Intro Real Analysis

Rosenlicht

April 2022

Contents

1   Notions from Set Theory

1.1   Sets and Elements. Subsets

Set theory supplies the basic language for all that follows, so we fix the most common conventions. We will not attempt to "define" the word set; informally a set is simply a collection of objects, called its *elements*. Typical sets are the students enrolled at a particular university, the straight lines in the plane, or the real numbers. Elements of a set may themselves be sets: an element of the set of straight lines in the plane is the set of points on that line. We use

capital letters for sets, lower case letters for their elements, and the symbol $\in$ for membership. Thus

$$x \in S$$

means that x is a member of the set S, while $x \notin S$ abbreviates "x is not an element of S". A set is completely determined by its elements; explicitly $X = Y$ if and only if every x satisfies $x \in X \Leftrightarrow x \in Y$. The difference between a set and its elements is crucial: the set $\{x\}$ consists of a single element, and must not be confused with x itself.

The notation $\{x_1, x_2, x_3\}$ lists the indicated elements, while set-builder notation

$$\{x \in S : x \text{ satisfies a stated property}\}$$

collects all elements of S with the desired property. For example

$$\{x \in \mathbb{R} : 0 \le x \le 1\}, \qquad \{n : n \text{ is a positive integer}\}$$

are two subsets of the real numbers $\mathbb{R}$. The symbol $\varnothing$ denotes the empty set, the unique set with no elements.

Definition 1.1 (Subset). *If X and S are sets we write $X \subseteq S$ when every element of X is also an element of S; that is*

$$X \subseteq S \quad \Longleftrightarrow \quad \forall x \, (x \in X \Rightarrow x \in S).$$

If $X \subseteq S$ and $X \ne S$ we write $X \subsetneq S$ and call X a proper subset of S.

The relation $X \subseteq S$ is always true when $X = S$, and is vacuously true when $X = \varnothing$ because there are no elements that violate the implication. A frequent source of confusion is mixing up the symbols $\subseteq$ and $\in$; for example $\{2,3\} \subseteq \{1,2,3\}$, whereas $2 \in \{1,2,3\}$.

1.2   Operations on Sets

The standard operations on sets are intersection, union, and complement.

Definition 1.2 (Intersection and Union). *If X and Y are sets their intersection and union are*

$$X \cap Y = \{x : x \in X \text{ and } x \in Y\}, \qquad X \cup Y = \{x : x \in X \text{ or } x \in Y\}.$$

The word "or" is inclusive: x belongs to $X \cup Y$ when it lies in X, or in Y, or in both sets.

Definition 1.3 (Complement and Difference). *If X is a subset of a fixed "ambient" set S, the complement of X in S is*

$$S \setminus X = \{x \in S : x \notin X\},$$

and we often abbreviate this to X^c once S is understood. For any two sets we also write

$$X - Y = X \cap Y^c = \{x \in X : x \notin Y\}$$

and call this the set difference or relative complement of Y in X.

The relations among these operations mirror the familiar algebraic identities. For instance, if $X, Y \subseteq S$ then

$$X^c \cap Y^c = (X \cup Y)^c, \qquad X^c \cup Y^c = (X \cap Y)^c,$$

which are the De Morgan laws.

Sets X and Y are *disjoint* if $X \cap Y = \varnothing$; a family of sets is disjoint if every two distinct members are disjoint. Intersections and unions extend to any finite number of sets by iterating the preceding definitions, and to arbitrary families via index sets. If I is an indexing set and $\{X_i\}_{i \in I}$ is a family of subsets of S, then

$$\bigcap_{i \in I} X_i = \{x : x \in X_i \text{ for every } i \in I\}, \qquad \bigcup_{i \in I} X_i = \{x : x \in X_i \text{ for at least one } i \in I\}.$$

In particular, if each $X_i \subseteq S$, then $(\bigcap_{i \in I} X_i)^c = \bigcup_{i \in I} X_i^c$.

Definition 1.4 (Ordered Pair and Cartesian Product). *The ordered pair (a, b) records a as the first component and b as the second; $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$. One possible set-theoretic realization is due to Kuratowski:*

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

Given sets X and Y the Cartesian product

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

is the set of all ordered pairs whose first coordinate comes from X and second from Y.

1.3 Functions

Definition 1.5 (Function). *If X and Y are sets, a function f from X to Y is a rule that assigns to each $x \in X$ a definite element $f(x) \in Y$. We encode this by writing $f : X \to Y$, and call X the domain and Y the codomain of f. Two functions $f, g : X \to Y$ are equal when $f(x) = g(x)$ for every $x \in X$.*

Functions may be described by listing the pairs $(x, f(x))$, by a formula (for example $f(x) = x^2 + 3x - 2$ on $\mathbb{R}$), or geometrically via their graphs. Any subset of the plane meeting each vertical line in at most one point is the graph of a real-valued function on a subset of $\mathbb{R}$. The defining rule need not be easy to compute: the function that sends each real number to the digit in its billionth decimal place is a legitimate example.

The graph of $f : X \to Y$ is the subset $\{(x, f(x)) : x \in X\}$ of $X \times Y$. Conversely, any subset of $X \times Y$ that contains exactly one ordered pair with first coordinate x for each $x \in X$ determines a function $X \to Y$. In this sense the data of the domain, codomain, and graph together determine the function.

If $f : X \to Y$ and $X' \subseteq X$, the restriction $f|_{X'} : X' \to Y$ is defined by $f|_{X'}(x) = f(x)$. If $Y \subseteq Y'$ we may view f as a function into Y' without changing

the rule. Given $f : X \to Y$ and $g : Y \to Z$, the composition $g \circ f : X \to Z$ is given by $(g \circ f)(x) = g(f(x))$. The identity map on X is denoted $i_X : X \to X$, $i_X(x) = x$.

Definition 1.6 (Injective, Surjective, Bijective). *A function $f : X \to Y$ is injective (or one-to-one) if $f(x_1) = f(x_2)$ implies $x_1 = x_2$. It is surjective (onto) if every $y \in Y$ equals $f(x)$ for some $x \in X$. A bijection is both injective and surjective, and hence establishes a one-to-one correspondence between X and Y.*

If $f : X \to Y$ is bijective there exists an inverse function $f^{-1} : Y \to X$ satisfying $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$. For an arbitrary function f and subsets $X' \subseteq X$, $Y' \subseteq Y$,

$$f(X') = \{f(x) : x \in X'\}, \qquad f^{-1}(Y') = \{x \in X : f(x) \in Y'\}$$

are called the image of X' and the inverse image (or preimage) of Y', respectively. When f is bijective these two uses of f^{-1} are compatible: $f^{-1}(\{y\})$ is the singleton containing $f^{-1}(y)$.

1.4 Finite and Infinite Sets

Let $\mathbb{N} = \{1, 2, 3, \dots\}$ denote the positive integers.

Definition 1.7 (Finite Set). *A set X is finite if either $X = \varnothing$ or there exists $n \in \mathbb{N}$ together with a bijection between X and $\{1, 2, \dots, n\}$. In that case we write $|X| = n$ and call n the number of elements of X.*

Any subset of a finite set is finite; if the subset is proper, then it contains fewer elements than the ambient set. A set is *infinite* when it is not finite: intuitively we can keep choosing distinct elements without ever exhausting the set. The following characterization is often useful.

Theorem 1.1. *A set X is infinite if and only if it can be put into a one-to-one correspondence with a proper subset of itself.*

Proof. If X is finite, any proper subset has fewer elements, so no bijection can exist. Conversely, suppose X is infinite. Choose distinct elements $x_1, x_2, x_3, \dots$ in X, and let $Y = X \setminus \{x_1, x_2, x_3, \dots\}$. Define $f : X \to \{x_2, x_3, \dots\} \cup Y$ by $f(x_k) = x_{k+1}$ for $k \geq 1$ and $f(y) = y$ for $y \in Y$. This f is a bijection from X onto a proper subset of X. $\qquad\square$

The natural numbers also provide a convenient definition of sequences. An n-tuple of elements of a set X is a function from $\{1, 2, \dots, n\}$ to X; we usually write it as $(x_1, x_2, \dots, x_n)$. An infinite sequence (or simply a sequence) of elements of X is a function from $\mathbb{N}$ to X, often denoted $(x_1, x_2, x_3, \dots)$.

2 The Real Number System

2.1 The Field Properties

Analysis rests on the exact algebraic and order structure of the real numbers. We begin by recording the axioms they satisfy.

Definition 2.1 (Group). *A* group *is a nonempty set G equipped with a binary operation $\cdot$ such that:*

1. *$x \cdot y \in G$ for all $x, y \in G$ (closure).*

2. *$(x \cdot y) \cdot z = x \cdot (y \cdot z)$ for all $x, y, z \in G$ (associativity).*

3. *There exists $e \in G$ with $e \cdot x = x \cdot e = x$ for every $x \in G$ (identity).*

4. *Each $x \in G$ has $x^{-1} \in G$ satisfying $x \cdot x^{-1} = x^{-1} \cdot x = e$ (inverse).*

The group is Abelian *if additionally $x \cdot y = y \cdot x$ for all x, y.*

Definition 2.2 (Field). *A* field *is a set $\mathbb{F}$ together with two binary operations, addition and multiplication, such that*

1. *$(\mathbb{F}, +)$ is an Abelian group with identity $0_{\mathbb{F}}$.*

2. *$(\mathbb{F} \setminus \{0_{\mathbb{F}}\}, \cdot)$ is an Abelian group with identity $1_{\mathbb{F}}$.*

3. *Multiplication distributes over addition:*

$$x \cdot (y + z) = x \cdot y + x \cdot z, \qquad (x + y) \cdot z = x \cdot z + y \cdot z.$$

Definition 2.3 (Real Numbers). *The* real number system $\mathbb{R}$ *is the unique (up to isomorphism) complete ordered field. Practically we view $\mathbb{R}$ as a set endowed with addition, multiplication, and an order relation satisfying the axioms listed in this section.*

Besides $\mathbb{R}$, the rationals $\mathbb{Q}$ and complexes $\mathbb{C}$ are important fields. Finite examples also exist (for instance $\mathbb{Z}/p\mathbb{Z}$ when p is prime), but they lack the order and completeness required for analysis.

Field axioms. Rosenlicht enumerates seven explicit properties; we collect them in two groups.

(F1) **Commutativity.** $a + b = b + a$ and $ab = ba$.

(F2) **Associativity.** $(a + b) + c = a + (b + c)$ and $(ab)c = a(bc)$.

(F3) **Identities.** There exist 0 and 1 with $a + 0 = a$ and $a \cdot 1 = a$.

(F4) **Inverses.** Every a has $-a$ with $a + (-a) = 0$; each $a \neq 0$ has a^{-1} with $aa^{-1} = 1$.

(F5) **Distributivity.** $a(b+c) = ab + ac$ and $(a+b)c = ac + bc$.

These axioms imply all of the familiar algebraic rules. We spell out the basic consequences because they are used constantly.

Lemma 2.1 (Cancellation). *If $(G, \cdot)$ is a group and $xy = xz$, then $y = z$; likewise $yx = zx$ implies $y = z$.*

Proof. Multiply $xy = xz$ on the left by x^{-1} to obtain $y = z$. $\square$

Corollary 2.0.1. *The identity element of a group is unique, inverses are unique, and $(x^{-1})^{-1} = x$.*

Proposition 2.1 (Algebraic identities). *Let $(\mathbb{F}, +, \cdot)$ be a field. For any $x, y \in \mathbb{F}$ the following hold.*

1. *$0 \cdot x = x \cdot 0 = 0$ and $(-1)x = x(-1) = -x$.*

2. *$xy = 0$ implies $x = 0$ or $y = 0$.*

3. *$(-x)(-y) = xy$ and $-(x+y) = (-x) + (-y)$.*

4. *If $x, y \neq 0$ then $(xy)^{-1} = y^{-1}x^{-1}$.*

Proof. Each statement is a short application of the axioms and Lemma 2.1. For instance,
$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0,$$
so subtracting $x \cdot 0$ from both sides shows $x \cdot 0 = 0$. The other items are similar. $\square$

2.2 Order

To do analysis, the field must also be ordered.

Definition 2.4 (Ordered field). *An* ordered field *is a field endowed with a relation $<$ such that for all x, y, z:*

1. *Exactly one of $x < y$, $x = y$, or $x > y$ holds (*trichotomy*).*

2. *If $x < y$, then $x + z < y + z$.*

3. *If $0 < z$ and $x < y$, then $xz < yz$.*

We write $x \leq y$ when $x < y$ or $x = y$ and denote the set of positive elements by $\mathbb{F}_{>0} = \{x : x > 0\}$. The order interacts with addition and multiplication exactly as intuition dictates.

Lemma 2.2 (Order manipulations). *If $x > 0$ and $y > 0$, then $x + y > 0$ and $xy > 0$. If $x < y$ and $z < w$, then $x + z < y + w$. Multiplying by a negative element reverses inequalities.*

In an ordered field we define intervals, rays, and the absolute value $|x| = \max\{x, -x\}$. These capture the geometric idea of distance on the number line.

2.3   Completeness and Bounds

The property that distinguishes $\mathbb{R}$ from $\mathbb{Q}$ is completeness.

Definition 2.5 (Bounds). *Let $S \subseteq \mathbb{R}$. A number u is an upper bound if $x \leq u$ for all $x \in S$. If u is an upper bound and $u \leq v$ for every upper bound v, then u is the supremum $\sup S$. Lower bounds and the infimum $\inf S$ are defined similarly.*

Theorem 2.1 (Least Upper Bound Axiom). *Every nonempty subset of $\mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$; every nonempty subset bounded below has an infimum.*

Theorem 2.2 (Monotone convergence). *If $\{a_n\}$ is increasing and bounded above, then $a_n \to \sup\{a_n\}$.*

Proof. Let $s = \sup\{a_n\}$. Given $\varepsilon > 0$, $s - \varepsilon$ is not an upper bound, so some N satisfies $a_N > s - \varepsilon$. For $n \geq N$ we have $s - \varepsilon < a_n \leq s$, proving $a_n \to s$. $\square$

Theorem 2.3 (Archimedean property). *For every $x > 0$ there exists $n \in \mathbb{N}$ with $n > x$.*

Proof. Suppose no such n exists. Then the set $\mathcal{A} = \{n : n \in \mathbb{N}\}$ is bounded above by x, so $s = \sup \mathcal{A}$ exists. But $s - 1$ is not an upper bound, so some $m \in \mathbb{N}$ satisfies $m > s - 1$, i.e., $m + 1 > s$. Since $m + 1 \in \mathbb{N}$, we contradict the definition of s. $\square$

Theorem 2.4 (Density of $\mathbb{Q}$). *If $a < b$ in $\mathbb{R}$, then there exists $q \in \mathbb{Q}$ with $a < q < b$.*

Proof. Consider $(b - a) > 0$. By the Archimedean property choose n with $n(b - a) > 1$. There exists an integer m such that $na < m < nb$, and then $m/n \in \mathbb{Q}$ satisfies $a < m/n < b$. $\square$

2.4   The Existence of Square Roots

Completeness quickly yields the fundamental fact that positive numbers possess square roots.

Theorem 2.5. *For every $a > 0$ there exists a unique $b > 0$ such that $b^2 = a$.*

Proof. Consider the set $S = \{x \in \mathbb{R} : x \geq 0 \text{ and } x^2 \leq a\}$. It is nonempty (since $0 \in S$) and bounded above (for example by $a + 1$). Let $b = \sup S$. We first show $b^2 \leq a$. Otherwise $b^2 > a$ and putting $\varepsilon = b^2 - a > 0$ we may take $\eta = \varepsilon/(2b)$; note that $b > 0$ because $a > 0$. Then $(b - \eta)^2 = b^2 - 2b\eta + \eta^2 = b^2 - \varepsilon + \eta^2 > a$, so every $x \in S$ satisfies $x < b - \eta$. That means $b - \eta$ is an upper bound smaller than b, contradicting the definition of b. Hence $b^2 \leq a$.

If $b^2 < a$, set $\varepsilon = a - b^2 > 0$ and choose $\eta = \min\{1, \varepsilon/(2b + 1)\}$. Then

$$(b + \eta)^2 = b^2 + 2b\eta + \eta^2 \leq b^2 + \eta(2b + 1) \leq b^2 + \varepsilon = a,$$

so $b + \eta \in S$. This contradicts b being an upper bound. Therefore $b^2 = a$.

Uniqueness follows from the order axioms: if $c > 0$ also satisfies $c^2 = a$, then $(b - c)(b + c) = 0$. Proposition 2.1 implies either $b = c$ or $b = -c$. The latter is impossible because $b, c > 0$, so $b = c$. $\qquad\square$

The existence of square roots supplies the positive branch of the absolute value and, together with the field axioms, completes the algebraic description of $\mathbb{R}$.

3 Metric Spaces

3.1 Definition and Examples

Metric spaces generalize familiar geometric notions of distance.

Definition 3.1 (Metric Space). *A metric space is a pair (M, d) consisting of a set M and a function $d \colon M \times M \to \mathbb{R}$ satisfying for all $x, y, z \in M$:*

1. *$d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$.*

2. *$d(x, y) = d(y, x)$.*

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

Lemma 3.1. *Nonnegativity is automatic: $d(x, y) = \frac{1}{2}(d(x, y) + d(y, x)) \geq \frac{1}{2} d(x, x) = 0$.*

Theorem 3.1 (General Triangle Inequality). *Given points $x_1, \ldots, x_n$ in (M, d),*

$$d(x_1, x_n) \leq \sum_{k=1}^{n-1} d(x_k, x_{k+1}).$$

Corollary 3.1.1 (Reverse Triangle Inequality). *For all $x, y, z \in M$, $|d(x, z) - d(z, y)| \leq d(x, y)$.*

These inequalities turn the metric into a powerful bookkeeping device for estimates.

Examples.

1. On $\mathbb{R}^n$ the Euclidean metric $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ arises from the dot product. The Cauchy–Schwarz inequality

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$$

   implies $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, which provides the triangle inequality for d.

2. The *taxicab* (or ℓ^1) metric on $\mathbb{R}^n$ is $d_1(\mathbf{x}, \mathbf{y}) = \sum |x_i - y_i|$; each coordinate satisfies the one-dimensional triangle inequality, so the sum does as well.

3. The *discrete* metric on an arbitrary set M is $d(x, y) = 0$ when $x = y$ and $d(x, y) = 1$ otherwise.

Subspaces. If $E \subseteq M$, the restriction of d to $E \times E$ makes E into a metric space; this is the *subspace metric*.

3.2 Open Sets, Closed Sets, and Neighborhoods

For $p \in M$ and $r > 0$ define the open ball $B_r(p) = \{x : d(x,p) < r\}$ and the closed ball $\overline{B}_r(p) = \{x : d(x,p) \leq r\}$.

Definition 3.2. *A subset $U \subseteq M$ is* open *if for every $x \in U$ there exists $r > 0$ with $B_r(x) \subseteq U$. A subset $F \subseteq M$ is* closed *if its complement is open. The empty set and M are both open and closed.*

Open sets collect points together with a "cushion" around them. Closed sets are precisely those that contain all limit points of sequences within them (see Proposition 3.2).

Proposition 3.1. *Open balls are open, closed balls are closed, and finite intersections of open (resp. closed) balls are open (resp. closed).*

Proof. If $x \in B_r(p)$ set $\rho = r - d(x,p) > 0$. Whenever $y \in B_\rho(x)$, the triangle inequality gives $d(y,p) \leq d(y,x) + d(x,p) < \rho + d(x,p) = r$, so $y \in B_r(p)$. The closed case follows by considering complements. $\square$

Theorem 3.2. *Let $\{U_\alpha\}_{\alpha \in A}$ be open subsets of M and $F_1, \ldots, F_n$ closed subsets.*

1. $\bigcup_{\alpha \in A} U_\alpha$ *is open.*

2. $\bigcap_{k=1}^{n} U_k$ *is open for any finite subcollection.*

3. $\bigcap_{\alpha \in A} F_\alpha$ *is closed and $\bigcup_{k=1}^{n} F_k$ is closed.*

Limit points, interior, and closure.

Definition 3.3. *A point x is a* limit point *of $E \subseteq M$ if every ball about x intersects $E \setminus \{x\}$. The set of all limit points is denoted E'. The closure $\overline{E} = E \cup E'$ is the smallest closed set containing E, and the* interior $\mathrm{int}(E)$ *is the largest open set contained in E.*

Proposition 3.2. *E is closed if and only if it contains its limit points; equivalently $E = \overline{E}$.*

Proof. If E is closed and $x \in E'$ then every ball around x meets E, hence x cannot lie in the open complement. Conversely, if E contains its limit points, E^c is open: given $x \in E^c$, there exists $r > 0$ with $B_r(x) \cap E = \varnothing$, so $B_r(x) \subseteq E^c$. $\square$

Proposition 3.3. *In $\mathbb{R}$ the open balls are intervals (a,b), while closed balls are $[a,b]$. In $\mathbb{R}^n$ rectangular boxes $\prod_{i=1}^{n}(a_i,b_i)$ are open, their closures $\prod[a_i,b_i]$ are closed, and both are bounded.*

3.3 Boundedness and Extremal Points

Definition 3.4 (Bounded Set). *A subset S is* bounded *if $S \subseteq \overline{B}_r(p)$ for some p and r.*

Balls witness boundedness, so finite unions of bounded sets are bounded. In $\mathbb{R}^n$ closed rectangles are bounded because each coordinate difference is bounded.

Theorem 3.3. *A nonempty closed subset of $\mathbb{R}$ that is bounded above has a maximum; one bounded below has a minimum.*

3.4 Convergent Sequences

Definition 3.5 (Convergence). *A sequence $\{p_n\}$* converges *to p if for every $\varepsilon > 0$ there exists N with $d(p_n, p) < \varepsilon$ for $n \geq N$.*

Limits are unique. A sequence converges to p if and only if every subsequence has a further subsequence converging to p. Open and closed sets can be characterized via sequences:

Proposition 3.4.   *1. U is open if and only if no sequence entirely outside U converges to a point of U.*

2. *F is closed if and only if whenever $p_n \in F$ and $p_n \to p$, then $p \in F$.*

3.5 Completeness

Definition 3.6 (Cauchy Sequence). *A sequence $\{p_n\}$ is* Cauchy *if for every $\varepsilon > 0$ there exists N with $d(p_n, p_m) < \varepsilon$ for all $m, n \geq N$. A metric space is* complete *if every Cauchy sequence converges.*

Proposition 3.5. *$\mathbb{R}^n$ with the Euclidean metric is complete.*

Proof. Each coordinate of a Cauchy sequence is Cauchy in $\mathbb{R}$, hence convergent. The coordinatewise limit lies in $\mathbb{R}^n$ and is the limit of the vector sequence. $\square$

Theorem 3.4 (Completion criterion). *A subspace $E \subseteq M$ is complete if and only if it is closed in M (when M itself is complete).*

3.6 Compactness

Definition 3.7 (Compactness). *A subset $K \subseteq M$ is* compact *if every open cover $\{U_\alpha\}$ of K admits a finite subcover. Equivalently, K is compact if every sequence in K has a convergent subsequence with limit in K (sequential compactness).*

Theorem 3.5 (Heine–Borel). *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proposition 3.6. *Closed subsets of compact sets are compact, and continuous images of compact sets are compact.*

3.7 Connectedness

Definition 3.8 (Connected Set)**.** *A metric space M is* connected *if it cannot be written as the union of two disjoint nonempty open sets. A subset $E \subseteq M$ is* connected *if it is connected with the subspace metric.*

Proposition 3.7. *Intervals in $\mathbb{R}$ are connected, and any connected subset of $\mathbb{R}$ is an interval. Products of connected sets are connected. Continuous images of connected sets are connected.*

Connectedness ensures that intermediate values cannot be skipped; this feeds directly into the Intermediate Value Theorem in the next chapter.

4 Continuous Functions

Continuous functions will carry the weight of the topological results developed in Chapter 3. Rosenlicht treats functions between arbitrary metric spaces; the real- and vector-valued cases fall out as specializations.

4.1 Definition of Continuity. Examples

Let (E, d) and (E', d') be metric spaces.

Definition 4.1 (Continuity at a point)**.** *A function $f \colon E \to E'$ is* continuous *at $p_0 \in E$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d(p, p_0) < \delta$ implies $d'(f(p), f(p_0)) < \varepsilon$. It is* continuous *if it is continuous at every point.*

The definition admits two equivalent reformulations.

Proposition 4.1 (Neighborhood formulation)**.** *The following are equivalent for $f \colon E \to E'$ and $p_0 \in E$.*

1. *f is continuous at p_0.*

2. *For every open set $U \subseteq E'$ with $f(p_0) \in U$, there exists an open set $V \subseteq E$ containing p_0 such that $f(V) \subseteq U$.*

3. *For every sequence $p_n \to p_0$ in E, we have $f(p_n) \to f(p_0)$ in E'.*

Corollary 4.0.1. *f is continuous if and only if the preimage of every open set is open (equivalently, the preimage of every closed set is closed).*

These criteria prove continuity for the functions encountered in calculus: polynomials are continuous everywhere, rational functions are continuous on their domain, absolute value and maximum/minimum are continuous combinations of basic operations, and so on.

4.2 Continuity and Limits

Continuity aligns perfectly with function limits.

Theorem 4.1 (Sequential characterization)**.** *If $f\colon E \to E'$ and $p \in E$, then f is continuous at p if and only if for every sequence $p_n \to p$ we have $f(p_n) \to f(p)$.*

Theorem 4.2 (Composition and restriction)**.** *If $f\colon E \to E'$ is continuous at $p \in E$ and $g\colon E' \to E''$ is continuous at $f(p)$, then $g \circ f$ is continuous at p. If $A \subseteq E$ carries the subspace metric and f is continuous on E, then the restriction $f|_A$ is continuous on A.*

Proposition 4.2 (Gluing lemma)**.** *If $E = F \cup G$ with F, G closed subsets and $f\colon E \to E'$ agrees with continuous maps on F and on G, then f is continuous.*

Continuity also respects monotone bijections between intervals: if $U, V \subseteq \mathbb{R}$ are intervals and $f\colon U \to V$ is strictly increasing and onto, then both f and f^{-1} are continuous.

4.3 Rational Operations and Vector-Valued Functions

The algebra of continuous real-valued functions mimics the algebra of real numbers.

Theorem 4.3 (Algebra of continuous functions)**.** *Let $f, g\colon E \to \mathbb{R}$ be continuous and let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be continuous. Then the functions $f + g$, $f - g$, fg, and $\varphi \circ f$ are continuous. If $g(p) \neq 0$ for all p, then f/g is continuous.*

Corollary 4.3.1. *Polynomials and rational functions (on the domain where the denominator is nonzero) are continuous. If $\mathbf{f}\colon E \to \mathbb{R}^n$ is given by $\mathbf{f}(p) = (f_1(p), \ldots, f_n(p))$, then $\mathbf{f}$ is continuous if and only if each coordinate function f_k is continuous.*

4.4 Continuous Functions on Compact Metric Spaces

Compactness translates topological information into quantitative statements.

Theorem 4.4 (Continuous images of compact sets)**.** *If $K \subseteq E$ is compact and $f\colon E \to E'$ is continuous, then $f(K) \subseteq E'$ is compact.*

Specializing to real-valued functions produces the classical extreme value and boundedness results.

Corollary 4.4.1. *If K is compact and $f\colon K \to \mathbb{R}$ is continuous, then f is bounded and attains both its maximum and minimum on K.*

Theorem 4.5 (Heine–Cantor)**.** *Continuous functions on compact metric spaces are uniformly continuous.*

Proof. Suppose f is not uniformly continuous. Then there exists $\varepsilon_0 > 0$ and points $p_n, q_n \in K$ with $d(p_n, q_n) \to 0$ but $|f(p_n) - f(q_n)| \geq \varepsilon_0$. Compactness provides a convergent subsequence $p_{n_k} \to p$; the paired points q_{n_k} also converge to p. Continuity contradicts the fixed gap ε_0. $\qquad\square$

Uniform continuity guarantees that Cauchy sequences of function values behave well. In particular, the sup metric

$$d_\infty(f, g) = \sup_{p \in K} d'(f(p), g(p))$$

turns $C(K, E')$ (continuous maps $K \to E'$) into a metric space. When K is compact and E' is complete, $C(K, E')$ is complete under d_∞: a uniformly Cauchy sequence of continuous functions converges uniformly to a continuous function.

4.5   Continuous Functions on Connected Metric Spaces

Connectedness rules out "jumps".

Theorem 4.6 (Image of a connected set). *If E is connected and $f\colon E \to E'$ is continuous, then $f(E)$ is connected.*

In $\mathbb{R}$ the connected subsets are intervals; applying Theorem 4.6 yields the Intermediate Value Theorem.

Corollary 4.6.1 (Intermediate Value Theorem). *If $f\colon [a, b] \to \mathbb{R}$ is continuous and y lies between $f(a)$ and $f(b)$, then there exists $c \in [a, b]$ with $f(c) = y$.*

Further consequences include: a continuous function on an interval that never changes sign is either strictly positive, strictly negative, or identically zero; a continuous bijection from a compact interval to an interval is monotone with a continuous inverse.

4.6   Sequences of Functions

Finally we examine sequences of continuous functions. Let (f_n) be functions $E \to E'$.

Definition 4.2 (Uniform convergence). *f_n converges uniformly to f if for every $\varepsilon > 0$ there exists N with $d'(f_n(p), f(p)) < \varepsilon$ for all $p \in E$ whenever $n \geq N$.*

Theorem 4.7 (Uniform limits preserve continuity). *If each f_n is continuous and $f_n \to f$ uniformly on E, then f is continuous.*

Proof. Given $\varepsilon > 0$, choose N such that $d'(f_n(p), f(p)) < \varepsilon/3$ for all p and $n \geq N$. Continuity of f_N provides $\delta > 0$ so that $d(p, q) < \delta$ implies $d'(f_N(p), f_N(q)) < \varepsilon/3$. For $d(p, q) < \delta$,

$$d'(f(p), f(q)) \leq d'(f(p), f_N(p)) + d'(f_N(p), f_N(q)) + d'(f_N(q), f(q)) < \varepsilon,$$

so f is continuous. $\qquad\square$

Uniform convergence is naturally encoded by the sup metric on $C(K, E')$ discussed earlier: $f_n \to f$ uniformly if and only if $d_\infty(f_n, f) \to 0$. When K is compact and E' complete, $C(K, E')$ is itself complete under d_∞.

Examples include Fourier-type partial sums converging uniformly on compact subsets, or piecewise linear approximations to continuous paths. This perspective will be revisited in the later chapters on series and approximations.

5   Differentiation

Differentiation returns us to single-variable calculus, but armed with the metric machinery from the previous chapters we can present the proofs succinctly.

5.1   Definition of the Derivative

Let $U \subseteq \mathbb{R}$ be open and $f \colon U \to \mathbb{R}$.

Definition 5.1 (Derivative)**.** *f is differentiable at $x_0 \in U$ if the limit*

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

exists. The function f' defined on the set of points where the limit exists is the derivative of f.

Equivalently, differentiability means there exists a linear function $L(x) = f(x_0) + f'(x_0)(x - x_0)$ such that

$$\lim_{x \to x_0} \frac{f(x) - L(x)}{x - x_0} = 0.$$

The graph of f near x_0 therefore hugs its tangent line.

Proposition 5.1. *If f is differentiable at x_0, then f is continuous at x_0.*

Proof. Using the definition with ε–δ notation,

$$|f(x) - f(x_0)| \leq |x - x_0| \left| \frac{f(x) - f(x_0)}{x - x_0} \right|,$$

and the quotient tends to $|f'(x_0)|$, so $f(x) \to f(x_0)$. $\qquad\square$

5.2   Rules of Differentiation

The derivative behaves linearly and obeys the product and chain rules familiar from calculus.

Theorem 5.1 (Algebraic rules)**.** *If f, g are differentiable at x_0 and $c \in \mathbb{R}$, then so are $f + g$, cf, and fg, with*

$$(f+g)'(x_0) = f'(x_0) + g'(x_0), \qquad (cf)'(x_0) = cf'(x_0), \qquad (fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

If $g(x_0) \neq 0$, then $\left(\frac{f}{g} \right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$.

Theorem 5.2 (Chain rule). *If f is differentiable at x_0 and g is differentiable at $f(x_0)$, then $g \circ f$ is differentiable at x_0 with*

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

Using these rules we recover the standard derivatives: polynomials differentiate via the binomial theorem, rational functions differentiate wherever they are defined, and so on.

5.3 The Mean Value Theorem

We now reach the central estimate for differentiable functions.

Theorem 5.3 (Rolle). *If $f\colon [a,b] \to \mathbb{R}$ is continuous, differentiable on (a,b), and satisfies $f(a) = f(b)$, then there exists $c \in (a,b)$ with $f'(c) = 0$.*

Proof. $[a,b]$ is compact, so f attains a maximum and minimum. If either extremum occurs at an interior point c, then near c the function cannot increase or decrease, forcing $f'(c) = 0$. Otherwise f is constant. $\qquad\square$

Theorem 5.4 (Mean Value Theorem). *If $f\colon [a,b] \to \mathbb{R}$ is continuous and differentiable on (a,b), then there exists $c \in (a,b)$ such that*

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

Proof. Apply Rolle's theorem to $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$ after noting that $g(a) = g(b)$. $\qquad\square$

Corollary 5.4.1. *If $f'(x) = 0$ for all x in an interval I, then f is constant on I.*

Corollary 5.4.2. *If $f'(x) \geq 0$ (respectively > 0) on an interval, then f is nondecreasing (respectively strictly increasing).*

Corollary 5.4.3 (Cauchy mean value theorem). *If f, g are continuous on $[a,b]$ and differentiable on (a,b) with $g'(x) \neq 0$, then there exists c such that*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

The mean value theorem underlies standard inequalities, e.g., the Lipschitz bound $|f(b) - f(a)| \leq M|b - a|$ when $|f'| \leq M$ on the interval, and L'Hospital's rule for $0/0$ forms (derived using Cauchy's version).

5.4 Taylor's Theorem

Repeated differentiation grants finer control via Taylor polynomials.

Definition 5.2 (Higher derivatives). *If $f^{(n-1)}$ exists on an open interval, we say f is n times differentiable when $f^{(n-1)}$ is differentiable; the new derivative is $f^{(n)}$.*

Theorem 5.5 (Taylor's formula with remainder). *Suppose f has $n+1$ continuous derivatives on $[a,b]$. For any $x \in [a,b]$ there exists ξ between x and a such that*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Proof. Define P_n to be the nth-degree Taylor polynomial at a. Consider

$$g(t) = f(t) - P_n(t) - \frac{f^{(n+1)}(x)}{(n+1)!}(t-a)^{n+1}.$$

Rolle's theorem applied repeatedly to appropriate combinations of g shows that $g^{(n+1)}$ vanishes somewhere between a and x, yielding the claimed ξ. $\square$

Corollary 5.5.1 (Estimates). *If $|f^{(n+1)}(t)| \leq M$ on $[a,b]$, then*

$$\left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leq \frac{M}{(n+1)!}|x-a|^{n+1}.$$

Taylor's theorem recovers power series expansions for elementary functions and provides precise error estimates for approximations.

6 Riemann Integration

Riemann integration assigns areas to graphs of real-valued functions on intervals and underpins the rest of analysis.

6.1 Definition and Examples

Let $[a,b]$ be a closed interval.

Definition 6.1 (Partition). *A partition of $[a,b]$ is a finite sequence $P = \{x_0, \ldots, x_N\}$ with $a = x_0 < x_1 < \cdots < x_N = b$. Its mesh is $|P| = \max_{1 \leq i \leq N}(x_i - x_{i-1})$.*

Definition 6.2 (Riemann sum). *If $f : [a,b] \to \mathbb{R}$ and P is a partition, a Riemann sum of f with respect to P is*

$$S(P,f) = \sum_{i=1}^{N} f(\xi_i)(x_i - x_{i-1}),$$

where each $\xi_i \in [x_{i-1}, x_i]$.

Definition 6.3 (Riemann integrability). *f is Riemann integrable on $[a, b]$ if there exists $A \in \mathbb{R}$ such that for every $\varepsilon > 0$ there is $\delta > 0$ with $|S(P, f) - A| < \varepsilon$ whenever $|P| < \delta$. The number A is the Riemann integral $\int_a^b f(x)\,dx$.*

Upper and lower Darboux sums provide a more practical criterion. For a partition P let

$$U(P, f) = \sum_{i=1}^{N} \left(\sup_{x \in [x_{i-1}, x_i]} f(x) \right) (x_i - x_{i-1}), \qquad L(P, f) = \sum_{i=1}^{N} \left(\inf_{x \in [x_{i-1}, x_i]} f(x) \right) (x_i - x_{i-1}).$$

Then f is integrable if and only if $\sup_P L(P, f) = \inf_P U(P, f)$; the common value equals the integral.

Continuous functions on $[a, b]$, monotone functions, and piecewise continuous functions are all integrable. Functions with jump discontinuities at finitely many points are integrable; the sign function is not integrable at a removable discontinuity unless we restrict the interval.

6.2 Linearity and Order Properties

Theorem 6.1 (Linearity). *If f, g are integrable on $[a, b]$ and $c \in \mathbb{R}$, then cf and $f + g$ are integrable, and*

$$\int_a^b cf(x)\,dx = c \int_a^b f(x)\,dx, \qquad \int_a^b (f + g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

Theorem 6.2 (Order properties). *If $f \le g$ on $[a, b]$, then $\int_a^b f \le \int_a^b g$. Moreover*

$$\left| \int_a^b f(x)\,dx \right| \le \int_a^b |f(x)|\,dx \le (b - a) \sup_{x \in [a, b]} |f(x)|.$$

Theorem 6.3 (Additivity). *If $a < c < b$ and f is integrable on $[a, b]$, then f is integrable on $[a, c]$ and $[c, b]$ and*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

6.3 Existence Criteria

To show f is integrable it suffices to show the upper and lower sums can be made arbitrarily close.

Proposition 6.1. *If f is bounded on $[a, b]$ and for every $\varepsilon > 0$ there exists $\delta > 0$ with $U(P, f) - L(P, f) < \varepsilon$ whenever $|P| < \delta$, then f is integrable.*

Continuity ensures this condition: given $\varepsilon > 0$, uniform continuity of f provides δ so that each subinterval of P with mesh $< \delta$ has oscillation $< \varepsilon/(b-a)$, forcing $U(P, f) - L(P, f) < \varepsilon$.

Monotone functions are integrable because the oscillation on each subinterval is at most the difference of endpoint values, which can be controlled by making the mesh small.

6.4 The Fundamental Theorem of Calculus

Integration and differentiation are inverse processes for continuous functions.

Theorem 6.4 (First Fundamental Theorem). *If f is integrable on $[a,b]$ and we define $F(x) = \int_a^x f(t)\, dt$, then F is continuous on $[a,b]$ and differentiable at every point where f is continuous; moreover $F'(x) = f(x)$ at such points.*

Proof. Continuity of F follows from the inequality $|F(x) - F(y)| \leq \int_y^x |f| \leq M|x - y|$ for $M = \sup |f|$. Differentiability uses the Mean Value Theorem and the fact that f is continuous at the point in question. □

Theorem 6.5 (Second Fundamental Theorem). *If f is continuous on $[a,b]$ and F is any antiderivative of f, then*

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

Proof. By Theorem 6.4, $F(x) = \int_a^x f(t)\, dt$ is an antiderivative. Any other antiderivative differs by a constant, so evaluating at $x = b$ gives the desired equality. □

The Fundamental Theorem reduces many definite integrals to antiderivatives and shows that continuous functions possess antiderivatives.

6.5 Logarithms and Exponentials

As an application, we construct the logarithm and exponential functions via integrals.

Definition 6.4. *For $x > 0$ define*

$$\log x = \int_1^x \frac{1}{t}\, dt.$$

Then $\log$ is continuous, strictly increasing, and differentiable with $\log'(x) = 1/x$. Set $\exp$ to be the inverse function: $\exp(\log x) = x$. Extend $\exp$ to all real numbers via $\exp(x) = \sum_{n=0}^{\infty} x^n/n!$ if desired; it satisfies $\exp'(x) = \exp(x)$ and $\exp(\log x) = x$.

Proposition 6.2. *The function $\log$ satisfies $\log(xy) = \log x + \log y$, $\log(1) = 0$, and $\log(x^r) = r \log x$ for rational r. Exponentiation obeys $\exp(x + y) = \exp(x)\exp(y)$, $\exp(0) = 1$, and $\exp$ is the unique solution of $y' = y$, $y(0) = 1$.*

These definitions agree with the familiar power-series constructions and yield the usual properties rigorously from the Riemann integral.

7 Interchange of Limit Operations

Limit processes rarely occur in isolation. We often encounter sequences of functions for which we wish to integrate, differentiate, or otherwise take limits term by term. This chapter records the situations where interchanging these operations is justified.

7.1 Integration and Differentiation of Sequences of Functions

Let $\{f_n\}$ be functions on a closed interval $[a, b]$.

Theorem 7.1 (Integration of uniformly convergent sequences)**.** *If each f_n is Riemann integrable on $[a, b]$ and $f_n \to f$ uniformly, then f is integrable and*

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

Proof. Uniform convergence implies the difference $|f_n - f|$ can be made uniformly small, hence the difference of integrals is bounded by $(b-a)\sup|f_n - f|$, which tends to zero. $\qquad\square$

Theorem 7.2 (Term-by-term differentiation)**.** *Suppose each f_n is differentiable on $[a, b]$, that $(f_n(x_0))$ converges for some $x_0 \in [a, b]$, and that f_n' converges uniformly on $[a, b]$ to a function g. Then f_n converges uniformly to a differentiable function f, and $f' = g$.*

Proof. Integrating f_n' from x_0 to x shows $f_n(x) = f_n(x_0) + \int_{x_0}^x f_n'(t)\, dt$. Uniform convergence of the derivatives implies uniform convergence of the integrals, hence the f_n converge uniformly to $f(x) = \lim f_n(x_0) + \int_{x_0}^x g(t)\, dt$, whose derivative is g. $\qquad\square$

7.2 Infinite Series

Series are sequences written additively. Let $\sum_{n=1}^{\infty} a_n$ denote the series with partial sums $s_n = \sum_{k=1}^n a_k$.

Definition 7.1. *The series $\sum a_n$ converges if the sequence (s_n) converges. It converges absolutely if $\sum |a_n|$ converges.*

Theorem 7.3 (Comparison tests)**.** *If $|a_n| \le b_n$ and $\sum b_n$ converges, then $\sum a_n$ converges absolutely. If $a_n \ge b_n \ge 0$ and $\sum b_n$ diverges, then so does $\sum a_n$.*

Theorem 7.4 (Alternating series test)**.** *If (a_n) is decreasing to 0 and the terms alternate in sign, then $\sum (-1)^{n-1} a_n$ converges, and the truncation error is at most a_{n+1} in magnitude.*

Definition 7.2. *A series of functions $\sum f_n$ converges* uniformly *on a set E if its sequence of partial sums converges uniformly. The Weierstrass M-test states that if $|f_n(x)| \leq M_n$ on E and $\sum M_n$ converges, then $\sum f_n$ converges uniformly.*

Uniform convergence allows termwise integration and differentiation via Theorems 7.1 and 7.2.

7.3 Power Series

Power series provide a major class of function series.

Definition 7.3. *A power series centered at a is $\sum_{n=0}^{\infty} c_n(x-a)^n$. The radius of convergence R is the unique extended real number such that the series converges absolutely for $|x-a| < R$ and diverges for $|x-a| > R$.*

Theorem 7.5 (Cauchy–Hadamard formula). *The radius of convergence satisfies $1/R = \limsup_{n\to\infty} |c_n|^{1/n}$, with the conventions $R = \infty$ if the limsup is 0 and $R = 0$ if the limsup is ∞.*

Theorem 7.6 (Termwise operations). *Inside the disk $|x-a| < R$ a power series may be differentiated and integrated term by term:*

$$\left(\sum_{n=0}^{\infty} c_n(x-a)^n\right)' = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1}, \qquad \int \sum_{n=0}^{\infty} c_n(x-a)^n dx = C + \sum_{n=0}^{\infty} \frac{c_n}{n+1}(x-a)^{n+1}.$$

Both series share the same radius of convergence.

7.4 The Trigonometric Functions

The sine and cosine functions arise naturally from power series.

Definition 7.4. *Define*

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

These series converge for all $x \in \mathbb{R}$.

Proposition 7.1. *The functions $\sin$ and $\cos$ are entire (analytic everywhere) and satisfy*

$$(\sin x)' = \cos x, \qquad (\cos x)' = -\sin x, \qquad \sin^2 x + \cos^2 x = 1.$$

Theorem 7.7 (Addition formulas). *For all $x, y \in \mathbb{R}$,*

$$\sin(x+y) = \sin x \cos y + \cos x \sin y, \qquad \cos(x+y) = \cos x \cos y - \sin x \sin y.$$

Proof. Expand the power series for $\sin(x+y)$ and $\cos(x+y)$ and rearrange terms; absolute convergence justifies the algebra. $\square$

7.5 Differentiation under the Integral Sign

Consider a function $F(t,x)$ defined on $[a,b] \times I$, where $I \subseteq \mathbb{R}$ is an interval, and define

$$\Phi(x) = \int_a^b F(t,x)\,dt.$$

Theorem 7.8 (Leibniz rule)**.** *If $F(t,x)$ and $\partial F/\partial x$ are continuous on $[a,b] \times I$, then Φ is differentiable on I with*

$$\Phi'(x) = \int_a^b \frac{\partial F}{\partial x}(t,x)\,dt.$$

Proof. Apply the mean value theorem in the x-variable, uniformly in t, and use Theorem 7.1. $\qquad\square$

This formula is frequently quoted as "differentiating under the integral sign." It allows parameter-dependent integrals to be handled by calculus on the parameter.

8 The Method of Successive Approximations

Iterative methods can solve a variety of problems—root finding, implicit equations, and differential equations. Rosenlicht distills the essential fixed-point principle and shows how it applies widely.

8.1 The Fixed Point Theorem

Theorem 8.1 (Banach fixed point theorem)**.** *Let (E,d) be a metric space and $F\colon E \to E$ satisfy $d(F(p),F(q)) \leq \lambda d(p,q)$ for some $\lambda < 1$. Then F has a unique fixed point p^* (i.e., $F(p^*) = p^*$), and for any starting point p_0 the iterates $p_{n+1} = F(p_n)$ converge to p^*.*

Proof. The contraction condition ensures $d(p_{n+1},p_n) \leq \lambda^n d(p_1,p_0)$, so (p_n) is Cauchy and hence convergent if E is complete. Passing to the limit in $p_{n+1} = F(p_n)$ yields $F(p^*) = p^*$. Uniqueness follows by applying the contraction inequality to two fixed points. $\qquad\square$

8.2 Simplest Case of the Implicit Function Theorem

Consider an equation $F(x,y) = 0$ with F continuous and continuously differentiable near (x_0,y_0).

Theorem 8.2 (Implicit function, one variable)**.** *If $\frac{\partial F}{\partial y}(x_0,y_0) \neq 0$, then there exists a neighborhood of x_0 on which the equation $F(x,y) = 0$ defines a unique differentiable function $y = \phi(x)$ with $\phi(x_0) = y_0$.*

Proof. Rewrite the equation as $y = y_0 - \frac{F(x,y)}{\partial F/\partial y(x_0,y_0)}$ + error and show this defines a contraction in y for x near x_0. The fixed point of the resulting map is the desired $\phi(x)$. Differentiability follows by implicit differentiation. $\qquad\square$

8.3 Existence and Uniqueness for ODEs

The fixed-point method also yields solutions to first-order ordinary differential equations.

Theorem 8.3 (Picard–Lindelöf)**.** *Let $f(x, y)$ be continuous in a rectangle $|x - x_0| \leq a$, $|y - y_0| \leq b$, and suppose f satisfies a Lipschitz condition in y: $|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$. Then the initial value problem $y' = f(x, y)$, $y(x_0) = y_0$ has a unique solution on some interval around x_0.*

Proof. Integrate the differential equation to obtain the integral equation $y(x) = y_0 + \int_{x_0}^{x} f(t, y(t)) \, dt$. Define an operator T on the space of continuous functions by $(Ty)(x) = y_0 + \int_{x_0}^{x} f(t, y(t)) \, dt$. Under the Lipschitz condition and for $|x - x_0|$ small enough, T is a contraction with respect to the sup metric, so it has a unique fixed point, which is the desired solution. $\square$

These iterative arguments justify the standard root-finding schemes (Newton's method), guarantee the solvability of implicit equations, and provide existence and uniqueness for differential equations under natural hypotheses.

9 Partial Differentiation

Functions of several variables inherit many properties from one-variable calculus, but extending definitions carefully avoids pitfalls.

9.1 Definitions and Basic Properties

Let $U \subseteq \mathbb{R}^n$ be open and $f \colon U \to \mathbb{R}$.

Definition 9.1 (Partial derivatives)**.** *The ith partial derivative of f at $a \in U$ is*

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a)}{h},$$

if the limit exists.

Having all partial derivatives exist does not guarantee good behavior. The correct higher-dimensional analogue of differentiability is best phrased in terms of linear approximations.

Definition 9.2 (Differentiability)**.** *f is differentiable at a if there exists a linear map $L \colon \mathbb{R}^n \to \mathbb{R}$ such that*

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

This L is the differential df_a, and it is given by $L(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i$ when f is differentiable with existing partials.

Theorem 9.1. *If f is differentiable at a, then all partial derivatives exist at a and f is continuous at a.*

9.2 Higher Derivatives

Higher-order derivatives are defined by iterating the differential, provided the lower-order derivatives are differentiable.

Definition 9.3. *If the first partial derivatives of f exist and are continuous on U, then f is said to be C^1. Inductively, f is C^k if its partial derivatives up to order k exist and are continuous.*

Mixed partial derivatives often coincide.

Theorem 9.2 (Equality of mixed partials)**.** *If the second-order partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ are continuous near a, then they are equal at a.*

Higher-dimensional Taylor expansions mirror the one-variable case by replacing powers with multi-index notation.

Theorem 9.3 (Taylor's theorem in $\mathbb{R}^n$)**.** *If f is C^{k+1} on a convex set containing a and $a + h$, then*

$$f(a + h) = f(a) + \sum_{|\alpha|=1}^{k} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + R_{k+1}(h),$$

where $|R_{k+1}(h)| \leq C\|h\|^{k+1}$ for some constant C.

9.3 Implicit Function Theorem

The implicit function theorem generalizes the one-dimensional result.

Theorem 9.4 (Implicit function theorem)**.** *Let $F \colon \mathbb{R}^{m+n} \to \mathbb{R}^n$ be C^1 near (x_0, y_0) with $F(x_0, y_0) = 0$. If the Jacobian matrix $\frac{\partial F}{\partial y}(x_0, y_0)$ is invertible, then there exist neighborhoods U of x_0 and V of y_0 and a unique C^1 function $\phi \colon U \to V$ such that $F(x, \phi(x)) = 0$ for all $x \in U$.*

Proof. Apply the contraction mapping principle to the map $y \mapsto y - \left(\frac{\partial F}{\partial y}\right)^{-1} F(x, y)$ with x fixed near x_0. The fixed point depends smoothly on x because the derivative of F is continuous. $\square$

This theorem yields local solvability of systems of equations and underpins change-of-variables formulas in multiple integrals, as well as implicit definitions of surfaces and manifolds.

10 Multiple Integrals

Riemann integration extends naturally to higher dimensions and supports integration over regions with curved boundaries.

10.1 Riemann Integration on Rectangles

Let $I = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$.

Definition 10.1 (Partitions and Riemann sums). *A partition P of I consists of partitions of each coordinate interval. Its mesh is $|P| = \max_i \max_j (x_{i,j} - x_{i,j-1})$. A Riemann sum for $f\colon I \to \mathbb{R}$ looks like*

$$S(P, f) = \sum f(\xi_{i_1, \ldots, i_n}) \prod_{k=1}^{n} (x_{k, i_k} - x_{k, i_k - 1}),$$

with $\xi_{i_1, \ldots, i_n}$ chosen from each sub-rectangle.

Upper and lower sums are defined using suprema/infima over each sub-rectangle. Integrability means the sup of the lower sums equals the inf of the upper sums; the common value is $\int_I f$.

10.2 Existence and Volume

Continuous functions on I are integrable. More generally, if f is bounded and continuous except on a set of measure zero, then f is integrable. We integrate over a region E with piecewise smooth boundary by integrating over a containing rectangle and setting $f = 0$ outside E.

The volume of E is $\int_E 1$. Regions with negligible boundary can be approximated by finite unions of rectangles, so their volumes exist.

10.3 Iterated Integrals

Fubini's theorem allows multi-dimensional integrals to be computed as repeated one-dimensional integrals.

Theorem 10.1 (Fubini). *If f is integrable on $[a, b] \times [c, d]$, then*

$$\int_a^b \left(\int_c^d f(x, y) \, dy \right) dx = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy.$$

The same holds in higher dimensions by iterating the process.

10.4 Change of Variables

Smooth coordinate changes scale integrals by the Jacobian determinant.

Theorem 10.2 (Change of variables). *Let $T\colon U \to V$ be a C^1 bijection between open subsets of $\mathbb{R}^n$ with Jacobian determinant J_T. For integrable f on V,*

$$\int_V f(y) \, dy = \int_U f(T(x)) \, |J_T(x)| \, dx.$$

This identity justifies polar, cylindrical, and spherical coordinates, as well as more general transformations used in evaluating multi-dimensional integrals.