$$

exists in S and $\alpha = \inf B$. In particular, $\inf B$ exists in S.

Proposition 1.14 The axioms for addition imply the following statements.

a) If $x + y = x + z$ then $y = z$

b) If $x + y = x$ then $y = 0$

c) If $x + y = 0$, then $y = -x$

d) $-(-x) = x$

Proposition 1.15 The axioms for multiplication imply the following statements.

a) If $x \neq 0$ $xy = xz$ then $y = z$

b) If $x \neq 0$ $xy = x$ then $y = 1$

c) If $x \neq 0$ $xy = 1$, then $y = 1/x$

d) If $x \neq 0$ $1/(1/x) = x$

Proposition 1.16 The field axioms imply the following statements, for any $x, y, z \in F$

a) $0x = 0$

b) If $x \neq 0$ and $y \neq 0$ then $xy \neq 0$

c) $(-x)y = -(xy) = x(-y)$

d) $(-x)(-y) = xy$

Proposition 1.18 The following statements are true in every ordered field.

a) If $x > 0$ then $-x < 0$ and vice versa

b) If $x > 0$ and $y < z$ then $xy < xz$

c) If $x < 0$ and $y < z$ then $xy > xz$

d) If $x \neq 0$ then $x^2 > 0$. In particular, $1 > 0$

e) If $0 < x < y$ then $0 < 1/y < 1/x$.

Theorem 1.19 There exists an ordered field $\mathbb{R}$ which has the least-upper-bound property. Moreover $\mathbb{R}$ contains $\mathbb{Q}$ as a subfield.

Theorem 1.20 a (Archimedean Property) If $x \in \mathbb{R}$, $y \in \mathbb{R}$, and $x > 0$, then there is a positive integer n such that $nx > y$.

b ($\mathbb{Q}$ is dense in $\mathbb{R}$) If $x, y \in \mathbb{R}$ and $x < y$, then there exists a $p \in \mathbb{Q}$ such that $x < p < y$. In other words, between and two real numbers there is a rational one.

Theorem 1.21 For every real $x > 0$ and every integer $n > 0$ there is one and only one positive real y such that $y^n = x$. This number is written $\sqrt[n]{x}$.

Corollary If a and b are positive real numbers and n is a positive integer, then $(ab)^{1/n} = a^{1/n}b^{1/n}$.

Theorem 1.25 These definitions of addition and multiplication turn the set of all complex numbers into a field with $(0, 0)$ and $(1, 0$ in the role of 0 and 1.

Theorem 1.26 For any real numbers $a, b \in \mathbb{R}$ we have

$$(a, 0) + (b, 0) = (a + b, 0) \qquad (a, 0)(b, 0) = (ab, 0)$$

Theorem 1.28 $i^2 = -1$

Theorem 1.29 If a and b are real, then $(a, b) = a + bi$

Theorem 1.31 If z and w are complex, then

a $\overline{z + w} = \overline{z} + \overline{w}$

b $\overline{zw} = \overline{z} \cdot \overline{w}$

c $z + \overline{z} = 2\Re(z)$, $z - \overline{z} = 2i\Im(z)$

d $z\bar{z}$ is real and positive (except when $z = 0$.)

Theorem 1.33 Let z and w are complex. then

a $|z| > 0$

b $|\bar{z}| = |z|$

c $|zw| = |z||w|$

d $|\Re(z)| \leqslant |z|$

e $|z + w| \leqslant |z| + |w|$.

Theorem 1.35 (Schwarz Inequality) If $a_1, \ldots, a_n$ and $b_1, \ldots b_n$ are complex numbers, then

$$\left|\sum_{j=1}^{n} a_j \bar{b}_j\right|^2 \leqslant \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$

Theorem 1.37 Suppose $\mathbf{x}, \mathbf{y}, \mathbf{z} \in R^k$ and $\alpha \in \mathbb{R}$. Then

a $|\mathbf{x}| \geqslant 0$;

b $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = \mathbf{0}$

c $|\alpha\mathbf{x}| = |\alpha||\mathbf{x}|$

d $|\mathbf{x} \cdot \mathbf{y}| \leqslant |\mathbf{x}||\mathbf{y}|$

e $|\mathbf{x} + \mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|$

f $|\mathbf{x} - \mathbf{z}| \leqslant |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}|$

Chapter 2 Basic Topology

Definitions

Function Consider two sets A and B, whose elements may be any objects whatsoever, and suppose that with each element x of A there is associated, in some manner, an element of B, which we denote by $f(x)$. Then f is said to be a *function* from A to B (or a *mapping* of A into B). The set A is called the *domain* of f (we also say f is defined on A), and the elements $f(x)$ are called the *values* of f. The set of all values of f is called the *range* of f.

One-to-One, Onto Let A and B be two sets and let f be a mapping of A into B. If $E \subset A$, $f(E)$ is defined to be the set of all elements $f(x)$ for $x \in E$. We call $f(E)$ the *image* of E under f. In this notation, $f(A)$ is the range of f. It is clear that $f(A) \subset B$. If $f(A) = B$, we say that f maps *onto* B. If $E \subset B$, $f^{-1}(E)$ denotes the set of all $x \in A$ such that $f(x) \in E$. We call $f^{-1}(E)$ the *inverse image* of E under f. If $y \in B$, $f^{-1}(y)$ is the set of all $x \in A$ such that $f(x) = y$. If, for each $y \in B$, $f^{-1}(y)$ consists of at most one element of A, then f is said to be a one-to-one mapping of A into B. This may also be expressed as follows: f is a 1-1 mapping of A into B provided that $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$, $x_1, x_2 \in A$.

Correspondence/ Equivalent If there exists a 1-1 mapping of A *onto* B, we say that A and B can be put in 1-1 *correspondence*, or that A and B have the same *cardinal number*, or, briefly, that A and B are *equivalent*, and we write $A \sim B$. This relation has the following properties:

a) It is reflexive: $A \sim A$

b) It is symmetric: If $A \sim B$, then $B \sim A$.

c) It is transitive: If $A \sim B$ and $B \sim C$, then $A \sim C$.

Any relation with these three properties is called an *equivalence relation*.

Finite For any positive integer n, let J_n be the set whose elements are the integers $1, 2, \ldots, n$; let J be the set consisting of all positive integers. For any set A,

a) we say A is *finite* if $A \sim J_n$ for some n (the empty set is also considered to be finite).

b) A is infinite if A is not finite

c) A is *countable* if $A \sim J$

d) A is *uncountable* if A is neither finite nor countable

e) A is *at most countable* if A is finite or countable

Sequence A *sequence* is a function f defined on the set J of all positive integers. If $f(n) = x_n$, for $n \in J$, it is customary to denote the sequence f by the symbol $\{x_n\}$. The values of f are called the *terms* of the sequence. If A is a set and if $x_n \in A$ for all $n \in J$, then $\{x_n\}$ is said to be a *sequence in* A, or a *sequence of elements of* A.

Subsets/ Family of Sets Let A and Ω be sets, and suppose that with each element α of A there is associated a subset of Ω which we denote by E_α. The set whose elements are set E_α will be denoted by $\{E_\alpha\}$. We shall call these a collection of sets or *family of sets.*

Union (From above) The *union* of the sets E_α is defined to be the set S such that $x \in S$ if and only if $x \in E_\alpha$ for at least one $\alpha \in A$. We use notation

$$S = \bigcup_{\alpha \in A} E_\alpha$$

If A consists of integers, we write one of the two following:

$$S = \bigcup_{i=1}^{n} E_i \qquad \text{or} \qquad S = E_1 \cup E_2 \cup \cdots \cup E_n$$

If A is the set of all positive integers, the usual notation is:

$$S = \bigcup_{i=1}^{\infty} E_i$$

Intersection The *intersection* of the sets E_α is defined to be the set P such that $x \in P$ if and only if $x \in E_\alpha$ for every $\alpha \in A$. We use notation:

$$P = \bigcap_{\alpha \in A} E_\alpha \quad \text{or} \quad P = \bigcap_{i=1}^{n} E_i = E_1 \cap E_2 \cap \ldots E_n \quad \text{or} \quad P = \bigcap_{i=1}^{\infty} E_i$$

If $A \cap B \neq \varnothing$ the we say that A and B *intersect*; otherwise they are *disjoint.*

Metric Space A set X, whose elements we shall call *points*, is said to be a *metric space* if with any two points p and q of X there is associated a real number $d(p, q)$ called the *distance* from p to q, such that:

a) $d(p, q) > 0$ if $p \neq q$; $d(p, p) = 0$.

b) $d(p, q) = d(q, p)$

c) $d(p, q) \leqslant d(p, r) + d(r, q)$ for any $r \in X$.

Segment By the *segment* (a, b) we mean the set of all real numbers x such that $a < x < b$.

Interval By the *interval* $[a, b]$ we mean the set of all real numbers x such that $a \leqslant x \leqslant b$.

K-Cell If $a_i < b_i$ for $i = 1, \ldots, k$, the set of all points $\mathbf{x} = (x_1, \ldots, x_k)$ in $\mathbb{R}^k$ whose coordinates satisfy the inequalities $a_i \leqslant x_i \leqslant b_i$, $(1 \leqslant i \leqslant k)$ is called a *k-cell.*

Ball If $\mathbf{x} \in \mathbb{R}^k$ and $r > 0$, the open (or closed) *ball* B with center at $\mathbf{x}$ and radius r is defined to be the set of all $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y} - \mathbf{x}| < r$ (or $|\mathbf{y} - \mathbf{x}| \leqslant r$).

Convex We call a set $E \subset \mathbb{R}^k$ *convex* if $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in E$ whenever $\mathbf{x} \in E$, $y \in E$, and $0 < \lambda < 1$.

Neighborhood Let X be a metric space. A *neighborhood* of p is a set $N_r(p)$ consisting of all q such that $d(p,q) < r$, for some $r > 0$. The number r is called the *radius* of $N_r(p)$.

Limit Point A point p is a *limit point* of the set E if *every* neighborhood of p contains a point $q \neq p$ such that $q \in E$.

Isolated Point If $p \in E$ and p is not a limit point of E, then p is called an *isolated point* of E.

Closed E is *closed* if every limit point of E is a point of E.

Interior A point p is an *interior* point of E if there is a neighborhood N of p such that $N \subset E$.

Open E is *open* if every point of E is an interior point of E.

Complement The *complement* of E (denoted E^c) is the set of all points $p \in X$ such that $p \notin E$

Perfect E is *perfect* if E is closed and if every point of E is a limit point of E.

Bounded E is *bounded* if there is a real number M and a point $q \in X$ such that $d(p,q) < M$ for all $p \in E$.

Dense E is *dense* in X if every point of X is a limit point of E, or a point of E (or both).

Closure If X is a metric space, if $E \subset X$, and if E' denotes the set of all limit points of E in X, then the *closure* of E is the set $\overline{E} = E \cup E'$.

Open Cover By an *open cover* of a set E in a metric space X we mean a collection $\{G_\alpha\}$ of open subsets of X such that $E \subset \bigcup_\alpha G_\alpha$.

Compact A subset K of a metric space X is said to be *compact* if every open cover of K contains a *finite* subcover. More explicitly, the requirement is that if $\{G_\alpha\}$is an open cover of K, then there are finitely many indices $\alpha_1, \dots, \alpha_n$ such that $K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$.

Separated Two subsets A and B of a metric space X are said to be *separated* if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty, i.e. if no point of A lies in the closure of B and no point of B lies in the closure of A

Connected A set $E \subset X$ is said to be *connected* if E is not a union of two nonempty separated sets.

Theorems

Theorem 2.8 Every infinite subset of a countable set A is countable.

Theorem 2.12 Let $\{E_n\}$, $n = 1, 2, 3, \ldots$, be a sequence of countable sets, and put

$$S = \bigcup_{n=1}^{\infty} E_n$$

Then S is countable.

Corollary Suppose A is at most countable, and, for every $\alpha \in A$, B_α is at most countable. Put

$$T = \bigcup_{\alpha \in A} B_\alpha$$

Then T is at most countable.

Theorem 2.13 Let A be a countable set, and let B_n be the set of all n-tuples $(a_1, \ldots, a_n)$, where $a_k \in A$, $(k = 1, 2, \ldots, n)$, and the elements $a_1, \ldots, a_n$ need not be distinct. Then B_n is countable.

Corollary The set of all rational numbers is countable.

Theorem 2.14 Let A be the set of all sequences whose elements are the digits 0 and 1. This set A is uncountable.

Theorem 2.19 Every neighborhood is an open set.

Theorem 2.20 If P is a limit point of a set E, then every neighborhood of p contains infinitely many points of E.

Corollary A finite point set has no limit points.

Theorem 2.22 Let $\{E_\alpha\}$ be a (finite or infinite) collection of sets E_α. Then

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c)$$

Theorem 2.23 A set E is open if and only if its complement is closed.

Corollary A set F is closed if and only if its complement is open.

Theorem 2.24 a) For any collection $\{G_\alpha\}$ of open sets, $\bigcup_\alpha G_\alpha$ is open.

b) For any collection $\{F_\alpha\}$ of closed sets, $\bigcap_\alpha F_\alpha$ is closed.

c) For any finite collection $G_1, \ldots, G_n$ of open sets, $\bigcap_{i=1}^n G_i$ is open.

d) For any finite collection $F_1, \ldots, F_n$ of closed sets, $\bigcup_{i=1}^n F_i$ is closed.

Theorem 2.27 If X is a metric space and $E \subset X$, then

a) $\overline{E}$ is closed,

b) $E = \overline{E}$ if an only if E is closed

c) $\overline{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$.

Theorem 2.28 Let E be a nonempty set of real numbers which is bounded above. Let $y = \sup E$. Then $y \in \overline{E}$. Hence $y \in E$ if E is closed.

Theorem 2.30 Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.

Theorem 2.33 Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.

Theorem 2.34 Compact subsets of metric spaces are closed.

Theorem 2.35 Closed subsets of compact sets are compact.

Corollary If F is closed and K is compact, then $F \cap K$ is compact.

Theorem 2.36 If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X such that the intersection of every finite sub-collection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is nonempty.

Corollary If $\{K_n\}$ is a sequence of nonempty compact sets such that $K_n \supset K_{n+1}$ ($n = 1, 2, 3, \dots$), then $\bigcap_1^\infty K_n$ is not empty.

Theorem 2.37 If E is an infinite subset of a compact set K, then E has a limit point in K.

Theorem 2.38 If $\{I_n\}$ is a sequence of intervals in $\mathbb{R}$ such that $I_n \supset I_{n+1}$ ($n = 1, 2, 3, \dots$), then $\bigcap_1^\infty I_n$ is not empty.

Theorem 2.39 Let k be a positive integer. If $\{I_n\}$ is a sequence of k-cells such that $I_n \subset I_{n+1}$ ($n = 1, 2, 3, \dots$), then $\bigcap_1^\infty I_n$ is not empty.

Theorem 2.40 Every k-cell is compact.

Theorem 2.41 If a set E in $\mathbb{R}^k$ has one of the following three properties, then it has the other two.

a) E is closed and bounded

b) E is compact

c) Every infinite subset of E has a limit point in E.

Theorem 2.42 (Weierstrass) Every bounded infinite subset of $|R^k$ has a limit point in $\mathbb{R}^k$.

Theorem 2.43 Let P be a nonempty perfect set in $\mathbb{R}^k$. Then P is uncountable.

Corollary Every interval $[a, b]$ is uncountable. In particular, the set of all real numbers is uncountable.

Theorem 2.47 A subset E of the real line $\mathbb{R}$ is connected if and only if it has the following property: if $x \in E$ and $y \in E$, and $x < z < y$, then $z \in E$.

Chapter 3 Numerical Sequences

Definitions

Converge/ Diverge A sequence $\{p_n\}$ in a metric space X is said to *converge* if there is a point $p \in X$ with the following property: for every $\varepsilon > 0$ there is an integer N such that $n \geqslant N$ implies that $d(p_n, p) < \varepsilon$. In this case we also say that $\{p_n\}$ converges to p, or that p is the limit of $\{p_n\}$, and we write $p_n \to p$, or

$$\lim_{n\to\infty} p_n = p$$

If $\{p_n\}$ does no converge, it is said to *diverge*.

Range/ Bounded The set of all points p_n is the *range* of $\{p_n\}$. The range of a sequence may be a finite set, or it may be infinite. The sequence $\{p_n\}$ is said to be *bounded* if it's range is bounded.

Subsequence Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3 < \ldots$. Then the sequence $\{p_{n_i}\}$ is called a *subsequence* of $\{p_n\}$. If $\{p_{n_i}\}$ converges, its limit is called a *subsequential limit* of $\{p_n\}$.

It is clear that $\{p_n\}$ converges to p is and only if every subsequence of $\{p_n\}$ converges to P.

Cauchy Sequence A sequence $\{p_n\}$ in a metric space X is said to be a *Cauchy sequence* if for every $\varepsilon > 0$ there is an integer N such that $d(p_n, p_m) < \varepsilon$ if $n \geqslant N$ and $m \geqslant N$.

Diameter Let E be a nonempty subset of a metric space X, and let S be the set of all real numbers of the form $d(p, q)$, with $p \in E$ and $q \in E$. The sup of S is called the *diameter of E*.

Complete A metric space in which every Cauchy sequence converges is said to be *complete.*

Monotonic A sequence $\{s_n\}$ of real numbers is said to be

a) *monotonically increasing* if $s_n \leqslant s_{n+1}$ $(n = 1, 2, 3, \ldots)$

b) *monotonically decreasing* if $s_n \geqslant s_{n+1}$ $(n = 1, 2, 3, \ldots)$

The class of monotonic sequences consists of the increasing and the decreasing sequences.

Convergence to Infinity Let $\{s_n\}$ be a sequence of real numbers with the following property: For every real M there is an integer N such that $n \geqslant N$ implies $s_n \geqslant M$. We then write: $s_n \to +\infty$.

Similarly, if for every real M there is an integer N such that $n \geqslant N$ implies $s_n \leqslant M$. We then write: $s_n \to -\infty$.

Upper/ Lower Limits Let $\{s_n\}$ be a sequence of real numbers. Let E be the set of numbers X such that $s_{n_k} \to x$ for some subsequence $\{s_{n_k}\}$. This set E contains all subsequential limits as defined above. Let:

$$s^* = \sup E \qquad s_* = \inf E$$

The numbers s^* and s_* are called the *upper* and *lower limits* of $\{s_n\}$; we use the notation:

$$\limsup_{n\to\infty} s_n = s^* \qquad \liminf_{n\to\infty} s_n = s_*$$

Series Given a sequence $\{a_n\}$, we use the notation $\sum_{n=p}^{q} a_n$ to denote the sum $a_p + a_{p+1} + \cdots + a_q$. With $\{a_n\}$ we associate a sequence $\{s_n\}$ where $s_n = \sum_{k=p}^{n} a_k$. For $\{s_n\}$ we also use the symbolic expression $a_1 + a_2 + a_3 + \ldots$ or, more concisely,

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

We call this an *infinite series*, or just a *series*. The numbers s_n are called the *partial sums* of the series. If $\{s_n\}$ converges to s, we say that the series *converges*, and write

$$\sum_{n=1}^{\infty} a_n = s.$$

The number s is called the sum of the series; but it should be clearly understood that s is the *limit of a sequence of sums*, and is not obtained by simple addition.

If $\{s_n\}$ diverges, the series is said to *diverge*.

e $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Power Series Given a sequence $\{c_n\}$ of complex numbers, the series

$$\sum_{n=0}^{\infty} c_n z^n$$

is called a *power series*. The numbers c_n are called the *coefficients* of the series; z is a complex number.

Absolute Convergence The series $\sum a_n$ is said to *converge absolutely* if the series $\sum |a_n|$ converges.

Product Given $\sum a_n$ and $\sum b_n$, we put

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

and call $\sum c_n$ the *product* of the two given series. Equivalently,

$$\text{product} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k}$$

Rearrangement Let $\{k_n\}$ $n = 1, 2, 3, \dots$ be a sequence in which every positive integer appears once and only once (that is, $\{k_n\}$ is a one-to-one function from J onto J, Putting $a'_n = a_{k_n}$ $(n = 1, 2, 3, \dots)$, we say that $\sum a'_n$ is a *rearrangement* of $\sum a_n$.

Theorems

Theorem 3.2 Let $\{p_n\}$ be a sequence in a metric space X.

a) $\{p_n\}$ converges to $p \in X$ if and only if every neighborhood of p contains p_n for all but finitely many n.

b) If $p \in X$, $p' \in X$, and if $\{p_n\}$ converges to p and to p', then $p' = p$.

c) If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.

d) If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$in E such that $p = \lim_{n\to\infty} p_n$

Theorem 3.3 Suppose $\{s_n\}$, $\{t_n\}$ are complex sequences, and $\lim_{n\to\infty} s_n = s$, $\lim_{n\to\infty} t_n = t$. Then,

a) $\lim_{n\to\infty}(s_n + t_n) = s + t$;

b) $\lim_{n\to\infty} cs_n = cs$, $lim_{n\to\infty}(c + s_n) = c + s$;

c) $\lim_{n\to\infty} s_n t_n = st$

d) $\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s}$, provided $s_n \neq 0$ and $s \neq 0$.

Theorem 3.4 Suppose $\mathbf{x}_n \in \mathbb{R}^k$ and $\mathbf{x}_n = (\alpha_{1,n}, \alpha_{2,n}, \dots, \alpha_{k,n})$. Then $\{\mathbf{x}_n\}$ converges to $\mathbf{x} = (\alpha_1, \alpha_2, \dots, \alpha_k)$ if and only if

$$\lim_{n\to\infty} \alpha_{j,n} = \alpha_j \qquad (1 \leqslant j \leqslant k)$$

Suppose $\{\mathbf{x}_n\}$, $\{\mathbf{y}_n\}$ are sequences in $\mathbb{R}^k$, $\{\beta_n\}$ is a sequence of real numbers, and $\mathbf{x}_n \to \mathbf{x}$, $\mathbf{y}_n \to \mathbf{y}$, $\beta_n \to \beta$. Then

$$\lim_{n\to\infty} (\mathbf{x}_n + \mathbf{y}_n) = \mathbf{x} + \mathbf{y} \qquad \lim_{n\to\infty} (\mathbf{x}_n \cdot \mathbf{y}_n) = \mathbf{x} \cdot \mathbf{y} \qquad \lim_{n\to\infty} (\beta_n \mathbf{x}_n) = \beta \mathbf{x}$$

Theorem 3.6 a) If $\{p_n\}$ is a sequence in a compact metric space X, then some subsequence of $\{p_n\}$ converges to a point of X.

b) Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.

Theorem 3.7 The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subset of X.

Theorem 3.10 a) If $\overline{E}$ is the closure of a set E in a metric space X, then

$$\operatorname{diam} \overline{E} = \operatorname{diam} E$$

b) If K_n is a sequence of compact sets in X such that $K_n \supset K_{n+1}$ and if

$$\lim_{n\to\infty} \operatorname{diam} K_n = 0,$$

then $\bigcap_1^\infty K_n$ consists of exactly one point.

Theorem 3.11 a) In any metric space X, every convergent sequence is a Cauchy sequence.

b) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some point of X.

c) In $\mathbb{R}^k$, every Cauchy sequence converges

Corollary All Compact metric spaces and all Euclidean spaces are complete.

Corollary Every Closed subset E of a complete metric space X is complete.

Theorem 3.14 Suppose $\{s_n\}$ is monotonic. Then $\{s_n\}$ converges if and only if it is bounded.

Theorem 3.17 Let $\{s_n\}$ be a sequence of real numbers. Let E and s^* be as defined above. Then s^* has the following two properties:

a) $s^* \in E$

b) If $x > s^*$, there is an integer N such that $n \geqslant N$ implies $s_n < x$.

Moreover, s^* is the only number with these properties. Furthermore, the analogous result is true for s_*.

Theorem 3.19 If $s_n \leqslant t_n$ for $n \geqslant N$, where N is fixed, then:

$$\liminf_{n\to\infty} s_n \leqslant \liminf_{n\to\infty} t_n \qquad \limsup_{n\to\infty} s_n \leqslant \limsup_{n\to\infty} t_n$$

Theorem 3.20 a) If $p > 0$, the $\lim_{n\to\infty} \frac{1}{n^p} = 0$

b) If $p > 0$, the $\lim_{n\to\infty} \sqrt[n]{p} = 1$

c) $\lim_{n\to\infty} \sqrt[n]{n} = 1$

d) If $p > 0$ and $\alpha \in \mathbb{R}$, then $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$

e) If $|x| < 1$, then $\lim_{n\to\infty} x^n = 0$.

Theorem 3.22 $\sum a_n$ converges if and only if for every $\varepsilon > 0$ there is an integer N such that

$$\left|\sum_{k=n}^{m} a_k\right| \leqslant \varepsilon$$

if $m \geqslant n \geqslant N$. In particular, by taking $m = n$, $|a_n| \leqslant \varepsilon$.

Theorem 3.23 If $\sum a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Theorem 3.24 A series of nonnegative terms converges if and only if its partial sums form a bounded sequence.

Theorem 3.25 (Comparison Test) a) If $|a_n| \leqslant c_n$ for $n \geqslant N_0$, where N_0 is some fixed integer, and if $\sum c_n$ converges, then $\sum a_n$ converges.

b) If $a_n \geqslant d_n \geqslant 0$ for $n \geqslant N_0$, and if $\sum d_n$ diverges, then $\sum a_n$ diverges.

Theorem 3.26 (Geometric Series) If $0 \leqslant x < 1$, then

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

If $x \geqslant 1$ this series diverges.

Theorem 3.27 (Cauchy Condensation Test) Suppose $a_1 \geqslant a_2 \geqslant a_3 \geqslant \cdots \geqslant 0$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the series

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \dots$$

converges.

Theorem 3.28 (p-Test) $\sum \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \leqslant 1$.

Theorem 3.29 If $p > 1$

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$$

converges. If $p \leqslant 1$, the series diverges.

Thoerem 3.31 $\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e$.

Theorem 3.32 e is irrational.

Theorem 3.33 (Root Test) Given $\sum a_n$, put $\alpha - \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Then:

a) if $\alpha < 1$, $\sum a_n$ converges

b) if $\alpha > 1$, $\sum a_n$ diverges

c) if $\alpha = 1$, the test gives no information

Theorem 3.34 (Ratio Test) The series $\sum a_n$

a) Converges if $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$

b) Diverges if $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \geqslant 1$ for all $n \geqslant n_0$, where n_0 is some fixed integer.

Theorem 3.37 For any sequence $\{c_n\}$ of positive numbers,

$$\liminf_{n\to\infty} \frac{c_{n+1}}{c_n} \leqslant \liminf_{n\to\infty} \sqrt[n]{c_n} \qquad \limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

Theorem 3.39 Given the power series $\sum c_n z^n$, put

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|}, \qquad R = \frac{1}{\alpha}$$

If $\alpha = 0$, then $R = +\infty$; if $\alpha = +\infty$, $R = 0$. (Note, R is called the *radius of convergence*). Then $\sum c_n z^n$ converges if $|z| < R$, and diverges if $|z| > R$.

Theorem 3.41 Given two sequence $\{a_n\}$, $\{b_n\}$, put $A_n = \sum_{k=0}^{n} a_k$ if $n \geqslant 0$; put $A_{-1} = 0$. Then, if $0 \leqslant p \leqslant q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

Theorem 3.42 Suppose

a) the partial sums A_n of $\sum a_n$ form a bounded sequence;

b) $b_0 \geqslant b_1 \geqslant b_2 \geqslant \ldots$

c) $\lim_{n\to\infty} b_n = 0$

Then $\sum a_n b_n$ converges.

Theorem 3.43 (Alternative Series Test) Suppose

a) $|c_1| \geqslant |c_2| \geqslant |c_3| \geqslant \ldots$;

b) $c_{2m-1} \geqslant 0$, $c_{2m} \leqslant 0$ $(m = 1, 2, 3, \ldots)$

c) $\lim_{n\to\infty} c_n = 0$

Then $\sum c_n$ converges

Theorem 3.44 Suppose the radius of convergence of $\sum c_n z^n$ is 1, and suppose $c_0 \geqslant c_1 \geqslant c_2 \geqslant \ldots$, $\lim_{n\to\infty} c_n = 0$. Then $\sum c_n z^n$ converges at every point on the circle $|z| = 1$, except possibly at $z = 1$.

Theorem 3.45 If $\sum a_n$ converges absolutely, then $\sum a_n$ converges.

Theorem 3.47 If $\sum a_n = A$ and $\sum b_n = B$, then $\sum(a_n + b_n) = A + B$ and $\sum c a_n = cA_n$ for any fixed c.

Theorem 3.50 Suppose

a) $\sum_{n=0}^{\infty} a_n$ converges absolutely

b) $\sum_{n=0}^{\infty} a_n = A$

c) $\sum_{n=0}^{\infty} b_n = B$

d) $c_n = \sum_{k=0}^{n} a_k b_{n-k} \qquad (n = 0, 1, 2, \ldots)$

Then $\sum_{n=0}^{\infty} c_n = AB$.

Theorem 3.51 If the series $\sum a_n = A$, $\sum b_n = B$, $\sum c_n = C$, and $c_n = a_0 b_n + \cdots + a_n b_0$, then $C = AB$.

Theorem 3.54 Let $\sum a_n$ be a series of real numbers which converges, but not absolutely. Suppose $-\infty \leqslant \alpha \leqslant \beta \leqslant \infty$. Then there exists a rearrangement $\sum a'_n$ with partial sums s'_n such that

$$\liminf_{n\to\infty} s'_n = \alpha \qquad \limsup_{n\to\infty} s'_n = \beta$$

Theorem 3.55 If $\sum a_n$ is a series of complex numbers which converges absolutely, then every rearrangement of $\sum a_n$ converges, and they all converge to the same sum.

Chapter 4 Continuity

Definitions

Limit of a Function Let X and Y be metric spaces; suppose $E \subset X$, f maps E into Y and p is a limit point of E. We write $f(x) \to q$ as $x \to p$ or

$$\lim_{x \to p} f(x) = q$$

if there is a point $q \in Y$ with the following property: Fore very $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), q) < \varepsilon$ for all points $x \in E$ for which $0 < d_X(x, p) < \delta$.

Sum/ Difference/ Product/ Quotient of Function Suppose we have two complex functions, f and g, both defined on E. By $f + g$ we mean the function which assigns to each point x of E the number $f(x) + g(x)$. Similarly we define $f - g$, fg, f/g for $g(x) \neq 0$. If f assigns to each point of x of E the same number c, f is said to be a constant function, or constant, and we write $f = c$. If f and g are real functions then $f(x) \geqslant g(x)$ is the same as $f \geqslant g$. The same holds for $\mathbf{f}, \mathbf{g} : E \to \mathbb{R}^K$.

Continuous Suppose X an Y are metric spaces, $E \subset X$, $p \in E$, and $f : E \to Y$. Then f is said to be *continuous at* p if for every $\varepsilon > 0$ there exists a δ such that

$$d_Y(f(x), f(p)) < \varepsilon$$

for all points $x \in E$ for which $d_X(x, p) < \delta$.

If f is continuous at every point of E, then f is said to be *continuous on* E.

Bounded A mapping $\mathbf{f}$ of a set E into $\mathbb{R}^k$ is said to be *bounded* if there is a real number M such that $|\mathbf{f}(x)| \leqslant M$ for all $x \in E$.

Uniformly Continuous Let f be a mapping of a metric space X into a metric space Y. We say that f is *uniformly continuous* on X if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(p), f(q)) < \varepsilon$$

for all p and q in X for which $d_X(p, q) < \delta$.

One-Sided Limit of a Function Let f be defined on (a, b). Consider any point x such that $a \leqslant x < b$. We write:

$$f(x+) = q$$

if $f(t_n) \to q$ as $n \to \infty$, for all sequences $\{t_n\}$ in (x, b) such that $t_n \to x$. To obtain the definition of $f(x-)$, for $a < x \leqslant b$, we restrict ourselves to sequences $\{t_n\}$ in (a, x).

It is clear that at any point $x \in (a, b)$ the $\lim_{t \to x} f(t)$ exists if and only if

$$f(x+) = f(x-) = \lim_{t \to x} f(t).$$

Simple Discontinuity Let f be defined on (a,b). If f is discontinuous at a point x and if $f(x+)$ and $f(x-)$ exists, then f is said to have a discontinuity of the *first kind*, or a *simple discontinuity*, at x. Otherwise the discontinuity is said to be of the *second kind*.

There are two ways in which a function can have a simple discontinuity:

$$f(x+) \neq f(x-) \qquad \text{or} \qquad f(x+) = f(x-) \neq f(x).$$

Monotonic Let f be real on (a,b). Then f is said to be *monotonically increasing* on (a,b) if $a < x < y < b$ implies $f(x) \leqslant f(y)$. If the last inequality is reversed, we obtain the definition of a *monotonically decreasing* function. The class of monotonic functions consists of both the increasing and the decreasing functions.

Neighborhood of Infinity For any real c, the set of real numbers x such that $x > c$ is called a neighborhood of $+\infty$ and is written $(c, +\infty)$. Similarly, the set $(-\infty, c)$ is a neighborhood of $-\infty$.

Limit at Infinity Let f be a real function defined on $E \subset \mathbb{R}$. We say that $f(t) \to A$ as $t \to x$, where A and x are in the extended real number system, if for every neighborhood U of A there is a neighborhood V of x such that $V \cap E$ is not empty, and such that $f(t) \in U$ for all $t \in V \cap E$, $t \neq x$.

Theorems

Theorem 4.2 Let X and Y be metric spaces; suppose $E \subset X$, f maps E into Y and p is a limit point of E. Then $\lim_{x\to p} f(x) = q$ if and only if $\lim_{n\to\infty} f(p_n) = q$ for every sequence $\{p_n\}$ in E such that $p_n \neq p$ $\lim_{n\to\infty} p_n = p$.

Corollary If f has a limit at p, this limit is unique.

Theorem 4.4 Suppose $E \subset X$, a metric space, p is a limit point of E, f and g are complex functions on E, and

$$\lim_{x\to p} f(x) = A \qquad \lim_{x\to p} g(x) = B$$

Then:

a) $\lim_{x\to p}(f+g)(x) = A + B$

b) $\lim_{x\to p}(fg) = AB$

c) $\lim_{x\to p}\left(\frac{f}{g}\right) = \frac{A}{B}$

Theorem 4.6 Suppose X an Y are metric spaces, $E \subset X$, $p \in E$ such that p is a limit point of E, and $f : E \to Y$. Then f is continuous at P if and only if $\lim_{x\to p} f(x) = f(p)$.

Theorem 4.7 Suppose X, Y, Z are metric spaces, $E \subset X$, f maps E into Y. g maps the range of f, $f(E)$, into Z, and H is the mapping of E into Z defined by

$$h(x) = g(f(x)) \qquad (x \in E)$$

If f is continuous at a point $p \in E$ and if g is continuous at the point $f(p)$, then h is continuous at P. The function h is called the *composition* or the *composite* of f and g, most commonly noted: $h = g \circ f$.

Theorem 4.8 A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y.

Corollary A mapping f of a metric space X into a metric space Y is continuous if and only if $f^{-1}(C)$ is closed in X for every closed set C in Y.

Theorem 4.9 Let f and g be complex continuous functions on a metric space X. Then $f + g$, fg, f/g are continuous on X.

Theorem 4.10 a) Let $f_1, \ldots, f_k$ be real functions on a metric space X, and let $\mathbf{f}$ be the mapping of X into $\mathbb{R}^k$ defined by $\mathbf{f}(x) = (f_1(x), \ldots, f_k(x))$, then $\mathbf{f}$ is continuous if and only if each of the functions f_i is continuous.

b) If $\mathbf{f}$ and $\mathbf{g}$ are continuous mappings of X into $\mathbb{R}^k$, then $\mathbf{f}+\mathbf{g}$ and $\mathbf{f}\cdot\mathbf{g}$ are continuous on X.

The functions f_i are called the *components* of $\mathbf{f}$.

Theorem 4.14 Suppose f is a continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact.

Theorem 4.15 If $\mathbf{f}$ is a continuous mapping of a compact metric space X into $\mathbb{R}^k$, then $\mathbf{f}(X)$ is closed and bounded. Thus, $\mathbf{f}$ is bounded.

Theorem 4.16 Suppose f is a continuous real function on a compact metric space X, and

$$M = \sup_{p \in X} f(p) \qquad m = \inf_{p \in X} f(p)$$

Then there exists points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$. That is, f attains it's maximum and minimum.

Theorem 4.17 Suppose f is a continuous 1-1 mapping of a compact metric space X onto a metric space Y. Then the inverse mapping f^{-1} defined on Y by

$$f^{-1}(f(x)) = x \qquad (x \in X)$$

is a continuous mapping of Y onto X.

Theorem 4.19 Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is uniformly continuous on X.

Theorem 4.20 Let E be a non-compact set in $\mathbb{R}$. Then

a) there exists a continuous function of E which is not bounded;

b) there exists a continuous and bounded function on E which has no maximum.

c) If, in addition, E is bounded then: there exists a continuous function on E which is not uniformly continuous.

Theorem 4.22 If f is a continuous mapping of a metric space X into a metric space Y, and if E is a connected subset of X, then $f(E)$ is connected.

Theorem 4.23 (Intermediate Value Theorem) Let f be a continuous real function on the interval $[a,b]$. If $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a,b)$ such that $f(x) = c$.

Theorem 4.29 Let f be monotonically increasing on (a,b). Then $f(x+)$ and $f(x-)$ exists at every point of x of (a,b). More precisely,

$$\sup_{a<t<x} f(t) = f(x-) \leqslant f(x) \leqslant f(x+) = \inf_{x<t<b} f(t)$$

Furthermore, if $a < x < y < b$, then

$$f(x+) \leqslant f(y-)$$

. Analogous results hold for monotonically decreasing functions.

Corollary Monotonic functions have no discontinuities of the second kind.

Theorem 4.30 Let f be monotonic on (a,b). Then the set of points of (a,b) at which f is discontinuous is at most countable.

Theorem 4.34 Let f and g be defined on $E \subset \mathbb{R}$ Suppose

$$f(t) \to A \qquad g(t) \to B \qquad \text{as } t \to x$$

Then

a) $f(t) \to A'$ implies $A' = A$

b) $(f+g)(x) \to A + B$

c) $(fg)(t) \to AB$

d) $(f/g)(t) \to A/B$

Note: $\infty - \infty$, $0 \cdot \infty$, ∞/∞, $A/0$ are not defined.

Chapter 5 Differentiation

Definitions

Deivative Let f be defined (and real-valued) on $[a, b]$. For any $x \in [a, b]$ form the quotient

$$\phi(t) = \frac{f(t) - f(x)}{t - x} \qquad (a < t < b, t \neq x)$$

and define

$$f'(x) = \lim_{t \to x} \phi(t)$$

provided this limit exists.

We thus associate with the function f and a function f' whose domain is the set of points x at which the above limit exists; f' is called the *derivative* of f.

Differentiable If f' is defined at a point x we say that f is *differentiable* at x. If f' is defined at every point of a set $E \subset [a, b]$, we say that f is *differentiable* on E.

Local Maximum Let f be a real function defined on a metric space X. We say that f has a *local maximum* at a point $p \in X$ if there exists $\delta > 0$ such that $f(q) \leqslant f(p)$ for all $q \in X$ with $d(p.q) < \delta$. (Local minima are defined likewise.)

Higher Order Derivatives If f has a derivative f' on an interval. and if f' is itself differentiable, we denote the derivative of f' by f'' and call f'' the *second derivative* of f. Continuing in this manner, we obtain functions

$$f,\ f',\ f'',\ f^{(3)},\ f^{(4)},\ \ldots,\ f^{(n)}$$

each of which is the derivative of the preceding one. $f^{(n)}$ is called the *nth derivative*, or the derivative of order n, of f.

Note, In order for $f^{(n)}(x)$ to exist at a point x, $f^{(n-1)}(t)$ must exist in a neighborhood of x (or in a one-sided neighborhood, if x is an endpoint of the interval on which f is defined), and $f^{(n-1)}(x)$ must be differentiable at x.

Theorems

Theorem 5.2 Let f be defined on $[a, b]$. If f is differentiable at a point $x \in [a, b]$, then f is continuous at x.

Theorem 5.3 Suppose f and g are defined on $[a, b]$ and are differentiable at a point $x \in [a, b]$. Then $f + g$, fg, f/g are differentiable at x, and:

a) $(f + g)'(x) = f'(x) + g'(x)$

b) $(fg)'(x) = f'(x)g(x) + g'(x)f(x)$

c) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - g'(x)f(x)}{g(x)^2}$

Theorem 5.5 (Chain Rule) Suppose f is continuous on $[a,b]$, $f'(x)$ exists at some point $x \in [a,b]$, g is defined on an interval I which contains the range of f, and g is differentiable at the point $f(x)$. If

$$h(t) = g(f(t)) \qquad (a \leqslant t \leqslant b)$$

then h is differentiable at x, and

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

Theorem 5.8 (Rolle's Theorem) Let f be defined on $[a,b]$; if f has a local maximum at a point $x \in (a,b)$, and if $f'(x)$ exists, then $f'(x) = 0$. (The analogous statement for local minima also holds.)

Theorem 5.9 (Generalized Mean Value Theorem) If f and g are continuous real functions on $[a,b]$ which are differentiable in (a,b), then there is a point $x \in (a,b)$ at which:

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x)$$

Theorem 5.10 (Mean Value Theorem) If f is a real continuous function on $[a,b]$ which is differentiable in (a,b), then there is a point $x \in (a,b)$ at which

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

Theorem 5.11 Suppose f is differentiable in (a,b).

a) If $f'(x) \geqslant 0$ for all $x \in (a,b)$, then f is monotonically increasing.
b) If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant.
c) If $f'(x) \leqslant 0$ for all $x \in (a,b)$, then f is monotonically decreasing.

Theorem 5.12 (Intermediate Value Theorem for Derivatives) Suppose f is a real differentiable function on $[a,b]$ and suppose $f'(a) < \lambda < f'(b)$. Then there is a point $x \in (a,b)$ such that $f'(x) = \lambda$.

Corollary If f is differentiable on $[a,b]$, then f' cannot have any simple discontinuities on $[a,b]$.

Theorem 5.13 (L'Hospital's Rule) Suppose f and g are real and differentiable in (a,b), and $g'(x) \neq 0$ for all $x \in (a,b)$, where $-\infty \leqslant a < b \leqslant +\infty$. Suppose

$$\frac{f'(x)}{g'(x)} \to A \qquad as \qquad x \to a.$$

If

$$f(x) \to 0 \qquad and \qquad g(x) \to 0 \qquad as \qquad x \to a$$

or if

$$g(x) \to +\infty \qquad as \qquad x \to a,$$

then

$$\frac{f(x)}{g(x)} \to A \qquad as \qquad x \to a$$

The analogous statement is also true if $x \to b$ or if $g(x) \to -\infty$.

Theorem 5.15 (Taylor's Theorem) Suppose f is a real function on $[a, b]$, n is a positive integer, $f^{(n-1)}$ is continuous on $[a, b]$, $f^{(n)}(t)$ exists for every $t \in (a, b)$. Let α, β be distinct points of $[a, b]$, and define:

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k$$

Then there exists a point x between α and β such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n.$$

Theorem 5.19 Suppose $\mathbf{f}$ is a continuous mapping of $[a, b]$ into $\mathbb{R}^k$ and $\mathbf{f}$ is differentiable in (a, b). Then there exists $x \in (a, b)$ such that

$$|\mathbf{f}(b) - \mathbf{f}(a)| \leqslant (b - a)|\mathbf{f}'(x)|.$$

Chapter 6 The Riemann-Stieltjes Integral

Definitions

Partition Let $[a, b]$ be a given interval. By a *partition* P of $[a, b]$ we mean a finite set of points $x_0, x_1, \ldots, x_n$, where

$$a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b.$$

We write

$$\Delta x_i = x_i - x_{i-1} \quad (i = 1, \ldots, n).$$

Integral Components Suppose f is a bounded real function on $[a, b]$. Corresponding to each partition P of $[a, b]$ we put:

$$M_i = \sup f(x) \quad (x_{i-1} \leqslant x \leqslant x_i), \qquad m_i = \inf f(x) \quad (x_{i-1} \leqslant x \leqslant x_i),$$

$$U(P, f) = \sum_{i=1}^{n} M_i \Delta x_i, \qquad L(P, f) = \sum_{i=1}^{n} m_i \Delta x_i,$$

$$\overline{\int_a^b} f\, dx = \inf_P U(P, f), \qquad \underline{\int_a^b} f\, dx = \sup_P L(P, f)$$

The last two are called the *upper* and *lower Riemann integrals* of f over $[a, b]$ respectively.

Riemann Integrable If the upper and lower integrals are equal we say that f is *Riemann-integrable* on $[a, b]$, we write $f \in \mathscr{R}$ (that is, $\mathscr{R}$ denotes the set of Riemann-integrable functions), and we denote the common value of the upper and lower integrals by:

$$\int_a^b f\, dx \quad \text{or by} \quad \int_a^b f(x)\, dx$$

Alpha Let α be a monotonically increasing function on $[a, b]$. Corresponding to each partition P of $[a, b]$ we write

$$\Delta \alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \quad (i = 1, \ldots, n).$$

Note, $\Delta \alpha_i \geqslant 0$.

Integral Components Suppose f is a bounded real function on $[a, b]$. Corresponding to each partition P of $[a, b]$ we put:

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta \alpha_i, \qquad L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta \alpha_i,$$

$$\overline{\int_a^b} f\, d\alpha = \inf_P U(P, f, \alpha), \qquad \underline{\int_a^b} f\, d\alpha = \sup_P L(P, f, \alpha)$$

Riemann- Stieltjes Integrable If the upper and lower integrals are equal we denote their common value by

$$\int_a^b f\,d\alpha \quad \text{or by} \quad \int_a^b f(x)\,d\alpha.$$

This is the *Riemann- Stieltjes integral* of f with respect to α over $[a,b]$. If this exists, we say that f is integrable with to α and we write $f \in \mathscr{R}(\alpha)$

Refinement We say that the partition P^* is a *refinement* of P is $P^* \supset P$.

Common Refinement Given two partitions, P_1 and P_2, we say that $P*$ is their *common refinement* if $P^* = P_1 \cup P_2$.

Unit Step Function The *unit step function* I is defined by:

$$I(x) = \begin{cases} 0 & (x \leqslant 0), \\ 1 & (x > 0). \end{cases}$$

Vector-valued Functions Let $f_1, \dots, f_k$ be real functions on $[a,b]$ and let $\mathbf{f} = (f_1, \dots, f_k)$ be the corresponding mapping of $[a,b]$ into $\mathbb{R}^k$. If α increases monotonically on $[a,b]$, to say that $\mathrm{f} \in \mathscr{R}(\alpha)$ means that $f_j \in \mathscr{R}(\alpha)$ for $j = 1, \dots, k$. If this is the case, we define

$$\int_a^b \mathbf{f}\,d\alpha = \left(\int_a^b f_1\,d\alpha, \dots, \int_a^b f_k\,d\alpha \right).$$

In other words, $\int \mathbf{f}\,d\alpha$ is the point in $\mathbb{R}^k$ whose jth coordinate is $\int f_j\,d\alpha$.

Curve/ Arc/ Closed Curve/ Length/ Rectifiable A continuous mapping γ of an interval $[a,b]$ into $\mathbb{R}^k$ is called a *curve* in $\mathbb{R}^k$. To emphasize the parameter interval $[a,b]$, we may also say that γ is a curve on $[a,b]$. If γ is one-to-one, γ is called an *arc*. If $\gamma(a) = \gamma(b)$, γ is said to be a *closed curve*. It should be noted that we define a curve to be a *mapping*, not a point set.

We associate to each partition $P = \{x_0, \dots, \mathrm{x}_n\}$ of $[a,b]$ and to each curve γ on $[a,b]$ the number

$$\Lambda(P,\gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1}|.$$

The ith term in this sum is the distance between the points. Hence $\Lambda(P,\gamma)$ is the length of a polygonal path with vertices at $\gamma(x_0), \gamma(x_1), \dots, \gamma(x_n)$, in this order. As our partition becomes finer and finer, this polygon approaches the range of γ more closely. Thus, *length* of γ is:

$$\Lambda(\gamma) = \sup \Lambda(P,\gamma),$$

If $\Lambda(\gamma) < \infty$ we say that γ is *rectifiable*.

Theorems

Theorem 6.4 If P^* is a refinement of P, then

$$L(P, f, \alpha) \leqslant L(P^*, f, \alpha) \quad \text{and} \quad U(P^*, f, \alpha) \leqslant U(P, f, \alpha).$$

Theorem 6.5 $\underline{\int_a^b} f \, d\alpha \leqslant \overline{\int_a^b} f \, d\alpha$

Theorem 6.6 $f \in \mathscr{R}(\alpha)$ on $[a, b]$ if and only if for every $\varepsilon > 0$ there exists a partition P such that:

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Theorem 6.7 a) If $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ holds for some P and some ε, then it holds for every refinement of P.

b) If $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ holds for $P = \{x_0, \dots, x_n\}$ and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \delta\alpha_i < \varepsilon$$

c) If $f \in \mathscr{R}(\alpha)$ and the hypotheses of (b) hold, then

$$\left| \sum_{i=1}^{n} f(t_i)\delta\alpha_i - \int_a^b f \, d\alpha \right| < \varepsilon$$

Theorem 6.8 If f is continuous on $[a, b]$ then $f \in \mathscr{R}(\alpha)$ on $[a, b]$.

Theroem 6.9 If f is monotonic on $[a, b]$, and if α is continuous on $[a, b]$, then $f \in \mathscr{R}(\alpha)$.

Theorem 6.10 Suppose f is bounded on $[a, b]$, f has only finitely many points of discontinuity on $[a, b]$ and α is continuous at every point at which f is discontinuous. Then $f \in Ra$.

Theorem 6.11 (Composition of Functions) Suppose $f \in \mathscr{R}(\alpha)$ on $[a, b]$, $m \leqslant f \leqslant M$, ϕ is continuous on $[m, M]$, and $h(x) = \phi(f(x))$ on $[a, b]$. Then $h \in \mathscr{R}(\alpha)$ on $[a, b]$.

Theorem 6.12 (Properties of Integrals) a) If $f_1, f_2 \in \mathscr{R}(\alpha)$ on $[a, b]$, then $f_1 + f_2 \in \mathscr{R}(\alpha)$, $cf \in \mathscr{R}(\alpha)$ for every constant c, and

$$\int_a^b (cf_1 + f_2) \, d\alpha = c \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha$$

b) If $f_1(x) \leqslant f_2(x)$ on $[a, b]$, then

$$\int_a^b f_1 \, d\alpha \leqslant \int_a^b f_2 \, d\alpha.$$

c) If $f \in \mathscr{R}(\alpha)$ on $[a,b]$ and $a < c < b$, then $f \in \mathscr{R}(\alpha)$ on $[a,c]$ and $f \in \mathscr{R}(\alpha)$ on $[c,b]$, and

$$\int_a^c f\,d\alpha + \int_c^b f\,d\alpha = \int_a^b f\,d\alpha.$$

d) If $f \in \mathscr{R}(\alpha)$ on $[a,b]$ and if $|f(x)| \leqslant M$ on $[a,b]$, then

$$\left|\int_a^b f\,d\alpha\right| \leqslant M[\alpha(a) - \alpha(b)].$$

e) If $f \in \mathscr{R}(\alpha_1)$ and $f \in \mathscr{R}(\alpha_2)$, then $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ and

$$\int_a^b f\,d(\alpha_1 + \alpha_2) = \int_a^b f\,d(\alpha_1) + \int_a^b f\,d(\alpha_2);$$

If $f \in \mathscr{R}(\alpha)$ and c is positive constant, then $f \in \mathscr{R}(c\alpha)$ and

$$\int_a^b f\,d(c\alpha) = c\int_a^b f\,d\alpha$$

Theorem 6.13 If $f \in \mathscr{R}(\alpha)$ and $f \in \mathscr{R}(\alpha)$ on $[a,b]$, then:

a) $fg \in \mathscr{R}(\alpha)$;

b) $|f| \in \mathscr{R}(\alpha)$ and $\left|\int_a^b f\,d\alpha\right| \leqslant \int_a^b |f|\,d\alpha$.

Theorem 6.15 If $a < s < b$, f is bounded on $[a,b]$, f is continuous at s and $\alpha = I(x - s)$, then

$$\int_a^b f\,d\alpha = f(s)$$

Theorm 6.16 Suppose $c_n \geqslant 0$ for $1, 2, 3, \ldots, \sum c_n$ converges, $\{s_n\}$ is a sequence of distinct points in (a,b) and

$$\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n)$$

Let f be continuous on $[a,b]$. Then

$$\int_a^b f\,d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Theorem 6.17 Assume α increases monotonically and $\alpha' \in \mathscr{R}$ on $[a,b]$. Let f be a bounded real function on $[a,b]$. Then $f \in \mathscr{R}(\alpha)$ if and only if $f\alpha' \in \mathscr{R}$. In that case

$$\int_a^b f\,d\alpha = \int_a^b f(x)\alpha'(x)\,dx.$$

Theorem 6.18 (Change of Variables) Suppose φ is a strictly increasing continuous function that maps an interval $[A, B]$ onto $[a, b]$. Suppose α is a monotonically increasing function on $[a, b]$ and $f \in \mathscr{R}(\alpha)$ on $[a, b]$. Define β and g on $[A, B]$ by

$$\beta(y) = \alpha(\varphi(y)), \qquad g(y) = f(\varphi(y)).$$

Then $g \in \mathscr{R}(\beta)$ and

$$\int_A^B d\, d\beta = \int_a^b f\, d\alpha$$

Theorem 6.20 Let $f \in \mathscr{R}$ on $[a, b]$, for $a \leqslant x \leqslant b$, put

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

Then F is continuous on $[a, b]$; furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0, and $F'(x_0) = f(x_0)$.

The Fundamental Theorem of Calculus If $f \in \mathscr{R}$ on $[a, b]$ and if there is a differentiable function F on $[a, b]$ such that $F' = F$, then

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

Theorem 6.22 (Integration by Parts) Suppose F and G are differentiable functions on $[a, b]$, $F' = f \in \mathscr{R}$, and $G' = g \in \mathscr{R}$. Then

$$\int_a^b F(x)g(x)\, dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\, dx.$$

Theorem 6.24 If $\mathbf{f}$ and $\mathbf{F}$ map $[a, b]$ into $\mathbb{R}^k$, if $\mathbf{f} \in \mathscr{R}$ on $[a, b]$, and if $\mathbf{F}' = \mathbf{f}$, then

$$\int_a^b \mathbf{f}(t)\, dt = \mathbf{F}(b) - \mathbf{F}(a).$$

Theorem 6.25 If $\mathbf{f}$ maps $[a, b]$ into $\mathbb{R}^k$ and if $\mathbf{f} \in \mathscr{R}(\alpha)$ for some monotonically increasing function α on $[a, b]$, then $|\mathbf{f}| \in \mathscr{R}(\alpha)$ and

$$\left| \int_a^b \mathbf{f}\, d\alpha \right| \leqslant \int_a^b |\mathbf{f}|\, d\alpha.$$

Theorem 6.27 If γ' is continuous on $[a, b]$, then γ is rectifiable, and

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\, dt.$$

Chapter 7 Sequences and Series of Functions

Definitions

Limit/ pointwise convergence/ sum Suppose $\{f_n\}$, $n = 1, 2, 3, \ldots$, is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in E$. We an then define a function f by

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad (x \in E)$$

Under these circumstances we say that $\{f_n\}$ converges on E, that f is the *limit*, or the *limit function*, of $\{f_n\}$, and that $\{f_n\}$ converges to f *pointwise* on E. Similarly, if $\sum f_n(x)$ converges for every $x \in E$, and if we define:

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (x \in E)$$

the function f is called the *sum* of the series σf_n.

Uniform Convergence We say that a sequence of functions $\{f_n\}$, $n = 1, 2, 3, \ldots$, converges *uniformly* on E to a function f is for every $\varepsilon > 0$ there is an integer N such that $n \geqslant N$ implies the following for all $x \in E$:

$$\mathrm{f}_n(x) - f(x) \leqslant \varepsilon$$

Supremum Norm If X is a metric space, $\mathscr{C}(X)$ will denote the set of all complex=values, continuous, bounded functions with domain X. We associate each $f \in \mathscr{C}(X)$ with its *supremum norm*

$$||f|| = \sup_{x \in X} |f(x)|.$$

Since f is assumed to be bounded, $||f|| \leqslant \infty$. It is obvious that $||f|| = 0$ if and only if $f(x) = 0$ for every $x \in X$. If $h = f + g$:

$$|h(x)| \leqslant |f(x)| + |g(x)| \leqslant ||f|| + ||g||$$

for all $x \in X$; hence

$$||f + g|| \leqslant ||f|| + ||g||.$$

$\mathscr{C}(X)$ as a Metric Space If we define the distance between $f \in \mathscr{C}(X)$ and $g \in CX$ to be $||f - g||$, it follows that the Axioms for a metric are satisfied. Therefore, A sequence $\{f_n\}$ converges to f with respect to the metric of $\mathbb{C}$ if and only if $f_n \rightrightarrows f$ on X.

Pointwise Bounded Let $\{f_n\}$ be a sequence of functions defined on a set E. We say that $\{f_n\}$ is *pointwise bounded* on E if the sequence $\{f_n(x)\}$ is bounded for every $x \in E$, that is, if there exists a finite-valued function ϕ defined on E such that

$$|f_n(x)| < \phi(x) \qquad (x \in E,\ n = 1, 2, 3, \ldots).$$

(If $\{f_n\}$ is pointwise bounded on E and E_1 is a countable subset of E it is always possible to find a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E_1$.)

Uniformly Bounded We say that $\{f_n\}$ is *uniformly bounded* on E if there exists a number M such that

$$|f_n(x)| < M \qquad (x \in E,\ n = 1, 2, 3, \ldots).$$

(If $\{f_n\}$ is a uniformly bounded sequence of continuous functions on a compact set E, there need not exists a subsequence which converges pointwise on E.)

Equicontinuous A family $\mathscr{F}$ of complex functions f defined on a set E in a metric space X is said to be *equicontinuous* on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon$$

whenever $d(x, y) < \delta$, $x, y \in E$, $f \in \mathscr{F}$. (Note: Every member of an equicontinuous family is uniformly continuous.)

Algebra A family $\mathscr{A}$ of complex functions defined on a set E is said to be an *algebra* if: (i) $f + g \in \mathscr{A}$, (ii) $fg \in \mathscr{A}$ (iii) $cf \in \mathscr{A}$. for all $f, g \in \mathscr{A}$ and for all complex constants c, that is $\mathscr{A}$ is closed under addition, multiplication, and scalar multiplication.

Uniformly Closed If $\mathscr{A}$ has the property that $f \in \mathscr{A}$ whenever $f_n \in A$ $(n = 1, 2, 3, \ldots)$ and $f_n \rightrightarrows f$ on E, then $\mathscr{A}$ is said to be *uniformly closed.*

Uniform Closure Let $\mathscr{B}$ be the set of all functions which are limits of uniformly convergent sequences of members of $\mathscr{A}$. Then $\mathscr{B}$ is called the *uniform closure* of $\mathscr{A}$.

Seperate Points Let $\mathscr{A}$ be a family of functions on a set E. Then $\mathscr{A}$ is said to *separate points* on E if every pair of distinct points $x_1, x_2 \in E$ there corresponds a function $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$.

Vanishes At No Point If to each $x \in E$ there corresponds a function $g \in \mathscr{A}$ such that $g(x) \neq 0$, we say that $|A$ vanishes at no point of E.

Theorems

Theorem 7.8 (Cauchy) The sequence of functions $\{f_n\}$, defined on E, converges uniformly on E if and only if for every $\varepsilon > 0$ there exists an integer N such that $m \geqslant N$, $n \geqslant N$, $x \in E$ implies:

$$|f_n(x) - f_m(x)| \leqslant \varepsilon$$

Theorem 7.9 Suppose

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad (x \in E)$$

Put

$$M_n = \sup_{x \in E} |f_n(x) - f(x)|$$

Then $\mathbf{f}_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n \to infty$.

Theorem 7.10 (M-test) Suppose $\{f_n\}$ is a sequence of functions defined on E, and suppose

$$|f_n(x)| \leqslant M_n \qquad (x \in E,\ n = 1, 2, 3, \ldots).$$

Then $\sum f_n$ converges uniformly on E if $\sum M_n$ converges.

Theorem 7.11 Suppose $f_n \to f$ uniformly on a set E in a metric space. Let x be a limit point of E, and suppose that

$$\lim_{t \to x} f_n(t) = A_n \qquad (n = 1, 2, 3, \ldots).$$

Then $\{A_n\}$ converges, and

$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n.$$

In other words, the conclusion is that:

$$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t)$$

Theorem 7.12 If $\{f_n\}$ is a sequence of continuous functions on E and if $f_n \rightrightarrows f$ on E, then f is continuous on E.

Theorem 7.13 Suppose K is compact, and

a) $\{f_n\}$ is a sequence of continuous functions on K

b) $\{f_n\}$ converges pointwise to a continuous function f on K,

c) $f_n(x) \geqslant f_{n+1}(x)$ for all $x \in K$, $n = 1, 2, 3 \ldots$ ($\{f_n\}$ is a decreasing sequence)

Then $f_n \rightrightarrows f$ on K.

Theorem 7.15 The aforementioned metric makes $\mathscr{C}(X)$ a complete metric space.

Theorem 7.16 Let α be monotonically increasing on $[a, b]$, for $n = 1, 2, 3, \ldots$, and suppose $f_n \rightrightarrows f$ on $[a, b]$. Then $f \in \mathscr{R}(\alpha)$ on $[a, b]$ and

$$\int_a^b f\, d\alpha = \lim_{n \to \infty} \int_a^b f_n\, d\alpha.$$

Corollary If $f_n \in \mathscr{R}(\alpha)$ on $[a, b]$ and if

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \qquad (a \leqslant x \leqslant b)$$

the series converging uniformly on $[a, b]$, then

$$\int_a^b f\, d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n\, d\alpha$$

In other words, the series may be integrated term by term.

Theorem 7.17 Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a,b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a,b]$. In $\{f_n'\}$ converges uniformly on $[a,b]$, then $\{f_n\}$ converges uniformly on $[a,b]$, to a function f, and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \qquad (a \leqslant x \leqslant b).$$

Theorem 7.18 There exists a real continuous function on the real line which is nowhere differentiable.

Theorem 7.23 If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.

Theorem 7.24 If K is a compact metric space, if $f_n \in \mathscr{C}(K)$ for $n = 1, 2, 3, \backslash dots$, and if $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.

Theorem 7.25 If K is compact, if $\mathrm{f}_n \in \mathscr{C}(K)$ for $n = 1, 2, 3, dots$, and if $\{f_n\}$ is pointwise bounded and equicontinuous on K, then

a) $\{f_n\}$ is uniformly bounded on K,

b) $\{f_n\}$ contains a uniformly convergent subsequence.

Theorem 7.26 (Stone-Weierstrass Theorem) If f is a continuous complex 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 $[a,b]$. If f is real, the P_n may be taken real.

Corollary 7.27 For every interval $[-a,a]$ there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and such that

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

uniformly on $[-a,a]$.

Theorem 7.29 Let $\mathscr{B}$ be the uniform closure of an algebra $\mathscr{A}$ of bounded functions. Then $\mathscr{B}$ is a uniformly closed algebra.

Theorem 7.31 Suppose $\mathscr{A}$ is an algebra of functions on a set E, $\mathscr{A}$ separates points on E, and $|A$ vanishes at no points of E. Suppose x_1, x_2 are distinct points of E, and c_1, c_2 are constants. Then $\mathscr{A}$ contains a function of f such that:

$$f(x_1) = c_1 \qquad f(x_2) = c_2$$

Theorem 7.32 Let $\mathscr{A}$ be an algebra of real continuous functions on a compact set K. IF $\mathscr{A}$ separates points on K and if $\mathscr{A}$ vanishes at no point of K, then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all real continuous functions on K.

Theorem 7.33 Suppose $\mathscr{A}$ is a self-adjoint algebra of complex continuous functions on a compact set K, $\mathscr{A}$ separates points on K, and $\mathscr{A}$ vanishes at no point of K. Then the uniform closure $\mathscr{B}$ of $|A$ consists of all complex continuous functions on K. In other words, $\mathscr{A}$ is dense in $\mathscr{C}(K)$.

Chapter 8 Some Special Functions

Definitions

Analytic Functions Functions of the form

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

Exponential Function Define

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

The ratio test shows that this series converges for every complex z. Note:

$$E(z)E(w) = \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = E(z+w)$$

Thus, $E(z)E(-z) = 1$. Further,

$$E'(z) = \lim_{h \to 0} \frac{E(z+h) - E(z)}{h} = \lim_{h \to 0} \frac{E(z+h) - 1}{h} E(z)$$

Let $E(1) = e$. So $E(n) = E(1+1+1+\cdots+1) = E(1)E(1)\ldots E(1) = e^n$. This holds for any $n \in \mathbb{Q}$. Furthermore, $E(x) = e^x = \sup e^p$ ($p < x$, p rational).

Triginometric Functions Define the following:

$$C(x) = \frac{1}{2}[E(ix) + E(-ix)] \qquad S(x) = \frac{1}{2i}[E(ix) - E(-ix)]$$

Note: $E(ix) = C(x) + iS(x)$. Further,

$$C'(x) = -S(x) \qquad S'(x) = C(X)$$

Ultimately equivalent to cos and sin.

Trigonometric Polynomial A trigonometric polynomial is a finite sum of the form

$$f(x) = a_0 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) \qquad (x real),$$

where $a_0, \ldots, a_N, b_1, \ldots, b_N$ are complex numbers. The above identities can also be written in the form

$$f(x) = \sum_{-N}^{N} c_n e^{inx}$$

It follows that:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx}\, dx = \begin{cases} 1 & (if n = 0) \\ 0 & (if n = \pm 1, \pm 2, \ldots) \end{cases}$$

Fourier Coefficients If f is an integrable function on $[-\pi, \pi]$, the numbers c_m defines by:

$$c_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}\,dx$$

for all integers m are called the *Fourier coefficients* of f,

Fourier Series The series:

$$\sum_{-\infty}^{\infty} c_n e^{inx}$$

formed with the Fourier coefficients is called the *Fourier series* of f.

Orthogonal System of Functions/ Orthonormal Let $\{\phi_n\}$ $(n = 1, 2, 3, \ldots)$ be a sequence of complex functions on $[a, b]$ such that

$$\int_a^b \phi_n(x)\overline{\phi_m(x)}\,dx = 0 \qquad (n \neq m).$$

Then $\{\phi_n\}$ is said to be an *orthogonal system of functions* on $[a, b]$. If in addition:

$$\int_a^b |\phi_b(x)|^2\,dx = 1$$

for all n, $\{\phi_n\}$ is said to be *orthonormal*.

Gamma Function For $0 < x < \infty$

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

The integral converges for these x. (When $x < 1$, both 0 and ∞ have to be looked at.)

Theorems

Theorem 8.1 Suppose the series

$$\sum_{n=0}^{\infty} c_n x^n$$

converges for $|x| < R$, and define

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < R).$$

Then the series converges uniformly on $[-R+\varepsilon, R-\varepsilon]$ no matter which $\varepsilon > 0$ is chosen. The function f is continuous and differentiable in $(-R, R)$.

Corollary Under the hypotheses of Theorem 8.1, f has derivatives of all orders in $(-R, R)$, which are given by:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\dots(n-k+1)c_n x^{n-k}.$$

In particular, $f^{(k)}(0) = k!c_k$.

Theorem 8.2 Suppose $\sum c_n$ converges. Put

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < 1).$$

Then

$$\lim_{x \to 1} f(x) = \sum_{n=0}^{\infty} c_n.$$

Theorem 8.3 Given a double sequence $\{a_{ij}\}$, $i = 1, 2, 3, \dots, j = 1, 2, 3, \dots$, suppose that:

$$\sum_{j=1}^{\infty} |a_{ij}| = b_i$$

and $\sum b_i$ converges. Then

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$$

Theorem 8.4 (Taylor's Theorem) Suppose

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < R).$$

If $-R < a < R$, then f can be expanded in a power series about the point $x - a$ which converges in $|x - a| < R - |a|$, and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \qquad (|x-a| < R - |a|)$$

Theorem 8.5 Suppose the series $\sum a_n x^n$ and $\sum b_n x^n$ converge in the segment $S = (-R, R)$. Let E be the set of all $x \in S$ at which

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

If E has a limit point in S, then $a_n = b_n$ for $n = 0, 1, 2, \dots$. Hence the above equation holds for all $x \in S$.

Theorem 8.6 Let e^x be defined on R^1 as it is above. Then:

a) e^x is continuous and differentiable for all x;

b) $(e^x)' = e^x$;

c) e^x is a strictly increasing function of x and $e^x > 0$;

d) $e^{x+y} = e^x e^y$;

e) $e^x \to +\infty$ as $x \to +\infty$, $e^x \to 0$ as $x \to -\infty$;

f) $\lim_{x\to+\infty} x^n e^{-x} = 0$ for every n.

Theorem 8.7 a) The function E is periodic, with period $2\pi i$.

b) The functions C and S are periodic with period 2π.

c) If $0 < t < 2\pi$ then $E(it) \neq 1$.

d) If z is a complex number with $|z| = 1$, there is a unique $r \in [0, 2\pi)$ such that $E(it) = z$.

Theorem 8.8 Suppose $a_0, \ldots, a_n$ are complex number, $n \geqslant 1$, $a_n \neq 0$,

$$P(z) = \sum_0^n a_k z^k.$$

Then $P(z) = 0$ for some complex number z.

Theorem 8.11 Let $\{\phi_n\}$ be orthonormal on $[a, b]$. Let

$$s_n(x) = \sum_{m=1}^n c_m \phi_m(x)$$

be the nth partial sum of the Fourier series of f, and suppose

$$t_n(x) = \sum_{m=1}^n \gamma_m \phi_m(x).$$

Then

$$\int_a^b |f - s_n|^2 \, dx \leqslant \int_a^b |f - t_n|^2 \, dx,$$

and equality holds if and only if $\gamma_m = c_m$.

Corollary If $f(x) = 0$ for all x in some segment J, then $\lim_{s_N}(f; x) = 0$ for every $x \in J$.

Theorem 8.15 If f is continuous (with period 2π) and if $\varepsilon > 0$, then there is a trigonometric polynomial P such that $|P(x) - f(x)| < \varepsilon$ for all real x.

Theorem 8.16 (Parseval's Theorem) Suppose f and g are Riemann-integrable functions with period 2π, and

$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx} \qquad g(x) \sim \sum_{-\infty}^{\infty} \gamma_n e^{inx}.$$

Then

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 \, dx = 0$$

$$\frac{1}{2\pi} int_{-\pi}^{\pi} f(x)\overline{g(x)} \, dx = \sum_{-\infty}^{\infty} c_n \overline{\gamma}_n$$

$$\frac{1}{2\pi} int_{-\pi}^{\pi} |f(x)^2 \, dx = \sum_{-\infty}^{\infty} |c_n|^2$$

Theorem 8.18 a) The function equation:

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

holds if $0 < x < \infty$.

b) $\Gamma(n+1) = n!$ for $n = 1, 2, 3, \ldots$.

c) $\log \Gamma$ is convex on $(0, \infty)$.

Theorem 8.19 If f is a positive function on $(0, \infty)$ such that

a) $f(x+1) = xf(x)$

b) $f(1) = 1$

c) $\log f$ is convex

then $f(x) = \Gamma(x)$.

Theorem 8.20 If $x > 0$ and $y > 0$, then

$$\int_0^1 f^{x-1}(1-t)^{y-1} \, dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

This integral is the so called *beta function* $B(x, y)$.

Stirling's Formula This provides a simple approximate expression for $\Gamma(x+1)$ when x is large (hence for $n!$ when n is large). The formula is

$$\lim_{x\to\infty} \frac{\Gamma(x+1)}{(x/e)^x \sqrt{2\pi x}} = 1$$

Chapter 9 Functions of Several Variables

Definitions

Vector Space A nonempty set $X \subset \mathbb{R}^n$ is a *vector space* if $\mathbf{x}+\mathbf{y} \in X$ and $c\mathbf{x} \in X$ for all $\mathbf{x}, \mathbf{y} \in X$, $c \in \mathbb{R}$.

Linear Combination If $\mathbf{x}_1, \dots, \mathbf{x}_k \in R^n$ and $c_1, \dots, c_k$ are scalars, the vector

$$c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k$$

is called a *linear combination* of $\mathbf{x}_1, \dots, \mathbf{x}_k$.

Span If $S \subset R^n$ and if E is the set of all linear combinations of elements of S we say that S *spans* E, or that E is *the span* of S. Observe that every span is a vector space.

Independent/ Dependent A set consisting of vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ (we shall use the notation $\{x_1, \dots, \mathbf{x}_k\}$ for such a set) is said to be *independent* if the relation $c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = 0$ implies that $c_1 = \cdots = c_k = 0$. Otherwise $\{x_1, \dots, \mathbf{x}_k\}$ is said to be *dependent*.

Dimension If a vector space X contains an independent set of r vectors but contains no independent set of $r+1$ vectors, we say that X has *dimension* r, and write: $\dim X = r$.

Basis/ Coordinates/ Standard Basis An independent subset of a vector space X which spans X is called a *basis* of X. Observe that if $B = \{x_1, \dots, \mathbf{x}_r\}$ is the basis of X, then every $\mathbf{x} \in X$ has a unique representation of the for $\mathbf{x} = \sum c_j\mathbf{x}_j$. Such a representation exists since V spans X, and it is unique since B is independent. The numbers $c_1, \dots, c_r$ are called the *coordinates* of $\mathbf{x}$ with respect to the basis B. The most familiar example of a basis is the set $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, where $\mathbf{e}_j$ is the vector in $\mathbb{R}^n$ whose jth coordinate is 1 and whose other coordinates are all 0. If $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = (x_1, \dots, x_n)$, then $\mathbf{x} = \sum x_j\mathbf{e}_j$. We shall call $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ the *standard basis*

Linear Transformation A mapping A of a vector space X into a vector space Y is said to be a *linear transformation* if

$$A(c\mathbf{x}_1 + \mathbf{x}_2) = cA(\mathbf{x}_1) + A(x_2)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and all scalars c. Note that $A\mathbf{x} = A(\mathbf{x})$. Further, a linear transformation A of X into Y is completely determined by its action on any basis.

Linear Operators A linear transformations of X into X are often called *linear operators* on X.

Invertible If A is a linear operator on X which (i) is one-to-one and (ii) maps X onto X, we say that A is invertible. In this case we can define an operator A^{-1} on X by requiring that $A^{-1}(A\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in X$.

Set of Linear Transformation Let $L(X, Y)$ be the *set of all linear transformations* of the vector space X into the vector space Y. Instead of $L(X, X)$ we shall simply write $L(X)$. If $A_1, A_2 \in L(X, Y)$ and if c_1, c_2 are scalars, define $c_1A_1 + c_2A_2$ by

$$(c_1A_1 + c_2A_2)\mathbf{x} = c_1A_1\mathbf{x} + c_2A_2\mathbf{x}$$

Clearly $c_1A_1 + c_2A_2 \in L(X, Y)$

Product If X, Y, Z are vector spaces and if $A \in L(X, Y)$ and $B \in L(Y, Z)$, we define their *product* BA to be the composition of A and B:

$$(BA)\mathbf{x} = B(A\mathbf{x}) \qquad (\mathbf{x} \in X)$$

Then $BA \in L(X, Z)$. Note that BA need not be the same as AB, even if $X = Y = Z$.

Norm For $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, define the *norm* $||A||$ of A to be the sup of all numbers $|A\mathbf{x}|$, where $\mathbf{x}$ ranges over all vectors in $\mathbb{R}^n$ with $|\mathbf{x}| \leqslant 1$. Observe that the inequality

$$|A\mathbf{x}| \leqslant ||A|||\mathbf{x}|$$

holds for all $\mathbf{x} \in \mathbb{R}^n$. Also, if λ is such the $|A\mathbf{x}| \leqslant \lambda|\mathbf{x}|$ for all $\mathbf{x} \in \mathbb{R}^n$, then $||A|| \leqslant \lambda$

Matrices Omitted, trivial.

Differentiable Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, and $\mathbf{x} \in E$. If there exists a linear transformation A of $\mathbb{R}^n$ into $\mathbb{R}^m$ such that

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

then we say that $\mathbf{f}$ is *differentiable* at $\mathbf{x}$, and we write:

$$\mathbf{f}'(\mathbf{x}) = A$$

If $\mathbf{f}$ is differentiable at every $x \in E$ we say that $\mathbf{f}$ is differentiable in E.

If $|\mathbf{h}|$ is small enough then $\mathbf{x} + \mathbf{h} \in E$, since E is open. Thus $\mathbf{f}(\mathbf{x} + \mathbf{h})$ is defined, $\mathbf{f}(\mathbf{x} + \mathbf{h}) \in \mathbb{R}^m$, and since $\mathscr{A} \in L(\mathbb{R}^n, \mathbb{R}^m)$, $A\mathbf{h} \in \mathbb{R}^m$. Thus

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(x) - A\mathbf{h} \in \mathbb{R}^m.$$

Notes a)

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

can be rewritten in the form:

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

where the remainder $r(\mathbf{h})$ satisfies: $\lim_{\mathbf{h}\to 0} \frac{|r(\mathbf{h})|}{|h|} = 0$. That is, for fixed $\mathbf{x}$ and small $\mathbf{h}$ the left side is approximately equal to $\mathbf{f}'(\mathbf{x})\mathbf{h}$, that is, to the value of the linear transformation applied to $\mathbf{h}$.

b) If f is differentiable in E then $\mathbf{f}'(\mathbf{x})$ is a function that maps E into $L(\mathbb{R}^n, \mathbb{R}^m)$.

c) f is continuous at any point at which f is differentiable.

d) The aforementioned derivative in part (a) is called the differential of f at **x**, or the total derivative, to distinguish it from the partial derivatives.

Components Consider $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. The *components* of f are the real functions $f_1, \ldots, f_m$ defined by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u}_i \qquad (\mathbf{x} \in E)$$

Partial Derivative For $\mathbf{x} \in E$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, we define:

$$(D_j f_i)(\mathbf{x}) = \lim_{t \to 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t},$$

provided the limit exists. Writing $f_i(x_1, \ldots, x_n)$ in place of $f_i(\mathbf{x})$ we see that $D_j f_i$ is the derivative of f_i with respect to x_j, keeping the other variables fixed. The notation $\frac{\partial f_i}{\partial x_j}$ is therefore often used in place of $D_j f_i$, and $D_j f_i$ is called a *partial derivative.*

Continuously Differentiable A differentiable mapping f of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is said to be *continuously differentiable* in E if $\mathbf{f}'$ is a continuous mapping of E into $L(\mathbb{R}^n, \mathbb{R}^m)$. More explicitly, it is required that to every $\mathbf{x} \in E$ and to every $\varepsilon > 0$ corresponds a $\delta > 0$ such that

$$\|\mathbf{f}'(\mathbf{y}) - \mathbf{f}'(\mathbf{x})\| < \varepsilon$$

if $\mathbf{y} \in E$ and $|\mathbf{x} - \mathbf{y}| < \delta$.

Contraction Let X be a metric space, with metric d. If φ maps X into x and if there is a number $c < 1$ such that

$$d(\varphi(x), \varphi(y)) \leqslant cd(x, y)$$

for all $x, y \in X$, then φ is said to be a *contraction* of X into X.

Fixed Point For $\varphi : X \to X$ a point $x \in X$ such that $\phi(x) = x$ is called a *fixed point.*

Notation for Implicit Function Theorem If $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$, let us write $(\mathbf{x}, \mathbf{y})$ for the point (or vector)

$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}$$

In what follows, the first entry in $(\mathbf{x}, \mathbf{y})$ or in a similar symbol will always be a vector in $\mathbb{R}^n$ and the second a vector in $\mathbb{R}^m$.

Every $A \in L(R^{n+m}, R^n)$ can be split into two linear transformations A_x and A_y defined by

$$A)x\mathbf{h} = A(\mathbf{h}, 0), \qquad A_y\mathbf{k} = A(0, \mathbf{k})$$

for any $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{k} \in \mathbb{R}^m$. Then $A_x \in L(\mathbb{R}^n)$ and $A_y \in L(R^m, R^n)$, and

$$A(\mathbf{h}, \mathbf{k}) = A_x\mathbf{h} + a_y\mathbf{k}.$$

Null Space The *null space* of A, $\mathscr{N}(A)$, is the set of all $\mathbf{x} \in X$ at which $A\mathbf{x} = \mathbf{0}$. It is clear that $\mathscr{N}(A)$ is a vector space in X.

Range The *range* of A, $\mathscr{R}(A)$, is a vector space in Y.

Rank The *rank* of A is defined to be the dimension of $\mathscr{R}(A)$.

Projection Let X be a vector space. An Operator $P \in L(X)$ is said to be a *projection* in X if $P^2 = P$.

More explicitly, the requirement is that $P(P\mathbf{x}) = P\mathbf{x}$ for every $\mathbf{x} \in X$. IN other words, p fixes every vector in its range $\mathscr{R}(P)$. Some elementary properties:

a) If P is a projection in X, then every $\mathbf{x} \in X$ has a unique representation of the form $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where $\mathbf{x}_1 \in \mathscr{R}(P)$, $\mathbf{x}_2 \in \mathscr{N}(P)$.

b) If X is a finite-dimensional vector space and if X_1 is a vector space in X, then there is a projection P in X with $\mathscr{R}(P) = X_1$.

Determinants If $(j_1, \ldots, j_n)$ is an ordered n-tuples, define

$$s(j_1, \ldots, j_n) = \prod_{p<q} sgn(j_q - j_p)$$

where sgn is the sign. Let $[A]$ be the matrix of a linear operator A on $\mathbb{R}^n$, relative to the standard basis $\{e_1, \ldots, e_n\}$, with entries a_{ij} in the ith row and jth column.

$$\det[A] = \sum s(j_1, \ldots, j_n) a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

The sum extends over n-tuples of integers. Let $\mathbf{x}_i$ be the ith column vector of A.

$$\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det[A].$$

Jacobians If $\mathbf{f}$ maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and if $\mathbf{f}$ is differentiable at a point $\mathbf{x} \in E$, the determinant of the linear operator $\mathbf{f}'(\mathbf{x})$ is called the *Jacobian* of $\mathbf{f}$ at $\mathbf{x}$:

$$J_{\mathbf{f}}(\mathbf{x}) = \det \mathbf{f}'(\mathbf{x})$$

For $(y_1, \ldots, y_n) = \mathbf{f}(x_1, \ldots, x_n)$, we shall also use the notation:

$$\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}$$

Second-order Partial Derivatives Suppose f is a real function defined in an open set $E \subset \mathbb{R}^n$ with partial derivatives $D_1 f, \ldots, D_n f$. If the functions $D_j f$ are themselves differentiable, then the *second-order partial derivatives* of f are defined by

$$D_{ij} f = D_i D_j f \qquad (i, j = 1, \ldots, n)$$

Theorems

Theorem 9.2 Let r be a positive integer. If a vector space X is spanned by a set of r vectors, then $\dim X \leqslant r$.

Corollary $\dim R^n = n$

Theorem 9.3 Suppose X is a vector space, and $\dim X = n$.

a) A set E of n vectors in X spans X if and only if E is independent.

b) X has a basis, and every basis consists of n vectors.

c) If $1 \leqslant r \leqslant n$ and $\{\mathbf{y}_1, \ldots \mathbf{y}_r\}$ is an independent set in X, then X has a basis containing $\{\mathbf{y}_1, \ldots \mathbf{y}_r\}$.

Theorem 9.5 A linear operator A on a finite-dimensional vector space X is one-to-one if and only if the range of A is all of X.

Theorem 9.7 a) If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $||A|| < \infty$ and A is a uniformly continuous mapping of $\mathbb{R}^n$ into $\mathbb{R}^m$.

b) If $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ and c is a scalar, then

$$||A + B|| \leqslant ||A|| + ||B||, \qquad ||cA|| = |c|||A||$$

With the distance between A and B defined as $||A - B||$, $L(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space.

c) If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^n, \mathbb{R}^k)$ then

$$||BA|| \leqslant ||B||||A||$$

Theorem 9.8 Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.

a) If $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and

$$||B - A|| \cdot ||A^{-1}|| < 1$$

then $B \in \Omega$

b) Ω is an open subset of $L(\mathbb{R}^n)$, and the mapping $A \to A^{-1}$ is continuous on Ω.

Theorem 9.12 Suppose E and $\mathbf{f}$ are as in the definition of differentiable, $\mathbf{x} \in E$, and the following holds with $A = A_1$ and $A = A_2$:

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

Then $A_1 = A_2$.

Theorem 9.15 (Chain Rule) Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, $\mathbf{f}$ is differentiable at $\mathbf{x}_0 \in E$, $\mathbf{g}$ maps an open set containing $\mathbf{f}(E)$ into $\mathbb{R}^k$, and $\mathbf{g}$ is differentiable at $\mathbf{f}(\mathbf{x}_0)$. Then the mapping $\mathbf{F}$ of E into $\mathbb{R}^k$ defined by

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

is differentiable at $\mathbf{x}_0$, and

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0).$$

Theorem 9.17 Suppose $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$, and $\mathbf{f}$ is differentiable at a point $\mathbf{x} \in E$. Then the partial derivatives $(D_j f_i)(\mathbf{x})$ exists, and

$$\mathbf{f}'(\mathbf{x})\mathbf{e}_j = \sum_{i=1}^{m}(D_j f_i)(\mathbf{x})\mathbf{u}_i \qquad (1 \leqslant j \leqslant n).$$

Theorem 9.19 Suppose $\mathbf{f}$ maps a convex open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, $\mathbf{f}$ is differentiable in E, and there is a real number M such that $||\mathbf{f}'(x)|| \leqslant M$ for every $\mathbf{x} \in E$. Then

$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| \leqslant M|\mathbf{b} - \mathbf{a}|$$

Corollary If, in addition $\mathbf{f}'(\mathbf{x}) = 0$ for all $\mathbf{x} \in E$, then $\mathbf{f}$ is constant.

Theorem 9.21 Suppose $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$. Then $\mathbf{f} \in \mathscr{C}'(E)$ if and only if the partial derivatives $D_j f_i$ exists and are continuous on E for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$.

Theorem 9.23 (Contraction Mapping Principle) If X is a complete metric space, and if φ is a contraction of X into X, then there exists one and only one $x \in X$ such that $\varphi(x) = x$.

Theorem 9.24 (Inverse Function Theorem) Suppose $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $\mathbf{f}'(\mathbf{a})$ is invertible for some $\mathbf{a} \in E$, and $\mathbf{b} = \mathbf{f}(\mathbf{a})$. Then

a) there exists open sets U and V in $\mathbb{R}^n$ such that $\mathbf{a} \in U$, $\mathbf{b} \in V$, $\mathbf{f}$ is one-to-one on U, and $\mathbf{f}(U) = V$;

b) if $\mathbf{g}$ is the inverse of $\mathbf{f}$, defined in V by

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x} \qquad (\mathbf{x} \in U),$$

then $\mathbf{g} \in \mathscr{C}'(V)$.

Theorem 9.25 If $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and if $\mathbf{f}'(x)$ is invertible for every $\mathbf{x} \in E$, then $\mathbf{f}(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$. In other words, $\mathbf{f}$ is an *open mapping* of E into $\mathbb{R}^n$

Theorem 9.27 If $A \in L(R^{m+n}, R^n)$ and if A_x is invertible, there there corresponds to every $\mathbf{k} \in \mathbb{R}^m$ a unique $\mathbf{h} \in R^n$ such that $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$. This $\mathbf{h}$ can be computed from $\mathbf{k}$ by the formula:

$$\mathbf{h} = -(A_x)^{-1}A_y\mathbf{k}$$

Theorem 9.28 (Implicit Function Theorem) Let $\mathbf{f}$ be a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^{n+m}$ into $\mathbb{R}^n$, such that $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ for some point $(\mathbf{a}, \mathbf{b}) \in E$.

Put $A = \mathbf{f}'(\mathbf{a}, \mathbf{b})$ and assume that A_x is invertible. (That is, the Jacobian, the determinant of the $n \times n$ matrix A_x, is nonzero.) Then there exists open sets $U \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$, with $(\mathbf{a}, \mathbf{b}) \in U$ and $\mathbf{b} \in W$m having the following property:

To every $\mathbf{y} \in W$ corresponds a unique $\mathbf{x}$ such that

$$(\mathbf{x}, \mathbf{y}) \in U \quad \text{and} \quad \mathbf{f}(\mathbf{x}, \mathbf{y}) = 0.$$

If this $\mathbf{x}$ is defined to be $\mathbf{g}(\mathbf{y})$, the $\mathbf{g}$ is a $\mathscr{C}'$-mapping of W into $\mathbb{R}^n$ $\mathbf{g}(\mathbf{b}) = \mathbf{a}$,

$$\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0} \qquad (y \in W),$$

and

$$\mathbf{g}'(\mathbf{b}) = -(A_x)^{-1} A_y$$

Theorem 9.32 Suppose m, n, r are nonnegative integers, $m \geqslant r$, $n \geqslant r$, $\mathbf{F}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and $\mathbf{F}'(\mathbf{x})$ has rank r for every $\mathbf{x} \in E$. Fix $\mathbf{a} \in E$, put $A = \mathbf{F}'(\mathbf{a})$, let Y_1 be the range of A, and let P be a projection in $\mathbb{R}^m$ whose range is Y_1. Let Y_2 be the null space of P. Then there are open sets U and V in $\mathbb{R}^n$ with $\mathbf{a} \in U$, $U \subset E$, and there is a 1-1 $\mathscr{C}'$-mapping $\mathbf{H}$ of V onto U (whose inverse is also of class $\mathscr{C}'$) such that

$$\mathbf{F}(\mathbf{H}(\mathbf{x})) = A\mathbf{x} + \varphi(A\mathbf{x}) \qquad (x \in V)$$

where φ is a $\mathscr{C}'$-mapping of the open set $A(V) \subset Y_1$ into Y_2.

Theorem 9.34 a) If I is the identity operator on $\mathbb{R}^n$, then

$$\det[I] = \det(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$$

b) det is a linear function of each of the column vectors $\mathbf{x}_j$, if the others are held fixed.

c) If $[A]_1$ is obtained from $[A]$ by interchanging two columns, then $\det[A_1] = -\det[A]$.

d) If $[A]$ has two equal columns, then $\det[A] = 0$.

Theorem 9.35 If $[A]$ and $[B]$ are $n \times n$ matrices then:

$$\det([B][A]) = \det([B]) \det[A].$$

Theorem 9.36 A linear operator A in $\mathbb{R}^n$ is invertible if and only if $\det[A] \neq 0$.

Theorem 9.40 Suppose f is defined in an open set $E \subset \mathbb{R}^2$, and $D_1 f$ and $D_{21} f$ exist at every point of E. Suppose $Q \subset E$ is a closed rectangle with sides parallel to the coordinate axes, having (a, b) and $(a + h, b + k)$ as opposite vertices. Put

$$\Delta(f, Q) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)$$

Then there is a point (x, y) in the interior of Q such that $\Delta(f, Q) = hk(D_2 1 f)(x, y)$.

Theorem 9.41 Suppose f is defined in an open set $E \subset \mathbb{R}^2$, and $D_1 f$. $D_{21} f$. and $D_2 f$ exist at every point of E, and $D_{21} f$ is continuous at some point $(a,b) \in E$. Then $D_{12} f$ exists at (a,b) and $(D_{12} f)(a,b) = (D_{21} f)(a,b)$

Corollary $D_{21} f = D_{12} f$ if $f \in \mathscr{C}''(E)$

Theorem 9.42 Suppose

a) $\varphi(x,t)$ is defined for $a \leqslant x \leqslant b$, $c \leqslant t \leqslant d$;

b) α is an increasing function on $[a,b]$;

c) $\phi^t \in \mathscr{R}(\alpha)$ for every $t \in [c,d]$;

d) $c < s < d$, and to every $\varepsilon > 0$ corresponds a $\delta > 0$ such that $|(D_2\varphi)(x,t) - (D_2\varphi)(x,s)| < \varepsilon$ for all $x \in [a,b]$ and for all $t \in (s-\delta, s_\delta)$.

Define

$$f(t) = \int_a^b \varphi(x,t)\, d\alpha \qquad (c \leqslant t \leqslant d).$$

Then $(D_2\varphi)^s \in \mathscr{R}(\alpha)$, $f'(s)$ exists, and

$$f'(s) = \int_a^b (D_2\varphi)(x,s)\, d\alpha$$