Notes on Mathematics

John Peloquin

Version α

Contents

Introduction

This book contains reference notes on undergraduate mathematics, intended for review. None of the results are original, and only a few proofs are original. Most proofs are presented informally, with just enough information to convey the idea.

Currently there is some inconsistency in the presentation. Definitions are not included with the abstract algebra material (groups, rings, and fields), and neither definitions nor techniques are included with the logic and computability material. Also, the set theory material is presented in a detailed lecture format. The book will be normalized over time.

Part I

Algebra

Chapter 1

Vector Spaces

This chapter covers linear algebra from [1], with supplementary notes on quotient spaces and dual spaces from [7].

Throughout, $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$, and V and W denote vector spaces over $\mathbb{F}$.

1.1 Vector Spaces and Subspaces

Definitions

Definition 1.1.1. A *vector space* over a field $\mathbb{F}$ is a set V together with a binary operation $+ : V \times V \to V$ called *addition* and an external binary operation $\cdot : \mathbb{F} \times V \to V$ called *scalar multiplication*, satisfying the following properties:

Associativity $(u+v)+w = u+(v+w)$ for all $u, v, w \in V$, and $(\alpha\beta)u = \alpha(\beta u)$ for all $\alpha, \beta \in \mathbb{F}$ and $u \in V$.

Commutativity $u + v = v + u$ for all $u, v \in V$.

Identities There exists $0 \in V$ such that $u+0 = u$ for all $u \in V$, and $1u = u$ for all $u \in V$ where 1 is the multiplicative identity in $\mathbb{F}$.

Inverses For all $u \in V$ there exists $v \in V$ such that $u + v = 0$.

Distributivity $(\alpha+\beta)u = \alpha u + \beta u$ for all $\alpha, \beta \in \mathbb{F}$ and $u \in V$, and $\alpha(u + v) = \alpha u + \alpha v$ for all $\alpha \in \mathbb{F}$ and $u, v \in V$.

Definition 1.1.2. A subset U of V is called a *subspace* if it is a vector space with respect to the restrictions of the operations on V.

Definition 1.1.3. The *sum* of subspaces $U_1, \ldots, U_n$ of V is

$$U_1 + \cdots + U_n = \{ u_1 + \cdots + u_n \mid u_i \in U_i, i \in \{1, \ldots, n\} \}$$

Definition 1.1.4. If $U = U_1 + \cdots + U_n$ and for each $u \in U$ there exist unique $u_i \in U_i$ such that $u = u_1 + \cdots + u_n$, then U is called the *direct sum* of $U_1, \ldots, U_n$, written $U = U_1 \oplus \cdots \oplus U_n$.

Theorems

Theorem 1.1.1 (Characterization of direct sum)**.** *If $U_1, \dots, U_n$ are subspaces of V, then*

$$V = U_1 \oplus \cdots \oplus U_n$$

iff

- *(a)* $V = U_1 + \cdots + U_n$
- *(b) If* $0 = u_1 + \cdots + u_n$ *with* $u_i \in U_i$ *for* $i \in \{1, \dots, n\}$*, then* $u_i = 0$ *for* $i \in \{1, \dots, n\}$

Proof idea. For the forward direction, by definition of direct sum.

For the reverse direction, translate equality to nullity. If

$$v = u_1 + \cdots + u_n = w_1 + \cdots + w_n \quad (u_i, w_i \in U_i)$$

then $0 = (u_1 - w_1) + \cdots + (u_n - w_n)$, so $u_i = w_i$ for all i. □

Corollary 1.1.1. *$V = U_1 \oplus U_2$ iff $V = U_1 + U_2$ and $U_1 \cap U_2 = \{0\}$.*

Techniques

- Translating between equality and nullity.

1.2 Basis and Dimension

Definitions

Definition 1.2.1. A *linear combination* of vectors $(v_1, \dots, v_n)$ is a vector of the form

$$v = a_1 v_1 + \cdots + a_n v_n$$

where $a_1, \dots, a_n \in \mathbb{F}$.

Definition 1.2.2. The *span* of $(v_1, \dots, v_n)$ is the set of all linear combinations of $(v_1, \dots, v_n)$, denoted $\operatorname{span}(v_1, \dots, v_n)$.

Definition 1.2.3. A vector space is *finite-dimensional* if it equals the span of finitely many vectors, otherwise it is *infinite-dimensional.*

Definition 1.2.4. A list $(v_1, \dots, v_n)$ is *linearly independent* if when $a_1, \dots, a_n \in \mathbb{F}$ and

$$a_1 v_1 + \cdots + a_n v_n = 0$$

then $a_1 = \cdots = a_n = 0$; otherwise the list is *linearly dependent.*

Definition 1.2.5. A *basis* of a vector space is a linearly independent list which spans the space.

Definition 1.2.6. The *dimension* of a finite-dimensional vector space V is the length of a basis in V, and is denoted $\dim V$.

Theorems

Theorem 1.2.1 (Linear dependence)**.** *If* $(v_1,\dots,v_n)$ *is linearly dependent and* $v_1 \neq 0$*, there exists* $i \in \{2,\dots,n\}$ *such that*

(a) $v_i \in \operatorname{span}(v_1,\dots,v_{i-1})$

(b) $\operatorname{span}(v_1,\dots,v_{i-1},v_{i+1},\dots,v_n) = \operatorname{span}(v_1,\dots,v_n)$

Proof idea. By definition of linear dependence, choose $a_1,\dots,a_n \in \mathbb{F}$ with

$$a_1 v_1 + \cdots + a_n v_n = 0$$

Since $v_1 \neq 0$, there is maximum $i \in \{2,\dots,n\}$ with $a_i \neq 0$. Then

$$v_i = -\frac{a_1}{a_i} v_1 - \cdots - \frac{a_{i-1}}{a_i} v_{i-1} \in \operatorname{span}(v_1,\dots,v_{i-1})$$

Now in any linear combination of $v_1,\dots,v_n$, v_i can be eliminated. □

Applications. Throwing out redundant vectors in linearly dependent lists, lengths, dimension.

Remark. Intuitively, a list is linearly dependent iff one of the vectors in the list can be written in terms of the others.

Theorem 1.2.2 (Lengths)**.** *In a finite-dimensional vector space,*

(a) The length of any linearly independent list is less than or equal to the length of any spanning list.

(b) Any linearly independent list can be extended to a basis.

(c) Any spanning list can be reduced to a basis.

(d) Any two bases have the same length.

Proof idea. For (a), by right shifting the spanning list. In detail, prepend the first linearly independent vector to the spanning list. By linear dependence, one of the original spanning vectors can be thrown out while preserving the span. Since this process can be repeated for the rest of the linearly independent vectors, the lengths of the lists must have been as stated.

For (b), by extending to a maximal linearly independent list (using (a)).

For (c), by throwing out redundant vectors.

For (d), by using (a) twice. □

Applications. Bases, dimension.

Corollary 1.2.1 (Existence of basis)**.** *Every finite-dimensional vector space has a basis.*

Proof idea. By extending the empty list, or reducing a spanning list. □

Corollary 1.2.2 (Length criteria for basis)**.** *In a vector space of dimension* n*,*

(a) *Any linearly independent list of length n is a basis.*

(b) *Any spanning list of length n is a basis.*

Proof idea. By trivial expansion or reduction. □

Theorem 1.2.3 (Characterization of basis)**.** *If $(v_1, \dots, v_n)$ is a basis of V, then for every $v \in V$ there exist unique $a_1, \dots, a_n \in \mathbb{F}$ such that*

$$v = a_1 v_1 + \cdots + a_n v_n$$

Proof idea. Existence by spanning, and uniqueness by linear independence. □

Theorem 1.2.4 (Dimension of a subspace)**.** *If V is finite-dimensional and U is a subspace of V, then U is finite-dimensional and* $\dim U \leq \dim V$.

Proof idea. Since V is finite-dimensional, so is U, lest we could build arbitrarily long linearly independent lists in U, and hence in V. So U has a basis, which is linearly independent in V, and hence whose length is at most the length of a basis in V. □

Theorem 1.2.5 (Dimension of a sum)**.** *If U_1, U_2 are finite-dimensional subspaces,*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Proof idea. Fix a basis $(u_1, \dots, u_k)$ of $U_1 \cap U_2$. Extend it to a basis $(u_1, \dots, u_k, v_1, \dots, v_l)$ of U_1, and to a basis $(u_1, \dots, u_k, w_1, \dots, w_m)$ of U_2. Argue directly that

$$(u_1, \dots, u_k, v_1, \dots, v_l, w_1, \dots, w_m)$$

is a basis of $U_1 + U_2$, so

$$\begin{aligned}
\dim(U_1 + U_2) &= k + l + m \\
&= (k + l) + (k + m) - k \\
&= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)
\end{aligned}$$ □

Corollary 1.2.3 (Dimension of a direct sum)**.** *If $U_1, \dots, U_n$ are finite-dimensional subspaces,*

$$U_1 + \cdots + U_n = U_1 \oplus \cdots \oplus U_n \iff \dim(U_1 + \cdots + U_n) = \dim U_1 + \cdots + \dim U_n$$

Proof idea. By induction on n.

For $n = 1$, the result is trivial.

For $n > 1$, by the characterization of direct sums with two summands, dimension of sums, and induction hypothesis,

$$\begin{aligned}
\dim(U_1 \oplus \cdots \oplus U_n) &= \dim(U_1 \oplus \cdots \oplus U_{n-1}) + \dim U_n \\
&= (\dim U_1 + \cdots + \dim U_{n-1}) + \dim U_n \\
&= \dim U_1 + \cdots + \dim U_n
\end{aligned}$$

Conversely, suppose

$$\dim(U_1 + \cdots + U_n) = \dim U_1 + \cdots + \dim U_n$$

By dimension of sums, we know for $U = (U_1 + \cdots + U_{n-1}) \cap U_n$ that

$$\dim(U_1 + \cdots + U_n) = \dim(U_1 + \cdots + U_{n-1}) + \dim U_n - \dim U$$

So

$$\dim(U_1 + \cdots + U_{n-1}) = \dim U_1 + \cdots + \dim U_{n-1} + \dim U$$

But

$$\dim(U_1 + \cdots + U_{n-1}) \leq \dim U_1 + \cdots + \dim U_{n-1}$$

So $\dim U = 0$. By the induction hypothesis then, $U_1 + \cdots + U_{n-1} = U_1 \oplus \cdots \oplus U_{n-1}$, and

$$\dim((U_1 \oplus \cdots \oplus U_{n-1}) + U_n) = \dim(U_1 \oplus \cdots \oplus U_{n-1}) + \dim U_n$$

Now by the dimension of sums and the characterization of direct sums with two summands, $U_1 + \cdots + U_n = U_1 \oplus \cdots \oplus U_n$ as desired. □

Theorem 1.2.6 (Existence of direct sum)**.** *If V is finite-dimensional and U is a subspace of V, there exists a subspace W of V such that $V = U \oplus W$.*

Proof idea. Extend a basis $(u_1, \dots, u_m)$ of U to a basis $(u_1, \dots, u_m, w_1, \dots, w_n)$ of V, then set $W = \operatorname{span}(w_1, \dots, w_n)$. □

Techniques

- Translating between equality and nullity.
- Choosing a basis for manipulation, computation, etc.
- Induction on dimension.

1.3 Linear Transformations

Definitions

Definition 1.3.1. A *linear map* or *linear transformation* from V to W is a map $T : V \to W$ satisfying the following *linearity* properties:

Additivity $T(u + v) = Tu + Tv$ for all $u, v \in V$.

Homogeneity $T(\alpha u) = \alpha Tu$ for all $\alpha \in \mathbb{F}$ and $u \in V$.

The space of all linear maps from V to W is denoted $\operatorname{Hom}(V, W)$.

A linear map from V to V is called a *linear operator* on V. The space of all linear operators on V is denoted $\operatorname{Hom}(V)$.

Definition 1.3.2. If $T \in \operatorname{Hom}(V, W)$, the *null space* or *kernel* of T is

$$\ker T = \{ v \in V \mid Tv = 0 \}$$

The *nullity* of T is $\dim \ker T$.

Definition 1.3.3. If $T \in \operatorname{Hom}(V, W)$, the *range* or *image* of T is

$$\operatorname{range} T = \{ Tv \mid v \in V \}$$

The *rank* of T is $\dim \operatorname{range} T$.

Definition 1.3.4. A linear map $T \in \operatorname{Hom}(V, W)$ is *invertible* if there exists $S \in \operatorname{Hom}(W, V)$ such that $ST = I_V$ and $TS = I_W$.

Definition 1.3.5. V and W are *isomorphic* if there is an invertible linear map $T \in \operatorname{Hom}(V, W)$.

Definition 1.3.6. An m-by-n *matrix* M over $\mathbb{F}$ is a rectangular array of elements of $\mathbb{F}$ with m rows and n columns. Entry (i, j) of M is often denoted M_{ij}. The space of all m-by-n matrices over $\mathbb{F}$ is denoted $\operatorname{Mat}(m, n, \mathbb{F})$.

If $T \in \operatorname{Hom}(V, W)$, $(v_1, \ldots, v_n)$ is a basis of V and $(w_1, \ldots, w_m)$ is a basis of W, then the matrix of T with respect to these bases is

$$\mathrm{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_n)) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

where $Tv_i = a_{1,i} w_1 + \cdots + a_{m,i} w_m$ for $i \in \{1, \ldots, n\}$.

If $v \in V$,

$$\mathrm{M}(v, (v_1, \ldots, v_n)) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

where $v = a_1 v_1 + \cdots + a_n v_n$.

If the bases are understood or unimportant, they are omitted.

Theorems

Theorem 1.3.1 (Uniqueness)**.** *A linear map is uniquely determined by its action on a basis.*

Proof idea. By linearity.

Let $S, T \in \operatorname{Hom}(V, W)$, and let $(v_1, \ldots, v_n)$ be a basis of V on which S and T agree. For any $v \in V$, there exist $a_1, \ldots, a_n \in \mathbb{F}$ with $v = a_1 v_1 + \cdots + a_n v_n$, so

$$\begin{aligned} Sv &= S(a_1 v_1 + \cdots + a_n v_n) \\ &= a_1 S v_1 + \cdots + a_n S v_n \\ &= a_1 T v_1 + \cdots + a_n T v_n \\ &= T(a_1 v_1 + \cdots + a_n v_n) = Tv \end{aligned}$$

So $S = T$. □

Corollary 1.3.1. *If $(v_1,\dots,v_n)$ is a basis of V and $(w_1,\dots,w_n)$ are any vectors in W, there exists a unique linear map $T \in \operatorname{Hom}(V,W)$ with $Tv_i = w_i$ for $i \in \{1,\dots,n\}$.*

Proof idea. Define T by

$$T(a_1v_1+\cdots+a_nv_n) = a_1w_1+\cdots+a_nw_n \qquad \square$$

Applications. Defining linear maps by specifying their action on a basis.

Theorem 1.3.2 (Rank nullity)**.** *If V is finite-dimensional and $T \in \operatorname{Hom}(V,W)$,*

$$\dim V = \dim\ker T + \dim\operatorname{range} T$$

Proof idea. Extend a basis $(u_1,\dots,u_m)$ of $\ker T$ to a basis $(u_1,\dots,u_m,w_1,\dots,w_n)$ of V, and argue that $(Tw_1,\dots,Tw_n)$ is a basis of range T. □

Applications. Structure of vector spaces, behavior of linear maps.

Remark. Since the dimension of a finite-dimensional vector space characterizes its isomorphism type, this is just a disguised form of the first isomorphism theorem.

Corollary 1.3.2. *If $\dim V > \dim W$, there is no injective linear map from V to W.*

Corollary 1.3.3. *If $\dim V < \dim W$, there is no surjective linear map from V to W.*

Applications. Systems of linear equations.

Theorem 1.3.3 (Characterization of injectivity)**.** *A linear map is injective iff its null space is zero.*

Proof idea. By translating between equality and nullity. If T is a linear map,

$$Tv = Tw \iff Tv - Tw = 0 \iff T(v-w) = 0 \iff v - w \in \ker T \qquad \square$$

Theorem 1.3.4 (Characterization of invertibility)**.** *A linear map is invertible iff it is both injective and surjective.*

Proof idea. By a standard argument, a linear map T has a unique inverse *function* T^{-1} iff T is injective and surjective. Linearity of T^{-1} follows from linearity of T. □

Corollary 1.3.4. *If V is finite-dimensional and $T \in \operatorname{Hom}(V)$, the following are equivalent:*

(a) T is invertible

(b) T is injective

(c) T is surjective

Proof idea. By rank nullity and the characterization of injectivity, (b) $\iff$ (c). Now the result follows from the theorem. □

Theorem 1.3.5 (Dimension and isomorphism type)**.** *Two finite-dimensional vector spaces are isomorphic iff they have the same dimension.*

Proof. By rank nullity for the forward direction and by mapping a basis to a basis for the reverse direction. □

Applications. Determining isomorphism, calculating dimension.

Remark. The isomorphism is not natural in general, since it relies on choice of bases.

Theorem 1.3.6 (Matrices)**.** *If V and W are finite-dimensional, $T \in \operatorname{Hom}(V,W)$, and $v \in V$, then*

$$\mathrm{M}(Tv) = \mathrm{M}(T)\,\mathrm{M}(v)$$

relative to any fixed bases for V and W.

Proof idea. By definitions. □

Applications. Computation.

Theorem 1.3.7. *If V and W are finite-dimensional,*

$$\mathrm{M} : \operatorname{Hom}(V,W) \cong \operatorname{Mat}(\dim W, \dim V, \mathbb{F})$$

Proof idea. By direct argument. □

Corollary 1.3.5. *If V and W are finite-dimensional,*

$$\dim \operatorname{Hom}(V,W) = (\dim V)(\dim W)$$

Techniques

- Choosing a basis for manipulation, computation, etc.
- Using rank nullity to analyze the structure of vector spaces and behavior of linear maps.
- Translating between equality and nullity.

1.4 Quotient Spaces

Definitions

Definition 1.4.1. If U is a subspace of V, then for $v \in V$, the set

$$v + U = \{\, v + u \mid u \in U \,\}$$

is called the *coset* of U containing v.

Definition 1.4.2. If U is a subspace of V, then the *quotient space* of V modulo U, denoted V/U, is the set of cosets of U in V

$$V/U = \{\, v + U \mid v \in V \,\}$$

with the following laws of composition:

$$\begin{gathered}(v+U)+(w+U)=(v+w)+U\\ \alpha(v+U)=\alpha v+U\end{gathered}$$

for all $v, w \in V$ and $\alpha \in \mathbb{F}$.

Definition 1.4.3. If U is a subspace of V, the *canonical map* from V to V/U is the map $\pi : V \to V/U$ mapping $v \mapsto (v+U)$.

Theorems

Theorem 1.4.1 (Quotient space)**.** *If U is a subspace of V, then the quotient space V/U forms a vector space.*

Theorem 1.4.2 (First isomorphism theorem)**.** *If $T \in \operatorname{Hom}(V,W)$ is surjective and $U = \ker T$, the map $\overline{T} : V/U \to W$ mapping $(v+U) \mapsto Tv$ witnesses an isomorphism $V/U \cong W$. Moreover, $T = \overline{T}\pi$, where π is the canonical map.*

Proof. By direct computation,

$$\begin{aligned} v+U=w+U &\iff (v-w)+U=U\\ &\iff (v-w)\in U\\ &\iff T(v-w)=0\\ &\iff Tv-Tw=0\\ &\iff Tv=Tw \end{aligned}$$

for all $v, w \in V$, so $\overline{T}$ is well-defined and injective. Linearity and surjectivity of $\overline{T}$ are induced from T, so $\overline{T}$ is an isomorphism. Finally,

$$(\overline{T}\pi)v = \overline{T}(\pi(v)) = \overline{T}(v+U) = Tv$$ □

Applications. Induction on the order of finite vector spaces.

Techniques

- Induction on the order of finite vector spaces.

1.5 Eigenvalues and Eigenvectors

Definitions

Definition 1.5.1. If $T \in \operatorname{Hom}(V)$, a subspace U of V is *invariant* under T if $Tu \in U$ for all $u \in U$.

Definition 1.5.2. If $T \in \operatorname{Hom}(V)$, $v \in V$, and $\lambda \in \mathbb{F}$, v is an *eigenvector* of T corresponding to λ if $Tv = \lambda v$. If there is a nonzero eigenvector of T corresponding to λ, then λ is called an *eigenvalue* of T.

Remark. Note eigenvalues and eigenvectors can be zero, but corresponding to an eigenvalue there must always be a nonzero eigenvector.

Definition 1.5.3. If $T \in \operatorname{Hom}(V)$ and $p(z)$ is a polynomial over $\mathbb{F}$ of the form

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n$$

where $a_1, \ldots, a_n \in \mathbb{F}$, then the *polynomial operator* $p(T)$ is

$$p(T) = a_0 I + a_1 T + \cdots + a_n T^n$$

Definition 1.5.4. Let M be a square matrix. M is *diagonal* if $M_{ij} = 0$ whenever $i \neq j$. M is *upper-triangular* if $M_{ij} = 0$ whenever $i > j$. M is *block upper-triangular* if

$$M = \begin{bmatrix} M_1 & & * \\ & \ddots & \\ 0 & & M_n \end{bmatrix}$$

where each M_i is a square matrix called a *block*.

Theorems

Theorem 1.5.1 (Characterizations of eigenvalue)**.** *If $T \in \operatorname{Hom}(V)$ and $\lambda \in \mathbb{F}$, the following are equivalent:*

(a) *λ is an eigenvalue of T*

(b) $\ker(T - \lambda I)$ *is nonzero*

(c) *$T - \lambda I$ is not injective*

(d) *$T - \lambda I$ is not surjective*

(e) *$T - \lambda I$ is not invertible*

Applications. Computing eigenvalues (for example, by computing $\det(T - \lambda I)$).

Theorem 1.5.2 (Eigenvalues and invariant subspaces)**.** *A linear operator on a vector space has an eigenvalue iff it has a one-dimensional invariant subspace.*

Theorem 1.5.3 (Linear independence of eigenvectors)**.** *Nonzero eigenvectors of a linear operator corresponding to distinct eigenvalues are linearly independent.*

Proof idea. By contradiction, looking at a redundant vector.

Let $T \in \operatorname{Hom}(V)$, $\lambda_1, \ldots, \lambda_n$ distinct eigenvalues of T, and $v_1, \ldots, v_n$ corresponding nonzero eigenvectors. If $(v_1, \ldots, v_n)$ is linearly dependent, fix $k > 1$ least such that $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$. So $(v_1, \ldots, v_{k-1})$ is linearly independent but

$$v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}$$

for some $a_1, \ldots, a_{k-1} \in \mathbb{F}$ not all zero. Applying T to both sides gives

$$\lambda_k v_k = a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1}$$

But $\lambda_k v_k = a_1\lambda_k v_1 + \cdots + a_{k-1}\lambda_k v_k$, so

$$0 = a_1(\lambda_1 - \lambda_k)v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1}$$

Therefore $\lambda_i = \lambda_k$ for some $i \in \{1,\dots,k-1\}$, contradicting distinctness. □

Corollary 1.5.1. *T has at most $\dim V$ distinct eigenvalues.*

Corollary 1.5.2. *If T has $\dim V$ distinct eigenvalues, then T has a diagonal matrix with respect to a basis of eigenvectors.*

Theorem 1.5.4 (Polynomial operators)**.** *If $p, q \in \mathbb{F}[x]$, T is a linear operator, and $c \in \mathbb{F}$,*

(a) $(p+q)(T) = p(T) + q(T)$

(b) $(cp)(T) = cp(T)$

(c) $(pq)(T) = p(T)q(T)$

Proof idea. For (a) and (b), by substitution.

For (c), by substition and induction on $n = \deg q$. If $n \le 0$, the result is trivial. If $n > 0$, let c be the leading coefficient of $q(z)$. Then

$$(pq)(z) = p(z)q(z) = p(z)[q(z) - cz^n] + cp(z)z^n$$

Now $\deg[q(z) - cz^n] < n$, so by linearity and induction we have

$$\begin{aligned}(pq)(T) &= [p(z)[q(z) - cz^n]](T) + [cp(z)z^n](T)\\ &= p(T)[q(T) - cT^n] + c[p(z)z^n](T)\\ &= p(T)q(T) - c[p(T)T^n - [p(z)z^n](T)]\end{aligned}$$

It is easy to verify $p(T)T^n = [p(z)z^n](T)$, so $(pq)(T) = p(T)q(T)$. □

Corollary 1.5.3. *Polynomial operator substitution $p \to p(T)$ is linear.*

Corollary 1.5.4. *Polynomial operators over T commute.*

Theorem 1.5.5 (Eigenvalues over $\mathbb{C}$)**.** *Every linear operator on a nonzero, finite-dimensional complex vector space has an eigenvalue.*

Proof idea. By factoring a complex polynomial operator.

Let V be such a vector space and $T \in \mathrm{Hom}(V)$. Set $n = \dim V$ and fix $v \in V$, $v \neq 0$. Then the list $(v, Tv, \dots, T^n v)$ is linearly dependent, so there exist $a_0, \dots, a_n \in \mathbb{C}$ such that

$$a_0 v + a_1 Tv + \cdots + a_n T^n v = 0$$

Set $p(z) = a_0 + a_1 z + \cdots + a_n z^n$. By polynomial factorization over $\mathbb{C}$, there exist values $c, \lambda_1, \dots, \lambda_m \in \mathbb{C}$ with $c \neq 0$ and $m \ge 1$ such that $p(z) = c(z - \lambda_1)\cdots(z - \lambda_m)$. So

$$\begin{aligned}0 &= a_0 v + a_1 Tv + \cdots + a_n T^n v\\ &= (a_0 I + a_1 T + \cdots + a_n T^n)v\\ &= c(T - \lambda_1 I)\cdots(T - \lambda_m I)v\end{aligned}$$

Therefore some $(T - \lambda_i I)$ is not injective, and λ_i is an eigenvalue of T. □

Applications. Simplifying matrices.

Corollary 1.5.5 (Upper-triangular matrices over $\mathbb{C}$)**.** *Every linear operator on a nonzero, finite-dimensional complex vector space has an upper-triangular matrix with respect to some basis.*

Proof idea. By induction on the dimension of the space, using an eigenvalue in the induction step.

Let V be such a space and $T \in \operatorname{Hom}(V)$. Fix an eigenvalue λ of T. Then $U = \operatorname{range}(T - \lambda I)$ is invariant under T and $\dim U < \dim V$, so there is a basis of U with respect to which $T|_U$ has an upper-triangular matrix. Extend this to a basis of V with respect to which T has an upper-triangular matrix. □

Theorem 1.5.6 (Upper-triangular matrices)**.** *If a linear operator has an upper-triangular matrix with respect to some basis, then the operator is invertible iff there are no zeros along the diagonal of the matrix.*

Proof idea. By rank nullity.

Let $T \in \operatorname{Hom}(V)$ and suppose $(v_1, \dots, v_n)$ is a basis of V such that $\mathrm{M}(T, (v_1, \dots, v_n))$ is upper-triangular. If the k-th value along the diagonal is zero, T maps $\operatorname{span}(v_1, \dots, v_k)$ into $\operatorname{span}(v_1, \dots, v_{k-1})$ of smaller dimension, so T is not injective and not invertible. Conversely, if T is not invertible, then T is not surjective so there exists k least such that $v_k \notin \operatorname{range} T|_{\operatorname{span}(v_1,\dots,v_k)}$. It is easy to see that the k-th value along the diagonal must be zero. □

Corollary 1.5.6. *The eigenvalues of the operator are the values along the diagonal.*

Proof idea. Since for any $\lambda \in \mathbb{F}$,

$$\mathrm{M}(T - \lambda I, (v_1, \dots, v_n)) = \begin{bmatrix} \lambda_1 - \lambda & & * \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{bmatrix} \qquad \square$$

Theorem 1.5.7 (Eigenvalues over $\mathbb{R}$)**.** *Every linear operator on a nonzero, finite-dimensional real vector space has an invariant subspace of dimension* 1 *or* 2*.*

Proof idea. By factoring a real polynomial operator.

Note an invariant subspace of dimension 1 corresponds to a linear factor, and an invariant subspace of dimension 2 to an irreducible quadratic factor. □

Applications. Simplifying matrices.

Remark. The differences between complex and real operators stem largely from the differences between complex and real polynomial factorization.

Corollary 1.5.7. *Every linear operator on a real vector space of odd dimension has an eigenvalue.*

Proof idea. By induction on the dimension of the space, using an invariant subspace of dimension 1 or 2 in the induction step.

Let V be such a vector space and $T \in \operatorname{Hom}(V)$. Let U be an invariant subspace of dimension 1 or 2. If U has dimension 1, T has an eigenvalue. If U has dimension 2, write $V = U \oplus W$. Then W has odd dimension and is invariant under $P_{W,U}T|_W$, so by induction $P_{W,U}T|_W$ has an eigenvalue, which is also an eigenvalue of T. □

Corollary 1.5.8 (Block upper-triangular matrices over $\mathbb{R}$)**.** *Every linear operator on a nonzero, finite-dimensional real vector space has a block upper-trianguar matrix with respect to some basis whose blocks are either* 1*-by-*1 *or* 2*-by-*2 *with no eigenvalues.*

Proof idea. By induction on the dimension of the space, using an invariant subspace of dimension 1 or 2 in the induction step. □

Theorem 1.5.8 (Diagonalization)**.** *Let V be finite-dimensional, $T \in \operatorname{Hom}(V)$, and $\lambda_1, \ldots, \lambda_m$ the distinct eigenvalues of T. The following are equivalent:*

(a) *T has a diagonal matrix with respect to some basis of V*

(b) *V has a basis consisting of eigenvectors of T*

(c) *T has one-dimensional invariant subspaces $U_1, \ldots, U_n$ such that*

$$V = U_1 \oplus \cdots \oplus U_n$$

(d) $V = \ker(T - \lambda_1 I) \oplus \cdots \oplus \ker(T - \lambda_m I)$

(e) $\dim V = \dim\ker(T - \lambda_1 I) + \cdots + \dim\ker(T - \lambda_m I)$

Techniques

- Applying polynomial theory to polynomial operators.
- Induction on dimension.
- Simplifying matrices of operators for ease of analysis and computation.
- Using rank nullity to analyze the behavior of operators.
- Translating between equality and nullity.

1.6 Inner Product Spaces

Let V and W be inner product spaces over $\mathbb{F}$.

Definitions

Definition 1.6.1. An *inner product* on V is a function $\langle \quad , \quad \rangle : V \times V \to \mathbb{F}$ satisfying the following properties:

Positivity $\langle v, v\rangle \geq 0$ for all $v \in V$.

Definiteness $\langle v, v\rangle = 0$ iff $v = 0$.

Left additivity $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

Left homogeneity $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$ for all $\alpha \in \mathbb{F}$ and $u, v \in V$.

Conjugate symmetry $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

An *inner product space* is a vector space together with an inner product on the space.

Definition 1.6.2. A list $(v_1, \ldots, v_n)$ is *orthogonal* if $\langle v_i, v_j\rangle = 0$ when $i \neq j$.

Definition 1.6.3. The *norm* of a vector v is $\|v\| = \sqrt{\langle v, v\rangle}$.

Definition 1.6.4. A vector v is *normal* if $\|v\| = 1$.

Definition 1.6.5. A list $(v_1, \ldots, v_n)$ is *orthonormal* if it is orthogonal and contains only normal vectors.

Definition 1.6.6. If U is a subset of V, the *orthogonal complement* of U is

$$U^\perp = \{ v \in V \mid \langle u, v\rangle = 0 \text{ for all } u \in U\}$$

If $V = U \oplus U^\perp$, the *orthogonal projection* of V onto U is defined by $P_U v = u$ where $v - u \in U^\perp$.

Definition 1.6.7. A *linear functional* on V is a linear map $\varphi \in \mathrm{Hom}(V, \mathbb{F})$.

Definition 1.6.8. If $T \in \mathrm{Hom}(V, W)$, the *adjoint* of T is the linear map $T^* \in \mathrm{Hom}(W, V)$ such that

$$\langle T v, w\rangle = \langle v, T^* w\rangle$$

for all $v \in V$ and $w \in W$.

Definition 1.6.9. If M is an m-by-n matrix over $\mathbb{F}$, the *conjugate transpose* of M is the n-by-m matrix N defined by $N_{ij} = \overline{M_{ji}}$.

Theorems

Theorem 1.6.1 (Geometry)**.** *Let $u, v \in V$.*

(a) (Pythagorean theorem) If $\langle u, v\rangle = 0$, $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

(b) (Cauchy-Schwarz inequality) $|\langle u, v\rangle| \leq \|u\| \, \|v\|$, and equality holds iff one of u or v is a scalar multiple of the other.

(c) (Triangle inequality) $\|u + v\| \leq \|u\| + \|v\|$, and equality holds iff one of u or v is a nonnegative scalar multiple of the other.

(d) (Parallelogram equality) $\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$.

Proof idea. For (a), by expanding the inner product on the left.

For (b), by orthogonal decomposition and the pythagorean theorem. Assume without loss of generality that $u \neq 0$ and write

$$v = \frac{\langle v, u\rangle}{\|u\|^2} u + w$$

where $\langle u, w\rangle = 0$. Then

$$\|v\|^2 = \frac{|\langle u, v\rangle|^2}{\|u\|^2} + \|w\|^2 \geq \frac{|\langle u, v\rangle|^2}{\|u\|^2}$$

from which the desired inequality follows.

For (c), by the Cauchy-Schwarz inequality,

$$\begin{aligned}\|u+v\|^2 &= \|u\|^2 + 2\operatorname{Re}\langle u, v\rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

from which the desired inequality follows.

For (d), by expanding the inner products on the left. □

Theorem 1.6.2 (Computation with orthonormal lists)**.** *If* $(e_1,\dots,e_n)$ *is orthonormal and* $v = a_1e_1 + \cdots + a_ne_n$ *where* $a_1,\dots,a_n \in \mathbb{F}$*, then* $a_i = \langle v, e_i\rangle$ *for* $i \in \{1,\dots,n\}$*, and*

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$$

Proof idea. By properties of the inner product and orthonormality. □

Applications. Computing with orthonormal bases.

Corollary 1.6.1 (Linear independence of orthonormal lists)**.** *Orthonormal lists are linearly independent.*

Proof idea. If $a_1e_1 + \cdots + a_ne_n = 0$, then $|a_1|^2 + \cdots + |a_n|^2 = 0$, so $a_1 = \cdots = a_n = 0$. □

Theorem 1.6.3 (Gram-Schmidt)**.** *If* $(v_1,\dots,v_n)$ *is linearly independent, there exists an orthonormal list* $(e_1,\dots,e_n)$ *with* $\operatorname{span}(e_1,\dots,e_k) = \operatorname{span}(v_1,\dots,v_k)$ *for all* $k \in \{1,\dots,n\}$.

Proof idea. By recursively defining a new list, ensuring orthogonality and normality at each step.

For $i \in \{1,\dots,n\}$, define

$$e_i = \frac{v_i - \langle v_i, e_1\rangle e_1 - \cdots - \langle v_i, e_{i-1}\rangle e_{i-1}}{\|v_i - \langle v_i, e_1\rangle e_1 - \cdots - \langle v_i, e_{i-1}\rangle e_{i-1}\|}$$ □

Applications. Constructing orthonormal bases.

Corollary 1.6.2 (Existence of orthonormal bases)**.** *In a finite-dimensional inner product space, any orthonormal list can be extended to an orthonormal basis. In particular, every finite-dimensional inner product space has an orthonormal basis.*

Corollary 1.6.3 (Upper-triangular matrices over $\mathbb{C}$)**.** *Every linear operator on a nonzero, finite-dimensional complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.*

Theorem 1.6.4 (Orthogonal decomposition)**.** *If U is a finite-dimensional subspace of V, then $V = U \oplus U^{\perp}$.*

Proof idea. By computing with an orthonormal basis.

Fix an orthonormal basis $(e_1, \ldots, e_n)$ of U. For $v \in V$,

$$v = \big(\langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n\big) + \big(v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_n\rangle e_n\big)$$

which shows $v \in U + U^{\perp}$. By definiteness, $U \cap U^{\perp} = \{0\}$. □

Applications. Orthogonal projection.

Remark. Orthogonal decomposition (with a one-dimensional subspace) was already used in the proof of Cauchy-Schwarz. Note the technique used to construct a vector orthogonal to a list of vectors is from Gram-Schmidt.

Theorem 1.6.5 (Orthogonal projection)**.** *If U is a finite-dimensional subspace of V and $v \in V$, then*

$$\|v - P_U v\| \le \|v - u\|$$

for all $u \in U$, and equality holds iff $u = P_U v$.

Proof idea. By the pythagorean theorem, for $u \in U$,

$$\|v - P_U v\|^2 \le \|v - P_U v\|^2 + \|P_U v - u\|^2 = \|v - u\|^2$$ □

Applications. Minimization, approximation.

Theorem 1.6.6 (Linear functionals are given by inner products)**.** *If V is finite-dimensional and $\varphi \in \mathrm{Hom}(V, \mathbb{F})$ is a linear functional, there exists a unique $v \in V$ such that $\varphi(u) = \langle u, v\rangle$ for all $u \in V$.*

Proof idea. By computing with an orthonormal basis.

Fix an orthonormal basis $(e_1, \ldots, e_n)$ of V. For $u \in V$,

$$\begin{aligned}\varphi(u) &= \varphi(\langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n) \\ &= \langle u, e_1\rangle \varphi(e_1) + \cdots + \langle u, e_n\rangle \varphi(e_n) \\ &= \langle u, \overline{\varphi(e_1)} e_1\rangle + \cdots + \langle u, \overline{\varphi(e_n)} e_n\rangle \\ &= \langle u, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n\rangle\end{aligned}$$

So set $v = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n$. Uniqueness follows from definiteness. □

Applications. Adjoints.

Theorem 1.6.7 (Matrix of the adjoint)**.** *Let $T \in \operatorname{Hom}(V,W)$, $(e_1,\dots,e_n)$ an orthonormal basis of V, and $(f_1,\dots,f_m)$ an orthonormal basis of W. Then*

$$\mathrm{M}(T^*,(f_1,\dots,f_m),(e_1,\dots,e_n))$$

is the conjugate transpose of

$$\mathrm{M}(T,(e_1,\dots,e_n),(f_1,\dots,f_m))$$

Proof idea. By computing the matrix entries (with the orthonormal bases!),

$$\mathrm{M}(T)_{ij} = \langle Te_j, f_i\rangle = \langle e_j, T^* f_i\rangle = \overline{\langle T^* f_i, e_j\rangle} = \overline{\mathrm{M}(T^*)_{ji}} \qquad \square$$

Techniques

- Using orthogonal decomposition and projection, often in combination with geometrical results, in particular for minimization and approximation.
- Computing with orthonormal bases.
- Translating between equality and nullity using properties of the inner product.

1.7 Operators on Inner Product Spaces

Let V be an inner product space over $\mathbb{F}$.

Definitions

Definition 1.7.1. An operator $T \in \operatorname{Hom}(V)$ is *normal* if $T^*T = TT^*$.

Definition 1.7.2. An operator $T \in \operatorname{Hom}(V)$ is *self-adjoint* if $T = T^*$.

Definition 1.7.3. An operator $T \in \operatorname{Hom}(V)$ is *positive* if T is self-adjoint and $\langle Tv, v\rangle \geq 0$ for all $v \in V$.

Definition 1.7.4. An operator $T \in \operatorname{Hom}(V)$ is an *isometry* if $\|Tv\| = \|v\|$ for all $v \in V$.

Definition 1.7.5. The *singular values* of an operator $T \in \operatorname{Hom}(V)$ are the eigenvalues of $\sqrt{T^*T}$.

Definition 1.7.6. A square matrix M is *block diagonal* if

$$M = \begin{bmatrix} M_1 & & 0 \\ & \ddots & \\ 0 & & M_n \end{bmatrix}$$

where each M_i is a square block.

Theorems

Remark. It is useful to think intuitively of operators on complex inner product spaces as generalizations of complex numbers, which are just operators on the complex plane through multiplication. Normal operators correspond to complex numbers; adjoints, to conjugates; self-adjoint operators, to real numbers; positive operators, to nonnegative numbers; and isometries, to the unit circle. In addition, the polar decomposition is just polar form for operators.

Theorem 1.7.1 (Definiteness)**.** *Let $T \in \operatorname{Hom}(V)$.*

(a) If V is complex and $\langle Tv, v\rangle = 0$ for all $v \in V$, then $T = 0$.

(b) If T is self-adjoint and $\langle Tv, v\rangle = 0$ for all $v \in V$, then $T = 0$.

Proof idea. By expressing $\langle Tu, w\rangle$ for arbitrary $u, w \in V$ as a linear combination of inner products of the form $\langle Tv, v\rangle$, then taking $w = Tu$. □

Corollary 1.7.1. *If V is complex, T is self-adjoint iff $\langle Tv, v\rangle \in \mathbb{R}$ for all $v \in V$.*

Proof idea. By translating between equality and nullity. For $v \in V$,

$$\begin{aligned} \langle Tv, v\rangle \in \mathbb{R} &\iff \langle Tv, v\rangle = \overline{\langle Tv, v\rangle} \\ &\iff \langle Tv, v\rangle = \langle v, Tv\rangle \\ &\iff \langle Tv, v\rangle = \langle T^* v, v\rangle \\ &\iff \langle (T - T^*)v, v\rangle = 0 \end{aligned}$$ □

Theorem 1.7.2. *Eigenvalues of self-adjoint operators are real.*

Proof idea. Let $T \in \operatorname{Hom}(V)$, λ be an eigenvalue of T, and be a corresponding nonzero eigenvector v. Then

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

Therefore $\lambda = \overline{\lambda}$. □

Theorem 1.7.3 (Norm characterization of normality)**.** *If $T \in \operatorname{Hom}(V)$, then T is normal iff $\|Tv\| = \|T^* v\|$ for all $v \in V$.*

Proof idea. By translating between equality and nullity. For $v \in V$,

$$\begin{aligned} \|Tv\| = \|T^* v\| &\iff \langle Tv, Tv\rangle = \langle T^* v, T^* v\rangle \\ &\iff \langle T^* Tv, v\rangle = \langle TT^* v, v\rangle \\ &\iff \langle (TT^* - T^* T)v, v\rangle = 0 \end{aligned}$$

The result follows by noting that $TT^* - T^* T$ is self-adjoint. □

Theorem 1.7.4 (Spectral theorems)**.** *Let V be finite-dimensional and $T \in \operatorname{Hom}(V)$.*

(a) If V is complex, then V has an orthonormal basis consisting of eigenvectors of T iff T is normal.

(b) *If V is real, then V has an orthonormal basis consisting of eigenvectors of T iff T is self-adjoint.*

Proof idea. For (a), the forward direction is immediate by computing matrices. For the reverse direction, choose an orthonormal basis with respect to which $\mathrm{M}(T)$ is upper triangular, then use the relationship between $\mathrm{M}(T)$ and $\mathrm{M}(T^*)$ and the norm characterization of normality to argue that $\mathrm{M}(T)$ must be diagonal, and hence the basis vectors must be eigenvectors of T.

For (b), again the forward direction is immediate. For the reverse direction, first argue that T must have an eigenvalue by factoring an appropriate real polynomial operator in T[1] and noting that any irreducible quadratic factors are injective since T is self-adjoint. Then proceed by induction on $\dim V$, using the linear subspace generated by an eigenvector in an orthogonal decomposition in the induction step. □

Applications. Simplifying matrices.

Theorem 1.7.5 (Normal operators over $\mathbb{R}$)**.** *If V is finite-dimensional and real and $T \in \mathrm{Hom}(V)$, then T is normal iff V has an orthonormal basis with respect to which $\mathrm{M}(T)$ is block diagonal with 1-by-1 blocks and 2-by-2 blocks of the form*

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \qquad (a, b \in \mathbb{R}, b > 0)$$

Proof idea. The reverse direction is immediate by computing matrices.

For the forward direction, proceed by induction on $\dim V$. In the induction step, fix a subspace of V invariant under T of dimension 1 or 2 and use it in an orthogonal decomposition. In case of a 2-dimensional subspace, use the norm characterization of normality to argue that the block of $\mathrm{M}(T)$ corresponding to the subspace has the desired form. □

Applications. Simplifying matrices.

Theorem 1.7.6 (Characterizations of positive operators)**.** *If V is finite-dimensional and $T \in \mathrm{Hom}(V)$, the following are equivalent:*

(a) *T is positive*

(b) *T is self-adjoint and all the eigenvalues of T are nonnegative*

(c) *T has a positive square root*

(d) *T has a self-adjoint square root*

(e) *$T = S^* S$ for some $S \in \mathrm{Hom}(V)$*

Proof idea. For (a) $\Longrightarrow$ (b), by direct argument.

For (b) $\Longrightarrow$ (c), by the spectral theorem and then taking square roots along the diagonal of $\mathrm{M}(T)$.

For (c) $\Longrightarrow$ (d) $\Longrightarrow$ (e), by definitions.

For (e) $\Longrightarrow$ (a), by direct argument. □

[1] See the proof that an operator on a real vector space has an invariant subspace of dimension 1 or 2.

Corollary 1.7.2. *Positive operators on finite-dimensional inner product spaces have unique positive square roots.*

Proof idea. To prove uniqueness, note that the eigenvalues of any square root must be precisely the square roots of the eigenvalues of the operator, with corresponding eigenvectors. □

Theorem 1.7.7 (Characterizations of isometry in $\mathbb{C}$ and $\mathbb{R}$)**.** *Let V be finite-dimensional and $T \in \operatorname{Hom}(V)$.*

(a) If V is complex, T is an isometry iff V has an orthonormal basis with respect to which $\mathrm{M}(T)$ is diagonal with $|\mathrm{M}(T)_{ii}| = 1$ for all i.

(b) If V is real, T is an isometry iff V has an orthonormal basis with respect to which $\mathrm{M}(T)$ is block diagonal with blocks either 1-by-1 of the form $[\lambda]$ with $|\lambda| = 1$ or 2-by-2 of the form

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Proof idea. For the reverse directions, by computing with matrices. For the forward directions, by the characterizations of normal operators, noting the only possible eigenvalues of an isometry are ±1 and, for (b), that a real isometry of the form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

with $b > 0$ satisfies $a^2 + b^2 = 1$, so an appropriate θ exists by polar form of (a, b). □

Remark. An isometry is composed of reflections and rotations.

Theorem 1.7.8 (Polar decomposition)**.** *If V is finite-dimensional and $T \in \operatorname{Hom}(V)$, there is an isometry $S \in \operatorname{Hom}(V)$ such that $T = S\sqrt{T^*T}$.*

Proof idea. By extension. First define $S : \operatorname{range}\sqrt{T^*T} \to \operatorname{range} T$ by $\sqrt{T^*T}v \mapsto Tv$, then extend it to an operator on V through orthogonal decomposition of V. □

Theorem 1.7.9 (Singular value decomposition)**.** *If V is finite-dimensional, $T \in \operatorname{Hom}(V)$, and $s_1, \ldots, s_n$ are the singular values of T, there exist orthonormal bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ of V such that for all $v \in V$,*

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

In particular,

$$\mathrm{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_n)) = \begin{bmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{bmatrix}$$

Proof idea. By polar decomposition and the spectral theorem.

Write $T = S\sqrt{T^*T}$, S an isometry, and let $(e_1, \ldots, e_n)$ be an orthonormal basis of V consisting of eigenvectors of $\sqrt{T^*T}$ with corresponding eigenvalues $s_1, \ldots, s_n$. Then for $v \in V$,

$$Tv = s_1\langle v, e_1\rangle Se_1 + \cdots + s_n\langle v, e_n\rangle Se_n$$

Now set $f_i = Se_i$ for $i \in \{1, \ldots, n\}$. □

Techniques

- Thinking of operators on complex inner product spaces as complex numbers.
- Using induction on dimension, especially with orthogonal decomposition.
- Using the norm characterization of normality to analyze matrices.
- Using the spectral and related theorems to simplify matrices for analysis and computation.
- Using polar decomposition to break an arbitrary operator into simpler parts.
- Translating between equality and nullity.

1.8 Operators on Complex Vector Spaces

Definitions

Definition 1.8.1. If $T \in \operatorname{Hom}(V)$, $\lambda \in \mathbb{F}$, and $v \in V$, then v is a *generalized eigenvector* of T corresponding to λ if $(T-\lambda I)^k v = 0$ for some $k \geq 0$.

The space of all generalized eigenvectors of T corresponding to λ is called the *generalized eigenspace* of T corresponding to λ.

Definition 1.8.2. The *multiplicity* of an eigenvalue is the dimension of its corresponding generalized eigenspace.

Definition 1.8.3. An operator $T \in \operatorname{Hom}(V)$ is *nilpotent* if $T^k = 0$ for some $k \geq 0$.

Definition 1.8.4. If V is complex, $T \in \operatorname{Hom}(V)$, and $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ are the distinct eigenvalues of T with multiplicities $d_1, \ldots, d_m$, the *characteristic polynomial* of T is

$$p(z) = (z-\lambda_1)^{d_1} \cdots (z-\lambda_m)^{d_m}$$

Definition 1.8.5. If $T \in \operatorname{Hom}(V)$, the *minimal polynomial* of T is the monic polynomial of smallest degree such that $p(T) = 0$.

Theorems

Theorem 1.8.1 (Generalized eigenspace)**.** *If V is finite-dimensional, $T \in \operatorname{Hom}(V)$, and $\lambda \in \mathbb{F}$, the generalized eigenspace of T corresponding to λ is* $\ker(T-\lambda I)^{\dim V}$.

Proof idea. For arbitrary $S \in \operatorname{Hom}(V)$, argue that

$$\ker S \subseteq \ker S^2 \subseteq \cdots \subseteq \ker S^{\dim V} = \ker S^{\dim V+1} = \cdots$$

Then take $S = T - \lambda I$. □

Theorem 1.8.2 (Eigenvalue multiplicity)**.** *If V is finite-dimensional, $T \in \operatorname{Hom}(V)$, and $\lambda \in \mathbb{F}$, the number of times λ appears along the diagonal of any upper-triangular matrix for T is* $\dim\ker(T-\lambda I)^{\dim V}$.

Proof idea. By a tedious induction on $\dim V$. □

Applications. Multiplicity.

Corollary 1.8.1 (Eigenvalue multiplicity over $\mathbb{C}$)**.** *If V is complex,*

$$\dim V = \sum \textit{multiplicities of eigenvalues of } T$$

Proof idea. By choosing an upper-triangular matrix for T and summing eigenvalue multiplicities along the diagonal. □

Applications. Generalized eigenspace decomposition.

Theorem 1.8.3 (Generalized eigenspace decomposition over $\mathbb{C}$)**.** *If V is finite-dimensional and complex, $T \in \mathrm{Hom}(V)$, and $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ are the distinct eigenvalues of T with corresponding generalized eigenspaces $U_1, \dots, U_m$, then*

(a) $V = U_1 \oplus \cdots \oplus U_m$.

(b) U_i *is invariant under* T *for* $i \in \{1, \dots, m\}$.

(c) $(T - \lambda_i I)|_{U_i}$ *is nilpotent on* U_i *for* $i \in \{1, \dots, n\}$.

Proof idea. For (a), by multiplicity,

$$\begin{aligned} \dim V &= \dim U_1 + \cdots + \dim U_m \\ &= \dim(U_1 + \cdots + U_m) \end{aligned}$$

where the second equality holds since $U_1 + \cdots + U_m$ is invariant under T and contains all the eigenvalues of T.

For (b), by direct argument.

For (c), by definitions. □

Applications. Simplifying matrices, analyzing operators.

Remark. The generalized eigenspace decomposition shows that any operator on a complex vector space is composed of parts which are just nilpotent operators plus scalar multiples of the identity.

Theorem 1.8.4 (Block diagonal matrices over $\mathbb{C}$)**.** *If V is finite-dimensional and complex, $T \in \mathrm{Hom}(V)$, and $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ are the distinct eigenvalues of T, then there is a basis of V with respect to which* $\mathrm{M}(T)$ *is block diagonal of the form*

$$\mathrm{M}(T) = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$$

where each block A_i is upper-triangular of the form

$$A_i = \begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix}$$

Proof idea. By the generalized eigenspace decomposition and analysis of nilpotent operators.

Write $V = U_1 \oplus \cdots \oplus U_m$ where U_i is the generalized eigenspace corresponding to λ_i. Then $T - \lambda_i I$ is nilpotent on U_i, so by constructing an appropriate basis for U_i (starting with a basis of $\ker(T-\lambda_i I)$, extending it to a basis of $\ker(T-\lambda_i I)^2$, etc),

$$\mathrm{M}((T-\lambda_i I)|_{U_i}) = \begin{bmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{bmatrix}$$

Combine these bases to form a basis of V which gives $\mathrm{M}(T)$ the desired form. □

Theorem 1.8.5 (Jordan form over $\mathbb{C}$)**.** *If V is finite-dimensional and complex, $T \in \mathrm{Hom}(V)$, and $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ are the distinct eigenvalues of T, then there is a basis of V with respect to which $\mathrm{M}(T)$ is block diagonal of the form*

$$\mathrm{M}(T) = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$$

where each block A_i is an upper-triangular Jordan block of the form

$$A_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}$$

Proof idea. By the generalized eigenspace decomposition and analysis of nilpotent operators, as before, but this time choosing bases of the generalized eigenspaces more carefully.

By a tedious induction on $\dim V$, it can be shown that any nilpotent operator $N \in \mathrm{Hom}(V)$ has a matrix of the form

$$\mathrm{M}(N) = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}$$

with respect to a certain basis of V. Apply this result to the generalized eigenspace decomposition and piece things together as before. □

Remark. All of the canonical matrix forms we have constructed for operators on complex vector spaces are upper triangular, but with increasing amounts of zeros.

Theorem 1.8.6 (Cayley-Hamilton)**.** *If V is finite-dimensional and complex, $T \in \mathrm{Hom}(V)$, and p is the characteristic polynomial of T, then $p(T) = 0$.*

Proof idea. Choose a basis $(v_1, \ldots, v_n)$ with respect to which $\mathrm{M}(T)$ is upper-triangular with diagonal entries $\lambda_1, \ldots, \lambda_n$. By multiplicity, $p(z) = (z - \lambda_1) \cdots (z - \lambda_n)$, so

$$p(T) = (T - \lambda_1 I) \cdots (T - \lambda_n I)$$

Now argue by induction on i that $(T - \lambda_1 I) \cdots (T - \lambda_i I) v_i = 0$ for $i \in \{1, \ldots, n\}$ using the form of $\mathrm{M}(T)$, so $p(T) = 0$. □

Theorem 1.8.7 (Minimal polynomial)**.** *Let V be finite-dimensional and $T \in \mathrm{Hom}(V)$.*

(a) If p is a polynomial over $\mathbb{F}$, $p(T) = 0$ iff the minimal polynomial of T divides p.

(b) The roots of the minimal polynomial of T are precisely the eigenvalues of T.

Proof idea. For (a), by polynomial division with remainder.
For (b), by direct argument. □

Applications. Computing eigenvalues.

Corollary 1.8.2. *If V is complex, the degree of the minimal polynomial of T is at most* $\dim V$.[2]

Proof idea. By Cayley-Hamilton. □

Techniques

- Simplifying matrices of operators for ease of analysis and computation using the generalized eigenspace decomposition and derived canonical forms.
- Computing eigenvalues by computing and factoring minimal polynomials.

1.9 Operators on Real Vector Spaces

Definitions

Definition 1.9.1. If V is real and $T \in \mathrm{Hom}(V)$, $(\alpha, \beta) \in \mathbb{R}^2$ is an *eigenpair* of T if $\alpha^2 < 4\beta$ and $T^2 + \alpha T + \beta I$ is not injective.

Definition 1.9.2. If V is real, $T \in \mathrm{Hom}(V)$, and $\mathrm{M}(T)$ is block upper-triangular of the form

$$\mathrm{M}(T) = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$$

where each A_i is either a 1-by-1 block of the form $[\lambda_i]$ or a 2-by-2 block of the form $\begin{bmatrix} a_i & c_i \\ b_i & d_i \end{bmatrix}$ with no eigenvalues, the *characteristic polynomial* of T is

$$p(x) = p_1(x) \cdots p_m(x)$$

[2] This result also holds for real vector spaces, by the real Cayley-Hamilton theorem.

where

$$p_i(x) = \begin{cases} x - \lambda_i & \text{if } A_i \text{ is 1-by-1} \\ (x - a_i)(x - d_i) - b_i c_i & \text{if } A_i \text{ is 2-by-2} \end{cases}$$

Definition 1.9.3. The multiplicity of an eigenpair (α, β) of T is $\dim \ker(T^2 + \alpha T + \beta I)^{\dim V}/2$.

Theorems

Remark. Proofs of results here are analogous to the proofs of corresponding results for complex vector spaces, and so are omitted.

Theorem 1.9.1 (Eigenvalue and eigenpair multiplicity over $\mathbb{R}$)**.** *If V is finite-dimensional and real and $T \in \mathrm{Hom}(V)$, then for any block upper-triangular matrix for T consisting of 1-by-1 blocks and 2-by-2 blocks with no eigenvalues,*

(a) For $\lambda \in \mathbb{R}$, $[\lambda]$ appears as a 1-by-1 block $\dim \ker(T - \lambda I)^{\dim V}$ times.

(b) For $\alpha, \beta \in \mathbb{R}$ with $\alpha^2 < 4\beta$, $x^2 + \alpha x + \beta$ appears as the characteristic polynomial of a 2-by-2 block $\dim \ker(T^2 + \alpha T + \beta I)^{\dim V}/2$ times.

Applications. Multiplicity.

Corollary 1.9.1.

$$\dim V = \sum \textit{multiplicities of eigenvalues of } T + 2 \sum \textit{multiplicities of eigenpairs of } T$$

Remark. By the theorem, there is an invariance with respect to the characteristic polynomials of blocks occurring in block upper-triangular representations of real operators, but this is not true in general with respect to the blocks themselves.

Theorem 1.9.2 (Generalized eigenspace and eigenpair space decomposition over $\mathbb{R}$)**.** *If V is finite-dimensional and real, $T \in \mathrm{Hom}(V)$, $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ are the distinct eigenvalues of T with corresponding generalized eigenspaces $U_1, \ldots, U_m$, and $(\alpha_1, \beta_1), \ldots, (\alpha_M, \beta_M)$ are the distinct eigenpairs of T with corresponding eigenpair spaces $V_1, \ldots, V_M$, then*

(a) $V = U_1 \oplus \cdots \oplus U_m \oplus V_1 \oplus \cdots \oplus V_M$.

(b) U_i is invariant under T for $i \in \{1, \ldots, m\}$.

(c) V_j is invariant under T for $j \in \{1, \ldots, M\}$.

(d) $(T - \lambda_i I)|_{U_i}$ is nilpotent on U_i for $i \in \{1, \ldots, m\}$.

(e) $(T^2 + \alpha_j T + \beta_j I)|_{U_j}$ is nilpotent on V_j for $j \in \{1, \ldots, M\}$.

Theorem 1.9.3 (Cayley-Hamilton)**.** *If V is finite-dimensional and real, $T \in \mathrm{Hom}(V)$, and p is the characteristic polynomial of T, then $p(T) = 0$.*

Techniques

- Simplifying matrices of operators for ease of analysis and computation using the generalized eigenspace and eigenpair space decomposition.

1.10 Matrices

Definitions

Definition 1.10.1. If $T \in \operatorname{Hom}(V)$ and $n = \dim V$, the *trace* of T is -1 times the coefficient of z^{n-1} in the characteristic polynomial of T, and is denoted $\operatorname{tr} T$.

Definition 1.10.2. If $T \in \operatorname{Hom}(V)$ and $n = \dim V$, the *determinant* of T is $(-1)^n$ times the constant in the characteristic polynomial of T, and is denoted $\det T$.

Definition 1.10.3. If A is an n-by-n matrix, the *trace* of A is $\operatorname{tr} A = \sum A_{ii}$.

Definition 1.10.4. If A is an n-by-n matrix, the *determinant* of A is

$$\det A = \sum_{\pi \in \operatorname{perm} n} \operatorname{sign} \pi \, A_{\pi(1)1} \cdots A_{\pi(n)n}$$

Theorems

Theorem 1.10.1 (Change of basis)**.** *If $T \in \operatorname{Hom}(V)$ and $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ are bases of V, then*

$$\mathrm{M}(T, (u_1, \ldots, u_n)) = A^{-1} \mathrm{M}(T, (v_1, \ldots, v_n)) A$$

where $A = \mathrm{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$.

Proof idea. By computation. □

Applications. Showing that change of basis is just conjugation.

Theorem 1.10.2 (Trace of matrices)**.** *If A and B are n-by-n matrices over $\mathbb{F}$ and $c \in \mathbb{F}$,*

(a) $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$.

(b) $\operatorname{tr}(cA) = c \operatorname{tr} A$.

(c) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof idea. By direct computation.

For (c),

$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{n} (AB)_{ii} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ji} \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ji} A_{ij} \\ &= \sum_{j=1}^{n} (BA)_{jj} = \operatorname{tr}(BA)\end{aligned}$$

□

Theorem 1.10.3 (Trace of operators)**.** *If V is finite-dimensional, $S,T \in \operatorname{Hom}(V)$, and $c \in \mathbb{F}$,*

(a) $\operatorname{tr} S = \operatorname{tr} \mathrm{M}(S)$ *with respect to any basis of V.*

(b) $\operatorname{tr}(S+T) = \operatorname{tr} S + \operatorname{tr} T$.

(c) $\operatorname{tr}(cS) = c \operatorname{tr} S$.

(d) $\operatorname{tr}(ST) = \operatorname{tr}(TS)$.

Proof idea. For (a), first argue directly that it is true for a basis with respect to which T has an upper-triangular matrix (if $\mathbb{F} = \mathbb{C}$) or an appropriate block upper-triangular matrix (if $\mathbb{F} = \mathbb{R}$), then use change of basis and trace of the matrix product for the case of an arbitrary basis.

For (b)–(d), by (a) and properties of traces of matrices. □

Applications. Calculating traces of operators.

Theorem 1.10.4 (Determinant of matrices)**.** *If A and B are n-by-n matrices over $\mathbb{F}$,*

$$\det(AB) = (\det A)(\det B)$$

Theorem 1.10.5 (Determinant of operators)**.** *If V is finite-dimensional and $S,T \in \operatorname{Hom}(V)$,*

(a) $\det S = \det \mathrm{M}(S)$ *with respect to any basis of V.*

(b) $\det(ST) = (\det S)(\det T)$.

Proof idea. By the same ideas in the proof for traces of operators. □

Applications. Calculating determinants of operators.

Remark. The determinant is not a linear map on matrices or operators, although it is multilinear in the columns and rows of a matrix.

Theorem 1.10.6 (Determinant characterization of invertibility)**.** *If V is finite-dimensional and $T \in \operatorname{Hom}(V)$, T is invertible iff* $\det T \neq 0$.

Proof idea. By definition, $\det T$ is the product of all the eigenvalues of T (if $\mathbb{F} = \mathbb{C}$) together with possibly other nonzero values (if $\mathbb{F} = \mathbb{R}$), so $\det T \neq 0$ iff 0 is not an eigenvalue of T, which is true iff T is injective, which is true iff T is invertible. □

Applications. Determining invertibility of operators, etc.

Theorem 1.10.7 (Determinant and characteristic polynomial)**.** *If V is finite-dimensional and $T \in \operatorname{Hom}(V)$, the characteristic polynomial of T equals* $\det(zI - T)$.

Applications. Computing characteristic polynomials.

Techniques

- Using matrices to compute traces and determinants of operators.
- Using determinants to determine invertibility of operators, etc.

1.11 Linear Functionals and Dual Spaces

Definitions

Definition 1.11.1. The *dual space* of V is the vector space $V^* = \mathrm{Hom}(V, \mathbb{F})$ of linear functionals from V to $\mathbb{F}$.

Definition 1.11.2. If $\beta = (v_1, \dots, v_n)$ is a basis of V, the *i-th coordinate function* is $\pi_i^\beta : V \to \mathbb{F}$ mapping $v \mapsto \alpha_i$, where

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

Definition 1.11.3. If $\beta = (v_1, \dots, v_n)$ is a basis of V, then the *dual basis* to β in V^* is $\beta^* = (\pi_1^\beta, \dots, \pi_n^\beta)$.

Definition 1.11.4. If $T \in \mathrm{Hom}(V, W)$, the *dual* or *transpose* of T is $T^* : W^* \to V^*$ mapping $g \mapsto gT$.

Definition 1.11.5. If $v \in V$, the *dual* of v is $v^* : V^* \to \mathbb{F}$ mapping $g \mapsto g(v)$.

Theorems

Theorem 1.11.1 (Dual isomorphism)**.** *If V is finite-dimensional, then $V \cong V^*$.*

Proof idea. By dimension, since

$$\dim V^* = \dim \mathrm{Hom}(V, \mathbb{F}) = (\dim V)(\dim \mathbb{F}) = \dim V \qquad \square$$

Remark. Note this isomorphism is not natural since it relies on a choice of basis.

Theorem 1.11.2 (Dual basis)**.** *If $\beta = (v_1, \dots, v_n)$ is a basis of V, then $\beta^* = (\pi_1^\beta, \dots, \pi_n^\beta)$ is a basis of V^*. Every $f \in V^*$ can be written uniquely in the form*

$$f = \sum_{i=1}^{n} f(v_i)\pi_i^\beta$$

Proof idea. By direct computation. □

Applications. Computing in the dual space.

Theorem 1.11.3 (Dual transformation matrix)**.** *Let V and W have bases $\beta = (v_1, \dots, v_n)$ and $\gamma = (w_1, \dots, w_m)$, respectively. If $T \in \mathrm{Hom}(V, W)$ with dual T^*, then*

$$\mathrm{M}(T^*, \gamma^*, \beta^*) = \mathrm{M}(T, \beta, \gamma)^{\mathrm{t}}$$

Proof idea. By direct computation, if $\mathrm{M}(T) = (a_{ij})$,

$$\begin{aligned} T^*(\pi_j^\gamma) &= \pi_j^\gamma T \\ &= \sum_{i=1}^n (\pi_j^\gamma T v_i)\pi_i^\beta \\ &= \sum_{i=1}^n \pi_j^\gamma \big(\sum_{k=1}^m a_{ki} w_k\big)\pi_i^\beta \\ &= \sum_{i=1}^n \sum_{k=1}^m a_{ki}\pi_j^\gamma(w_k)\pi_i^\beta \\ &= \sum_{i=1}^n a_{ji}\pi_i^\beta \end{aligned}$$

So $\mathrm{M}(T^*)_{ij} = \mathrm{M}(T)_{ji}$, that is, $\mathrm{M}(T^*) = \mathrm{M}(T)^{\mathrm{t}}$. □

Applications. Computing the dual transformation.

Theorem 1.11.4 (Double dual isomorphism)**.** *If V is finite-dimensional, then the map $v \mapsto v^*$ witnesses a natural isomorphism $V \cong V^{**}$.*

Proof idea. The map is trivially linear, and injective since its kernel is trivial (note if $v^* = 0$, then for any basis β of V, the coefficients of v over β must all be zero). The rest follows by dimension. □

Remark. If V is infinite-dimensional, none of V, V^*, and V^{**} need be isomorphic.

Chapter 2

Groups

This chapter covers group theory from [4].

2.1 Groups and Subgroups

Theorems

Theorem 2.1.1 (Structure of cyclic groups). *Let $G = \langle x \rangle$ be a cyclic group of order n.*

(a) Every subgroup of G is cyclic. More specifically, if $H \leq G$, then $H = 1$ or $H = \langle x^d \rangle$, where d is least positive such that $x^d \in H$.

(b) If $n < \infty$,

(i) For each positive divisor d of n, G has a unique subgroup of order d, namely $H = \langle x^a \rangle$ where $a = n/d$.

(ii) For any $a \in \mathbb{Z}$, $\langle x^a \rangle = \langle x^{(a,n)} \rangle$.

(iii) For any $a, b \in \mathbb{Z}$, $\langle x^a \rangle \leq \langle x^b \rangle$ iff $(b,n)|(a,n)$.

(c) If $n = \infty$,

(i) For any $a, b \in \mathbb{Z}$ nonnegative with $a \neq b$, $\langle x^a \rangle \neq \langle x^b \rangle$.

(ii) For any $a \in \mathbb{Z}$, $\langle x^a \rangle = \langle x^{|a|} \rangle$.

(iii) For any $a, b \in \mathbb{Z}$, $\langle x^a \rangle \leq \langle x^b \rangle$ iff $b|a$.

Proof idea. For (a), by division with remainder.

For (b), by calculating orders of elements. For (i), $|H| = |x^a| = n/(a,n) = n/a = d$. If $K \leq G$ with $|K| = d$, then by (a) $K = \langle x^b \rangle$ with $n/(b,n) = |x^b| = d = n/a$, so $a|b$ and $K \leq H$. Since $|H| = |K|$, $K = H$. For (ii), similarly. For (iii), assume without loss of generality that $a|n$ and $b|n$ by (ii). If $\langle x^a \rangle \leq \langle x^b \rangle$, then $x^a = (x^b)^k = x^{bk}$ for some k, so $n/a = n/(a,n) = |x^a| = |x^{bk}| = n/(bk,n) = n/[b(k,n)]$, so $b|a$. The reverse direction is immediate.

For (c), by even simpler arguments. □

Remark. By this result, the subgroups of a cyclic group correspond bijectively to either the positive divisors of the order of the group (if the group is finite) or the positive integers (if the group is infinite), and all inclusions among subgroups are determined by divisibility. This gives us the complete structure for the group.

Techniques

- Using basic number theory for calculations with group elements.
- Using subgroup lattices to understand the structure of groups.

Remark. Take care when using subgroup lattices. Two nonisomorphic groups may have the same lattice. Also, if $N \trianglelefteq G$, the isomorphism type of G is not in general uniquely determined by the isomorphism types of factors G/N and N, so a group isomorphism type is not in general uniquely determined for a lattice even when types for its 'top' and 'bottom' relative to a fixed normal subgroup are.

2.2 Homomorphisms and Quotients

Theorems

Theorem 2.2.1 (Normality)**.** *If G is a group and $N \leq G$, the following are equivalent:*

(a) $N \trianglelefteq G$

(b) $gNg^{-1} \subseteq N$ *for all* $g \in G$

(c) $gN = Ng$ *for all* $g \in G$

(d) $N_G(N) = G$

(e) the coset space G/N forms a group under the usual operation

(f) N is the kernel of a homomorphism on G

Theorem 2.2.2 (Tower law)**.** *If G is a group and $H \leq K \leq G$, then*

$$|G:H| = |G:K||K:H|$$

Proof idea. The cosets of K partition G, and the number of G-cosets of H in any coset gK of K is equal to $|K:H|$, by the map $kH \mapsto gkH$ ($k \in K$). □

Corollary 2.2.1 (Lagrange)**.** *If G is finite, $|K|$ divides $|G|$ and the number of left [right] cosets of K in G is $|G|/|K|$.*

Applications. Getting combinatorial information about subgroups and ruling out possibilities.

Theorem 2.2.3 (Isomorphism theorems)**.** *Let G be a group.*

(a) If $\varphi : G \to H$ is a surjective homomorphism and $K = \ker\varphi$, then $G/K \cong H$.

(b) If $H, K \leq G$ and $H \leq N_G(K)$, then $HK/K \cong H/H \cap K$.

(c) *If $H, K \trianglelefteq G$ and $H \leq K$, then* $(G/H)/(K/H) \cong G/K$.

Proof idea. For (a), by the isomorphism $gK \mapsto \varphi(g)$.

For (b), by (a). Define $\varphi : H \to HK/K$ by $h \mapsto hK$. Then φ is surjective and $\ker\varphi = H \cap K$, so $H/H \cap K \cong HK/K$.

For (c), by (a). Define $\varphi : G/H \to G/K$ by $gH \mapsto gK$. Then φ is surjective and $\ker\varphi = K/H$, so $(G/H)/(K/H) \cong G/K$. □

Theorem 2.2.4 (Lattice theorem)**.** *If G is a group and $N \trianglelefteq G$, there is a bijection between the subgroups H of G containing N and the subgroups $\overline{H} = H/N$ of $\overline{G} = G/N$ which respects subgroup lattice structure. More specifically, for all $H, K \leq G$ with $N \leq H, K$,*

(a) $H \leq K$ *iff* $\overline{H} \leq \overline{K}$.

(b) $H \trianglelefteq K$ *iff* $\overline{H} \trianglelefteq \overline{K}$.

(c) $\overline{\langle H, K\rangle} = \langle \overline{H}, \overline{K}\rangle$.

(d) $\overline{H \cap K} = \overline{H} \cap \overline{K}$.

(e) *if* $H \leq K$, $|K : H| = |\overline{K} : \overline{H}|$.

Applications. Relating the structures of G and G/N, often in inductive arguments; showing that the lattice of G/N is the part of the lattice of G above N.

Techniques

- Determining subgroup lattice structure with the tower law (Lagrange).
- Proving normality:
  - Directly (conjugation).
  - Showing left cosets equal right cosets.
  - Determining normalizers (with help of the tower law, etc).
  - Exhibiting kernels.
  - Special cases (for example, subgroups of the center).
- Using induction on the order of a group:
  - Involving subgroups.
  - Involving quotients, often by subgroups of the center (since these are always normal).
- Factoring groups into parts using composition series (Jordan-Hölder).
- Classifying finite groups using the Jordan-Hölder program:
  - Classifying all finite simple groups.
  - Finding all ways to assemble simple groups into larger groups.

2.3 Group Actions

Theorems

Theorem 2.3.1 (Orbit partitioning)**.** *Let G act on A. The orbits of A under G partition A, so if $a_1, \dots, a_k$ are representatives from the orbits,*

$$|A| = |O_{a_1}| + \cdots + |O_{a_k}|$$

Proof idea. Orbits are equivalence classes. □

Applications. Getting combinatorial information from group actions.

Theorem 2.3.2 (Orbit-stabilizer)**.** *Let G act on A. If $a \in A$, $|O_a| = |G : G_a|$, so*

$$|G| = |G_a||O_a|$$

Proof idea. By the map $gG_a \mapsto g \cdot a$ $(g \in G)$ and Lagrange. □

Applications. Getting combinatorial information from group actions.

Corollary 2.3.1 (Cycle decomposition in S_n)**.** *If $\sigma \in S_n$, σ has a decomposition of the form*

$$\sigma = (a_1 \cdots a_{m_1}) \cdots (a_{m_{k-1}+1} \cdots a_{m_k})$$

where the cycles partition $\{1, \dots, n\}$. This decomposition is unique up to the order of the cycles and cyclic permutation of the numbers in the cycles.

Proof idea. Let σ act on $\{1, \dots, n\}$. Then the orbits uniquely determine the cycles. □

Applications. Cycle shape, conjugacy classes in S_n.

Theorem 2.3.3 (Cayley)**.** *If G is a group, G is isomorphic to some permutation group. In particular, if $|G| = n$, G is isomorphic to a subgroup of S_n.*

Proof idea. Let G act on itself by left multiplication. □

Applications. Getting permutations from abstract groups, showing that the study of abstract groups is (in a certain sense) equivalent to the study of permutation groups, getting combinatorial information about finite groups.

Theorem 2.3.4 (Class equation)**.** *If G is finite and $g_1, \dots, g_k \in G$ are representatives from the noncentral conjugacy classes in G, then*

$$|G| = |Z(G)| + \sum_{i=1}^{k} |G : C_G(g_i)|$$

Proof idea. Let G act on itself by conjugation, then apply the orbit partitioning and orbit-stabilizer theorems, noting that central conugacy classes are singletons. □

Applications. Getting combinatorial information about groups, normal subgroups (which are unions of conjugacy classes), centers, centralizers, etc. and ruling out possibilities.

Theorem 2.3.5 (Sylow)**.** *Let G be a group with $|G| = p^a m$ where p is prime not dividing m.*

(a) There exists a Sylow p-subgroup of G.

(b) If Q is a p-subgroup of G, there exists $g \in G$ and a Sylow p-subgroup P of G such that $Q \leq gPg^{-1}$. In particular, all Sylow p-subgroups of G are conjugate.

(c) If n_p is the number of Sylow p-subgroups of G, $n_p \equiv 1$ (mod p). If P is a Sylow p-subgroup of G, $n_p = |G : N_G(P)|$, so n_p divides m.

Proof idea. For (a), by induction on $|G|$. If p divides $|Z(G)|$, then (by induction on abelian groups) there exists $N \leq Z(G)$ with $|N| = p$. Now $N \trianglelefteq G$, so for $\overline{G} = G/N$, $|\overline{G}| = p^{a-1}m$. By induction, $\overline{G}$ has a subgroup $\overline{P} = P/N$ where $N \leq P \leq G$ and $|\overline{P}| = p^{a-1}$. Then $|P| = p^a$, so P is a Sylow p-subgroup of G. If p does not divide $|Z(G)|$, by the class equation there must exist $g \in G - Z(G)$ such that p^a divides $|C_G(g)|$. Since $|C_G(g)| < |G|$, by induction $C_G(g)$ has a Sylow p-subgroup which is also a Sylow p-subgroup of G.

For (b) and (c), by group action. Fix a Sylow p-subgroup P_1 of G and let $\mathscr{C} = \{P_1, \ldots, P_k\}$ be the set of all G-conjugates of P_1. Then Q acts on $\mathscr{C}$ by conjugation and partitions $\mathscr{C}$ into orbits of size $|Q : N_G(P_i) \cap Q| = |Q : P_i \cap Q|$ (for representative P_i). Taking $Q = P_1$, it follows $k \equiv 1$ (mod p). Moreover, there cannot be any p-subgroup Q not contained in some P_i lest $k \not\equiv 1$ (mod p), so (b) holds. Now $\mathscr{C}$ just consists of the Sylow p-subgroups of G, so $n_p = k$ and (c) follows. □

Applications. Breaking groups into simpler pieces (p-subgroups), finding normal subgroups, proving groups are not simple, classifying groups.

Techniques

- Getting information about groups using group actions:
  - On arbitrary sets.
  - By left multiplication on cosets and elements of the group.
  - By conjugation on subsets of the group.
- Applying combinatorial results (orbit partitioning, orbit-stabilizer, the class equation).
- Getting information about (quotients of) normalizers and centralizers from information about automorphism groups.
- Using induction on the order of a group.
- Restricting the source of group actions to subgroups about which we know more, to get additional combinatorial information.
- Breaking down groups with Sylow's theorems.
- Using permutation representations to prove groups isomorphic to subgroups of S_n and A_n.
- Using cycle shapes to determine conjugacy classes in S_n.

2.4 Direct Products and Semidirect Products

Theorems

Theorem 2.4.1 (Fundamental theorem of finite abelian groups)**.** *If G is an abelian group with $|G| = p_1^{a_1} \cdots p_k^{a_k}$, where $p_1, \ldots, p_k$ are pairwise distinct primes and $a_1, \ldots, a_k \geq 1$, there exist groups $A_1, \ldots, A_k$ such that*

(a) $G \cong \prod_{i=1}^{k} A_i$

(b) *For each factor A_i, $|A_i| = p_i^{a_i}$ and there exist $b_1 \geq \cdots \geq b_m \geq 1$ with $a_i = \sum_{j=1}^{m} b_j$ such that $A_i \cong \prod_{j=1}^{m} Z_{p_i^{b_j}}$.*

Moreover, this factorization is unique in the sense that if $G \cong \prod_{i=1}^{r} B_i$ with $|B_i| = p_i^{a_i}$ for all i, then $A_i \cong B_i$ for all i.

Proof idea. By the structure theorem for finitely generated modules over principal ideal domains, since finite abelian groups are just finitely generated $\mathbb{Z}$-modules.

Alternately, for existence, by induction. Since a finite abelian group is a direct product of Sylow subgroups,[1] assume without loss of generality that G is a p-group. First argue that an elementary abelian p-group E can be factored as $E = M \times \langle x \rangle$, where M is maximal and $x \neq 1$. Now let H be the subgroup of G consisting of p-th powers, and K the subgroup consisting of elements of order p. Note H and K are just the range and kernel, respectively, of the map $x \mapsto x^p$, and G/H and K are both elementary abelian. If possible, pull back a factorization of G/H into G and appeal to induction on a factor. Otherwise, appeal to the induction hypothesis on $G/K \cong H$, and then pull back p-th roots of the generators from the factorization of G/K. For uniqueness, also use induction. □

Applications. Classifying finite abelian groups.

Theorem 2.4.2 (Commutators)**.** *Let G be a group.*

(a) *For all $x, y \in G$, $xy = yx[x, y]$. In particular $xy = yx$ iff $[x, y] = 1$.*

(b) *For all $H \leq G$, $H \trianglelefteq G$ iff $[H, G] \leq H$.*

(c) G' char G.

(d) *For all $N \trianglelefteq G$, G/N is abelian iff $G' \leq N$. In particular, G/G' is the largest abelian quotient of G.*

Applications. Determining when elements commute.

Theorem 2.4.3 (Internal direct products)**.** *Let G be a group. If $H, K \trianglelefteq G$ and $H \cap K = 1$, then $HK \cong H \times K$.*

Proof idea. Each element in HK can be written uniquely in the form hk for some $h \in H$ and $k \in K$. Also, since both H and K are normal, $[h, k] \in H \cap K = 1$ for all $h \in H$ and $k \in K$, hence every element of H commutes with every element of K. It follows that the map $hk \mapsto (h, k)$ is the desired isomorphism. □

[1] See characterizations of finite nilpotence below.

Applications. Recognizing direct products.

Theorem 2.4.4 (Internal semidirect products)**.** *Let G be a group. If $H, K \leq G$, $H \trianglelefteq G$, and $H \cap K = 1$, then $HK \cong H \rtimes K$, where K acts on H by conjugation.*

Proof idea. Again, the map $hk \mapsto (h, k)$ is the desired isomorphism (but for different reasons!). □

Applications. Recognizing semidirect products.

Techniques

- Classifying finite abelian groups of order n using the fundamental theorem:

  1. Factor n into its prime factorization $n = p_1^{a_1} \cdots p_k^{a_k}$.
  2. For each i, determine all possible abelian groups of order $p_i^{a_i}$ by first determining all partitions of the number a_i of form $a_i = b_1 + \cdots + b_j$ with $b_1 \geq \cdots \geq b_j \geq 1$, then forming corresponding direct products

     $$Z_{p_i^{b_1}} \times \cdots \times Z_{p_i^{b_j}}$$

  3. Determine all possible abelian groups of order n by forming all direct products of the form $A_1 \times \cdots \times A_k$ where A_i is a group from the previous step with $|A_i| = p_i^{a_i}$.

- Classifying finite groups of order n using semidirect products:

  1. Show that every group G of order n has proper subgroups H and K which form an internal semidirect product $G = HK$ (for example, using Sylow's theorems).
  2. Determine all possible isomorphism types for H and K (inductively).
  3. For each pair H, K in the previous step, find all possible automorphism representations φ of K on H.
  4. For each triple H, K, φ in the previous step, form the semidirect product $H \rtimes_\varphi K$.
  5. Determine the distinct isomorphism types in the previous step.

- In inductive arguments involving finite abelian (or more generally, nilpotent) groups, taking quotients by maximal subgroups and leveraging the simple structure of the quotients.
- Taking quotients to impose relations.
- Proving normality using commutator subgroups.
- Using direct and semidirect products to construct examples.

2.5 Nilpotence, Solvability, Simplicity, and Freeness

Theorems

Theorem 2.5.1 (Finite nilpotence)**.** *If G is a finite group, the following are equivalent:*

(a) G is nilpotent

(b) Normalizers grow in G (that is, if $H < G$, then $H < N_G(H)$)

(c) Every Sylow subgroup of G is normal

(d) G is a direct product of Sylow subgroups

(e) Every maximal subgroup of G is normal

Proof idea. For (a) $\Longrightarrow$ (b), by induction (on the nilpotence class of G). If $Z(G) \leq H$, then $\overline{H} < \overline{G} = G/Z(G)$, so by induction $\overline{H} < N_{\overline{G}}(\overline{H}) = \overline{N_G(H)}$, and hence by the lattice theorem $H < N_G(H)$. If $Z(G) \not\leq H$, then $H < \langle H, Z(G)\rangle \leq N_G(H)$.

For (b) $\Longrightarrow$ (c), by taking normalizers. If P is a nonnormal Sylow subgroup of G, then $P \trianglelefteq N_G(P) < G$. But then P is characteristic in $N_G(P)$, so $P \trianglelefteq N_G(N_G(P))$, so $N_G(P) = N_G(N_G(P))$, contradicting that normalizers grow.

For (c) $\Longrightarrow$ (d), by internal direct product recognition.

For (d) $\Longrightarrow$ (a), by induction, since $G/Z(G)$ is also a product of Sylow subgroups.

Note (b) $\Longrightarrow$ (e) is immediate.

For (e) $\Longrightarrow$ (c), use Frattini's argument. If P is a nonnormal Sylow subgroup of G, then $P \trianglelefteq N_G(P) < M \lhd G$ for some maximal subgroup M of G. But by Frattini's argument, $G = MN_G(P) = M$, a contradiction. □

Applications. Finding large normalizers (for example, when proving non-simplicity), classifying nilpotent groups, taking quotients by maximal subgroups, etc.

Corollary 2.5.1 (p-groups)**.** *p-groups are nilpotent.*

Applications. Applying techniques for nilpotent groups to p-groups, and arbitrary finite groups via Sylow p-subgroups.

Theorem 2.5.2 (Universal property of free groups)**.** *If S is a set, G is a group, and $f : S \to G$ is any set map, there exists a unique homomorphism $\varphi : F(S) \to G$ extending f.*

Moreover, this property characterizes $F(S)$ up to a unique isomorphism which is the identity on S.

Proof idea. To satisfy the desired properties, φ must be defined by

$$\varphi(\prod s_i^{\alpha_i}) = \prod f(s_i)^{\alpha_i} \qquad (s_i \in S, \alpha_i \in \{-1,0,1\})$$

And since $F(S)$ is free of relations in S, φ is well defined. This establishes existence and uniqueness of φ.

If $F(S)$ and $F^*(S)$ both satisfy the universal property, then there exist two unique homomorphisms $\varphi : F(S) \to F^*(S)$ and $\varphi^* : F^*(S) \to F(S)$ extending the identity map on S. Then by the universal property again, $\varphi\varphi^*$ and $\varphi^*\varphi$ are the identity, so φ and φ^* are inverses and $F(S) \cong F^*(S)$. □

Techniques

- Factoring nilpotent groups into parts using upper [lower] central series, and using induction on nilpotence class.
- Factoring solvable groups into parts using commutator series, and using induction on solvable length.
- In inductive arguments involving finite nilpotent groups, taking quotients by maximal subgroups and leveraging the simple structure of the quotients.
- Proving a finite group is not simple:
    - Counting elements using Sylow's theorem and showing that some proper Sylow subgroup must be normal.
    - Permutation representations on cosets. In particular, exploiting a lower bound on possible indices for proper subgroups of the simple group, working directly inside S_n and A_n, etc.
    - Finding large normalizers of Sylow subgroups by working with multiple primes.
    - Finding large normalizers of intersections of Sylow subgroups.
- Defining group actions on algebraic (and induced geometric) structures to prove existence and uniqueness of simple groups.
- Taking quotients to impose relations.
- Defining homomorphisms using free groups and presentations.

Chapter 3

Rings

This chapter covers ring theory from [4].

3.1 Rings and Ideals

Theorems

Theorem 3.1.1 (Ideals)**.** *If R is a ring and I is an additive subgroup of R, the following are equivalent:*

(a) I is an ideal

(b) the additive coset space R/I forms a ring under the natural operations

(c) I is the kernel of a ring homomorphism on R

Theorem 3.1.2 (Isomorphism theorems)**.** *Let R be a ring.*

(a) If $\varphi : R \to S$ is a surjective ring homomorphism and $I = \ker\varphi$, then $R/I \cong S$.

(b) If I is a subring of R and J is an ideal of R, then $I + J/J \cong I/I \cap J$.

(c) If I and J are ideals of R with $I \subseteq J$, then $(R/I)/(J/I) \cong R/J$.

Proof idea. Apply the group-theoretic isomorphism theorems, then argue directly that the group isomorphisms are ring isomorphisms. □

Theorem 3.1.3 (Lattice theorem)**.** *If R is a ring and I is an ideal in R, there is a bijective correspondence between the subrings J of R containing I and the subrings $\overline{J} = J/I$ of $\overline{R} = R/I$ which respects subring lattice structure and preserves ideals.*

Proof idea. Apply the group-theoretic lattice theorem, then argue directly that the correspondence preserves ideals. □

Applications. Relating the structures of R and R/I.

Theorem 3.1.4 (Cancellation in integral domains)**.** *If R is an integral domain, $a, b, c \in R$, $a \neq 0$, and $ab = ac$, then $b = c$.*

Theorem 3.1.5 (Ideals in fields)**.** *If R is a commutative ring with identity $1 \neq 0$, then R is a field iff the only ideals in R are (0) and (1).*

Applications. Proving that rings are fields, or that ideals are trivial.

Theorem 3.1.6 (Quotients)**.** *Let R be a commutative ring with identity $1 \neq 0$ and I an ideal in R.*

(a) R/I is a field iff I is maximal.

(b) R/I is an integral domain iff I is prime.

Proof idea. By the lattice theorem, for (a) using the ideal structure of fields and for (b) using definitions. □

Applications. Proving that quotients are fields [integral domains], or that ideals are maximal [prime], constructing field extensions.

Corollary 3.1.1. *Maximal ideals are prime.*

Remark. The converse of the corollary is false in general, but it is true in principal ideal domains.

Theorem 3.1.7 (Fields of fractions)**.** *If R is an integral domain, there is a unique smallest field F containing R, in the sense that*

(i) F contains a subring isomorphic to R, and

(ii) if K is any field containing a subring isomorphic to R, then K also contains an extension of that subring isomorphic to F. Moreover, the extension is just the subfield of K generated by the subring.

Proof idea. Construct the field of 'fractions' over R by taking equivalence classes of pairs over R, then defining addition and multiplication in the natural ways. Embed R in the natural way. Argue that any field containing R must contain all the 'fractions', hence must contain F. □

Applications. Enabling one to do computations in a field over an integral domain, which might be more convenient (for example, when working with integers, or with polynomials whose coefficients lie in an integral domain, etc.).

Theorem 3.1.8 (Chinese remainder theorem)**.** *If R is a commutative ring with identity $1 \neq 0$ and $A_1, \ldots, A_k$ are pairwise comaximal ideals in R, then the projection map*

$$\varphi : R \to \prod_{i=1}^{k} R/A_i$$

is a surjective ring homomorphism with $\ker \varphi = A_1 \cap \cdots \cap A_k = A_1 \cdots A_k$, so

$$R/A_1 \cdots A_k \cong \prod_{i=1}^{k} R/A_i$$

Proof idea. By induction on $k \geq 2$.

For $k = 2$, use the fact that there exist $x \in A_1$ and $y \in A_2$ with $x + y = 1$ to show that φ is surjective and $A_1 \cap A_2 = A_1 A_2$, then appeal to the first isomorphism theorem.

For $k > 2$, argue that A_1 and $A_2 \cdots A_k$ are comaximal, then appeal to $k = 2$. □

Applications. Solving simultaneous congruences, etc.

Corollary 3.1.2. *The Euler φ function is multiplicative, that is*

$$\varphi(mn) = \varphi(m)\varphi(n) \qquad (m, n) = 1$$

Proof idea. If $(m, n) = 1$, then $\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, hence in particular

$$(\mathbb{Z}/mn\mathbb{Z})^\times \cong (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$$

The order on the left is $\varphi(mn)$, and the order on the right is $\varphi(m)\varphi(n)$. □

Techniques

- Using group-theoretic results to help establish ring-theoretic results.
- Using induction on the order of finite rings, involving ideals and quotients.
- Relating ideal structure to global structure.
- Computing in fraction fields.
- Using the chinese remainder theorem to solve simultaneous congruences, etc.

3.2 Euclidean Domains, Principal Ideal Domains, Unique Factorization Domains, and Integral Domains

Theorems

Theorem 3.2.1 (Ideals and greatest common divisors)**.** *If R is a ring and $a, b \in R$, then $d \in R$ is a greatest common divisor of a and b iff (d) is the smallest principal ideal containing (a, b).*

Applications. Translating between arithmetic properties and properties of ideals.

Corollary 3.2.1. *Greatest common divisors are unique up to units (i.e., they are associates).*

Theorem 3.2.2 (Euclidean algorithm)**.** *If R is a euclidean domain and $a, b \in R$ are nonzero with division $a = bq + r$ (where $q, r \in R$ with $r = 0$ or $N(r) < N(b)$), then*

$$(a, b) = (b, r)$$

In particular, if r_n is the last nonzero remainder in the euclidean algorithm with a and b (where $r_0 = b$), then $(a, b) = (r_n)$, so *(i)* r_n *is a greatest common divisor of a and b, and*

(ii) $r_n = ax + by$ *for some* $x, y \in R$.

Proof idea. By induction,

$$(a,b) = (b,r_1) = (r_1,r_2) = \cdots = (r_n,r_{n+1}) = (r_n)$$

where the last equality holds since $r_{n+1} = 0$. Now (i) and (ii) are immediate. □

Applications. Computing greatest common divisors (efficiently), computing multiplicative inverses of coprime elements, etc.

Theorem 3.2.3 (EDs are PIDs)**.** *Euclidean domains are principal ideal domains.*

Proof idea. By division with remainder.

If R is a euclidean domain and $I \subseteq R$ is a nonzero ideal, fix $b \in I$ nonzero with minimum norm $N(b)$. Then for any $a \in I$, there exist $q, r \in R$ with $a = bq + r$ where $r = 0$ or $N(r) < N(b)$. Now $r \in I$, so by minimality of $N(b)$, $r = 0$ and $a = bq \in (b)$. Thus $I = (b)$ is principal. □

Theorem 3.2.4 (Prime and maximal ideals)**.**

(a) In an integral domain, maximal ideals are prime.

(b) In a principal ideal domain, an ideal is prime iff it is maximal.

Proof idea. For (a), by previous results.

For (b), by direct computation. □

Theorem 3.2.5 (Primes and irreducibles)**.**

(a) In an integral domain, primes are irreducible.

(b) In a principal ideal domain, an element is prime iff it is irreducible.

(c) In a unique factorization domain, an element is prime iff it is irreducible.

Proof idea. For (a), by direct computation.

For (b), if p is irreducible, argue that (p) is maximal.

For (c), if p is irreducible and $p|ab$, factor both sides and argue that p must divide one of the irreducible factors in a or b. □

Theorem 3.2.6 (PIDs are UFDs)**.** *Principal ideal domans are unique factorization domains.*

Proof idea. For existence of factorizations, by a subdivision argument.

If R is a principal ideal domain and $a \in R$ cannot be factored into finitely many irreducibles, then a factorization tree for a can be arbitrarily extended in height. By recursion and the axiom of choice, there exists an infinite descending sequence of proper divisors

$$\cdots \mid a_k \mid \cdots \mid a_2 \mid a_1 \mid a$$

This means there is an infinite properly ascending chain of ideals

$$(a) \subset (a_1) \subset (a_2) \subset \cdots \subset (a_k) \subset \cdots$$

But this chain must eventually collapse since $I = \bigcup_{i=1}^{\infty} (a_i)$ is a principal ideal of form $I = (b)$, so $b \in (a_i)$ for some i and

$$I = (b) \subseteq (a_i) \subseteq (a_{i+1}) \subseteq \cdots \subseteq I$$

—a contradiction. Therefore factorizations exist.

For uniqueness of factorizations, use the primality of irreducibles to argue that the irreducibles in any two factorizations of an element must be associates. □

Theorem 3.2.7 (GCDs in UFDs)**.** *If R is a unique factorization domain and $a, b \in R$ with*

$$a = u \prod_{i=1}^{k} p_i^{e_i} \qquad b = v \prod_{i=1}^{k} p_i^{f_i}$$

where $u, v \in R$ are units, $p_1, \ldots, p_k \in R$ are primes, and $e_i, f_j \geq 0$, then

$$d = \prod_{i=1}^{k} p_i^{\min(e_i, f_i)}$$

is a greatest common divisor of a and b.

Proof idea. By unique factorization. □

Applications. Computing greatest common divisors from factorizations.

Theorem 3.2.8 (Fundamental theorem of arithmetic)**.** *$\mathbb{Z}$ is a euclidean domain, hence a princpal ideal domain and unique factorization domain. In particular, every integer can be factored uniquely as a product of primes.*

Proof idea. By induction. □

Applications. Arithmetic with the integers.

Remark. By definitions and results,

fields ⊆ euclidean domains (EDs)
⊆ principal ideal domains (PIDs) ⊆ unique factorization domains (UFDs)
⊆ integral domains (IDs) ⊆ commutative rings

Each of these inclusions is proper.

Techniques

- Translating between arithmetic and ideal properties.
- Computing with the euclidean algorithm.
- Computing with principal ideals.
- Factoring complex elements into simpler pieces.
- Subdivision to construct ascending chains of ideals and take limits.

3.3 Polynomials

Theorems

Let R be a commutative ring with identity $1 \neq 0$.

Theorem 3.3.1 (Reduction of coefficients)**.** *If I is an ideal in R, then*

$$R[x]/I[x] \cong (R/I)[x]$$

Proof idea. By the first isomorphism theorem. □

Applications. Diophantine equations, irreducibility criteria (e.g. Eisenstein).

Theorem 3.3.2 (Polynomial rings over integral domains)**.** *R is an integral domain iff $R[x]$ is an integral domain, in which case*

$$\deg p(x)q(x) = \deg p(x) + \deg q(x)$$

for all $p(x), q(x) \in R[x]$.

Proof idea. By looking at leading terms.

If R is an integral domain and $p(x), q(x) \in R[x]$ are nonzero with leading terms $a_n x^n$ and $b_m x^m$, respectively, then since a_n and b_m are nonzero, $a_n b_m$ is nonzero and the leading term of $p(x)q(x)$ is $a_n b_m x^{n+m}$. □

Applications. Computation of degree.

Corollary 3.3.1. *The units in $R[x]$ are just the units in R.*

Theorem 3.3.3 (Polynomial rings over fields)**.** *R is a field iff $R[x]$ is a euclidean domain with norm given by degree, in which case for all $a(x), b(x) \in R[x]$ with $b(x) \neq 0$, there exist unique $q(x), r(x) \in R[x]$ with*

$$a(x) = b(x)q(x) + r(x) \qquad \textit{where } \deg r(x) < \deg b(x)$$

Proof idea. For the forward direction, by induction on degree.

Let $a(x) = a_n x^n + \cdots + a_0$ and $b(x) = b_m x^m + \cdots + b_0$. If $n < m$, take $q(x) = 0$ and $r(x) = a(x)$. Otherwise note $a(x) - (a_n/b_m)x^{n-m}b(x)$ has degree less than n, so by induction there exist $s(x), r(x) \in R[x]$ such that

$$a(x) - \frac{a_n}{b_m}x^{n-m}b(x) = b(x)s(x) + r(x) \qquad \text{where } \deg r(x) < \deg b(x)$$

Take $q(x) = s(x) + (a_n/b_m)x^{n-m}$. To prove uniqueness, compute degrees.

For the reverse direction, note if $R[x]$ is a principal ideal domain then $R \cong R[x]/(x)$ is a field since (x) is prime and hence maximal in $R[x]$. □

Applications. Polynomial division with remainder, equivalence of roots and linear factors, irreducibility criteria, field extensions, etc.

Corollary 3.3.2. *If F is a field and $f(x) \in F[x]$, $F[x]/(f(x))$ is a field iff $f(x)$ is irreducible in $F[x]$.*

Proof idea. Because $F[x]$ is a principal ideal domain, $f(x)$ is irreducible iff $f(x)$ is prime iff $(f(x))$ is maximal. □

Applications. Constructing simple algebraic field extensions.

Corollary 3.3.3. *If F is a field and $f(x) \in F[x]$ has unique factorization*

$$f(x) = f_1(x)^{n_1} \cdots f_k(x)^{n_k}$$

where the $f_i(x)$ are pairwise distinct irreducibles, then

$$F[x]/(f(x)) \cong \prod_{i=1}^{k} F[x]/(f_i(x)^{n_i})$$

Proof idea. By the chinese remainder theorem. □

Theorem 3.3.4 (Gauss)**.** *If R is a unique factorization domain with fraction field F and $p(x) \in R[x]$ is reducible in $F[x]$, then $p(x)$ is reducible in $R[x]$. More specifically, if $p(x) = a(x)b(x)$ with $a(x), b(x) \in F[x]$, there exist $r, s \in F$ such that $ra(x), sb(x) \in R[x]$ and $p(x) = ra(x)sb(x)$.*

Conversely, if the greatest common divisor of the coefficients of $p(x)$ is 1 and $p(x)$ is reducible in $R[x]$, then $p(x)$ is reducible in $F[x]$.

Proof idea. For the forward direction, by clearing denominators in the factorization and then cancelling out each of the irreducible factors of the common denominator.

Let d be a common denominator of all of the coefficients of $a(x)$ and $b(x)$, so

$$dp(x) = a'(x)b'(x)$$

where $a'(x) \in R[x]$ and $b'(x) \in R[x]$ are R-multiples of $a(x)$ and $b(x)$, respectively. If d is not a unit, write $d = p_1 \cdots p_k$ where each p_i is irreducible, and hence prime, in R. For each i, $\overline{a'(x)b'(x)} = \overline{0}$ in the integral domain $(R/(p_i))[x]$, so p_i must divide one of $a'(x)$ or $b'(x)$ in $R[x]$. Therefore each p_i in the above equation can be cancelled out while keeping the equation in $R[x]$. It follows that $p(x)$ reduces in $R[x]$ into F-multiples of $a(x)$ and $b(x)$, as desired.

The reverse direction is trivial. □

Applications. Relating reducibility in $R[x]$ and $F[x]$, unique factorization in $R[x]$, transferring irreducibility criteria from $F[x]$ to $R[x]$, etc.

Corollary 3.3.4 (Polynomial rings over unique factorization domains)**.** *R is a unique factorization domain iff $R[x]$ is a unique factorization domain.*

Proof idea. By Gauss, pulling back unique factorization from $F[x]$ to $R[x]$. □

Theorem 3.3.5 (Linear factors and roots)**.** *If F is a field, $f(x) \in F[x]$ has a linear factor $(x - \alpha)$ iff $f(\alpha) = 0$.*

Proof idea. For the forward direction by substitution, and for the reverse direction by division with remainder. □

Corollary 3.3.5. *$f(x)$ has at most $\deg f(x)$ roots, even counting multiplicity.*

Corollary 3.3.6. *If $\deg f(x)$ is 2 or 3, $f(x)$ is irreducible in $F[x]$ iff it has no roots in F.*

Applications. Determining irreducibility of polynomials of low degree.

Theorem 3.3.6 (Rational roots)**.** *If R is a unique factorization domain with fraction field F and $p(x) = a_n x^n + \cdots + a_0 \in R[x]$ has a root $r/s \in F$ where $r, s \in R$ and $(r, s) = 1$, then $s|a_n$ and $r|a_0$ in R.*

Proof idea. By direct computation of $p(r/s)$ and clearing denominators. □

Applications. Finding rational roots, determining irreducibility of polynomials of low degree.

Theorem 3.3.7 (Reduction and irreducibility)**.** *If R is an integral domain, $f(x) \in R[x]$ is nonconstant monic, and I is a proper ideal in R such that $\overline{f(x)}$ cannot be properly factored in $(R/I)[x]$, then $f(x)$ is irreducible in $R[x]$.*

Applications. Determining irreducibility of polynomials.

Corollary 3.3.7 (Eisenstein)**.** *If R is an integral domain with fraction field F, $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in R[x]$, and P is a prime ideal in R such that $a_{n-1}, \ldots, a_0 \in P$ and $a_0 \notin P^2$, then $p(x)$ is irreducible in $R[x]$ and $F[x]$.*

Proof idea. By reducing coefficients mod P.

If $p(x) = a(x)b(x)$ in $R[x]$, then $\overline{a(x)b(x)} = \overline{x}^n$ in the integral domain $(R/P)[x]$. Therefore the constant terms of $\overline{a(x)}$ and $\overline{b(x)}$ must both be $\overline{0}$, that is, the constant terms of $a(x)$ and $b(x)$ must both be in P, so $a_0 \in P^2$—a contradiction. Irreducibility in $F[x]$ follows from Gauss. □

Applications. Determining irreducibility of polynomials.

Techniques

- Induction on degree of polynomials.
- Looking at leading and constant terms of products of polynomials.
- Division with remainder of polynomials.
- Reducing coefficients of polynomials.
- Transferring properties between R and $R[x]$:
  - R commutative $\Longleftrightarrow$ $R[x]$ commutative
  - R ID $\Longleftrightarrow$ $R[x]$ ID
  - R UFD $\Longleftrightarrow$ $R[x]$ UFD
  - R field $\Longleftrightarrow$ $R[x]$ PID $\Longleftrightarrow$ $R[x]$ ED

- Transferring properties between $R[x]$ and $F[x]$.
- Irreducibility criteria:
  - Roots (for polynomials of low degree)
  - Reducing coefficients
  - Eisenstein
  - Substitution

Chapter 4

Fields

This chapter covers field theory from [4].

4.1 Fields and Field Extensions

Theorems

Let F be a field.

Theorem 4.1.1 (Field homomorphisms)**.** *If R is a ring and $\varphi : F \to R$ is a nonzero ring homomorphism, then φ is injective.*

Proof idea. By the ideal structure of fields, $\ker \varphi = (0)$. □

Applications. Showing that we cannot study fields by taking quotients (as with groups or rings), but by embedding them in larger rings.

Theorem 4.1.2 (Tower law)**.** *If $F \subseteq L \subseteq K$ is a tower of fields, then*

$$[K : F] = [K : L][L : F]$$

In particular, $[L : F]$ divides $[K : F]$.

Proof idea. By direct computation with bases. □

Applications. Structure of extensions.

Theorem 4.1.3 (Composites of finite extensions)**.** *If K and L are finite over F (with both contained in some common extension), then KL is finite over F and*

$$[KL : F] \le [K : F][L : F]$$

If $[K : F]$ and $[L : F]$ are relatively prime, then equality holds.

Proof idea. By direct computation with bases and the tower law. □

Applications. Structure of extensions.

Theorem 4.1.4 (Prime subfield)**.** *F has a subfield isomorphic to either $\mathbb{F}_p$ (if $\operatorname{char} F = p$) or $\mathbb{Q}$ (if $\operatorname{char} F = 0$).*

Proof idea. Take the subfield generated by 1 in F. □

Applications. Viewing all fields as extension fields.

Theorem 4.1.5 (Existence of simple algebraic extensions)**.** *If $f(x) \in F[x]$ is nonconstant, there exists a field K extending F and containing a root α of $f(x)$.*

Proof idea. By construction, by adjoining a new element and taking a quotient to impose the root relation.

Assume without loss of generality that $f(x)$ is irreducible over F and set $K = F[x]/(f(x))$. Then K is a field containing an isomorphic copy of F, and if α denotes the image of x in K, $f(\alpha) = 0$ by construction. □

Applications. Adjoining a root of a polynomial, existence of splitting fields.

Theorem 4.1.6 (Structure of simple algebraic extensions)**.** *If $f(x) \in F[x]$ is irreducible over F with $n = \deg f(x)$, and $K = F[x]/(f(x))$ with α the image of x in K, then the elements $1, \alpha, \ldots, \alpha^{n-1}$ form a basis for K over F. In particular, $[K : F] = n$ and*

$$K = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} \mid a_0, \ldots, a_{n-1} \in F\}$$

Proof idea. For spanning, by division by $f(x)$ with remainder in $F[x]$, and for linear independence, by irreducibility of $f(x)$ over F. □

Applications. Computing in simple algebraic extensions, showing the finiteness of finitely generated algebraic extensions.

Theorem 4.1.7 (Structure of simple algebraic extensions)**.** *If $f(x) \in F[x]$ is irreducible over F and K is a field extending F and containing a root α of $f(x)$, then $F(\alpha) \cong F[x]/(f(x))$.*

Proof idea. By mapping.

Let $\varphi : F[x] \to F(\alpha)$ map $g(x) \mapsto g(\alpha)$. Then φ factors through $(f(x))$ to the field mapping

$$\overline{\varphi} : F[x]/(f(x)) \to F(\alpha)$$

Now $\overline{\varphi}$ is injective since it is nonzero, and it is surjective since its image is a subfield of K containing F and α, and hence $F(\alpha)$. □

Theorem 4.1.8 (Uniqueness of simple algebraic extensions)**.** *Let $\varphi : F \to F^*$ be an isomorphism, $f(x) \in F[x]$ be irreducible over F, and $f^*(x) \in F^*[x]$ correspond to $f(x)$ under φ. If α is any root of $f(x)$ in an extension K of F and β is any root of $f^*(x)$ in an extension K^* of F^*, then φ extends to an isomorphism $\Phi : F(\alpha) \to F^*(\beta)$:*

$$\begin{tikzcd}
\Phi: & F(\alpha) \arrow[r] \arrow[d, dash] & F^*(\beta) \arrow[d, dash] \\
\varphi: & F \arrow[r] & F^*
\end{tikzcd}$$

In particular, if α, β are any two roots of $f(x)$, $F(\alpha) \cong F(\beta)$.

Proof idea. By the structure of simple algebraic extensions.

Note φ naturally extends to the ring isomorphism $F[x] \to F^*[x]$ mapping $f(x) \mapsto f^*(x)$, so φ naturally extends to an isomorphism

$$F(\alpha) \cong F[x]/(f(x)) \cong F^*[x]/(f^*(x)) \cong F^*(\beta) \qquad \square$$

Applications. Algebraic indistinguishability of roots, uniqueness of splitting fields.

Theorem 4.1.9 (Minimal polynomial)**.** *If α is algebraic over F, there exists a unique monic irreducible polynomial $m_{\alpha,F}(x) \in F[x]$ such that $m_{\alpha,F}(\alpha) = 0$. If $f(x) \in F[x]$ and $f(\alpha) = 0$, then $m_{\alpha,F}(x)$ divides $f(x)$.*

Proof idea. Take $m_{\alpha,F}(x) \in F[x]$ monic of minimal degree with $m_{\alpha,F}(\alpha) = 0$. The division and uniqueness follow by division with remainder in $F[x]$. □

Applications. Structure of simple algebraic extensions, irreducibility criterion.

Theorem 4.1.10 (Finite simple extensions)**.** *The extension $F(\alpha)/F$ is finite iff α is algebraic over F, in which case*

$$[F(\alpha) : F] = \deg m_{\alpha,F}(x) = \deg_F \alpha$$

Proof idea. For the forward direction, by linear dependence. The $n+1$ elements $1, \alpha, \ldots, \alpha^n$ are linearly dependent, so there are nonzero $a_0, \ldots, a_n \in F$ such that

$$a_0 + a_1\alpha + \cdots + a_n\alpha^n = 0$$

Therefore α is algebraic over F.

For the reverse direction and degree equality, by the structure of simple algebraic extensions, since $F(\alpha) \cong F[x]/(m_{\alpha,F}(x))$. □

Applications. Relating finite and (finitely generated) algebraic extensions.

Corollary 4.1.1. *A finite extension is algebraic, and the degree of any element is at most the degree of the extension.*

Proof idea. If K/F is finite, then for any $\alpha \in K$, $F(\alpha)/F$ is finite and

$$\deg_F \alpha = [F(\alpha) : F] \le [K : F] \qquad \square$$

Theorem 4.1.11 (Finite extensions)**.** *An extension K/F is finite iff $K = F(\alpha_1, \ldots, \alpha_n)$ where $\alpha_1, \ldots, \alpha_n$ are algebraic over F, in which case*

$$[K : F] \le \prod_{i=1}^{n} \deg_F \alpha_i$$

Proof idea. For the forward direction, let $\alpha_1, \ldots, \alpha_n$ be a basis for K over F.

For the reverse direction, by induction using the tower law. Note the result holds in the simple case $n = 1$, and if $n > 1$,

$$K = F(\alpha_1, \ldots, \alpha_n) = F(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n)$$

so by induction $L = F(\alpha_1, \ldots, \alpha_{n-1})/F$ is finite and

$$[K:F] = [K:L][L:F] \leq \big(\deg_F \alpha_n\big)\Big(\prod_{i=1}^{n-1} \deg_F \alpha_i\Big) = \prod_{i=1}^{n} \deg_F \alpha_i \qquad \square$$

Remark. This result generalizes the result that a *simple* extension is finite iff it is algebraic by showing that an *arbitrary* extension is finite iff it is finitely generated algebraic. An algebraic extension need not be finite (for example, consider the field of real algebraic numbers).

Corollary 4.1.2 (Algebraic number fields)**.** *If K/F is arbitrary, the elements in K algebraic over F form a subfield of K.*

Proof idea. If α, β are algebraic over F, then $F(\alpha, \beta)$ is finite and hence algebraic over F and contains $\alpha \pm \beta$, $\alpha\beta$, and α/β (if $\beta \neq 0$). □

Corollary 4.1.3 (Algebraic extensions)**.** *If K is algebraic over L and L is algebraic over F, then K is algebraic over F.*

Proof idea. If $\alpha \in K$, α is the root of some polynomial $p(x) = a_n x^n + \cdots + a_0 \in L[x]$. Now

$$F \subseteq F(a_0, \ldots, a_n) \subseteq F(a_0, \ldots, a_n)(\alpha)$$

Each extension is finite, so by the tower law the full extension is finite and algebraic, so α is algebraic over F. □

Theorem 4.1.12 (Existence of splitting fields)**.** *If $f(x) \in F[x]$ is nonconstant, there exists a splitting field for $f(x)$ over F.*

Proof idea. By induction on the degree of $f(x)$, using existence of simple algebraic extensions. □

Applications. Adjoining all roots of a polynomial.

Theorem 4.1.13 (Uniqueness of splitting fields)**.** *Let $\varphi : F \to F^*$ be an isomorphism, $f(x) \in F[x]$ be nonconstant, and $f^*(x) \in F^*[x]$ correspond to $f(x)$ under φ. If E is any splitting field for $f(x)$ over F and E^* is any splitting field for $f^*(x)$ over F^*, then φ extends to an isomorphism $\Phi : E \to E^*$:*

$$\begin{tikzcd}
\Phi: & E \arrow[r] \arrow[d, dash] & E^* \arrow[d, dash] \\
\varphi: & F \arrow[r] & F^*
\end{tikzcd}$$

In particular, the splitting field for F is unique up to isomorphism.

Proof idea. By induction on the degree of $f(x)$, using uniqueness of simple algebraic extensions. □

Applications. Computing in familiar extensions, showing that separability is not an embedding property.

Theorem 4.1.14 (Existence of algebraic closures)**.** *F has an algebraic closure.*

Proof idea. Construct an algebraically closed field containing F one step at a time in countably many steps, at each step adjoining a new element for each nonconstant polynomial over the current field and taking a quotient to impose root relations, and then taking a union at the limit step. Finally, take the set of elements in the union which are algebraic over F.

Let $F_0 = F$ and for each $i \geq 0$ recursively define F_{i+1} from F_i as follows: introduce distinct indeterminates x_f for each nonconstant $f(x) \in F_i[x]$, and set

$$F_{i+1} = F_i[\ldots, x_f, \ldots]/M$$

where M is a maximal ideal containing the proper ideal $I = (\ldots, f(x_f), \ldots)$. Then F_{i+1} is an extension of F_i in which every nonconstant polynomial over F_i has a root. Take the union

$$K = \bigcup_{i=0}^{\infty} F_i$$

Then K is algebraically closed since any nonconstant polynomial over K has all its coefficients lying in some F_i, and hence a root lying in $F_{i+1} \subseteq K$. The set of elements in K algebraic over F is the closure of F. □

Theorem 4.1.15 (Uniqueness of algebraic closures)**.** *The algebraic closure of F is unique up to isomorphism.*

Theorem 4.1.16 (Fundamental theorem of algebra)**.** $\mathbb{C}$ *is algebraically closed.*

Remark. The previous two results show that when working with algebraic numbers over $\mathbb{Q}$, we might as well work in $\mathbb{C}$.

Theorem 4.1.17 (Separability and derivatives)**.** *A nonconstant polynomial $f(x) \in F[x]$ is separable iff $f(x)$ is relatively prime to its formal derivative $f'(x)$.*

Proof idea. By direct computation. □

Applications. Determining separability without going into the splitting field.

Theorem 4.1.18 (Separability)**.** *If either* char $F = 0$, *or* char $F = p$ *and F is finite, then a nonconstant polynomial over F is separable iff it is the product of distinct irreducible factors over F.*

Proof idea. If char $F = 0$, then by the formal derivative criterion any irreducible is separable, and distinct irreducibles never share a root (by the minimal polynomial).

If char $F = p$ and F is finite, suppose towards a contradiction that $f(x) \in F[x]$ is irreducible but inseparable. By the formal derivative criterion, $f'(x) = 0$. Since char $F = p$, $f(x) = q(x^p)$ for some $q(x) \in F[x]$. Since F is finite, by the Frobenius property every coefficient of $q(x)$ has a p-th root and $f(x)$ is actually a p-th power—a contradiction. □

Applications. Determining separability without going into the splitting field.

Techniques

- Relating ideal structure to global structure.
- Induction on degree of polynomials.
- Division with remainder of polynomials.
- Viewing all fields as extension fields.
- Using linear algebra (module theory) techniques to study field extensions.
- Using the tower law to determine structure of field extensions (cf. tower law for groups).
- Adjoining 'free' elements to fields and taking quotients by maximal ideals to construct extension fields satisfying desired relations (cf. free groups and group presentations).
- Using results about finite extensions to derive results about arbitrary algebraic extensions, by working with only finitely many elements at a time.
- Proceeding one step at a time in countably many steps.
- Doing computations involving algebraic elements in one large algebraically closed field.
- Using the minimal polynomial as an irreducibility criterion.
- Computing multiplicative inverses in field extensions:
  - Euclidean algorithm
  - Plugging and chugging into the minimal polynomial

Part II

Analysis

Chapter 5

Real Analysis

This chapter covers single-variable analysis from [8].

5.1 Real Numbers

Definitions

Basic order-theoretic and algebraic definitions assumed.

Definition 5.1.1. An ordered set S has the *least upper bound property* if every nonempty subset of S bounded above in S has a least upper bound (supremum) in S.

Theorems

Theorem 5.1.1 (Existence of $\mathbb{R}$)**.** *There exists an ordered field $\mathbb{R}$ (the real numbers) with the least upper bound property and containing $\mathbb{Q}$ as a subfield.*

Proof idea. Dedekind cuts or equivalence classes of Cauchy sequences over $\mathbb{Q}$. □

Applications. All subsequent theory.

Theorem 5.1.2 (Archimedean property of $\mathbb{R}$)**.** *For any $x, y \in \mathbb{R}$ with $x > 0$, there exists an integer $n > 0$ such that $nx > y$.*

Proof idea. By the least upper bound property.
If not, derive a contradiction from the supremum of the set $N = \{nx \mid n \in \mathbb{Z}\}$. □

Applications. Choosing integer bounds for real quantities as required in a variety of arguments, density of the rationals in the reals, etc.

Theorem 5.1.3 (Density of $\mathbb{Q}$ in $\mathbb{R}$)**.** *For any $x, y \in \mathbb{R}$ with $x < y$, there exists $q \in \mathbb{Q}$ with $x < q < y$.*

Proof idea. By the archimedean property.

Desire $m, n \in \mathbb{Z}$ with $n > 0$ such that $x < m/n < y$ or $nx < m < ny$. Choose n by the archimedean property so that $ny - nx = n(y-x) > 1$, to ensure that an integer will appear between nx and ny. Let m to be the least integer greater than nx (use the archimedean property, again) and argue that m/n works. □

Theorem 5.1.4 (n-th roots in $\mathbb{R}$)**.** *For any $x \in \mathbb{R}$ with $x \geq 0$, there exists a unique $y \in \mathbb{R}$ with $y \geq 0$ and $y^n = x$.*

Proof idea. By the least upper bound property.

Set $y = \sup\{y \mid y^n < x\}$. Using a bound for $b^n - a^n$ in terms of $a, b \in \mathbb{R}$ with $a < b$, argue by contradiction that neither $y^n < x$ nor $y^n > x$ can hold, lest y fails to be a least upper bound. □

Theorem 5.1.5 (Existence of $\mathbb{C}$)**.** *There exists an algebraically closed field $\mathbb{C}$ (the complex numbers) containing $\mathbb{R}$ as a subfield.*

Proof idea. Let complex numbers be ordered pairs of real numbers with addition and multiplication defined appropriately. Prove algebraic closure later. □

Theorem 5.1.6 (Cauchy-Schwarz)**.** *Given complex numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$,*

$$|\sum_{k=1}^{n} a_k \overline{b_k}|^2 \leq \sum_{k=1}^{n} |a_k| \cdot \sum_{k=1}^{n} |b_k|$$

Proof idea. Use a proof from linear algebra.

Write $\boldsymbol{a} = (a_1, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, \ldots, b_n)$. In the complex inner product space $\mathbb{C}^n$ with standard inner product $\boldsymbol{x} \bullet \boldsymbol{y} = \sum_{k=1}^{n} x_k \overline{y_k}$ and norm $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \bullet \boldsymbol{x}}$, the inequality can be expressed equivalently as

$$|\boldsymbol{a} \bullet \boldsymbol{b}|^2 \leq \|\boldsymbol{a}\|^2 \, \|\boldsymbol{b}\|^2$$

To prove this, assume $\boldsymbol{b} \neq \boldsymbol{0}$ and orthogonally decompose $\boldsymbol{a}$ in terms of $\boldsymbol{b}$ as

$$\boldsymbol{a} = \frac{\boldsymbol{a} \bullet \boldsymbol{b}}{\|\boldsymbol{b}\|^2} \boldsymbol{b} + \boldsymbol{w}$$

Then appeal to the Pythagorean theorem to obtain

$$\|\boldsymbol{a}\|^2 = \frac{(\boldsymbol{a} \bullet \boldsymbol{b})^2}{\|\boldsymbol{b}\|^2} + \|\boldsymbol{w}\|^2 \geq \frac{(\boldsymbol{a} \bullet \boldsymbol{b})^2}{\|\boldsymbol{b}\|^2}$$

from which the above inequality is immediate. □

Applications. Triangle inequality in $\mathbb{R}^k$ (allowing us to treat $\mathbb{R}^k$ as a metric space), mean value inequality in $\mathbb{R}^k$, integral bounds in $\mathbb{R}^k$, etc.

Techniques

- Least upper bound property.
- Archimedean property.

5.2 Basic Topology

Definitions

Definition 5.2.1. A *metric space* is a pair (X, d) where X is a set and $d : X \times X \to \mathbb{R}$ is a distance function satisfying positive definiteness, symmetry, and triangle inequality.

Definition 5.2.2. Assume a background metric space.

(a) A *neighborhood* of p is a set $N_\delta(p) = \{ q \mid d(p, q) < \delta \}$ for some $\delta > 0$.

(b) A point p is a *limit point* of a set E if every neighborhood of p contains some point of E other than p.

(c) A set E is *open* if every point in E has some neighborhood contained in E.

(d) A set E is *closed* if every limit point of E is in E.

(e) A set E is *bounded* if there is some $\delta > 0$ such that $d(p, q) < \delta$ for all points $p, q \in E$.

Definition 5.2.3. A set E in a metric space is *compact* if every open covering of E has a finite subcovering.

Definition 5.2.4. A set in a metric space is *connected* if it is not the union of two nonempty *separated* sets, that is, sets each of whose closure is disjoint from the other set.

Theorems

Theorem 5.2.1 (Basics of infinite cardinality)**.**

(a) Infinite subsets of countable sets are countable.

(b) Countable unions of countable sets are countable.

(c) Finite products of countable sets are countable.

(d) $\mathbb{Q}$ *is countable.*

(e) $\mathbb{R}$ *is uncountable.*

Proof idea. For (a), given a sequence for the set, recursively define a subsequence for the subset.

For (b), draw arrows.

For (c), draw arrows then use induction.

For (d), note that rationals can be represented by pairs of integers.

For (e), diagonalize on decimal expansions. □

Theorem 5.2.2. *If X is a metric space, $E \subseteq X$, and $p \in X$ a limit point of E, then every neighborhood of p contains infinitely many points of E.*

Proof idea. If not, take a neighborhood of p containing only finitely many points of E. Then the neighborhood of p whose radius is the minimum disance between these points and p contains no points of E other than possibly p—a contradiction. □

Theorem 5.2.3 (Basics of open and closed sets)**.** *Assume a background metric space.*

(a) A set is open [closed] iff its complement is closed [open].

(b) A union of open sets is open, and a finite intersection of open sets is open.

(c) An intersection of closed sets is closed, and a finite union of closed sets is closed.

(d) Every set has a closure.

Proof idea. Note (a) follows from definitions.

For (b), let S be a collection of open sets. To see that $\bigcup S$ is open, note if $p \in \bigcup S$, then $p \in E$ for some $E \in S$, so p is interior to $E \subseteq \bigcup S$, that is, p is interior to $\bigcup S$. If S is finite, say $S = \{E_1, \ldots, E_k\}$, then if $p \in \bigcap S$ there exist neighborhoods $N_1, \ldots, N_k$ of p with $N_i \subseteq E_i$ for $1 \leq i \leq k$. Now $N = \bigcap N_i$ is a neighborhood of p, and $N \subseteq \bigcap S$. This shows $\bigcap S$ is open.

Now (c) follows from (a) and (b).

For (d), given E, let $\overline{E}$ be the intersection of all closed sets containing E. It is easy to verify that $\overline{E} = E \cup E'$ where E' is the set of limit points of E. □

Theorem 5.2.4 (Invariance of compactness)**.** *Let X be a metric space and $E \subseteq Y \subseteq X$. Then E is compact relative to Y iff E is compact relative to X.*

Proof idea. Note a set F is open relative to Y iff $F = G \cap Y$ with G open relative to X.

Now suppose E is compact relative to Y. If $\{F_\alpha\}$ is a covering of E open relative to X, then $\{F_\alpha \cap Y\}$ is a covering of E open relative to Y. The latter contains a finite subcovering of E, and hence so does the former.

The converse is similar. □

Applications. Showing that compactness is not an embedding property, allowing us to speak meaningfully of a compact metric space.

Theorem 5.2.5. *A set E is compact iff every infinite subset of E has a limit point in E.*

Proof idea. If $F \subseteq E$ is infinite with no limit point in E, then for each $p \in E$ there is a neighobrhood N_p of p containing at most one point of F. But then $\{N_p\}$ is an open covering of E with no finite subcovering, so E is not compact.

Suppose towards a contradiction that every infinite subset of E has a limit point in E but E is not compact. Fix an open covering $\{E_\alpha\}$ with no finite subcovering. Note E must be infinite and bounded. Recursively define a descending chain of subsets. First fix $\delta > 0$ and $p_1 \in E$ with $N_1 = N_\delta(p_1) \supseteq E$. Now assume a point p_k and neighborhood N_k of p_k with radius $\delta/2^{k-1}$ has been defined such that

(i) $p_k \neq p_1, \ldots, p_{k-1}$.

(ii) $p_k \in \bigcap_{i=1}^k N_i$.

(iii) $\bigcap_{i=1}^k N_i$ has no finite subcovering.

Write $P_k = \bigcap_{i=1}^k N_i$.

To define N_{k+1}, set $r = \delta/2^k$ and proceed as follows: choose $q_1 \in P_k - \{p_1, \ldots, p_k\}$ and consider $N_r(q_1)$. If $N_r(q_1) \cap P_k$ has no finite subcovering, set $q = q_1$ and stop.

Otherwise, there must be $q_2 \in P_k - (N_r(q_1) \cup \{p_1, \ldots, p_k\})$. Now repeat this process with q_2... Eventually there must be some $q \in P_k - \{p_1, \ldots, p_k\}$ such that $N_r(q) \cap P_k$ has no finite subcovering, lest P_k will be covered by finitely many r-neighborhoods each of whose restrictions to P_k has finite subcovering, so P_k has finite subcovering—contradicting (iii). Set $p_{k+1} = q$ and $N_{k+1} = N_r(q)$. Then (i), (ii), (iii) hold at $k+1$.

By recursion, this yields a descending chain

$$P_1 \supseteq P_2 \supseteq \cdots \supseteq P_k \supseteq \cdots$$

where for all k, $P_k \subseteq N_k$ and P_k has no finite subcovering. Also, $\{p_i\}$ is an infinite sequence with $p_k \in P_k$ for all k. Let p be a limit point for $\{p_i\}$. Then p lies in some E_α. But then by choosing k large enough, we have $P_k \subseteq N_k \subseteq E_\alpha$—contradicting that P_k has no finite subcovering. □

Applications. Showing 'finiteness' of compactness, Bolzano-Weierstrass theorem.

Theorem 5.2.6. *If E is compact, then E is closed and bounded.*

Proof idea. To see that E is bounded, fix a radius $r > 0$ and for each $p \in E$ let N_p be a neighborhood of p with radius r. Then $\{N_p\}$ is an open covering of E, and has a finite subcovering $\{N_{p_1}, \ldots, N_{p_k}\}$. Let d be the maximum distance between pairs of the points $p_1, \ldots, p_k$. Then it is immediate that the distance between any two points in E is bounded by $M = d + 2r$, so E is bounded.

To see that E is closed, observe that E^c is open. Given $q \in E^c$, choose for each $p \in E$ a neighborhood N_p of p and a neighborhood M_p of q such that N_p and M_p are disjoint. Again, $\{N_p\}$ is an open covering of E with a finite subcovering $\{N_{p_1}, \ldots, N_{p_k}\}$. Now $M = \bigcap M_{p_i}$ is a neighborhood of q disjoint from E, so q is interior to E^c. □

Applications. Showing 'finiteness' of compactness, Heine-Borel theorem, extreme value theorem.

Remark. The converse is false! But it is true in $\mathbb{R}^k$ (see the Heine-Borel theorem).

Theorem 5.2.7. *Closed subsets of compact sets are compact.*

Proof idea. Add the complement of the subset to any open covering. □

Theorem 5.2.8 (Descending chains of compact sets)**.** *Let $\{K_i\}$ be a sequence of nonempty compact sets with $K_i \supseteq K_{i+1}$ for all $i \geq 1$. Then $\bigcap K_i$ is nonempty.*

Proof idea. Recursively construct an infinite sequence of points and take a limit.

Suppose $\bigcap K_i$ is empty. Fix an arbitrary point $p_1 \in K_1$. Now assuming that p_k is defined, observe that there must exist some K_i with $p_k \notin K_i$. Choose $p_{k+1} \in K_i$. By recursion, $\{p_k\}$ is an infinite sequence of points in K_1. By compactness of K_1, there exists a limit point p of $\{p_k\}$ in K_1. Claim that $p \in \bigcap K_i$. Indeed, if not, there is some K_i with $p \notin K_i$. But then p is not a limit point of K_i, so p cannot be a limit point of $\{p_k\}$, almost all of whose members are in K_i—a contradiction. □

The remaining results concern the topology of $\mathbb{R}^k$.

Theorem 5.2.9 (Descending chains of k-cells in $\mathbb{R}^k$). *Let $\{I_i\}$ be a sequence of k-cells in $\mathbb{R}^k$ with $I_i \supseteq I_{i+1}$ for all $i \geq 1$. Then $\bigcap I_i$ is nonempty.*

Proof idea. By the least upper bound property.

First prove case $k = 1$, taking the least upper bound of the left endpoints of the intervals. Then reduce case $k > 1$ to case $k = 1$ on each coordinate. □

Applications. The previous two results are useful in arguments where descending chains are constructed (for example, by subdivision), often to derive a contradiction with a limit element. Heine-Borel theorem, Bolzano-Weierstrass theorem, etc.

Theorem 5.2.10 (Compactness of k-cells in $\mathbb{R}^k$). *In $\mathbb{R}^k$, k-cells are compact.*

Proof idea. If not, choose a witness k-cell and covering, and recursively subdivide the k-cell to generate an infinite descending chain of k-cells none of which has a finite subcovering. By the previous theorem, there is a point in the intersection, which is covered by some open set. But then sufficiently small k-cells in the chain are also covered by the set—a contradiction. □

Remark. By the above, this theorem is actually equivalent to the previous theorem. The proof technique of recursive subdivision is very powerful and reusable.

Corollary 5.2.1 (Heine-Borel). *In $\mathbb{R}^k$, a set is compact iff it is closed and bounded.*

Proof idea. The forward direction is known. For the reverse direction, note such a set is contained in a compact k-cell, hence is compact. □

Applications. Extreme value theorem.

Corollary 5.2.2 (Bolzano-Weierstrass). *Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.*

Proof idea. Such a set is contained in a compact k-cell, hence has a limit point. □

Theorem 5.2.11 (Connectedness in $\mathbb{R}$). *$E \subseteq \mathbb{R}$ is connected iff for every $x, y \in E$ and $z \in \mathbb{R}$, if $x < z < y$ then $z \in E$.*

Proof idea. If $x, y \in E$ and $z \in \mathbb{R}$ with $x < z < y$ but $z \notin E$, then

$$E = ((-\infty, z) \cap E) \cup (E \cap (z, \infty))$$

is a union of nonempty separated sets, so E is not connected.

Conversely, if $E = X \cup Y$ with X and Y nonempty and separated, take $x \in X$ and $y \in Y$ and, assuming $x < y$, consider $z = \sup(X \cap [x, y])$. □

Applications. Intermediate value theorem.

Techniques

- Diagonalization.
- Covering compact sets with finitely many neighborhoods.
- Subdivision to construct descending chains and take limit points.

5.3 Sequences and Series of Numbers

Definitions

Sequences

Definition 5.3.1. A *sequence* in a set X is a function $x : \mathbb{N} \to X$, usually denoted $\{x_n\}$ with $x_n = x(n)$. If $\sigma : \mathbb{N} \to \mathbb{N}$ is increasing, the sequence $x \circ \sigma$ is a *subsequence* of x.

A sequence is *finite*, *infinite*, etc. according as its range satisfies these properties.

Definition 5.3.2. In a metric space, a sequence $\{x_n\}$ *converges* to a point p if for all $\epsilon > 0$, there exists N such that for all $n \geq N$, $d(x_n, p) < \epsilon$. In this case p is called a *limit* of $\{x_n\}$, denoted $p = \lim x_n$ or $x_n \to p$. If $\{x_n\}$ does not have a limit, it *diverges*.

A limit of a subsequence of $\{x_n\}$ is a *subsequential limit* of $\{x_n\}$.

Definition 5.3.3. In a metric space, a sequence $\{x_n\}$ is a *Cauchy sequence* if for all $\epsilon > 0$, there exists N such that for all $m, n \geq N$, $d(x_m, x_n) < \epsilon$.

Definition 5.3.4. A metric space in which every Cauchy sequence converges is *complete*.

Definition 5.3.5. In $\mathbb{R}$, a sequence $\{x_n\}$ is *monotonically increasing* if $x_n \leq x_{n+1}$ for all n, and *monotonically decreasing* if $x_n \geq x_{n+1}$ for all n. A sequence is *monotonic* if it is monotonically increasing or monotonically decreasing.

Definition 5.3.6. In $\mathbb{R}$, for a sequence $\{x_n\}$, if for all M there exists N such that for all $n \geq N$, $x_n > M$, then $\{x_n\}$ *diverges to* $+\infty$, written $x_n \to +\infty$. Analogously, we may have $\{x_n\}$ *diverges to* $-\infty$, written $x_n \to -\infty$.

Definition 5.3.7. In $\mathbb{R}$, for a sequence $\{x_n\}$, the set of *extended subsequential limits* of $\{x_n\}$ includes the set of subsequential limits of $\{x_n\}$ together with possibly $+\infty$ or $-\infty$ according as $\{x_n\}$ has subsequences diverging to $+\infty$ or $-\infty$, respectively.

Let E be the set of extended subsequential limits of $\{x_n\}$. Then define

$$\liminf x_n = \inf E \qquad \limsup x_n = \sup E$$

with supremum and infimum taken in $\mathbb{R} \cup \{\pm\infty\}$. These are called the *lower limit* and *upper limit* of $\{x_n\}$, respectively.

Series

Definition 5.3.8. Let $\{a_n\}$ be a sequence in $\mathbb{C}$. If $s_n = \sum_{k=0}^{n} a_k$, then $\{s_n\}$ is called the *series* of $\{a_n\}$, denoted $\sum a_n$. Each s_n is a *partial sum* of $\{a_n\}$, and each a_n is a *term* of $\sum a_n$. If $\sum a_n$ converges to s, we also informally identify $\sum a_n$ with s.

Definition 5.3.9. The series $\sum a_n$ converges *absolutely* if $\sum |a_n|$ converges. $\sum a_n$ converges *conditionally* if $\sum a_n$ converges but not absolutely.

Definition 5.3.10. The series $\sum a_n$ is *alternating* if terms have alternating sign (in $\mathbb{R}$).

Definition 5.3.11. If $\sum a_n$ is a series and $\pi : \mathbb{N} \to \mathbb{N}$ is a bijection, $\sum a_{\pi(n)}$ is a *rearrangement* of $\sum a_n$.

Definition 5.3.12. The series $\sum c_n z^n$ (a function of z) is a *power series*.

Theorems

Sequences

Theorem 5.3.1 (Basics of limits of sequences)**.** *Assume a background metric space.*

(a) *Limits of sequences are unique.*

(b) *p is the limit of $\{x_n\}$ iff every neighborhood of p contains x_n for almost all n.*

(c) *p is a limit point of a set E iff p is the limit of an infinite sequence in E.*

(d) *p is a limit point of the range of $\{x_n\}$ iff p is the limit of an infinite subsequence of $\{x_n\}$.*

(e) *Convergent sequences are bounded.*

(f) *The set of subsequential limits of a sequence is closed.*

Proof idea. For (a), observe two limits must be arbitrarily close, and hence equal.

Note (b) follows from definitions.

For (c), if p is a limit point of E, choose pairwise distinct $x_k \in N_{\delta_k}(p) \cap E$ for $\delta_k = 1/k$ with $k = 1, 2, \ldots$; use (b) for the converse.

Note (d) follows from (c).

Note (e) follows from (b).

For (f), a limit point of subsequential limits is itself a subsequential limit. □

Theorem 5.3.2 (Limits and field operations in $\mathbb{C}$)**.** *Let $\{a_n\}, \{b_n\}$ be convergent sequences in $\mathbb{C}$.*

(a) $\lim(a_n + b_n) = \lim a_n + \lim b_n$

(b) $\lim(a_n b_n) = \lim a_n \cdot \lim b_n$

(c) $\lim(1/a_n) = 1/\lim a_n$ *provided $a_n \neq 0$ for $n = 1, 2, \ldots$ and* $\lim a_n \neq 0$.

Proof idea. Write $a = \lim a_n$ and $b = \lim b_n$.

For (a), write $(a_n + b_n) - (a + b) = (a_n - a) + (b_n - b)$ and let $n \to \infty$.

For (b), write $a_n b_n - ab = (a_n - a)(b_n - b) + a(b_n - b) + b(a_n - a)$ and let $n \to \infty$.

For (c), write $1/a_n - 1/a = (a - a_n)/a_n a$ and let $n \to \infty$.

(Formally, use epsilons!) □

Theorem 5.3.3 (Limits in $\mathbb{R}^k$)**.** *Let $\{\boldsymbol{x}_n\}, \{\boldsymbol{y}_n\}$ be sequences in $\mathbb{R}^k$.*

(a) *If $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,k})$ and $\boldsymbol{x} = (x_1, \ldots, x_k)$, then*

$$\boldsymbol{x} = \lim \boldsymbol{x}_n \quad \Longleftrightarrow \quad x_i = \lim x_{n,i} \ (1 \leq i \leq k)$$

(b) *If $\{\boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_n\}$ are convergent and $\alpha = \lim \alpha_n$ in $\mathbb{R}$, then*

(i) $\lim(\alpha_n \boldsymbol{x}_n) = \alpha \lim \boldsymbol{x}_n$

(ii) $\lim(\boldsymbol{x}_n + \boldsymbol{y}_n) = \lim \boldsymbol{x}_n + \lim \boldsymbol{y}_n$

(iii) $\lim(\boldsymbol{x}_n \bullet \boldsymbol{y}_n) = \lim \boldsymbol{x}_n \bullet \lim \boldsymbol{y}_n$

Proof idea. For (a), use epsilons and the definition of the norm in $\mathbb{R}^k$.

Note (b) follows from (a) and properties of limits in $\mathbb{R}$. □

Applications. Characterizing properties of $\mathbb{R}^k$-valued sequences [functions] in terms of properties of $\mathbb{R}$-valued component sequences [functions], and thereby extending results from $\mathbb{R}$ to $\mathbb{R}^k$ by applying them to components. For functions, this includes limits, continuity, differentiability, etc.

Theorem 5.3.4 (Limits and compactness)**.** *Every sequence in a compact metric space has a convergent subsequence.*

Proof idea. If the sequence is finite, the result is trivial. If the sequence is infinite, it has a limit point (of its range) by compactness, hence a subsequential limit. □

Corollary 5.3.1 (Bolzano-Weierstrass)**.** *Every bounded sequence in* $\mathbb{R}^k$ *has a convergent subsequence.*

Proof idea. Such a sequence is contained in a compact k-cell. □

Remark. These results are essentially reformulations of results from the previous chapter into the language of sequences.

Theorem 5.3.5 (Cauchy sequences)**.** *Assume a background metric space.*

(a) Convergent sequences are Cauchy.

(b) If the space is compact, Cauchy sequences are convergent (that is, compact spaces are complete).

Proof idea. For (a), if points are getting arbitrarily close to a limit, they are getting arbitrarily close to each other.

For (b), choose a convergent subsequence, then argue all points of the sequence are getting arbitrarily close to the subsequential limit. □

Corollary 5.3.2 (Completeness of $\mathbb{R}^k$)**.** $\mathbb{R}^k$ *is complete.*

Proof idea. Cauchy sequences are bounded, so contained in a compact k-cell. □

Applications. Cauchy criterion for convergence of a sequence [series].

Theorem 5.3.6 (Monotonic sequences in $\mathbb{R}$)**.** *If a sequence is monotonic, it is convergent iff it is bounded.*

Proof idea. The forward direction is known. For the reverse direction, use the least upper bound or greatest lower bound property. □

Applications. Convergence criterion for series of nonnegative terms.

Theorem 5.3.7 (Lower and upper limits in $\mathbb{R}\cup\{\pm\infty\}$)**.** *Let* $\{x_n\}$ *be a sequence in* $\mathbb{R}$ *and*

$$\alpha = \liminf x_n \qquad \beta = \limsup x_n$$

(a) α *and* β *are extended subsequential limits of* $\{x_n\}$*.*

(b) If $\gamma < \alpha$, then $\gamma < x_n$ for almost all n.

(c) If $\gamma > \beta$, then $\gamma > x_n$ for almost all n.

Moreover, α and β are unique in satisfying these properties.

Proof idea. For (a), note that either α and β are extended limit points of the set of extended subsequential limits of $\{x_n\}$, or else they are the only elements in the set, hence are extended subsequential limits themselves.

For (b), if $x_n \le \gamma$ for infinitely many n, then $\{x_n\}$ has an extended subsequential limit $\lambda \le \gamma$, so $\alpha \le \lambda \le \gamma < \alpha$—a contradiction.

For (c), argue similarly.

Uniqueness is immediate. □

Series

All series are in $\mathbb{C}$ unless otherwise implied. Ordering of terms (e.g. nonnegativity) always implies those terms are in $\mathbb{R}$.

Theorem 5.3.8 (Cauchy criterion)**.** *The series $\sum a_n$ converges iff for all $\epsilon > 0$, there exists N such that for all $n \ge m \ge N$,*

$$|\sum_{k=m}^{n} a_k| < \epsilon$$

Proof idea. This just means the sequence of partial sums is Cauchy. □

Applications. Proving convergence without reference to a limit.

Corollary 5.3.3 (Divergence test)**.** *If $\sum a_n$ converges, then* $\lim a_n = 0$.

Applications. Proving divergence.

Remark. The converse is false! Consider the harmonic series $\sum 1/n$.

Theorem 5.3.9. *A series of nonnegative terms converges iff its sequence of partial sums is bounded.*

Proof idea. The forward direction is known, and the reverse direction holds since the partial sums monotonically increase. □

Theorem 5.3.10 (Geometric series)**.** *If $0 \le x < 1$, $\sum x^n = 1/(1-x)$. If $x \ge 1$, $\sum x^n$ diverges.*

Proof idea. Write $s_n = \sum_{k=0}^{n} x^k = 1 + \cdots + x^n$. Then $xs_n = s_{n+1} - 1$, so if $x \ne 1$,

$$s_n = \frac{1 - x^{n+1}}{1 - x}$$

Thus $s_n \to 1/(1-x)$ as $n \to \infty$.

If $x = 1$, $s_n = n \to \infty$ as $n \to \infty$, so divergence follows from the divergence test. □

Applications. Often used with the comparison test for convergence.

Theorem 5.3.11 (Comparison test)**.** *Let $\sum a_n$ and $\sum b_n$ be series.*

(a) *If $|a_n| \le b_n$ for almost all n and $\sum b_n$ converges, then $\sum a_n$ converges.*

(b) *If $a_n \ge b_n \ge 0$ for almost all n and $\sum b_n$ diverges, then $\sum a_n$ diverges.*

Proof idea. For (a), use the Cauchy criterion.

Note (b) follows from (a). □

Applications. Proving convergence and divergence, root test, ratio test.

Corollary 5.3.4 (Root test)**.** *Let $\sum a_n$ be a series and $\lambda = \limsup \sqrt[n]{|a_n|}$.*

(a) *If $\lambda < 1$, then $\sum a_n$ converges.*

(b) *If $\lambda > 1$, then $\sum a_n$ diverges.*

Proof idea. For (a), by comparison with a geometric series. If $\lambda < 1$, fix $\lambda < \beta < 1$. Then by the upper limit property for λ, $\sqrt[n]{|a_n|} < \beta$ for almost all n, that is, $|a_n| < \beta^n$ for almost all n. But $\sum \beta^n$ is a convergent geometric series, so $\sum a_n$ converges by the comparison test.

For (b), by the divergence test. If $\lambda > 1$, there are infinitely many a_n with $|a_n| > 1$, so $\lim a_n \neq 0$. □

Corollary 5.3.5 (Ratio test)**.** *Let $\sum a_n$ be a series and $\lambda = \limsup |\frac{a_{n+1}}{a_n}|$.*

(a) *If $\lambda < 1$, then $\sum a_n$ converges.*

(b) *If $|a_{n+1}/a_n| \ge 1$ for almost all n, then $\sum a_n$ diverges.*

Proof idea. For (a), by comparison with a geometric series. If $\lambda < 1$, fix $\lambda < \beta < 1$ and N such that $|a_{n+1}/a_n| < \beta$ for all $n \ge N$. Then

$$|a_{N+k}| < \beta |a_{N+k-1}| < \beta^2 |a_{N+k-2}| < \cdots < \beta^k |a_N|$$

for all $k \ge 1$. Convergence follows by comparison with $\sum \beta^n$.

For (b), by the divergence test. □

Remark. The comparison, root, and ratio tests are tests for absolute convergence. The root test is better than the ratio test. The root and ratio tests give no information when $\lambda = 1$, as illustrated by the series $\sum 1/n$ and $\sum 1/n^2$.

Theorem 5.3.12 (Cauchy subseries theorem)**.** *Let $\sum a_n$ be a series with $a_1 \ge a_2 \ge \cdots \ge 0$. Then $\sum a_n$ converges iff the series $\sum 2^n a_{2^n}$ converges.*

Proof idea. Since both of the series have nonnegative terms, it is sufficient to show equivalence for boundedness of partial sums.

Show that for any given partial sum of one series, you can go far enough out in the other series to surpass it (up to a multiple of 2). Therefore one sequence is unbounded iff the other sequence is unbounded. □

Applications. $\sum \frac{1}{n^k}$ converges for $k > 1$ and diverges for $k \le 1$.

Theorem 5.3.13 (Series and field operations)**.** *Let $\sum a_n$ and $\sum b_n$ be series.*

(a) If $\sum_{n=0}^{\infty} a_n$ is convergent, then for all $N \ge 0$, $\sum_{n=N+1}^{\infty} a_n$ is convergent and

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{N} a_n + \sum_{n=N+1}^{\infty} a_n$$

(b) If $\sum a_n$ and $\sum b_n$ are convergent, then

$$\sum (a_n + b_n) = \sum a_n + \sum b_n$$

(c) If $\sum a_n$ is absolutely convergent, $\sum b_n$ is convergent, and $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, then

$$\sum c_n = \left(\sum a_n\right)\left(\sum b_n\right)$$

(d) If $\sum a_n$ has bounded partial sums and $b_1 \ge b_2 \ge \cdots \ge 0$ and $\lim b_n = 0$, then $\sum a_n b_n$ converges.

Proof idea. For (a) and (b), use the addition rule for limits of the partial sums.

For (c), let A_n and A be the partial sum and sum of $\sum a_n$, respectively; similarly use B_n, B for $\sum b_n$ and C_n for $\sum c_n$. Then

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

Set $\delta_n = a_0(B - B_n) + \cdots + a_n(B - B_0)$. Now argue using absolute convergence of $\sum a_n$ and convergence of $\sum b_n$ that $\delta_n \to 0$ as $n \to \infty$, so $C_n \to AB$ as $n \to \infty$.

For (d), use summation by parts to express a partial sum of $\sum a_n b_n$ in terms of a partial sum of $\sum a_n$ and terms of $\sum b_n$, then use the Cauchy criterion. □

Corollary 5.3.6 (Alternating series)**.** *Let $\sum c_n$ be an alternating series with $|c_1| \ge |c_2| \ge \cdots$ and $\lim |c_n| = 0$. Then $\sum c_n$ converges.*

Proof idea. Use (c) with $a_n = (-1)^{n+1}$ and $b_n = |c_n|$. □

Theorem 5.3.14 (Rearrangements)**.** *Let $\sum a_n$ be a series.*

(a) If $\sum a_n$ is conditionally convergent and $-\infty \le \alpha \le \beta \le +\infty$, then there exists a rearrangement $\sum a_n'$ with

$$\liminf s_n' = \alpha \qquad \limsup s_n' = \beta$$

where $\{s_n'\}$ is the sequence of partial sums.

(b) If $\sum a_n$ is absolutely convergent, every rearrangement converges (absolutely) to the same sum.

Proof idea. For (a), argue that the sum of the positive terms in $\sum a_n$ must diverge to $+\infty$, and the sum of the negative terms to $-\infty$, lest $\sum a_n$ is absolutely convergent.

Now choose sequences $\{\alpha_n\}$ and $\{\beta_n\}$ with $\alpha_n \to \alpha$ and $\beta_n \to \beta$ and $\alpha_n < \beta_n$ for all n. Construct a sequence of partial sums for a rearrangement as follows: first take positive terms from $\sum a_n$ (in their original order) until the partial sum is just greater than β_1; then take negative terms until the partial sum is just less than α_1; and so on... This process can be carried out indefinitely since we have enough positive and negative terms. The resulting rearrangement has the desired properties.

For (b), use the Cauchy criterion, and go sufficiently far out in the terms of any rearrangement. □

Theorem 5.3.15 (Radius of convergence)**.** *Let $\sum c_n z^n$ be a power series and $R = 1/\limsup \sqrt[n]{|c_n|}$. Then the series converges (absolutely) for $|z| < R$ and diverges for $|z| > R$.*

Proof idea. By the root test. □

Techniques

- Lower and upper limit properties.
- Proving convergence or divergence of sequences:
  - Cauchy criterion.
  - Monotonicity.
  - Arithmetic (sums, products, etc.).
- Proving convergence or divergence of series:
  - Known forms (geometric series, alternating series, etc.).
  - Divergence test.
  - Cauchy criterion.
  - Comparison test.
  - Root test.
  - Ratio test.
  - Boundedness (for nonnegative terms).
  - Cauchy subseries (for nonnegative, monotonically decreasing terms).
  - Arithmetic (sums, summation by parts, products, etc.).
  - Using properties of power series (continuity, etc.).
- Reducing sequences in $\mathbb{R}^k$ to sequences (of components) in $\mathbb{R}$.

5.4 Continuity

Definitions

Definition 5.4.1. Let X, Y be metric spaces, $E \subseteq X$, and $f : E \to Y$. Let $p \in X$ be a limit point of E and $q \in Y$. If for every $\epsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(x), q) < \epsilon$

for all $x \in E$ with $0 < d_X(x,p) < \delta$, then q is a *limit* of f at p, denoted $q = \lim_{x\to p} f(x)$ or $f(x) \to q$ as $x \to p$.

Definition 5.4.2. Let X, Y be metric spaces, $E \subseteq X$, and $f : E \to Y$. Let $p \in E$. Then f is *continuous at* p if for every $\epsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \epsilon$ for all $x \in E$ with $d_X(x,p) < \delta$.

If f is not continuous at p (but defined at p), f is *discontinuous at* p.

If f is continuous at every point of E, then f is *continuous (on E)*.

Definition 5.4.3. Let X, Y be metric spaces, $E \subseteq X$, and $f : E \to Y$. Then f is *uniformly continuous (on E)* if for every $\epsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ for all $x, y \in E$ with $d_X(x,y) < \delta$.

Definition 5.4.4. Let Y be a metric space and $f : (a,b) \to Y$. Let $p \in (a,b)$ and $q \in Y$. If for every $\epsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(x), q) < \epsilon$ for all $x \in (p-\delta, p)$, then q is a *left-hand limit* of f at p, denoted $q = f(p-)$. Analogously for *right-hand limit* $f(p+)$.

Theorems

Theorem 5.4.1 (Characterization of limits by sequences)**.** *Let X, Y be metric spaces, $E \subseteq X$, and $f : E \to Y$. Let $p \in X$ be a limit point of E and $q \in Y$. Then*

$$\lim_{x\to p} f(x) = q \qquad \Longleftrightarrow \qquad \begin{array}{l} \displaystyle\lim_{n\to\infty} f(x_n) = q \\ \textit{for all sequences } \{x_n\} \textit{ in } E \\ \textit{with } \lim x_n = p \textit{ and } x_n \neq p \end{array}$$

Proof idea. Use epsilons. □

Corollary 5.4.1. *Limits of functions are unique.*

Proof idea. By uniqueness of sequential limits. □

Remark. Uniqueness also holds for extended limits of functions on $\mathbb{R}$.

Theorem 5.4.2 (Limits and field operations in $\mathbb{C}$)**.** *Let X be a metric space, $E \subseteq X$, $f : E \to \mathbb{C}$, and $g : E \to \mathbb{C}$. If f and g have limits at $p \in X$, then*

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

(b) $\lim_{x\to p}(fg)(x) = \lim_{x\to p} f(x) \cdot \lim_{x\to p} g(x)$

(c) $\lim_{x\to p}(f/g)(x) = \lim_{x\to p} f(x) / \lim_{x\to p} g(x)$ *if* $\lim_{x\to p} g(x) \neq 0$.

Proof idea. By the above theorem and properties of sequential limits in $\mathbb{C}$. □

Remark. Analogous results hold for extended limits of functions on $\mathbb{R}$, wherever the expressions on the right are defined (determinate forms).

Theorem 5.4.3 (Limits and operations in $\mathbb{R}^k$)**.** *Let X be a metric space, $E \subseteq X$, $\boldsymbol{f} : E \to \mathbb{R}^k$, and $\boldsymbol{g} : E \to \mathbb{R}^k$. If $\boldsymbol{f}$ and $\boldsymbol{g}$ have limits at $p \in X$ and $\alpha \in \mathbb{R}$, then* *(a)* $\lim_{x\to p}(\alpha \boldsymbol{f})(x) = \alpha \lim_{x\to p} \boldsymbol{f}(x)$

(b) $\lim_{x\to p}(\boldsymbol{f}+\boldsymbol{g})(x) = \lim_{x\to p} \boldsymbol{f}(x) + \lim_{x\to p} \boldsymbol{g}(x)$

(c) $\lim_{x\to p}(\boldsymbol{f}\bullet\boldsymbol{g})(x) = \lim_{x\to p} \boldsymbol{f}(x) \bullet \lim_{x\to p} \boldsymbol{g}(x)$

Proof idea. By the above theorem and properties of sequential limits in $\mathbb{R}^k$. □

Theorem 5.4.4 (Characterizations of continuity)**.** *Let X, Y be metric spaces, $E \subseteq X$, and $f : E \to Y$. Let $p \in E$ be a limit point of E.*

(a) f is continuous at p iff $\lim_{x\to p} f(x) = f(p)$.

(b) f is continuous iff f^{-1} takes open [closed] sets to open [closed] sets.

Proof idea. For (a), use epsilons.

For (b), use epsilons for open sets, then take complements for closed sets. □

Theorem 5.4.5 (Continuity and composition)**.** *Let $f : X \to Y$ be continuous and $g : Y \to Z$ continuous on $f(X)$. Then $g \circ f$ is continuous.*

Proof idea. Use nested epsilons. □

Theorem 5.4.6 (Continuity and field operations in $\mathbb{C}$)**.** *Let $f : X \to \mathbb{C}$ and $g : X \to \mathbb{C}$ be continuous. Then*

(a) $f + g$ is continuous.

(b) fg is continuous.

(c) f/g is continuous if $g(x) \neq 0$ on X.

Proof idea. At isolated points there is nothing to prove. At limit points, use the limit characterization of continuity and properties of limits in $\mathbb{C}$. □

Theorem 5.4.7 (Continuity in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f} : X \to \mathbb{R}^k$ and $\boldsymbol{g} : X \to \mathbb{R}^k$.*

(a) If $\boldsymbol{f} = (f_1, \ldots, f_k)$, then $\boldsymbol{f}$ is continuous iff f_i is continuous for $1 \le i \le k$.

(b) If $\boldsymbol{f}$ and $\boldsymbol{g}$ are continuous and $\alpha \in \mathbb{R}$, then

(i) $\alpha\boldsymbol{f}$ is continuous.

(ii) $\boldsymbol{f} + \boldsymbol{g}$ is continuous.

(iii) $\boldsymbol{f} \bullet \boldsymbol{g}$ is continuous.

Proof idea. For (a), at isolated points there is nothing to prove; at limit points, use the limit characterization of continuity together with the characterization of limits (of sequences) in $\mathbb{R}^k$.

For (b), use (a) and properties of continuity in $\mathbb{R}$. □

Theorem 5.4.8 (Continuity preserves compactness)**.** *Let X be compact and $f : X \to Y$ continuous. Then $f(X)$ is compact.*

Proof idea. Since f^{-1} takes open sets to open sets, any open covering of $f(X)$ can be pulled back to an open covering of X, yielding a finite subcovering. □

Corollary 5.4.2 (Extreme value theorem in $\mathbb{R}$)**.** *Let X be compact, $f : X \to \mathbb{R}$ continuous. Set*

$$\alpha = \inf f(X) \qquad \beta = \sup f(X)$$

Then there exist $x, y \in X$ with $f(x) = \alpha$ and $f(y) = \beta$.

Proof idea. By Heine-Borel, $f(X)$ is closed and bounded, so $\alpha, \beta \in f(X)$. □

Applications. Mean value theorem, algebraic closure of $\mathbb{C}$, etc.

Theorem 5.4.9 (Uniform continuity)**.** *Let X be compact and $f : X \to Y$ continuous. Then f is uniformly continuous.*

Proof idea. Given $\epsilon > 0$, for each point in X fix a neighborhood whose extension witnesses continuity of f at the point for ϵ. By compactness, X is covered by finitely many such neighborhoods. Let $\delta > 0$ be a fraction of the minimum radius. It follows that uniformity holds for ϵ with δ. □

Applications. Riemann integrability of continuous functions.

Theorem 5.4.10 (Continuity preserves connectedness)**.** *Let X be connected and $f : X \to Y$ continuous. Then $f(E)$ is connected.*

Proof idea. Since f^{-1} takes closed sets to closed sets. □

Corollary 5.4.3 (Intermediate value theorem in $\mathbb{R}$)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. If* $\min\{f(a), f(b)\} < y < \max\{f(a), f(b)\}$*, then there exists $x \in (a, b)$ with $f(x) = y$.*

Proof idea. By the characterization of connectedness in $\mathbb{R}$. □

Theorem 5.4.11 (Monotonic functions on $\mathbb{R}$)**.** *If $f : (a, b) \to \mathbb{R}$ increases monotonically, then for $x \in (a, b)$,*

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

If $a < x < y < b$, then $f(x+) \leq f(y-)$.

Analogously if f decreases monotonically.

Proof idea. Use epsilons. □

Applications. Monotonic functions have only jump discontinuities, and at most countably many.

Techniques

- Reducing function limits to sequence limits.
- Proving continuity:
  - Characterizations (limit, open and closed sets).
  - Closure properties (composites, sums, products, etc.).

- Using continuity:
  - Characterizations.
  - Uniform continuity.
  - Preservation of compactness.
  - Preservation of connectedness.
- Reducing $\mathbb{R}^k$-valued functions to $\mathbb{R}$-valued (component) functions.

5.5 Differentiability

Definitions

Definition 5.5.1. Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$ and $x \in [a,b]$. If the limit

$$\lim_{t\to x} \frac{\boldsymbol{f}(t) - \boldsymbol{f}(x)}{t - x}$$

exists, $\boldsymbol{f}$ is *differentiable at* x, and the limit is denoted $\boldsymbol{f}'(x)$.

The function $\boldsymbol{f}'$ induced is called the *derivative* of $\boldsymbol{f}$. If $\boldsymbol{f}'$ is defined on $E \subseteq [a,b]$, $\boldsymbol{f}$ is *differentiable on* E.

This is continued with $\boldsymbol{f}'', \boldsymbol{f}'''$, etc. and more generally $\boldsymbol{f}^{(n)}$ for $n \geq 0$.

Definition 5.5.2. Let X be a metric space, $f : X \to \mathbb{R}$, and $p \in X$. Then $f(p)$ is a *local maximum* of f at p if there exists $\delta > 0$ such that $f(q) \leq f(p)$ for all $q \in X$ with $d(p,q) < \delta$. Analogously for *local minimum*.

Theorems

Theorem 5.5.1 (Differentiability implies continuity in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$. If $x \in [a,b]$ and $\boldsymbol{f}$ is differentiable at x, then $\boldsymbol{f}$ is continuous at x.*

Proof idea. Note for $t \neq x$,

$$\boldsymbol{f}(t) - \boldsymbol{f}(x) = \frac{\boldsymbol{f}(t) - \boldsymbol{f}(x)}{t - x} \cdot (t - x) \to 0 \quad \text{as} \quad t \to x \qquad \square$$

Remark. The converse is false! For an extreme counterexample, see the Weierstrass everywhere continuous but nowhere differentiable function.

Theorem 5.5.2 (Chain rule in $\mathbb{R}$)**.** *Let $f : [a,b] \to [c,d]$ and $g : [c,d] \to \mathbb{R}$. If $x \in [a,b]$ and f is differentiable at x and g is differentiable at $f(x)$, then*

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

Proof idea. By differentiability of f at x, there exists a function $\delta(t)$ such that

$$f(t) - f(x) = [f'(x) + \delta(t)](t - x)$$

where $\delta(t) \to 0$ as $t \to x$ and $\delta(x) = 0$. Similarly, by differentiability of g at $y = f(x)$, there exists a function $\epsilon(u)$ such that

$$g(u) - g(y) = [g'(y) + \epsilon(u)](u - y)$$

where $\epsilon(u) \to 0$ as $u \to y$ and $\epsilon(y) = 0$. (Note ϵ is continuous at y!)

Now write $h = g \circ f$. Then

$$\begin{aligned} h(t) - h(x) &= g(f(t)) - g(f(x)) \\ &= [g'(f(x)) + \epsilon(f(t))][f(t) - f(x)] \\ &= [g'(f(x)) + \epsilon(f(t))][f'(x) + \delta(t)](t - x) \end{aligned}$$

Now divide by $t - x$ and let $t \to x$. (Note this relies on continuity of ϵ at y since it might be that $f(t) = y$ infinitely often as $t \to x$.) □

Theorem 5.5.3 (Derivatives and field operations in $\mathbb{C}$)**.** *Let $f : [a,b] \to \mathbb{C}$ and $g : [a,b] \to \mathbb{C}$ be differentiable at $x \in [a,b]$. Then*

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

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

(c) $(f/g)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2} \quad (g(x) \neq 0)$

Proof idea. Express difference quotients of the new functions in terms of difference quotients of the given functions, then take limits.

For (a), this is trivial.

For (b), similarly to sequence products note

$$\begin{aligned} f(t)g(t) - f(x)g(x) = [f(t) - f(x)][g(t) - g(x)] \\ + [f(t) - f(x)]g(x) + [g(t) - g(x)]f(x) \end{aligned}$$

Now divide by $t - x$ and let $t \to x$.

For (c), note

$$\frac{f(t)}{g(t)} - \frac{f(x)}{g(x)} = \frac{[f(t) - f(x)]g(x) - [g(t) - g(x)]f(x)}{g(t)g(x)}$$

Now divide by $t - x$ and let $t \to x$. □

Theorem 5.5.4 (Derivatives in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$ and $\boldsymbol{g} : [a,b] \to \mathbb{R}^k$.*

(a) If $\boldsymbol{f} = (f_1, \ldots, f_k)$, $\boldsymbol{f}$ is differentiable at $x \in [a,b]$ iff f_i is differentiable at x for $1 \le i \le k$, in which case

$$\boldsymbol{f}'(x) = (f_1'(x), \ldots, f_k'(x))$$

(b) If $\boldsymbol{f}$ and $\boldsymbol{g}$ are differentiable at $x \in [a,b]$ and $\alpha \in \mathbb{R}$, then

(i) $(\alpha \boldsymbol{f})'(x) = \alpha \boldsymbol{f}'(x)$

(ii) $(\boldsymbol{f}+\boldsymbol{g})'(x) = \boldsymbol{f}'(x) + \boldsymbol{g}'(x)$

(iii) $(\boldsymbol{f}\bullet\boldsymbol{g})'(x) = \boldsymbol{f}'(x)\bullet\boldsymbol{g}(x) + \boldsymbol{f}(x)\bullet\boldsymbol{g}'(x)$

Proof idea. For (a), use the characterization of limits (of sequences) in $\mathbb{R}^k$.
For (b), use (a) and properties of derivatives in $\mathbb{R}$. □

Theorem 5.5.5 (Fermat)**.** *Let $f : [a,b] \to \mathbb{R}$. If f has a local maximum or local minimum at $x \in (a,b)$, and f is differentiable at x, then $f'(x) = 0$.*

Proof idea. Look at left-hand and right-hand limits separately. □

Theorem 5.5.6 (Rolle)**.** *Let $f : [a,b] \to \mathbb{R}$. If $f(a) = f(b)$ and f is continuous on $[a,b]$ and differentiable on (a,b), there exists $x \in (a,b)$ with $f'(x) = 0$.*

Proof idea. By the extreme value theorem, there exists $x \in (a,b)$ such that $f(x)$ is maximum or minimum for f, so $f'(x) = 0$. □

Theorem 5.5.7 (Mean value theorems in $\mathbb{R}$)**.** *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b).*

(a) (Cauchy) There exists $x \in (a,b)$ with

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

(b) There exists $x \in (a,b)$ with

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

Proof idea. For (a), set

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

so $h'(x) = 0$ for $x \in (a,b)$ iff (a) holds. Now apply Rolle's theorem to h.
Now (b) follows from (a) with $g(x) = x$. □

Applications. Relating values of functions to values of derivatives, Taylor's theorem, L'Hospital's rule, change of variable for Riemann-Stieltjes integration, fundamental theorem of calculus for Riemann integration, etc.

Remark. Note the mean value theorem asserts the existence of a point where the instantaneous rate of change of a differentiable function equals its average rate of change (on an interval).

Remark. The mean value theorem fails in $\mathbb{R}^k$! But see the mean value inequality.

Corollary 5.5.1 (Monotonicity in $\mathbb{R}$)**.** *Let $f : (a,b) \to \mathbb{R}$ be differentiable.*

(a) f is monotonically increasing iff $f'(x) \ge 0$ for all $x \in (a,b)$.

(b) f is constant iff $f'(x) = 0$ for all $x \in (a,b)$.

(c) f is monotonically decreasing iff $f'(x) \le 0$ for all $x \in (a,b)$.

Proof idea. By the mean value theorem. □

Theorem 5.5.8 (Taylor)**.** *Let $f : [a,b] \to \mathbb{R}$. Suppose $n > 0$ and $f^{(n-1)}$ is continuous on $[a,b]$ and differentiable on (a,b). If $\alpha,\beta \in (a,b)$ and $\alpha \neq \beta$, there exists x between α,β such that*

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

where

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

is the Taylor polynomial of degree $(n-1)$ for f at α.

Proof idea. By repeated application of Rolle's theorem.

Define C using $f(\beta) = P(\beta) + C(\beta-\alpha)^n$, so x between α,β is desired satisfying $f^{(n)}(x) = n!C$. Define

$$h(x) = f(x) - P(x) - C(x-\alpha)^n$$

so the result holds for x iff $h^{(n)}(x) = 0$.

Note $h^{(k)}(\alpha) = 0$ for $0 \leq k \leq n-1$ and $h(\beta) = 0$. By Rolle's theorem, there exists x_1 between α,β with $h^{(1)}(x_1) = 0$; then there exists x_2 between α, x_1 with $h^{(2)}(x_2) = 0$; etc. Finally, there exists x_n with $h^{(n)}(x_n) = 0$, so the result holds with $x = x_n$. □

Applications. Approximating differentiable functions with error bound.

Remark. Taylor's theorem generalizes the mean value theorem, which is just case $n = 1$ of Taylor's theorem.

Theorem 5.5.9 (L'Hospital)**.** *Let $f : (a,b) \to \mathbb{R}$ and $g : (a,b) \to \mathbb{R}$ be differentiable with $-\infty \leq a < b \leq +\infty$, where $g'(x) \neq 0$ in (a,b). Let $a \leq c \leq b$. If $f(x), g(x) \to 0$ as $x \to c$, or $g(x) \to \pm\infty$ as $x \to c$, and if*

$$\frac{f'(x)}{g'(x)} \to A \quad \textit{as} \quad x \to c \qquad (-\infty \leq A \leq +\infty)$$

Then $f(x)/g(x) \to A$ as $x \to c$.

Proof idea. By Cauchy's mean value theorem.

If $A < +\infty$, use the theorem to relate values of the derivative quotient to values of the function quotient and show that for all r with $A < r$, there is some neighborhood of c in which $f(x)/g(x) < r$.

If $A > -\infty$, do the inverse. □

Applications. Evaluating limits yielding indeterminate forms such as $\frac{0}{0}$, $\frac{\infty}{\infty}$, etc.

Remark. L'Hospital's rule fails in $\mathbb{R}^k$!

Theorem 5.5.10 (Intermediate value theorem for derivatives in $\mathbb{R}$)**.** *Let $f : [a,b] \to \mathbb{R}$ be differentiable. If $\min\{f'(a), f'(b)\} < y < \max\{f'(a), f'(b)\}$, there exists $x \in (a,b)$ with $f'(x) = y$.*

Proof idea. Set $h(x) = f(x) - yx$. Then $h'(a)$ and $h'(b)$ differ in sign, so by the extreme value theorem h has a maximum or minimum at some $x \in (a, b)$ where $h'(x) = 0$, that is, $f'(x) = y$. □

Theorem 5.5.11 (Mean value inequality in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f} : [a, b] \to \mathbb{R}^k$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $x \in (a, b)$ with*

$$\left\|\boldsymbol{f}'(x)\right\| \geq \frac{\left\|\boldsymbol{f}(b) - \boldsymbol{f}(a)\right\|}{b - a}$$

Proof idea. Construct an $\mathbb{R}$-valued function from $\boldsymbol{f}$ using the dot product and apply the mean value theorem in $\mathbb{R}$, then use Cauchy-Schwarz.

In detail, set $\boldsymbol{z} = \boldsymbol{f}(b) - \boldsymbol{f}(a)$, and define $\varphi(t) = (\boldsymbol{z} \bullet \boldsymbol{f})(t)$. By the mean value theorem applied to φ, there exists $x \in (a, b)$ with

$$\boldsymbol{z} \bullet \boldsymbol{f}'(x) = \frac{\boldsymbol{z} \bullet \boldsymbol{f}(b) - \boldsymbol{z} \bullet \boldsymbol{f}(a)}{b - a} = \frac{\boldsymbol{z} \bullet \boldsymbol{z}}{b - a} = \frac{\|\boldsymbol{z}\|^2}{b - a}$$

By Cauchy-Schwarz, $\left\|\boldsymbol{z}\right\| \left\|\boldsymbol{f}'(x)\right\| \geq \boldsymbol{z} \bullet \boldsymbol{f}'(x)$, so

$$\left\|\boldsymbol{f}'(x)\right\| \geq \frac{\|z\|}{b - a} = \frac{\left\|\boldsymbol{f}(b) - \boldsymbol{f}(a)\right\|}{b - a}$$ □

Applications. Bounding change in a function by values of its derivative, as in the proof of the theorem on uniform convergence and differentiation.

Techniques

- Proving differentiability and calculating derivatives:
  - Limit of difference quotient.
  - Closure properties (composites, sums, products, etc.).
- Relating values of functions to values of their derivatives using mean value theorems.
- Translating problems about arbitrary values of derivatives into problems about zeros of derivatives, then using extrema or mean value theorems to find zeros.
- Approximating functions using Taylor series expansions.
- Reducing $\mathbb{R}^k$-valued functions to $\mathbb{R}$-valued functions.

5.6 Integrability

Definitions

Definition 5.6.1. A *partition* of the interval $[a, b]$ is a set $P = \{x_0, \ldots, x_n\}$ with

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b$$

For P, $\Delta x_i = x_i - x_{i-1}$ for $1 \leq i \leq n$.

Definition 5.6.2. Let P, Q, R be partitions of an interval. If $P \subseteq R$, then R is a *refinement* of P. If $P \cup Q \subseteq R$, then R is a *common refinement* of P and Q.

Definition 5.6.3. Let $f : [a,b] \to \mathbb{R}$ be bounded, $\alpha : [a,b] \to \mathbb{R}$ monotonically increasing. Let $P = \{x_0, \dots, x_n\}$ be a partition of $[a,b]$. Define

$$\begin{aligned} m_i &= \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} \\ M_i &= \sup\{f(x) \mid x \in [x_{i-1}, x_i]\} \qquad\qquad (1 \le i \le n) \\ \Delta\alpha_i &= \alpha(x_i) - \alpha(x_{i-1}) \end{aligned}$$

Now define

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

These are the *lower* and *upper Riemann-Stieltjes sums* of f with respect to α over P.

Now define

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

where the limits are taken over all partitions P of $[a,b]$. These are the *lower* and *upper Riemann-Stieltjes integrals* of f with respect to α over $[a,b]$. In this context, f is called an *integrand* and α an *integrator.*

If the upper and lower integrals are equal, denote their common value by

$$\int_a^b f\,d\alpha$$

the *Riemann-Stieltjes integral* of f with respect to α over $[a,b]$. In this case, say f is *integrable (in the Riemann-Stieltjes sense)* with respect to α over $[a,b]$, and write $f \in \mathscr{R}(\alpha)$ (on $[a,b]$).

If $\alpha = x$, omit α from the notation and speak merely of Riemann sums, integrals, integrability, etc. Write $f \in \mathscr{R}$ if $f \in \mathscr{R}(x)$.

Definition 5.6.4. Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$ be bounded, $\alpha : [a,b] \to \mathbb{R}$ monotonically increasing. If $\boldsymbol{f} = (f_1, \dots, f_k)$, write $\boldsymbol{f} \in \mathscr{R}(\alpha)$ if $f_i \in \mathscr{R}(\alpha)$ for $1 \le i \le k$, and in this case define

$$\int_a^b \boldsymbol{f}\,d\alpha = \Big(\int_a^b f_1\,d\alpha, \dots, \int_a^b f_k\,d\alpha\Big)$$

Theorems

Theorem 5.6.1 (Ordering of sums in $\mathbb{R}$)**.** *Let $f : [a,b] \to \mathbb{R}$ be bounded and $\alpha : [a,b] \to \mathbb{R}$ monotonically increasing. For any partition P of $[a,b]$ and refinement P^* of P,*

$$L(P, f, \alpha) \le L(P^*, f, \alpha) \le \underline{\int_a^b} f\,d\alpha \le \overline{\int_a^b} f\,d\alpha \le U(P^*, f, \alpha) \le U(P, f, \alpha)$$

Proof idea. The outer four inequalities are immediate from definitions.

For the inner inequality, note by taking a common refinement that for partitions Q and R, $L(Q,f,\alpha) \le U(R,f,\alpha)$. Keeping R fixed and letting Q vary shows $\underline{\int}_a^b f\,d\alpha \le U(R,f,\alpha)$. Now letting R vary shows $\underline{\int}_a^b f\,d\alpha \le \overline{\int}_a^b f\,d\alpha$. □

Theorem 5.6.2 (Cauchy criterion for integrability in $\mathbb{R}$)**.** *Let $f:[a,b]\to\mathbb{R}$ be bounded and $\alpha:[a,b]\to\mathbb{R}$ monotonically increasing. Then $f\in\mathscr{R}(\alpha)$ iff for every $\epsilon>0$ there exists a partition P of $[a,b]$ such that*

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

Proof idea. By ordering of sums. □

Applications. Proving functions integrable.

Theorem 5.6.3 (Sample points in $\mathbb{R}$)**.** *Let $f:[a,b]\to\mathbb{R}$ be bounded and $\alpha:[a,b]\to\mathbb{R}$ monotonically increasing. Suppose $\epsilon>0$ and $P=\{x_0,\dots,x_n\}$ is a partition of $[a,b]$ with*

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

Let $s_i,t_i\in[x_{i-1},x_i]$ for $1\le i\le n$. Then

(a) $\displaystyle\sum_{i=1}^{n}|f(s_i)-f(t_i)|\Delta\alpha_i<\epsilon$

(b) $\displaystyle|\sum_{i=1}^{n}f(s_i)\Delta\alpha_i-\int_a^b f\,d\alpha|<\epsilon \quad \textit{if } f\in\mathscr{R}(\alpha)$

Proof idea. By ordering of sums since $m_i\le f(s_i),f(t_i)\le M_i$ for $1\le i\le n$. □

Applications. Integration arguments requiring the use of sample points, like those involving application of the mean value theorem (change of variable, fundamental theorem of calculus). Partially illustrating the relationship between our definition of integrability and a definition in terms of limits of sums using sample points.

Theorem 5.6.4 (Continuity implies integrability in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f}:[a,b]\to\mathbb{R}^k$ be bounded and $\alpha:[a,b]\to\mathbb{R}$ monotonically increasing.*

(a) If $\boldsymbol{f}$ is continuous, then $\boldsymbol{f}\in\mathscr{R}(\alpha)$.

(c) If $\boldsymbol{f}$ is discontinuous at only finitely many points and α is continuous at these points, then $\boldsymbol{f}\in\mathscr{R}(\alpha)$.

Proof idea. Assume without loss of generality that $k=1$, by the characterization of continuity in $\mathbb{R}^k$ and definition of the integral in $\mathbb{R}^k$.

For (a), use uniform continuity of $\boldsymbol{f}$ to bound $\Delta\boldsymbol{f}$ across intervals of a partition. In detail, since $\boldsymbol{f}$ is continuous on a compact set, $\boldsymbol{f}$ is uniformly continuous. Given $\epsilon>0$, set $M=\alpha(b)-\alpha(a)$ and choose $\delta>0$ such that $\left\|\boldsymbol{f}(s)-\boldsymbol{f}(t)\right\|<\epsilon/(M+1)$ when $s,t\in[a,b]$ and $|s-t|<\delta$. Now choosing a partition P with intervals of length $<\delta$, $U(P,\boldsymbol{f},\alpha)-L(P,\boldsymbol{f},\alpha)<\epsilon$. It follows that $\boldsymbol{f}\in\mathscr{R}(\alpha)$ by the Cauchy criterion.

For (b), divide and conquer. Use uniform continuity of α on intervals around points of discontinuity of $\boldsymbol{f}$ to bound $\Delta\alpha$ there, and use uniform continuity of $\boldsymbol{f}$ elsewhere to bound $\Delta\boldsymbol{f}$. □

Remark. The converse is false! Also, compare that continuity implies integrability but does *not* imply differentiability.

Theorem 5.6.5 (Monotonicity and integrability in $\mathbb{R}$)**.** *If $f : [a,b] \to \mathbb{R}$ is monotonic and $\alpha : [a,b] \to \mathbb{R}$ monotonically increasing and continuous, then $f \in \mathscr{R}(\alpha)$.*

Proof idea. Swap f and α in the proof that continuity implies integrability (inverting signs appropriately if f is monotonically decreasing). □

Theorem 5.6.6 (Integrability of composites in $\mathbb{R}$)**.** *Let $f : [a,b] \to \mathbb{R}$ and $\alpha : [a,b] \to \mathbb{R}$ with $f \in \mathscr{R}(\alpha)$. If $A \le f \le B$ and g is continuous on $[A,B]$, then $h = g \circ f \in \mathscr{R}(\alpha)$.*

Proof idea. Divide and conquer, using uniform continuity of g to bound Δh where possible, and integrability of f to bound $\Delta\alpha$ elsewhere.

In detail, given $\epsilon > 0$, set $M = \alpha(b) - \alpha(a)$ and choose $\delta > 0$ such that $|g(s) - g(t)| < \epsilon/2(M+1)$ whenever $s, t \in [A,B]$ and $|s - t| < \delta$. Now fix N with $|g| < N$ and choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that

$$U(P,f,\alpha) - L(P,f,\alpha) = \sum_{i=1}^{n} [M_i - m_i]\Delta\alpha_i < \frac{\delta\epsilon}{4N}$$

Let I be the set of indices i with $M_i - m_i < \delta$, and J the remainder. Then

$$\begin{aligned} U(P,h,\alpha) - L(P,h,\alpha) &\le \frac{\epsilon}{2(M+1)} \sum_{i\in I} \Delta\alpha_i + 2N \sum_{i\in J} \Delta\alpha_i \\ &\le \frac{\epsilon}{2(M+1)} M + 2N \frac{\epsilon}{4N} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

It follows that $h \in \mathscr{R}(\alpha)$ by the Cauchy criterion. □

Theorem 5.6.7 (Properties of the integral)**.** *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$ and $\boldsymbol{g} : [a,b] \to \mathbb{R}^k$ be bounded, $\alpha : [a,b] \to \mathbb{R}$ and $\beta : [a,b] \to \mathbb{R}$ monotonically increasing, and $c \in \mathbb{R}$.*

(a) (Linearity in integrand in $\mathbb{R}^k$) If $\boldsymbol{f} \in \mathscr{R}(\alpha)$ and $\boldsymbol{g} \in \mathscr{R}(\alpha)$, then

$$\int_a^b (\boldsymbol{f} + \boldsymbol{g})\, d\alpha = \int_a^b \boldsymbol{f}\, d\alpha + \int_a^b \boldsymbol{g}\, d\alpha$$
$$\int_a^b (c\boldsymbol{f})\, d\alpha = c \int_a^b \boldsymbol{f}\, d\alpha$$

(b) (Linearity in integrator in $\mathbb{R}^k$) If $\boldsymbol{f} \in \mathscr{R}(\alpha)$ and $\boldsymbol{f} \in \mathscr{R}(\beta)$ and $c \ge 0$, then

$$\int_a^b \boldsymbol{f}\, d(\alpha + \beta) = \int_a^b \boldsymbol{f}\, d\alpha + \int_a^b \boldsymbol{f}\, d\beta$$
$$\int_a^b \boldsymbol{f}\, d(c\alpha) = c \int_a^b \boldsymbol{f}\, d\alpha$$

(c) *(Additivity in endpoints in $\mathbb{R}^k$) If $\boldsymbol{f} \in \mathscr{R}(\alpha)$ and $a < c < b$, then*

$$\int_a^b \boldsymbol{f}\, d\alpha = \int_a^c \boldsymbol{f}\, d\alpha + \int_c^b \boldsymbol{f}\, d\alpha$$

(d) *(Preservation of order in $\mathbb{R}$) If $k = 1$, $\boldsymbol{f} \le \boldsymbol{g}$, and $\boldsymbol{f}, \boldsymbol{g} \in \mathscr{R}(\alpha)$, then*

$$\int_a^b \boldsymbol{f}\, d\alpha \le \int_a^b \boldsymbol{g}\, d\alpha$$

(e) *(Boundedness in $\mathbb{R}$) If $k = 1$, $m \le \boldsymbol{f} \le M$, and $\boldsymbol{f} \in \mathscr{R}(\alpha)$, then*

$$m[\alpha(b) - \alpha(a)] \le \int_a^b \boldsymbol{f}\, d\alpha \le M[\alpha(b) - \alpha(a)]$$

Proof idea. Assume without loss of generality that $k = 1$.

For the first part of (a), note for any partition P that

$$L(P, \boldsymbol{f}, \alpha) + L(P, \boldsymbol{g}, \alpha) \le L(P, \boldsymbol{f} + \boldsymbol{g}, \alpha) \le U(P, \boldsymbol{f} + \boldsymbol{g}, \alpha) \le U(P, \boldsymbol{f}, \alpha) + U(P, \boldsymbol{g}, \alpha)$$

Since $\boldsymbol{f}$ and $\boldsymbol{g}$ are integrable, there are partitions making the upper and lower sums for $\boldsymbol{f}$ and $\boldsymbol{g}$ arbitrarily close, and hence making the upper and lower sums for $\boldsymbol{f} + \boldsymbol{g}$ arbitrarily close, so $\boldsymbol{f} + \boldsymbol{g}$ is integrable. Moreover, this inequality and the ordering of sums together show that $\int(\boldsymbol{f} + \boldsymbol{g})$ and $\int \boldsymbol{f} + \int \boldsymbol{g}$ are arbitrarily close, hence equal.

The second part of (a), as well as (b) and (c), are similar.

For (d), any lower sum for $\boldsymbol{f}$ is bounded above by the corresponding lower sum for $\boldsymbol{g}$, so also by $\int \boldsymbol{g}$. Since $\int \boldsymbol{f}$ is the least upper bound of lower sums for $\boldsymbol{f}$, $\int \boldsymbol{f} \le \int \boldsymbol{g}$.

Now (e) follows from (d) and integration of constants. □

Corollary 5.6.1. *Let $\alpha : [a, b] \to \mathbb{R}$ be monotonically increasing.*

(a) *If $f : [a, b] \to \mathbb{R}$, $g : [a, b] \to \mathbb{R}$, and $f, g \in \mathscr{R}(\alpha)$, then $fg \in \mathscr{R}(\alpha)$.*

(b) *If $\boldsymbol{f} : [a, b] \to \mathbb{R}^k$ and $\boldsymbol{f} \in \mathscr{R}(\alpha)$, then $\|\boldsymbol{f}\| \in \mathscr{R}(\alpha)$ and*

$$|\int_a^b \boldsymbol{f}\, d\alpha| \le \int_a^b |\boldsymbol{f}|\, d\alpha$$

Proof idea. For (a), note $fg = [(f + g)^2 - (f - g)^2]/4$, so integrability follows from linearity and integrability of composites (with t^2).

For (b), similarly if $\boldsymbol{f} = (f_1, \ldots, f_k)$, then $\|\boldsymbol{f}\| = \left(f_1^2 + \cdots + f_k^2\right)^{1/2} \in \mathscr{R}(\alpha)$. For the bound, expand into component integrals, massage, and apply Cauchy-Schwarz. In detail, by linearity in integrands,

$$\left\|\int \boldsymbol{f}\right\|^2 = (\int \boldsymbol{f}) \bullet (\int \boldsymbol{f}) = \Sigma(\int f_i)(\int f_i) = \Sigma \int (\int f_i) f_i = \int \Sigma(\int f_i) f_i$$

By Cauchy-Schwarz,

$$\Sigma(\int f_i) f_i = (\int \boldsymbol{f}) \bullet \boldsymbol{f} \le \left\|\int \boldsymbol{f}\right\| \|\boldsymbol{f}\|$$

By order preservation then,

$$\int \Sigma(\int f_i) f_i \le \int \left\|\int \boldsymbol{f}\right\| \|\boldsymbol{f}\| = \left\|\int \boldsymbol{f}\right\| \int \|\boldsymbol{f}\|$$

Therefore $\left\|\int \boldsymbol{f}\right\| \le \int \|\boldsymbol{f}\|$. □

Theorem 5.6.8 (Change of variable in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$ be bounded and $\alpha : [a,b] \to \mathbb{R}$ monotonically increasing.*

(a) If $\alpha' \in \mathscr{R}$, then $\boldsymbol{f} \in \mathscr{R}(\alpha)$ iff $\boldsymbol{f}\alpha' \in \mathscr{R}$, in which case

$$\int_a^b \boldsymbol{f}\, d\alpha = \int_a^b \boldsymbol{f}\alpha'\, dx$$

(b) If $\varphi : [A,B] \to [a,b]$ is a strictly increasing bijection, $\beta = \alpha \circ \varphi$, and $\boldsymbol{g} = \boldsymbol{f} \circ \varphi$, then $\boldsymbol{f} \in \mathscr{R}(\alpha)$ iff $\boldsymbol{g} \in \mathscr{R}(\beta)$, in which case

$$\int_A^B \boldsymbol{g}\, d\beta = \int_a^b \boldsymbol{f}\, d\alpha$$

Proof idea. Assume without loss of generality that $k = 1$.

For (a), use the mean value theorem to relate sums for the two integrals. In detail, given $\epsilon > 0$, fix M with $|\boldsymbol{f}| < M$ and choose P with $U(P, \alpha') - L(P, \alpha') < \epsilon/M$, so for all sample points $s_i, t_i \in [x_{i-1}, x_i]$,

$$\sum_{i=1}^n |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i < \frac{\epsilon}{M}$$

By the mean value theorem, fix $s_i \in [x_{i-1}, x_i]$ with $\Delta\alpha_i = \alpha'(s_i)\Delta x_i$. From the above it follows that for any $t_i \in [x_{i-1}, x_i]$,

$$|\sum_{i=1}^n \boldsymbol{f}(t_i)\Delta\alpha_i - \sum_{i=1}^n \boldsymbol{f}(t_i)\alpha'(t_i)\Delta x_i| < \epsilon$$

From this it follows that

$$|L(P, \boldsymbol{f}, \alpha) - L(P, \boldsymbol{f}\alpha')| \le \epsilon \qquad \text{and} \qquad |U(P, \boldsymbol{f}, \alpha) - U(P, \boldsymbol{f}\alpha')| \le \epsilon$$

Finally, taking refinements of P, it follows that

$$|\underline{\int}_a^b \boldsymbol{f}\, d\alpha - \underline{\int}_a^b \boldsymbol{f}\alpha'\, dx| \le \epsilon \qquad \text{and} \qquad |\overline{\int}_a^b \boldsymbol{f}\, d\alpha - \overline{\int}_a^b \boldsymbol{f}\alpha'\, dx| \le \epsilon$$

Since ϵ was arbitrary, these corresponding upper and lower integrals must be equal, from which the result follows.

For (b), note partitions of $[A,B]$ and $[a,b]$ correspond naturally under φ, and corresponding sums for $\boldsymbol{g}$ with β and $\boldsymbol{f}$ with α are equal. □

Applications. Computing integrals by making variable substititions.

The remaining results apply only to the Riemann integral.

Theorem 5.6.9 (Fundamental theorem of calculus in $\mathbb{R}^k$)**.** *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^k$ with $\boldsymbol{f} \in \mathscr{R}$.*

(a) *Define $\boldsymbol{F} : [a,b] \to \mathbb{R}^k$ by*

$$\boldsymbol{F}(x) = \int_a^x \boldsymbol{f}(t)\,dt$$

Then $\boldsymbol{F}$ is continuous, and if $\boldsymbol{f}$ is continuous at x_0, $\boldsymbol{F}'(x_0) = \boldsymbol{f}(x_0)$.

(b) *If $\boldsymbol{F} : [a,b] \to \mathbb{R}^k$ and $\boldsymbol{F}' = \boldsymbol{f}$, then*

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

Proof idea. Assume without loss of generality that $k = 1$, by the characterizations of continuity and differentiability in $\mathbb{R}^k$ and the definition of the integral in $\mathbb{R}^k$.

For (a), bound integrals. In detail, fix M with $|\boldsymbol{f}| < M$ and note for $a \le x \le y \le b$,

$$|\boldsymbol{F}(y) - \boldsymbol{F}(x)| = |\int_x^y \boldsymbol{f}\,dt| < M(y - x)$$

Therefore $\Delta\boldsymbol{F}$ is small when Δx is small, so $\boldsymbol{F}$ is uniformly continuous. Now note if $a \le x \le x_0 \le y \le b$ with $x < y$, then

$$|\frac{\boldsymbol{F}(y) - \boldsymbol{F}(x)}{y - x} - \boldsymbol{f}(x_0)| = \frac{1}{y - x}|\int_x^y [\boldsymbol{f}(t) - \boldsymbol{f}(x_0)]\,dt|$$

By continuity of $\boldsymbol{f}$ at x_0, the quantity on the right can be made arbitrarily small by taking x, y arbitrarily close to x_0, so $\boldsymbol{F}'(x_0) = \boldsymbol{f}(x_0)$.

For (b), use the mean value theorem. In detail, given $\epsilon > 0$, choose a partition P such that $U(P, \boldsymbol{F}') - L(P, \boldsymbol{F}') < \epsilon$, so for any sample points $s_i \in [x_{i-1}, x_i]$,

$$|\sum_{i=1}^n \boldsymbol{F}'(s_i)\Delta x_i - \int_a^b \boldsymbol{f}| < \epsilon$$

By the mean value theorem, fix $s_i \in [x_{i-1}, x_i]$ with $\boldsymbol{F}'(s_i)\Delta x_i = \boldsymbol{F}(x_i) - \boldsymbol{F}(x_{i-1})$. Then $\sum \boldsymbol{F}'(s_i)\Delta x_i = \sum[\boldsymbol{F}(x_i) - \boldsymbol{F}(x_{i-1})] = \boldsymbol{F}(b) - \boldsymbol{F}(a)$, so

$$|\boldsymbol{F}(b) - \boldsymbol{F}(a) - \int_a^b \boldsymbol{f}| < \epsilon$$

Since ϵ was arbitrary, equality holds. □

Applications. Showing that differentiation and integration are inverse processes, computing derivatives using integrands, computing integrals using antiderivatives, translating results between the languages of differentiation and integration, etc.

Corollary 5.6.2 (Integration by parts in $\mathbb{R}$)**.** *Let $f : [a,b] \to \mathbb{R}$, $g : [a,b] \to \mathbb{R}$ be differentiable with $f', g' \in \mathscr{R}$. Then*

$$\int_a^b fg' = (fg)(b) - (fg)(a) - \int_a^b f'g$$

Proof idea. By the product rule for differentiation and the fundamental theorem of calculus, both applied to fg. □

Techniques

- Proving integrability and calculating integrals:
  - Cauchy criterion.
  - Continuity.
  - Monotonicity.
  - Closure properties (continuous composites, sums, products, etc.).
  - Fundamental theorem.
  - Integration by parts.
  - Change of variable.
- Uniform continuity to bound change in a function across a set.
- Divide and conquer.
- Mean value theorem to relate values of functions to values of their derivatives.
- Fundamental theorem of calculus to relate differentiation and integration.
- Reducing $\mathbb{R}^k$-valued functions to $\mathbb{R}$-valued (component) functions.

5.7 Sequences and Series of Functions

All functions are $\mathbb{C}$-valued unless otherwise implied.

Definitions

Definition 5.7.1. Let E be a set. A sequence $\{f_n\}$ of functions on E *converges (pointwise)* to the function f on E if

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

In this case, f is the *(pointwise) limit* of $\{f_n\}$ on E, denoted $f = \lim f_n$ or $f_n \to f$.

As usual, the series $\sum f_n$ is just the sequence of partial sums $s_n = \sum_{i=1}^n f_i$, and if $\sum f_n$ converges we informally identify $\sum f_n$ with the limit, the *(pointwise) sum.*[1]

Definition 5.7.2. Let E be a set. A sequence $\{f_n\}$ of functions on E *converges uniformly* to the function f on E if for every $\epsilon > 0$, there exists N such that for all $n \geq N$ and $x \in E$,

$$|f_n(x) - f(x)| < \epsilon$$

In this case, f is a *uniform limit.*

A uniform limit of $\sum f_n$ is called a *uniform sum.*

[1] Note there are at least three ways one might decide to interpret a statement of the form $f = \sum f_n$, namely (i) that f is a function mapping numbers to numbers, or (ii) that f is a function mapping numbers to sequences of partial sums of numbers, or (iii) that f is a sequence of partial sums of functions. We are formally following (iii), but also informally following (i).

Definition 5.7.3. Let X be a metric space. Then $\mathscr{C}(X)$ denotes the set of all bounded, continuous functions on X.

For $f \in \mathscr{C}(X)$, define the *supremum norm* of f by

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

For $f, g \in \mathscr{C}(X)$, define the distance between f and g by $\|f - g\|$, so $\mathscr{C}(X)$ forms a metric space.

Definition 5.7.4. A set $\mathscr{A}$ of functions on a set E is an *algebra* if it is closed under addition, scalar multiplication, and multiplication. (If considering only $\mathbb{R}$-valued functions, scalars are assumed in $\mathbb{R}$.)

The *uniform closure* $\overline{\mathscr{A}}$ of $\mathscr{A}$ consists of all functions which are uniform limits of sequences in $\mathscr{A}$. $\mathscr{A}$ is *uniformly closed* if $\mathscr{A} = \overline{\mathscr{A}}$.

$\mathscr{A}$ *separates points* on E if for all $x, y \in E$ with $x \neq y$, there exists $f \in \mathscr{A}$ with $f(x) \neq f(y)$.

$\mathscr{A}$ *vanishes nowhere* on E if for all $x \in E$, there exists $f \in \mathscr{A}$ with $f(x) \neq 0$.

$\mathscr{A}$ is *self-adjoint* if $\overline{f} \in \mathscr{A}$ for all $f \in \mathscr{A}$.

Theorems

Theorem 5.7.1 (Characterizations of uniform convergence)**.** *Let $\{f_n\}$ be a sequence of functions defined on a set E.*

(a) *(Cauchy criterion) $\{f_n\}$ converges uniformly iff for every $\epsilon > 0$, there exists N such that for all $m, n \geq N$ and $x \in E$,*

$$|f_m(x) - f_n(x)| < \epsilon$$

(b) *(Supremum limit criterion) If f is defined on E and $M_n = \sup_{x \in E} |f_n(x) - f(x)|$, then $f_n \to f$ uniformly iff $M_n \to 0$ as $n \to \infty$.*

Proof idea. For (a), the forward direction immediate. For the reverse direction, note for each $x \in E$ that the sequence $\{f_n(x)\}$ is Cauchy and hence converges to some value $f(x)$ (by completeness of $\mathbb{C}$), and trivially $f_n \to f$ uniformly.

Note (b) is immediate from definitions. □

Applications. Proving uniform convergence without reference to a limit.

Corollary 5.7.1. *$\sum f_n$ converges uniformly iff for every $\epsilon > 0$, there exists N such that for all $n \geq m \geq N$ and $x \in E$,*

$$|\sum_{i=m}^{n} f_i(x)| < \epsilon$$

Corollary 5.7.2. *$f_n \to f$ in $\mathscr{C}(X)$ iff $f_n \to f$ uniformly.*

Theorem 5.7.2 (Weierstrass)**.** *Let $\{f_n\}$ be a sequence of functions. If $|f_n| \leq M_n$ for all n and $\sum M_n$ converges, then $\sum f_n$ converges uniformly.*

Proof idea. By the Cauchy criterion. □

Applications. Proving uniform convergence.

Remark. Uniform convergence of function sequences [series] is like convergence of numerical sequences [series].

Theorem 5.7.3 (Uniform convergence and continuity)**.** *If $\{f_n\}$ is a sequence of continuous functions and $f_n \to f$ uniformly, then f is continuous.*

Proof idea. Use continuous functions close to f.

In detail, let p be a point and $\epsilon > 0$. Choose N with $|f - f_N| < \epsilon/3$. By continuity of f_N at p, there exists $\delta > 0$ such that $|f_N(x) - f_N(p)| < \epsilon/3$ whenever $d(x,p) < \delta$. Therefore

$$|f(x) - f(p)| \le |f(x) - f_N(x)| + |f_N(x) - f_N(p)| + |f_N(p) - f(p)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

whenever $d(x,p) < \delta$. So f is continuous at p. □

Corollary 5.7.3. *If $\{f_n\}$ is a sequence of continuous functions and $\sum f_n$ converges uniformly, then $\sum f_n$ is continuous.*

Corollary 5.7.4. *$\mathscr{C}(X)$ is a complete metric space.*

Proof idea. A Cauchy sequence in $\mathscr{C}(X)$ is convergent by the Cauchy criterion, and the limit is uniform and hence bounded and continuous. □

Theorem 5.7.4 (Uniform convergence and integration)**.** *Let $\alpha : [a,b] \to \mathbb{R}$. If $\{f_n\}$ is a sequence of functions defined on $[a,b]$ with $f_n \in \mathscr{R}(\alpha)$ for all n, and $f_n \to f$ uniformly, then $f \in \mathscr{R}(\alpha)$ and*

$$\int_a^b \lim f_n \, d\alpha = \lim \int_a^b f_n \, d\alpha$$

Proof idea. Assume without loss of generality that the functions are $\mathbb{R}$-valued, and use integrable functions close to f.

In detail, given $\epsilon > 0$, choose N with $|f - f_n| < \epsilon$ for $n \ge N$, so $f_n - \epsilon < f < f_n + \epsilon$ for $n \ge N$. Then by properties of sums,

$$\int_a^b f_n \, d\alpha - \epsilon[\alpha(b) - \alpha(a)] < \underline{\int_a^b} f \, d\alpha \le \overline{\int_a^b} f \, d\alpha < \int_a^b f_n \, d\alpha + \epsilon[\alpha(b) - \alpha(a)]$$

for $n \ge N$. Now let $\epsilon \to 0$. □

Corollary 5.7.5. *If $\alpha : [a,b] \to \mathbb{R}$ and $\{f_n\}$ is a sequence of functions defined on $[a,b]$ with $f_n \in \mathscr{R}(\alpha)$ for all n, and $\sum f_n$ converges uniformly, then $\sum f_n \in \mathscr{R}(\alpha)$ and*

$$\int_a^b \sum f_n \, d\alpha = \sum \int_a^b f_n \, d\alpha$$

Applications. Integrating certain series (like power series) term by term.

Theorem 5.7.5 (Uniform convergence and differentiation)**.** *If $\{f_n\}$ is a sequence of functions differentiable on $[a,b]$ which converges at some point, and $\{f_n'\}$ converges uniformly, then $\{f_n\}$ converges uniformly to a differentiable function and*

$$(\lim f_n)' = \lim f_n'$$

Proof idea. Use the mean value inequality to establish uniform convergence of $\{f_n\}$ from uniform convergence of $\{f_n'\}$ (and convergence of $\{f_n\}$ at a point).

Now look at difference quotients. Write $f = \lim f_n$. Fix x, and for $t \neq x$ let $\phi(t)$ be the difference quotient of f, and $\phi_n(t)$ the difference quotient of f_n, for all n. Then for $t \neq x$, ϕ_n is continuous for all n, and $\phi_n \to \phi$ uniformly. Thus we can interchange limit processes to obtain

$$f'(x) = \lim_{t\to x} \phi(t) = \lim_{t\to x} \lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f_n'(x) \qquad \square$$

Corollary 5.7.6. *If $\{f_n\}$ is a sequence of functions differentiable on $[a,b]$, $\sum f_n$ converges at some point, and $\sum f_n'$ converges uniformly, then $\sum f_n$ converges uniformly and*

$$\left(\sum f_n\right)' = \sum f_n'$$

Applications. Differentiating certain series (like power series) term by term.

Remark. Note we assume uniform convergence of the sequence of *derivatives*, not the sequence of functions.

Theorem 5.7.6 (Weierstrass)**.** *If $f : [a,b] \to \mathbb{C}$ is continuous, there exists a sequence $\{P_n\}$ of polynomials with coefficients in $\mathbb{C}$ such that $P_n \to f$ uniformly on $[a,b]$.*

The P_n may be assumed to have coefficients in $\mathbb{R}$ if f is $\mathbb{R}$-valued.

Proof idea. Assume without loss of generality that $[a,b] = [0,1]$ and $f(x) = 0$ for all $x \notin (0,1)$.

For each n, define a polynomial Q_n which forms a single 'bump' of unit area over $[-1,1]$, where bumps get narrower in the middle as $n \to \infty$. Then define $P_n(x)$ to 'accumulate' an approximate value for $f(x)$ by integrating over the bump Q_n:

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

Since the bumps get arbitrarily narrow as $n \to \infty$, this integral gets arbitrarily close to 'picking out' the single value $f(x)$ as $n \to \infty$. Therefore the approximation converges uniformly as $n \to \infty$. $\square$

Applications. Approximating continuous functions by polynomials.

Theorem 5.7.7 (Stone)**.** *Let K be a compact set.*

(a) Let $\mathscr{A}$ be an algebra of continuous $\mathbb{R}$-valued functions on K. If $\mathscr{A}$ separates points on K and vanishes nowhere on K, then the uniform closure of $\mathscr{A}$ consists of all continuous $\mathbb{R}$-valued functions on K.

(b) *Let $\mathscr{A}$ be an algebra of continuous $\mathbb{C}$-valued functions on K. If $\mathscr{A}$ is self-adjoint, separates points on K, and vanishes nowhere on K, then the uniform closure of $\mathscr{A}$ consists of all continuous $\mathbb{C}$-valued functions on K.*

Proof idea. For (a), proceed in steps.

First, argue that $|f| \in \overline{\mathscr{A}}$ whenever $f \in \overline{\mathscr{A}}$ using Weierstrass' theorem, and use this together with some algebra and induction to show that $\min(f_1, \dots, f_n) \in \overline{\mathscr{A}}$ and $\max(f_1, \dots, f_n) \in \overline{\mathscr{A}}$ whenever $f_1, \dots, f_n \in \overline{\mathscr{A}}$.

Then let $g : K \to \mathbb{R}$ be continuous. Given $\epsilon > 0$, find $f \in \overline{\mathscr{A}}$ with $|f - g| < \epsilon$ in two steps. For all $x, y \in K$, construct $f_{x,y} \in \overline{\mathscr{A}}$ with $f_{x,y}(x) = g(x)$ and $f_{x,y}(y) = g(y)$. For each fixed x, using continuity of $f_{x,y}$ and g at each $y \in K$, compactness of K, and maximization, construct a function $f_x \in \overline{\mathscr{A}}$ with $f_x(x) = g(x)$ and $g - \epsilon < f_x$ on K. Then by continuity of f_x and g at each $x \in K$, compactness of K, and minimization, construct $f \in \overline{\mathscr{A}}$ with $g - \epsilon < f < g + \epsilon$ on K, as desired.

For (b), note by self-adjointness that the $\mathbb{R}$-valued components of functions in $\mathscr{A}$ form a subalgebra of $\mathscr{A}$ satisfying the hypotheses of (a). Therefore (b) follows from (a). □

Applications. Approximating continuous (periodic) functions with trigonometric polynomials, etc.

Remark. Stone's theorem generalizes Weierstrass' theorem, which is the special case for algebras of polynomials.

Techniques

- Proving uniform convergence:
  - Cauchy criterion.
  - Supremum limit criterion (limit criterion in $\mathscr{C}(X)$).
  - Weierstrass M-test.
- Mean value theorem to relate values of functions to values of their derivatives.
- Approximating functions by integrating.
- Approximating functions with polynomials (etc.).
- Reducing $\mathbb{C}$-valued functions to $\mathbb{R}$-valued (component) functions.

5.8 Power Series and Fourier Series

Definitions

Definition 5.8.1. A function $f : \mathbb{C} \to \mathbb{C}$ is *analytic* if it has a power series representation, that is, if it can be written in the form $f(z) = \sum_{n=0}^{\infty} c_n (z - a)^n$.

Definition 5.8.2. The *exponential* function $\exp : \mathbb{C} \to \mathbb{C}$ is defined by

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Also, $e = \exp(1)$.

Definition 5.8.3. The *(natural) logarithm* function $\log : (0, +\infty) \to \mathbb{R}$ is defined by

$$\log(\exp(x)) = x \quad (x \in \mathbb{R}) \qquad \text{or} \qquad \exp(\log(y)) = y \quad (y > 0)$$

Definition 5.8.4. The *cosine* and *sine* functions $\cos : \mathbb{R} \to \mathbb{R}$ and $\sin : \mathbb{R} \to \mathbb{R}$ are defined by

$$\exp(ix) = \cos(x) + i \sin(x) \quad (x \in \mathbb{R})$$

Definition 5.8.5. A *trigonometric polynomial* is an expression of the form

$$f(x) = a_0 + \sum_{n=1}^{N} [a_n \cos(nx) + b_n \sin(nx)] \quad (a_i, b_j \in \mathbb{C}, x \in \mathbb{R})$$

or

$$f(x) = \sum_{n=-N}^{N} c_n \exp(inx) \quad (c_i \in \mathbb{C}, x \in \mathbb{R})$$

Definition 5.8.6. A *trigonometric series* is a function of the form

$$f(x) = \sum_{-\infty}^{\infty} c_n \exp(inx) = \lim_{N \to \infty} \sum_{n=-N}^{N} c_n \exp(inx) \quad (x \in \mathbb{R})$$

Definition 5.8.7. Let $f : \mathbb{R} \to \mathbb{C}$ have period 2π and $f \in \mathscr{R}$ on $[-\pi, \pi]$. The *(trigonometric) Fourier coefficients* of f are

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \exp(-inx)\, dx \quad (n \in \mathbb{Z})$$

The *(trigonometric) Fourier series* of f is given by

$$f(x) \sim \sum_{-\infty}^{\infty} c_n \exp(inx) \quad (x \in \mathbb{R})$$

and the partial sums are denoted $s_{N,f}(x)$.

Definition 5.8.8. Let $f : \mathbb{R} \to \mathbb{C}$ and $f \in \mathscr{R}$ on $[-\pi, \pi]$. The *mean-square norm* of f on $[-\pi, \pi]$ is given by

$$\|f\|_2 = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx}$$

Theorems

Theorem 5.8.1 (Convergence of power series in $\mathbb{C}$)**.** *Let $f(z) = \sum c_n z^n$ converge for $|z| < R$.*

(a) f converges absolutely for $|z| < R$.

(b) For any $0 < \epsilon \le R$, f converges uniformly on $|z| \le R - \epsilon$.

Proof idea. For (a), recall any power series converges absolutely within its radius of convergence (by the root test), and note R is at most the radius of convergence for f.

For (b), use the Weierstrass M-test. For $|z| \le R - \epsilon$,

$$|c_n z^n| = |c_n||z|^n \le |c_n|(R-\epsilon)^n$$

But $R - \epsilon < R$, so $\sum |c_n|(R-\epsilon)^n$ converges by part (a), and the result follows. □

Applications. Proving properties (continuity, differentiability, integrability, etc.) of power series from properties of the terms.

Remark. Uniform convergence does not occur in general on $|z| < R$.

Theorem 5.8.2 (Differentiability of power series in $\mathbb{C}$)**.** *Let $f(z) = \sum c_n z^n$ converge for $|z| < R$. Then f is differentiable term by term, that is,*

$$\Big(\sum_{n=0}^{\infty} c_n z^n\Big)' = \sum_{n=1}^{\infty} n c_n z^{n-1} \qquad (|z| < R)$$

Proof idea. By uniform convergence of power series and the theorem on uniform convergence and differentiation. □

Corollary 5.8.1. *f is infinitely differentiable term by term, that is,*

$$\Big(\sum_{n=0}^{\infty} c_n z^n\Big)^{(k)} = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) c_n z^{n-k} \qquad (|z| < R)$$

and $c_k = f^{(k)}(0)/k!$ for all k.

Proof idea. By induction using the theorem. □

Corollary 5.8.2 (Continuity of power series in $\mathbb{C}$)**.** *f is continuous on $|z| < R$.*

Corollary 5.8.3 (Uniqueness of power series representations in $\mathbb{C}$)**.** *f has a unique power series representation on $|z| < R$.*

Remark. Since $c_k = f^{(k)}/k!$, information can be transferred between derivatives of f and coefficients of its power series representation.

Theorem 5.8.3 (Integrability of power series in $\mathbb{C}$)**.** *Let $f(z) = \sum c_n z^n$ converge for $|z| < R$. Then for any $0 < \epsilon < R$, $f \in \mathscr{R}$ on $[-R+\epsilon, R-\epsilon]$, and*

$$\int_{-R+\epsilon}^{R-\epsilon} \sum c_n x^n \, dx = \Big[\sum \frac{c_n x^{n+1}}{n+1}\Big]_{-R+\epsilon}^{R-\epsilon}$$

Proof idea. By the fundamental theorem of calculus. □

Theorem 5.8.4 (Taylor)**.** *Let $f(x) = \sum c_n x^n$ converge on $(-R, R)$. If $a \in (-R, R)$, then*

$$f(x) = \sum \frac{f^{(n)}(a)}{n!}(x-a)^n \qquad (|x-a| < R - |a|)$$

Proof idea. Massage the power series expansion for f, apply the binomial theorem, then change the order of a double summation.

In detail, prove

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} c_n [a + (x-a)]^n \\ &= \sum_{n=0}^{\infty} c_n \sum_{k=0}^{n} \binom{n}{k} a^{n-k}(x-a)^k \\ &= \sum_{k=0}^{\infty} \Bigg[\sum_{n=k}^{\infty} \binom{n}{k} c_n a^{n-k}\Bigg](x-a)^k \\ &= \sum \frac{f^{(k)}(a)}{k!}(x-a)^k \end{aligned}$$

To justify the change in order of summation, use a lemma:

Lemma 5.8.1. *If $\{a_{ij}\}$ is a double sequence in $\mathbb{C}$ such that for all i, $\sum_{j=0}^{\infty} |a_{ij}| = b_i$, and $\sum_{i=0}^{\infty} b_i$ converges, then*

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{ij} = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} a_{ij}$$

Proof idea. By clever use of continuity.

Let the a_{ij} form a grid. For summing over rows, use a metric space $x_0, x_1, \dots, x_\infty$ with $x_n \to x_\infty$ as $n \to \infty$, and for each i define a row sum function f_i with

$$f_i(x_n) = \sum_{j=0}^{n} a_{ij} \qquad \text{and} \qquad f_i(x_\infty) = \sum_{j=0}^{\infty} a_{ij}$$

Then f_i is continuous at x_∞. Now define a column sum function $g(x) = \sum_{i=0}^{\infty} f_i(x)$. By the assumptions, $f_i \to g$ uniformly, so g is continuous at x_∞ and we can use this to interchange the order of row and column summation:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{ij} = g(x_\infty) = \lim_{x_n \to x_\infty} g(x_n) = \lim_{n\to\infty} \sum_{i=0}^{\infty}\sum_{j=0}^{n} a_{ij} = \lim_{n\to\infty} \sum_{j=0}^{n}\sum_{i=0}^{\infty} a_{ij} = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} a_{ij} \quad \square$$

The conditions of the lemma hold when $|x-a| < R - |a|$. □

Applications. Approximating analytic functions at arbitrary points within their radii of convergence, obtaining better approximations near those points.

Theorem 5.8.5 (Properties of the exponential in $\mathbb{C}$)**.**

(a) $\exp(z)$ *is defined on* $\mathbb{C}$.

(b) $\exp(z+w) = \exp(z)\exp(w)$ *for all* $z, w \in \mathbb{C}$.

(c) $\exp(z) \neq 0$ *for all* $z \in \mathbb{C}$, *and* $\exp(x) > 0$ *for all* $x \in \mathbb{R}$.

(d) $\exp(x)$ *is strictly increasing on* $\mathbb{R}$.

(e) $\exp(x) \to +\infty$ *as* $x \to +\infty$, *and* $\exp(x) \to 0$ *as* $x \to -\infty$.

(f) $x^n \exp(-x) \to 0$ *as* $x \to +\infty$.

(g) $\exp'(z) = \exp(z)$ *for all* $z \in \mathbb{C}$.

(h) $\exp(x) = e^x$ *for all* $x \in \mathbb{R}$.

Proof idea. By looking at the power series expansion $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$.

For (a), by the ratio test.

For (b), by absolute convergence and the binomial theorem,

$$\begin{aligned}\exp(z)\exp(w) &= \sum \frac{z^n}{n!} \sum \frac{w^n}{n!} \\ &= \sum \sum_{k=0}^{n} \frac{z^{n-k} w^k}{(n-k)!k!} \\ &= \sum \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^{n-k} w^k \\ &= \sum \frac{(z+w)^n}{n!} = \exp(z+w)\end{aligned}$$

Note in particular $\exp(z)\exp(-z) = \exp(0) = 1$ for all $z \in \mathbb{C}$, which together with the expansion shows (c)–(e).

For (f), $\exp(x) > x^{n+1}/(n+1)!$ for $x > 0$, so

$$x^n \exp(-x) = \frac{x^n}{\exp(x)} < \frac{(n+1)!}{x} \to 0 \text{ as } x \to +\infty$$

For (g), note

$$\lim_{h\to 0} \frac{\exp(z+h) - \exp(z)}{h} = \exp(z) \lim_{h\to 0} \frac{[\exp(h) - 1]}{h} = \exp(z)$$

For (h), use (b) and induction to establish the result for integers, then take roots to establish it for rationals, then take limits and appeal to continuity of $\exp(x)$ to establish it for reals. □

Theorem 5.8.6 (Properties of the logarithm in $\mathbb{R}$)**.**

(a) $\log(x)$ *is defined on* $(0, +\infty)$.

(b) $\log(xy) = \log(x) + \log(y)$ *for all* $x, y > 0$.

(c) $\log(x)$ *is strictly increasing on* $(0, +\infty)$.

(d) $\log(x) \to +\infty$ *as* $x \to +\infty$, *and* $\log(x) \to -\infty$ *as* $x \to 0$.

(e) $x^{-n}\log(x) \to 0$ *as* $x \to +\infty$ *for all* $n > 0$.

(f) $\log'(x) = 1/x$ *for all* $x > 0$.

Proof idea. By definition of $\log(x)$ and properties of $\exp(x)$. □

Corollary 5.8.4. $\log(x) = \displaystyle\int_1^x \frac{1}{t}\,dt$

Corollary 5.8.5 (Exponentiation in $\mathbb{R}$)**.** *For* $x, y \in \mathbb{R}$ *with* $x > 0$,

$$x^y = \exp(y\log(x))$$

Corollary 5.8.6 (Derivatives of powers in $\mathbb{R}$)**.** *For* $x, y \in \mathbb{R}$ *with* $x > 0$,

$$(x^y)' = yx^{y-1}$$

Theorem 5.8.7 (Properties of trigonometric functions in $\mathbb{C}$)**.**

(a) $\cos^2(x) + \sin^2(x) = 1$ *for all* $x \in \mathbb{R}$.

(b) $\cos'(x) = -\sin(x)$ *and* $\sin'(x) = \cos(x)$ *for all* $x \in \mathbb{R}$.

(c) *There exists unique* $\pi > 0$ *such that* $\pi/2$ *is least* $x > 0$ *with* $\cos(x) = 0$*; also,* $\cos(x)$ *and* $\sin(x)$ *have period* 2π *and* $\exp(z)$ *has period* $2\pi i$.

(d) *If* $z \in \mathbb{C}$ *and* $|z| = 1$*, there exists a unique* $x \in [0, 2\pi)$ *such that* $z = \exp(ix)$.

Proof idea. By definition of $\cos(x)$ and $\sin(x)$ and properties of $\exp(z)$.

For (a), by definitions

$$\cos^2(x) + \sin^2(x) = |\exp(ix)|^2 = \exp(ix)\overline{\exp(ix)} = \exp(ix)\exp(-ix) = 1$$

For (b), by the derivative of $\exp(z)$.

For (c), note $\cos(0) = 1$, and by looking at terms of the series there exist $x > 0$ with $\cos(x) < 0$, hence by continuity of $\cos(x)$ and the intermediate value theorem there is a least $x_0 > 0$ with $\cos(x_0) = 0$. Set $\pi = 2x_0$. By (a) and (b), $\sin(\pi/2) = 1$, hence $\exp(2\pi i) = 1$, and the periodicity properties follow.

For (d), write $z = x + iy$, then consider the possible quadrants of the complex plane in which x and y may lie and appeal to the intermediate value theorem on $\cos(x)$ or $\sin(x)$. □

Corollary 5.8.7 (Polar form in $\mathbb{C}$)**.** *If* $z \in \mathbb{C}$ *and* $z \neq 0$*, there exist unique* $r \geq 0$ *and* $\theta \in [0, 2\pi)$ *such that* $z = r\exp(i\theta)$.

Applications. Simplifying complex multiplication, showing that it can be interpreted geometrically as scaling and rotation in the complex plane (note if $z = re^{i\alpha}$ and $w = se^{i\beta}$, $zw = rse^{i(\alpha+\beta)}$).

Remark. The function $\exp(ix)$ parametrizes the unit circle in the complex plane over $[0, 2\pi)$. The functions $\cos(x)$ and $\sin(x)$, and the number π, can be shown to have their usual geometric properties.

Theorem 5.8.8 (Algebraic closure of $\mathbb{C}$)**.** *Let $P(z)$ be a nonconstant polynomial over $\mathbb{C}$. Then $P(z)$ has a root in $\mathbb{C}$.*

Proof idea. By contradiction, using polar form to 'control' multiplication in $\mathbb{C}$ and exhibit a problem if there is no root.

If $P(z)$ has no root, let $\mu = \inf_{z\in\mathbb{C}} |P(z)|$. Show that $P(z)$ takes on magnitude μ at some point z_0 by finding a closed disc outside of which $|P(z)|$ is large, then appealing to the extreme value theorem inside.

Simplify things a bit by writing $Q(z) = P(z + z_0)/P(z_0)$, so Q is a nonconstant polynomial, $Q(0) = 1$, and $|Q(z)| \geq 1$ for all $z \in \mathbb{C}$. Now using an upper bound for $Q(z)$ in terms of z, construct $z = re^{i\theta}$ such that $|Q(z)| < 1$—a contradiction. □

Applications. Factoring polynomials in $\mathbb{C}$ into linear factors, etc.

Theorem 5.8.9 (Properties of trigonometric polynomials)**.** *Let $f(x) = \sum_{-N}^{N} c_n e^{inx}$.*

(a) f has period 2π.

(b) $c_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-imx}\,dx$ for all $m \in \mathbb{Z}$.

(c) f has a unique trigonometric polynomial representation.

(d) f is $\mathbb{R}$-valued iff $\overline{c_n} = c_{-n}$ for $0 \leq n \leq N$.

Proof idea. For (a), by periodicity of the complex exponential function.

For (b), by the fundamental theorem of calculus and periodicity of the complex exponential function,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx} = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0 \end{cases} \qquad (n \in \mathbb{Z})$$

The rest follows.

Note (c) follows from (b), and (d) follows from (b) and (c). □

Theorem 5.8.10 (Trigonometric polynomial approximation)**.** *Let $f : \mathbb{R} \to \mathbb{C}$ be continuous with period 2π. Then there exists a sequence $\{P_n\}$ of trigonometric polynomials such that $P_n \to f$ uniformly.*

Proof idea. By Stone's Theorem.

Regard f and the trigonometric polynomials as functions on the unit circle in the complex plane. The trigonometric polynomials form a self-adjoint algebra which separates points and vanishes nowhere. □

Theorem 5.8.11 (Least-squares property of Fourier sums)**.** *Let $f : \mathbb{R} \to \mathbb{C}$ have period 2π and $f \in \mathscr{R}$ on $[-\pi, \pi]$. If*

$$t_N(x) = \sum_{-N}^{N} a_n e^{inx} \quad (a_i \in \mathbb{C}, x \in \mathbb{R})$$

then

$$\|f - s_{N,f}\|_2 \leq \|f - t_N\|_2$$

and equality holds iff $a_n = c_n$ for $-N \leq n \leq N$.

Proof idea. Expand and massage.

$$\begin{aligned}\int |f-t_N|^2 &= \int |f|^2 - \int f\overline{t_N} - \int \overline{f} t_N + \int |t_N|^2 \\ &= \int |f|^2 - 2\pi\Big(\sum c_n\overline{a_n} + \sum \overline{c_n} a_n - \sum a_n\overline{a_n}\Big) \\ &= \int |f|^2 + 2\pi\Big(\sum |c_n - a_n|^2 - \sum |c_n|^2\Big)\end{aligned}$$

The expression on the right is minimized iff $a_n = c_n$. □

Applications. Showing that among trigonometric polynomials, the partial sums of the Fourier series for f provide the best mean-square approximation to f.

Corollary 5.8.8 (Bessel inequality)**.**

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |s_{N,f}(x)|^2\,dx \le \sum_{-\infty}^{\infty} |c_n|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx$$

In particular, $c_n \to 0$ as $n \to \infty$.

Proof idea. Take $a_n = c_n$ in the last equality of the proof. □

Applications. Showing that the mean-square magnitudes of the partial sums of the Fourier series for f are bounded above by the mean-square magnitude of f.

Theorem 5.8.12 (Mean-square convergence of Fourier sums)**.** *Let $f : \mathbb{R} \to \mathbb{C}$ have period 2π and $f \in \mathscr{R}$ on $[-\pi,\pi]$. Then*

$$\lim_{N\to\infty} \left\|f - s_{N,f}\right\|_2 = 0$$

Proof idea. By trigonometric polynomial approximation.

In detail, given $\epsilon > 0$, construct continuous periodic g with $\left\|f-g\right\|_2 < \epsilon/3$, and choose a trigonometric polynomial P of degree N_0 with $\left\|g-P\right\|_2 < \epsilon/3$. Then by the least-squares property of Fourier sums for g,

$$\left\|g - s_{N,g}\right\|_2 < \epsilon/3 \quad (N \ge N_0)$$

By the Bessel inequality,

$$\left\|s_{N,f} - s_{N,g}\right\|_2 = \left\|s_{N,f-g}\right\|_2 \le \left\|f-g\right\|_2 < \epsilon/3$$

Therefore by the triangle inequality,

$$\left\|f - s_{N,f}\right\|_2 \le \left\|f-g\right\|_2 + \left\|g - s_{N,g}\right\|_2 + \left\|s_{N,g} - s_{N,f}\right\|_2 \le \epsilon \quad (N \ge N_0)$$

The result follows as $\epsilon \to 0$. □

Remark. Pointwise [uniform, etc.] convergence for Fourier series is more delicate than for power series, and we do not cover this. It is worth noting informally that convergence of a Fourier series at a point follows from ‘localized’ properties of the function at the point (e.g. differentiability), unlike with power series.

Techniques

- Getting information about analytic functions from power series expansions and conversely, including transferring information between coefficients and derivative values to relate local and global behavior.
- Approximating analytic functions using Taylor series expansions.
- Approximating continuous periodic functions using Fourier series expansions.

Chapter 6

Complex Analysis

This chapter covers complex analysis from [2], with supplementary information from [9].

6.1 Complex Numbers

Definitions

Basic definitions assumed.

Theorems

Theorem 6.1.1 (Existence of $\mathbb{C}$)**.** *There exists an algebraically closed field $\mathbb{C}$ (the complex numbers) containing $\mathbb{R}$ as a subfield.*

Proof idea. Let complex numbers be ordered pairs of real numbers with addition and multiplication defined appropriately. Prove algebraic closure later. □

Applications. All subsequent theory.

Theorem 6.1.2 (Polar form)**.** *If $z \neq 0$, there exist unique $r \geq 0$ and $\theta \in (-\pi, \pi]$ such that*

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta}$$

Proof idea. By geometry in the complex plane and Euler's identity (definition). □

Applications. Multiplication, geometry of multiplication as scaling and rotation.

Theorem 6.1.3 (Roots)**.** *If $z = re^{i\theta}$ and $n \geq 1$, the n distinct n-th roots of z are given by*

$$w_k = \sqrt[n]{r}\exp\left(\frac{\theta + 2k\pi}{n}\right) \quad (k = 0, \ldots, n-1)$$

Proof idea. Polar form.

In detail, if $w = se^{i\phi}$ is an n-th root of z, then

$$s^n e^{in\phi} = w^n = z = re^{i\theta}$$

so by geometry in the complex plane, $s^n = r$ and $n\phi = \theta + 2k\pi$ for some integer k. The distinct roots are obtained by letting $k = 0, \dots, n-1$. □

Techniques

- Using polar form for multiplication.

6.2 Analyticity

Definitions

Definition 6.2.1. Let f be defined in some deleted neighborhood of z_0 and w_0 be fixed. If for each $\epsilon > 0$ there exists $\delta > 0$ such that $|f(z) - w_0| < \epsilon$ whenever $0 < |z - z_0| < \delta$, then w_0 is called a *limit* of $f(z)$ as z approaches z_0, denoted

$$\lim_{z \to z_0} f(z) = w_0 \qquad \text{or} \qquad f(z) \to w_0 \text{ as } z \to z_0$$

Definition 6.2.2. A function f is *continuous* at a point z_0 if $\lim_{z \to z_0} f(z) = f(z_0)$.

Definition 6.2.3. A function f is *differentiable* at a point z_0 if

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists. This limit is called the *derivative* of f at z_0, and is denoted $f'(z_0)$.

This induces a function f', the first derivative of f. This can be continued with f'', f''', etc. In general, $f^{(n)}$ denotes the n-th derivative of f.

Definition 6.2.4. A function is *analytic* (or *holomorphic*) at a point z_0 if it is differentiable throughout some neighborhood of z_0. It is *entire* if it is analytic everywhere.

Definition 6.2.5. A real function $f(x, y)$ of real variables is *harmonic* in a domain if it has continuous first and second partial derivatives and

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

in the domain.

Definition 6.2.6. Let $u(x, y)$ and $v(x, y)$ be real functions of real variables. Then v is a *harmonic conjugate* of u in a domain if u and v are harmonic and satisfy the Cauchy-Riemann equations in the domain.

Theorems

Theorem 6.2.1 (Limits)**.** *Let $f = u + iv$, $z = x + iy$, $z_0 = x_0 + iy_0$, and $w_0 = u_0 + iv_0$. Then*

$$\lim_{z \to z_0} f(z) = w_0 \iff \lim_{(x,y)\to(x_0,y_0)} u(x,y) = u_0 \quad \textit{and} \quad \lim_{(x,y)\to(x_0,y_0)} v(x,y) = v_0$$

Proof idea. By the triangle inequality and definition of absolute value. □

Applications. Transferring properties of limits [continuity, etc.] in $\mathbb{R}$ to properties of limits [continuity, etc.] in $\mathbb{C}$.

Corollary 6.2.1 (Limits and field operations)**.** *If f and g have limits at z_0, then*

(a) $\lim_{z\to z_0}(f+g)(z) = \lim_{z\to z_0} f(z) + \lim_{z\to z_0} g(z)$.

(b) $\lim_{z\to z_0}(fg)(z) = \lim_{z\to z_0} f(z) \lim_{z\to z_0} g(z)$.

(c) $\lim_{z\to z_0}(f/g)(z) = \lim_{z\to z_0} f(z)/\lim_{z\to z_0} g(z)$ *if* $\lim_{z\to z_0} g(z) \neq 0$.

Corollary 6.2.2 (Continuity)**.** *Let $f = u + iv$ be defined in a neighborhood of $z_0 = x_0 + iy_0$. Then f is continuous at z_0 iff u and v are continuous at (x_0, y_0).*

Corollary 6.2.3 (Continuity and field operations)**.** *If f and g are continuous at z_0, then*

(a) $f + g$ is continuous at z_0.

(b) fg is continuous at z_0.

(c) f/g is continuous at z_0 if $g(z_0) \neq 0$.

Theorem 6.2.2 (Continuity and composition)**.** *If f is continuous at z_0 and g is continuous at $f(z_0)$, then $g \circ f$ is continuous at z_0.*

Proof idea. Nested epsilons. □

Since the definition of complex differentiability is *formally* identical to that of real differentiability, the following few theorems are proved just as for real functions.

Theorem 6.2.3 (Differentiability implies continuity)**.** *If f is differentiable at z_0, then f is continuous at z_0.*

Theorem 6.2.4 (Derivatives and field operations)**.** *If f and g are differentiable at z_0, then*

(a) $(f+g)'(z_0) = f'(z_0) + g'(z_0)$.

(b) $(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0)$.

(c) $(f/g)'(z_0) = \dfrac{f'(z_0)g(z_0) - f(z_0)g'(z_0)}{[g(z_0)]^2}$ *if* $g(z_0) \neq 0$.

Theorem 6.2.5 (Chain rule)**.** *If f is differentiable at z_0 and g is differentiable at $f(z_0)$, then*

$$(g \circ f)'(z_0) = g'(f(z_0))f'(z_0)$$

Theorem 6.2.6 (Cauchy-Riemann equations)**.** *If $f = u + iv$ is differentiable at $z_0 = x_0 + iy_0$, then u and v have first partial derivatives at (x_0, y_0) satisfying*

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \qquad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0)$$

and

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i\frac{\partial v}{\partial x}(x_0, y_0)$$

Proof idea. Look at the limit of the difference quotient, letting $z \to z_0$ two ways, first horizontally then vertically. □

Corollary 6.2.4. *If $z_0 = r_0 e^{i\theta_0}$, equivalently*

$$r\frac{\partial u}{\partial r}(r_0, \theta_0) = \frac{\partial v}{\partial \theta}(r_0, \theta_0) \qquad \frac{\partial u}{\partial \theta}(r_0, \theta_0) = -r\frac{\partial v}{\partial r}(r_0, \theta_0)$$

and

$$f'(z_0) = e^{-i\theta_0}\left[\frac{\partial u}{\partial r}(r_0, \theta_0) + i\frac{\partial v}{\partial r}(r_0, \theta_0)\right]$$

Proof idea. Multivariable chain rule, with $x = r\cos\theta$ and $y = r\sin\theta$. □

Applications. Proving properties of differentiability, proving non-differentiability.

Remark. The converse is false! For f to be complex differentiable at a point, u and v must be real differentiable there (see [9, §II.5]). This can be summarized as:

complex differentiability = real differentiability + Cauchy-Riemann equations

Remark. Compare that complex limits [continuity] can be fully characterized in terms of real limits [continuity] of components, whereas complex differentiability cannot be analogously characterized. Complex differentiability is stronger than (even multivariable) real differentiability of components, owing ultimately to the complex division present in the definition of the complex derivative.

Corollary 6.2.5. *If $f' = 0$ in a domain D, then f is constant on D.*

Proof idea. By the Cauchy-Riemann equations, and then the mean value theorem applied to the components of f. □

Theorem 6.2.7 (Sufficient condition for differentiability)**.** *If $f = u + iv$ and u and v have first partial derivatives which are continuous at z_0 and satisfy the Cauchy-Riemann equations at z_0 (in rectangular or polar form), then f is differentiable at z_0.*

Proof idea. Recall from multivariable analysis that continuity of the partials implies real differentiability of u and v. This together with the Cauchy-Riemann equations implies that the difference quotient of f has a limit. □

Applications. Proving differentiability.

Theorem 6.2.8 (Characterizations of analyticity)**.** *Let $f = u + iv$ be defined on a domain D. The following are equivalent:*[1]

(a) *f is analytic on D.*

(b) *u and v have continuous partial derivatives of all orders on D, and satisfy the Cauchy-Riemann equations on D.*

(c) *v is a harmonic conjugate of u on D.*

Proof idea. For (a) $\Longrightarrow$ (b), we know about the Cauchy-Riemann equations and prove the rest later.

For (b) $\Longrightarrow$ (c), recall from multivariable analysis that the mixed second partials of u and v are equal, so

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\Big(\frac{\partial u}{\partial x}\Big) = \frac{\partial}{\partial x}\Big(\frac{\partial v}{\partial y}\Big) = \frac{\partial}{\partial y}\Big(\frac{\partial v}{\partial x}\Big) = \frac{\partial}{\partial y}\Big(-\frac{\partial u}{\partial y}\Big) = -\frac{\partial^2 u}{\partial y^2}$$

So u is harmonic, and similarly for v.

For (c) $\Longrightarrow$ (a), by the sufficient condition for differentiability. □

Techniques

- Reducing limits and continuity of complex-valued functions to that of real-valued component functions.
- Proving differentiability (or non-differentiability) and calculating derivatives:
  - Limit of difference quotient.
  - Closure (composites, sums, products, etc.).
  - Cauchy-Riemann equations + real differentiability.
  - Harmonic conjugates.
- Using the Cauchy-Riemann equations to help apply results from real analysis to component functions.

6.3 Elementary Functions

Definitions

Definitions of real elementary functions assumed.

Definition 6.3.1. The *exponential function* is given by

$$e^z = e^x(\cos y + i \sin y)$$

where $z = x + iy$. It is also denoted $\exp z$.

[1] Another important characterization of analyticity involves *conformality,* that is, the preservation of angles between intersecting curves. See [9, §II.11–12].

Definition 6.3.2. The trigonometric functions cos and sin are given by

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}$$

The rest of the trigonometric functions are then defined as usual.

Definition 6.3.3. The hyperbolic functions cosh and sinh are given by

$$\cosh z = \frac{e^{z} + e^{-z}}{2} \qquad \sinh z = \frac{e^{z} - e^{-z}}{2}$$

The rest of the hyperbolic functions are then defined as usual.

Definition 6.3.4. For $z \neq 0$, the multivalued *argument function* $\arg z$ gives the angles of z from polar form. The *principal argument function* $\operatorname{Arg} z$ gives the unique angle in $(-\pi, \pi]$.

Definition 6.3.5. The multivalued *logarithmic function* is given by

$$\log z = \ln|z| + i \arg z \quad (z \neq 0)$$

Definition 6.3.6. The multivalued *generalized power function* is given by

$$z^c = \exp(c \log z) \quad (z \neq 0)$$

Definition 6.3.7. The multivalued *generalized exponential function* is given by

$$c^z = \exp(z \log c) \quad (c \neq 0)$$

Definition 6.3.8. A *branch* of a multivalued function f is a function f^* defined on some domain D which is analytic on D and takes on values of f in D.

A *branch cut* of f is a portion of a line or curve used to define a branch of f.

A *branch point* of f is a point common to all branch cuts of f.

Definition 6.3.9. The *principal branch* of the logarithmic [generalized power, generalized exponential] function uses angles in $(-\pi, \pi)$ and is denoted $\operatorname{Log} z$ [P.V. z^c, P.V. c^z].

Theorems

Basic properties of elementary functions assumed.

Remark. The following functions are all multivalued without suitable restrictions (e.g., to branches):

- arg (argument)
- log (logarithm)
- z^c (generalized power function, including n-th root function $z^{1/n}$)
- c^z (generalized exponential function)

- Inverse trigonometric functions
- Inverse hyperbolic functions

All of this owes ultimately to the periodicity of angles in the complex plane.

There is sometimes ambiguity. For instance, does $e^{1/n}$ denote the single real value $\sqrt[n]{e}$, or the set of n complex n-th roots of e? The former is conventionally the value of e^z at $z = 1/n$, while the latter is the value of $z^{1/n}$ at $z = e$. Also, note c^z is multivalued in general, while e^z is not (e^z is obtained from c^z by taking the principal branch with $c = e$). Ambiguity is resolved by context.

6.4 Integrals

Definitions

Definition 6.4.1. An *arc* (or *curve* or *path*) is a continuous function $\gamma : [a, b] \to \mathbb{C}$.

Definition 6.4.2. An arc is *simple* if it does not cross itself, *closed* if its endpoints meet, and *simple closed* if only its endpoints meet.

Definition 6.4.3. An arc is *smooth* if it is continuously differentiable everywhere and its derivative is nonzero everywhere except possibly at its endpoints.

Definition 6.4.4. A *contour* is a piecewise smooth arc.

Definition 6.4.5. A simple closed contour is *positively oriented* if it is parametrized in the counterclockwise direction, and *negatively oriented* if it is parametrized in the clockwise direction.

Definition 6.4.6. A domain is *simply connected* if every simple closed contour lying in the domain encloses only points in the domain. Otherwise it is *multiply connected.*

Remark. Intuitively, a domain is simply connected if it has no holes.

Theorems

Theorem 6.4.1 (Properties of contour integrals)**.**

(a) (Linearity) If f and g are piecewise continuous on the contour C and $z_0 \in \mathbb{C}$, then

$$\int_C f + g = \int_C f + \int_C g \qquad \int_C z_0 f = z_0 \int_C f$$

(b) (Additivity) If f is piecewise continuous on the contour $C_1 + C_2$, then

$$\int_{C_1+C_2} f = \int_{C_1} f + \int_{C_2} f$$

(c) *(Boundedness) If f is piecewise continuous on the contour C and $|f| \le M$ on C, then*

$$|\int_C f| \le ML$$

where L is the length of C. Such an M always exists, and if C is parametrized by $z(t)$ on $[a,b]$, then $L = \int_a^b |z'(t)|\,dt$.

Proof idea. By definition of the contour integral and results from real analysis. □

Theorem 6.4.2 (Fundamental theorem of calculus)**.** *Let f be continuous on a domain D.*

(a) *If integrals of f are independent of path in D and $z_0 \in D$, then*

$$F(z) = \int_{z_0}^{z} f \qquad (z \in D)$$

is analytic in D, and $F' = f$ in D.

(b) *If $F' = f$ in D, then integrals of f are independent of path in D. In particular, if $z_0, z_1 \in D$ and C is any contour in D from z_0 to z_1, then*

$$\int_C f = F(z_1) - F(z_0)$$

Proof idea. For (a), use a line segment from z_0 to z and the proof from real analysis.

For (b), use the definition of the contour integral and the fundamental theorem of calculus from real analysis. In detail, assume without loss of generality that C is a smooth arc parametrized by $z(t)$ on $[a,b]$. Then

$$\frac{d}{dt}F(z(t)) = F'(z(t))z'(t) = f(z(t))z'(t)$$

Therefore

$$\int_C f = \int_a^b f(z(t))z'(t)\,dt = F(z(b)) - F(z(a)) = F(z_1) - F(z_0) \qquad \square$$

Applications. Showing that integration and differentiation are inverse processes, calculating integrals and derivatives, etc.

Corollary 6.4.1. *Let f be continuous on a domain D. Then the following are equivalent:*

(a) *f has an antiderivative in D*

(b) *Integrals of f are independent of path in D.*

(c) *Integrals of f along closed contours in D vanish.*

Theorem 6.4.3 (Jordan)**.** *Let C be a simple closed contour. Then C is the boundary of two distinct domains, an interior which is bounded and an exterior which is unbounded.*[2]

[2] Not proved in [2].

Applications. Cauchy-Goursat, etc.

Theorem 6.4.4 (Cauchy-Goursat)**.** *If f is analytic on and inside a simple closed contour C, then*

$$\int_C f = 0$$

Proof idea. Let R be the region consisting of C and its interior. Break the integral over C up into a sum of integrals over adjacent squares covering R. Approximate these integrals by choosing inside each square a sample point at which the derivative of f well approximates difference quotients, and thus relates to values of f on the square. Show that these integrals can be made arbitrarily small by taking arbitrarily small squares, and hence vanish.

To show that suitable squares and sample points can always be chosen, use the analyticity of f in a subdivision argument. By the Jordan curve theorem, R can be covered with a single square. If no sample point exists in this square, subdivide it into four smaller squares; if sample points do not exist in each of these squares, subdivide the ones without a sample point; and so on. This process generates a descending tree of squares. If the process never ends, take a limit point in some descending chain. Since f is analytic at this point, difference quotients are well approximated nearby. But this includes small enough squares in the chain, so the process must end after all. □

Corollary 6.4.2 (Contours inside a contour)**.** *Let C be a simple closed contour. Let $C_0, \dots, C_k$ be simple closed contours inside C, of the same orientation as C, which are pairwise disjoint and whose interiors are pairwise disjoint. Then if f is analytic on all these contours and between C and the $C_0, \dots, C_k$, then*

$$\int_C f = \sum_{i=0}^{k} \int_{C_i} f$$

Proof idea. Draw polygonal lines connecting the contours $C, C_0, \dots, C_k$ in order to express $\int_C f + \sum \int_{-C_i} f$ as a sum of integrals about simple closed contours. □

Applications. Calculating integrals, deformation of path, Cauchy's residue theorem.

Corollary 6.4.3 (Deformation of path)**.** *Let C be a simple closed contour, and let C_0 be a simple closed contour inside C of the same orientation. If f is analytic on and between C and C_0, then*

$$\int_C f = \int_{C_0} f$$

Applications. Simplifying integrals, Cauchy's integral formula, Laurent series.

Corollary 6.4.4 (Analyticity implies antiderivative)**.** *If f is analytic on a simply connected domain D, f has an antiderivative on D.*

Proof idea. Use a stronger version of Cauchy-Goursat which states that whenever f is analytic on a simply connected domain, integrals of f around *arbitrary* closed contours inside the domain vanish.[3] Then apply the fundamental theorem. □

[3] Not proved in [2].

Theorem 6.4.5 (Cauchy integral formula)**.** *If f is analytic on and inside a positively oriented simple closed contour C and z_0 is inside C, then*

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz \qquad (n \geq 0)$$

Proof idea. Prove case $n = 0$, leaving the general case unproved.[4]

Use deformation of path to simplify the integral. Let C_ρ be a positively oriented circle of radius ρ about z_0, inside C. Then

$$\begin{aligned} \int_C \frac{f(z)}{z-z_0}\, dz - 2\pi i f(z_0) &= \int_{C_\rho} \frac{f(z)}{z-z_0}\, dz - \int_{C_\rho} \frac{f(z_0)}{z-z_0}\, dz \\ &= \int_{C_\rho} \frac{f(z)-f(z_0)}{z-z_0}\, dz \end{aligned}$$

By continuity of f at z_0, the last integral can be made arbitrarily small by letting $\rho \to 0$, so it vanishes. □

Applications. Relating local behavior of analytic functions and their derivatives at points to global behavior along contours, calculating function and derivative values, calculating integrals, calculating residues, Taylor series, Laurent series, etc.

Corollary 6.4.5 (Cauchy inequality)**.** *If f is analytic on and inside a positively oriented circle C of radius R about z_0, and $|f| \leq M$ on C, then*

$$|f^{(n)}(z_0)| \leq \frac{n!M}{R^n} \qquad (n \geq 0)$$

Corollary 6.4.6. *If f is analytic at a point, then f is infinitely analytic at the point.*

Corollary 6.4.7. *If $f = u + iv$ is analytic at a point, then u and v have continuous partial derivatives of all orders at the point.*

Remark. By the above, if f is analytic at a point, local behavior of f at the point is related to global behavior of f. Moreover, f can be infinitely differentiated and antidifferentiated at the point. In this way, analytic functions are like power series. This is confirmed by the fact that f is analytic at a point iff it has a Taylor series expansion there.

Theorem 6.4.6 (Morera)**.** *Let f be continuous on a domain D. If $\int_C f = 0$ for every closed contour C inside D, then f is analytic on D.*

Proof idea. By the fundamental theorem, f has an antiderivative on D. So f is the derivative of an analytic function on D, hence is analytic on D. □

Applications. A converse to Cauchy-Goursat (for continuous functions, on simply connected domains).

Theorem 6.4.7 (Liouville)**.** *If f is bounded and entire, then f is constant.*

[4] Not proved in [2] for $n > 2$.

Proof idea. By Cauchy's inequality, show that $f' = 0$ by letting $R \to \infty$. □

Applications. Fundamental theorem of algebra.

Theorem 6.4.8 (Fundamental theorem of algebra)**.** *If P is a nonconstant polynomial, then P has a root.*

Proof idea. By Liouville's theorem. If P has no root, then $1/P$ is bounded and entire, hence constant. But then P is constant, which is false. □

Applications. Algebraic closure of $\mathbb{C}$, factoring polynomials into linear factors, etc.

Theorem 6.4.9 (Maximum modulus principle)**.** *If f is nonconstant and analytic on a domain D, and then $|f|$ has no maximum value on D.*

Proof idea. Use Cauchy's integral formula and integral bounds to argue that the value of $|f|$ at a point in D is the arithmetic mean of the values of $|f|$ on any circle in D surrounding the point. So if $|f|$ has a local maximum in D, f is constant nearby.

Now if $|f|$ has a maximum value on D, any two points in D are connected by a sequence of overlapping neighborhoods on each of which f is constant, so f is constant on D. □

Corollary 6.4.8. *If f is nonconstant and continuous on a closed and bounded region R, and analytic interior to R, then $|f|$ attains a maximum value on the boundary of R but not interior to R.*

Proof idea. By the extreme value theorem from real analysis, $|f|$ attains a maximum value on R. By the maximum modulus principle, it must occur on the boundary. □

Techniques

- Evaluating contour integrals:
  - Definition (and results from real analysis).
  - Properties of integrals (linearity, additivity, etc.).
  - Fundamental theorem of calculus.
  - Cauchy-Goursat.
  - Deformation of path.
  - Cauchy's integral formula (and Cauchy's inequality).
- Using subdivision to construct descending chains and take limit points.
- Relating differentiation and integration using the fundamental theorem of calculus.
- Relating local properties of functions and their derivatives to global properties using Cauchy's integral formula, Cauchy's inequality, etc.
- Transferring local properties of functions to global properties using a sequence of overlapping neighborhoods in a connected domain.

6.5 Series

Definitions

Definitions related to numerical sequences and series assumed.

Definition 6.5.1. A *power series* is a function of the form $f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$.

Definition 6.5.2. A power series $f(z) = \sum a_n(z-z_0)^n$ *converges* at a point z_1 if $f(z_1)$ is defined, that is, if $\sum a_n(z_1-z_0)^n$ converges. Otherwise it *diverges* at z_1.

Definition 6.5.3. A power series $f(z) = \sum a_n(z-z_0)^n$ *converges absolutely* at a point z_1 if

$$\sum |a_n(z_1-z_0)^n|$$

converges.

Definition 6.5.4. A power series $f(z) = \sum a_n(z-z_0)^n$ *converges uniformly* on the set S if for every $\epsilon > 0$, there exists N such that for all $n \geq N$,

$$|f(z) - \sum_{k=0}^{n} a_k(z-z_0)^k| < \epsilon$$

for all $z \in S$.

Theorems

Theorem 6.5.1 (Properties of power series)**.** *Let*

$$f(z) = \sum a_n(z-z_0)^n$$

be an arbitrary power series about z_0.

(a) *(Radius of convergence) There exists a unique* R *with* $0 \leq R \leq +\infty$ *such that* f *converges on* $|z-z_0| < R$ *and diverges on* $|z-z_0| > R$.

(b) *(Absolute convergence)* f *converges absolutely inside its radius of convergence.*

(c) *(Uniform convergence)* f *converges uniformly on any closed set inside its radius of convergence.*

(d) *(Continuity)* f *is continuous inside its radius of convergence.*

(e) *(Integrability)* f *is integrable term by term along any contour* C *inside its radius of convergence, that is,*

$$\int_C \sum a_n(z-z_0)^n = \sum \int_C a_n(z-z_0)^n$$

(f) *(Differentiability)* f *is analytic inside its radius of convergence. Moreover,* f *is differentiable term by term there, that is,*

$$\frac{d}{dz}\sum_{n=0}^{\infty} a_n(z-z_0)^n = \sum_{n=1}^{\infty} a_n n(z-z_0)^{n-1}$$

(g) *(Uniqueness) f has a unique power series expansion about z_0.*

(h) *(Arithmetic) If $g(z) = \sum b_n(z-z_0)^n$ is another power series about z_0, then inside the smaller of the radii of convergence of f and g, power series for the functions $f+g$, fg, and f/g (when $g(z)$ is never zero) can be obtained by formally adding, multiplying, and dividing the power series for f and g.*

Proof idea. For (a)–(e), (g), and (h), use the same proofs from real analysis.

For (f), use Morera's theorem and Cauchy's integral formula. Note from (e) and the fact that each term $a_n(z-z_0)^n$ has an antiderivative that if C is a closed contour inside the radius of convergence, $\int_C f = 0$. Thus f is analytic by Morera's theorem. Term by term differentiability follows from an extension of (e) and Cauchy's integral formula. □

Theorem 6.5.2 (Taylor series)**.** *If f is analytic at z_0, then f has a Taylor series expansion about z_0. That is, there exists* $0 < R \leq +\infty$ *such that*

$$f(z) = \sum a_n(z-z_0)^n \qquad (|z-z_0| < R)$$

with coefficients

$$a_n = \frac{f^{(n)}(z_0)}{n!} \qquad (n \geq 0)$$

Moreover, these coefficients are unique.

Proof idea. For fixed z, use Cauchy's integral formula to express $f(z)$ as an integral. Then split the integrand up into a partial sum of a geometric series plus a remainder term. Use linearity and then Cauchy's integral formula again on the terms of the sum to obtain the Taylor coefficients. Finally, show that the error vanishes.

In detail, assume without loss of generality that $z_0 = 0$. Fix z with $|z| < R$, and let C be a positively oriented circle about the origin of radius R_0 where $|z| < R_0 < R$. Then

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(s)}{s-z}\,ds$$

Write

$$\frac{1}{s-z} = \frac{1}{s}\left[\frac{1}{1-(z/s)}\right] = \frac{1}{s}\left[\sum_{n=0}^{N-1}\left(\frac{z}{s}\right)^n - \frac{(z/s)^N}{1-(z/s)}\right] = \sum_{n=0}^{N-1}\frac{1}{s^{n+1}}z^n + \frac{z^N}{s^N(s-z)}$$

Then

$$\begin{aligned} f(z) &= \sum_{n=0}^{N-1}\left[\frac{1}{2\pi i}\int_C \frac{f(s)}{s^{n+1}}\,ds\right]z^n + \frac{z^N}{2\pi i}\int_C \frac{1}{s^N(s-z)}\,ds \\ &= \sum_{n=0}^{N-1}\frac{f^{(n)}(0)}{n!}z^n + \rho_N(z) \end{aligned}$$

where $\rho_N(z)$ is the remainder term, and $\rho_N(z) \to 0$ as $N \to \infty$.

Uniqueness follows from uniqueness of power series expansions. □

Applications. Manipulation, approximation, relating local and global behavior, calculating integrals, calculating residues, etc.

Theorem 6.5.3 (Laurent series)**.** *If f is analytic in the annular domain $R_0 < |z - z_0| < R_1$ and C is any positively oriented simple closed contour about z_0 inside the domain, then f has a Laurent series expansion in the domain that uses C. That is,*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} b_n (z - z_0)^{-n} \qquad (R_0 < |z - z_0| < R_1)$$

where

$$a_n = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz \quad (n \geq 0) \qquad b_n = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^{-n+1}}\, dz \quad (n \geq 1)$$

Moreover, these coefficients are unique.

Proof idea. For fixed z, draw circles in the annular domain and use Cauchy-Goursat and Cauchy's integral formula to express $f(z)$ as a sum of two integrals. As in the proof for Taylor series, break these integrals up into sums and use Cauchy's integral formula again to obtain the desired form. Use deformation of path to express the coefficients in terms of C.

Again, uniqueness follows from uniqueness of power series expansions. □

Applications. Manipulation, approximation, relating local and global behavior, calculating integrals, calculating residues, Cauchy's residue theorem, etc.

Remark. The Laurent series is a generalization of the Taylor series.

Techniques

- Obtaining Taylor [Laurent] coefficients using Cauchy's integral formula.
- Obtaining Taylor [Laurent] series expansions from existing Taylor [Laurent] series expansions using formal manipulations and uniqueness (and not using definitions!).
- Relating local properties of functions and their derivatives to global properties using Taylor [Laurent] series expansions.

6.6 Residues and Poles

Definition

Definition 6.6.1. If f is not analytic at z_0 but is analytic at a point in every neighborhood of z_0, then z_0 is a *singularity* of f. If f is also analytic in some deleted neighborhood of z_0, then z_0 is an *isolated singularity* of f.

Definition 6.6.2. If f has an isolated singularity at z_0, the *residue* of f at z_0, denoted $\operatorname{Res}_{z=z_0} f(z)$, is the coefficient of $(z - z_0)^{-1}$ in the Laurent series expansion of f about z_0, that is, $(1/2\pi i) \int_C f$ where C is any positively oriented simple closed contour about z_0 lying in a deleted neighborhood of z_0 in which f is analytic.

Definition 6.6.3. Let f have an isolated singularity at z_0, with Laurent series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n + \sum_{n=1}^{\infty} b_n(z-z_0)^{-n}$$

If $m \geq 0$ is least such that $b_n = 0$ for all $n > m$, then z_0 is a *pole of order* m of f. Otherwise, if $b_n \neq 0$ for arbitrarily large n, then z_0 is an *essential singularity* of f. A pole of order 0 is a *removable singularity*, and a pole of order 1 is a *simple pole*.

Definition 6.6.4. If f is analytic at z_0 and there exists $m > 0$ such that $f^{(n)}(z_0) = 0$ for $0 \leq n < m$ and $f^{(m)}(z_0) \neq 0$, then z_0 is a *zero of order* m of f.

Theorems

Theorem 6.6.1 (Cauchy residue theorem)**.** *If f is analytic at every point on and inside a positively oriented simple closed contour C except for a finite number of points $z_0, \ldots, z_n$ inside C, then*

$$\int_C f = 2\pi i \sum_{k=0}^{n} \operatorname{Res}_{z=z_k} f(z)$$

Proof idea. By Cauchy-Goursat and Laurent series expansions. □

Applications. Calculating integrals.

Theorem 6.6.2 (Characterization of poles)**.** *The function f has a pole of order $m > 0$ at z_0 iff*

$$f(z) = \frac{\phi(z)}{(z-z_0)^m}$$

where ϕ is analytic and nonzero at z_0. In this case,

$$\operatorname{Res}_{z=z_0} f(z) = \frac{\phi^{(m-1)}(z_0)}{(m-1)!}$$

Proof idea. Since ϕ is analytic at z_0 iff ϕ has a Taylor series expansion about z_0. □

Applications. Calculating residues at poles.

Theorem 6.6.3 (Characterization of zeros)**.** *If f is analytic at z_0, f has a zero of order $m > 0$ at z_0 iff*

$$f(z) = (z-z_0)^m g(z)$$

where g is analytic and nonzero at z_0.

Proof idea. By the Taylor series expansion. □

Applications. Relating local to global behavior, calculating residues at poles.

Corollary 6.6.1 (Isolation of zeros)**.** *If f is analytic and zero at z_0, f is zero throughout some neighborhood of z_0 or f is nonzero throughout some deleted neighborhood of z_0.*

Proof idea. By the Taylor series expansion.

In detail, if it is not true that f is zero throughout some neighborhood of z_0, then the coefficients in its Taylor series expansion about z_0 cannot all vanish. So f has a zero of some order $m > 0$ at z_0. But then by the theorem, and continuity, f is nonzero throughout some deleted neighrobhood of z_0. □

Corollary 6.6.2 (Uniqueness of analytic functions)**.** *Let f and g be analytic on some domain D. If f and g are equal on some line segment or subdomain in D, then f and g are equal on D.*

Proof idea. Assume without loss of generality that $g = 0$ and use isolation of zeros. If f is analytic at z_0 and zero on some line segment or subdomain containing z_0, then f is zero throughout some neighborhood of z_0. Now use connectedness of D. □

Theorem 6.6.4 (Zeros and poles)**.** *If f and g are analytic at z_0, f is nonzero at z_0, and g has a zero of order $m > 0$ at z_0, then f/g has a pole of order m at z_0.*

Proof idea. By the characterization of zeros and characterization of poles. □

Applications. Calculating residues at poles.

Corollary 6.6.3 (Simple poles)**.** *If f and g are analytic at z_0, f is nonzero at z_0, and g has a zero of order* 1 *at z_0, then f/g has a simple pole at z_0 and*

$$\operatorname{Res}_{z=z_0} \frac{f(z)}{g(z)} = \frac{f(z_0)}{g'(z_0)}$$

Techniques

- Evaluating contour integrals using Cauchy's residue theorem.
- Calculating residues:
  - Cauchy's integral formula.
  - Laurent series coefficients.
  - Poles.
- Relating local properties of functions and their derivatives to global properties using Taylor [Laurent] series expansions.
- Translating problems about uniqueness of power series into problems about zeros, and using isolation of zeros.
- Transferring local properties of functions to global properties using a sequence of overlapping neighborhoods in a connected domain.

6.7 Applications of Residues

Definition

Definition 6.7.1. A function is *meromorphic* in a domain if it is analytic throughout the domain except possibly at poles.

Definition 6.7.2. If f is continuous and nonzero on a positively oriented simple closed contour C, and K is the closed image contour, and $\Delta_C \arg f$ denotes the change in argument of an image point $f(z)$ after z makes one oriented traversal of C, then the *winding number (with respect to the origin)* of K is

$$\frac{1}{2\pi}\Delta_C \arg f$$

Remark. Intuitively, the winding number of a curve is the number of times it winds around the origin.

Theorems

Theorem 6.7.1 (Argument principle)**.** *Let C be a positively oriented simple closed contour. If f is analytic and nonzero on C and meromorphic inside C, then*

$$\frac{1}{2\pi}\Delta_C \arg f = Z - P$$

where Z is the number of zeros of f inside C and P is the number of poles of f inside C, both counting multiplicity (order).

Proof idea. By evaluating $\int_C f'/f$ in two ways, once directly using polar form and again using Cauchy's residue theorem. □

Applications. Calculating winding numbers, counting zeros and poles.

Corollary 6.7.1 (Rouché)**.** *If f and g are analytic on and inside a simple closed contour C and $|f| > |g|$ on C, then f and $f+g$ have the same number of zeros inside C.*

Proof idea. By the argument principle. Note f and $f+g$ are nonzero on C, so

$$Z_f = \frac{1}{2\pi}\Delta_C \arg f \qquad Z_{f+g} = \frac{1}{2\pi}\Delta_C \arg(f+g)$$

where Z_f and Z_{f+g} denote the number of zeros inside C of f and $f+g$, respectively. But

$$\begin{aligned}\Delta_C \arg(f+g) &= \Delta_C \arg f[1+g/f]\\ &= \Delta_C \arg f + \Delta_C \arg(1+g/f)\end{aligned}$$

But $|g|/|f| < 1$ on C, so $\Delta_C \arg(1+g/f) = 0$. □

Applications. Counting zeros.

Techniques

- Evaluating certain *real* integrals using appropriately constructed contours and Cauchy's residue theorem, including using indented paths for isolated real singularities, etc.
- Calculating winding numbers by counting zeros and poles, and conversely (argument principle).
- Counting zeros and poles:
  - Winding numbers (argument principle).
  - Rouché's theorem.

Part III

Foundations

Chapter 7

Logic

This chapter covers mathematical logic from [5].

7.1 Syntax

Theorems

Theorem 7.1.1. *If the symbol set S is at most countable, then the set T^S of terms over S and the set L^S of formulas over S are countable.*

Theorem 7.1.2 (Induction on terms and formulas)**.** *Let S be a symbol set.*

(a) Suppose $T \subseteq T^S$ and

(1) Every variable symbol is in T

(2) Every constant symbol in S is in T

(3) If $t_1, \ldots, t_n \in T$ and $f \in S$ is an n-ary function symbol, then $f t_1 \cdots t_n \in T$

Then $T = T^S$.

(b) Suppose $L \subseteq L^S$ and

(1) If $t_1, t_2 \in T^S$, then $t_1 \equiv t_2 \in L$

(2) If $t_1, \ldots, t_n \in T^S$ and $R \in S$ is an n-ary relation symbol, then $R t_1 \cdots t_n \in L$

(3) If $\varphi \in L$, then $\neg\varphi \in L$

(4) If $\varphi, \psi \in L$, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$, and $(\varphi \leftrightarrow \psi)$ are in L

(5) If $\varphi \in L$ and x is a variable symbol, then $\forall x\varphi$ and $\exists x\varphi$ are in L

Then $L = L^S$.

Proof idea. By induction on the length of derivations, T^S and L^S are the smallest sets satisfying their respective closure conditions. □

Applications. Inductive proofs on terms and formulas.

Theorem 7.1.3 (Unique readability for terms and formulas)**.** *Let S be a symbol set.*

(a) If $t \in T^S$, then exactly one of the following holds:

(1) $t = x$ for a unique variable symbol x

(2) $t = c$ for a unique constant symbol $c \in S$

(3) $t = f t_1 \cdots t_n$ for unique $f \in S$ and unique $t_1, \ldots, t_n \in T^S$

(b) If $\varphi \in L^S$, then exactly one of the following holds:

(1) $\varphi = t_1 \equiv t_2$ for unique $t_1, t_2 \in T^S$

(2) $\varphi = R t_1 \cdots t_n$ for unique $R \in S$ and unique $t_1, \ldots, t_n \in T^S$

(3) $\varphi = \neg\psi$ for unique $\psi \in L^S$

(4) φ is exactly one of $(\psi \wedge \chi)$, $(\psi \vee \chi)$, $(\psi \rightarrow \chi)$, or $(\psi \leftrightarrow \chi)$ for unique $\psi, \chi \in L^S$

(5) φ is exactly one of $\forall x\psi$ or $\exists x\psi$ for unique x and unique $\psi \in L^S$

Proof idea. By induction on terms and formulas. □

Applications. Recursive definitions on terms and formulas.

7.2 Semantics

Theorems

Theorem 7.2.1 (Coincidence)**.** *Let $\mathscr{I}_1 = (\mathscr{A}_1, \beta_1)$ be an S_1-interpretation and $\mathscr{I}_2 = (\mathscr{A}_2, \beta_2)$ be an S_2-interpretation on the same domain. Set $S = S_1 \cap S_2$.*

(a) If $t \in T^S$ and $\mathscr{I}_1$ and $\mathscr{I}_2$ agree on the symbols and variables occuring in t, then $\mathscr{I}_1(t) = \mathscr{I}_2(t)$.

(b) If $\varphi \in L^S$ and $\mathscr{I}_1$ and $\mathscr{I}_2$ agree on the symbols and free variables occuring in φ, then $\mathscr{I}_1 \models \varphi$ iff $\mathscr{I}_2 \models \varphi$.

Proof idea. Induction on terms and formulas. □

Applications. Justifying the intuition that only the interpretation of the symbols and *free* variables occurring in a formula are relevant in determining its truth, and so in particular that sentences express purely structural properties. Allowing us to define the notions of interpretation, consequence, satisfiability (etc.) without reference to a fixed symbol set.

Theorem 7.2.2 (Isomorphism)**.** *Let $\mathscr{A}$ and $\mathscr{B}$ be S-structures with $\pi : \mathscr{A} \cong \mathscr{B}$. Then for all $\varphi \in L^S_n$ and $a_1, \ldots, a_n \in A$,*

$$\mathscr{A} \models \varphi[a_1, \ldots, a_n] \iff \mathscr{B} \models \varphi[\pi(a_1), \ldots, \pi(a_n)]$$

Proof idea. Induction on terms and formulas. □

Corollary 7.2.1. *If $\mathscr{A} \cong \mathscr{B}$, then for all sentences φ, $\mathscr{A} \models \varphi$ iff $\mathscr{B} \models \varphi$.*

Applications. Justifying the intuition that the same structural statements are true in isomorphic structures.

Theorem 7.2.3 (Substructure)**.** *Let $\mathscr{A}$ and $\mathscr{B}$ be S-structures with $\mathscr{A} \subseteq \mathscr{B}$.*

(a) If $\varphi \in L_n^S$ is universal, then for all $a_1, \dots, a_n \in A$,

$$\mathscr{B} \models \varphi[a_1, \dots, a_n] \implies \mathscr{A} \models \varphi[a_1, \dots, a_n]$$

(b) If $\varphi \in L_n^S$ is existential, then for all $a_1, \dots, a_n \in A$,

$$\mathscr{A} \models \varphi[a_1, \dots, a_n] \implies \mathscr{B} \models \varphi[a_1, \dots, a_n]$$

Proof idea. Induction on formulas. □

Corollary 7.2.2. *Let $\mathscr{A} \subseteq \mathscr{B}$. Then for all universal sentences φ, if $\mathscr{B} \models \varphi$ then $\mathscr{A} \models \varphi$, and for all existential sentences φ, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

Applications. Motivating construction of universal axioms.

Theorem 7.2.4 (Substitution)**.** *Let $\mathscr{I}$ be an S-interpretation, and let $x_0, \dots, x_r$ be pairwise distinct variables and $t_0, \dots, t_r \in T^S$.*

(a) If $t \in T^S$, then

$$\mathscr{I}\left(t \frac{t_0 \cdots t_r}{x_0 \cdots x_r}\right) = \mathscr{I} \frac{\mathscr{I}(t_0) \cdots \mathscr{I}(t_r)}{x_0 \cdots x_r}(t)$$

(b) If $\varphi \in L^S$, then

$$\mathscr{I} \models \varphi \frac{t_0 \cdots t_r}{x_0 \cdots x_r} \iff \mathscr{I} \frac{\mathscr{I}(t_0) \cdots \mathscr{I}(t_r)}{x_0 \cdots x_r} \models \varphi$$

Proof idea. Induction on terms and formulas. □

Applications. Establishing the equivalence of syntactic and semantic substitution. Loosely speaking, it shows that you can equivalently substitute new meanings for existing terms, or substitute new terms having those meanings.

7.3 Proof

Theorems

Fix a symbol set S.

Theorem 7.3.1 (Correctness)**.** *If $\Phi \vdash \varphi$, then $\Phi \models \varphi$.*

Proof idea. Induction on derivable sequents (in the sequent calculus for S). □

Applications. Showing that formal proofs do not yield incorrect results.

Corollary 7.3.1 (Correctness)**.** *If Φ is satisfiable, then Φ is consistent.*

Proof idea. If Φ is inconsistent, then $\Phi \vdash \varphi$ and $\Phi \vdash \neg\varphi$ for some φ. But then $\Phi \models \varphi$ and $\Phi \models \neg\varphi$ by the theorem, so Φ is not satisfiable. □

Remark. Note the corollary is actually equivalent to the theorem. Indeed, if a set of formulas is consistent whenever it is satisfiable, then if $\Phi \not\vdash \varphi$, the set $\Phi \cup \{\neg\varphi\}$ is not consistent and hence not satisfiable, so $\Phi \models \varphi$.

Theorem 7.3.2 (Finiteness)**.**

(a) $\Phi \vdash \varphi$ *iff* $\Phi_0 \vdash \varphi$ *for some finite* $\Phi_0 \subseteq \Phi$.

(b) Φ *is consistent iff* Φ_0 *is consistent for all finite* $\Phi_0 \subseteq \Phi$.

(c) Φ *is inconsistent iff* Φ_0 *is inconsistent for some finite* $\Phi_0 \subseteq \Phi$.

Proof idea. By definition of proof. □

Applications. With correctness and completeness, yields compactness.

7.4 Completeness

Theorems

Theorem 7.4.1 (Henkin)**.** *Let* Φ *be a consistent set of formulas that is negation complete and contains witnesses. Then* Φ *is satisfiable.*

Proof idea. Construct the term interpretation $\mathscr{I}^{\Phi}$ based on provable consequences of Φ, then show by induction on formulas that $\mathscr{I}^{\Phi} \models \varphi$ iff $\Phi \vdash \varphi$. □

Applications. Completeness.

Theorem 7.4.2 (Completeness)**.** *If* Φ *is consistent, then* Φ *is satisfiable.*

Proof idea. Construct a consistent extension of Φ which is negation complete and contains witnesses, then appeal to Henkin's theorem.

Countable case. Suppose S is at most countable. If free(Φ) is finite, adjoin witness formulas to Φ for the countably many formulas in L^S beginning with an existential quantifier, using a 'new' variable symbol for the witness term at each step. Argue by induction that the extension is consistent. Then recursively extend the extension to obtain a maximally consistent set of formulas, which is negation complete.

If free(Φ) is infinite, replace free variables in Φ with constants to obtain a set Φ^* of sentences corresponding to Φ. By a simple substitution argument, for all $\Psi \subseteq \Phi$, Ψ is satisfiable iff its corresponding subset $\Psi^* \subseteq \Phi^*$ is satisfiable. Now by looking at corresponding finite subsets and using the above, Φ^* is consistent because Φ is consistent. Therefore Φ^* is satisfiable by the above, and hence Φ is satisfiable.

General case. Suppose S is arbitrary. Adjoin witness formulas to Φ recursively, at each step introducing new witness constants and constructing witness formulas for the previous step. Take the union. Again, argue by induction that the extension is consistent (this makes use of the countable case above), and then recursively extend to obtain a maximally consistent set of formulas.

□

Applications. Adequacy of proof calculus. Compactness. Lowenheim-Skolem.

Corollary 7.4.1 (Completeness)**.** *If* $\Phi \models \varphi$*, then* $\Phi \vdash \varphi$*.*

Proof idea. If $\Phi \models \varphi$, then $\Phi \cup \{\neg\varphi\}$ is not satisfiable and hence not consistent by the theorem, so $\Phi \vdash \varphi$. □

Remark. Note the corollary is actually equivalent to the theorem. Indeed, suppose $\Phi \models \varphi$ implies $\Phi \vdash \varphi$. If Φ is not satisfiable, $\Phi \models \psi$ is vacuously true for all ψ, so for (say) $\varphi = v_0 \equiv v_0$, $\Phi \models \varphi$ and $\Phi \models \neg\varphi$. But then $\Phi \vdash \varphi$ and $\Phi \vdash \neg\varphi$, so Φ is inconsistent.

Theorem 7.4.3 (Adequacy)**.**

(a) $\Phi \models \varphi$ *iff* $\Phi \vdash \varphi$*.*

(b) Φ *is consistent iff* Φ *is satisfiable.*

Proof idea. Correctness and completeness. □

Remark. By previous remarks, (a) and (b) are actually equivalent.

7.5 Compactness and Löwenheim-Skolem

Theorems

Theorem 7.5.1 (Compactness)**.** Φ *is satisfiable iff every finite* $\Phi_0 \subseteq \Phi$ *is satisfiable.*

Proof idea. Correctness, completeness, and finiteness (of proofs). □

Applications. Demonstrating expressive limitations of first-order logic. Proving that certain classes of structures are not axiomatizable in first-order logic (i.e., are not elementary or Δ-elementary). Proving that certain axioms cannot be formulated in first-order logic. Constructing non-standard models.

Remark. Compactness shows that first-order formulas cannot in general express that a property holds for *some unspecified finite number* of objects in a model (where *finite* is meant in the sense external to the model). For example, it is not possible with first-order formulas to express the properties *being a finite structure*, *being a torsion group*, *being a connected graph*, etc., all of which involve finiteness in this way. It is also generally not possible to express properties regarding all subsets (etc.) of a structure. These limitations open the door to non-standard models.

One way to 'circumvent' these limitations is to consider first-order models of set theory (ZFC), thereby bringing notions of *finite*, *subset*, etc. into the model itself. Note however this changes the meanings of the terms; the meanings *internal* to the model do not correspond to the meanings *external* to the model. Another approach is to consider more expressive logical systems.

Corollary 7.5.1. *If* Φ *has arbitrarily large finite models, then* Φ *has an infinite model.*

Proof idea. Set $\Psi = \Phi \cup \{\varphi_{\geq n} \mid n \geq 2\}$, where $\varphi_{\geq n}$ is a sentence stating that there exist at least n elements. Then Ψ is finitely satisfiable by hypothesis, so Ψ is satisfiable by compactness. Any model of Ψ is an infinite model of Φ. □

Theorem 7.5.2 ("Downward" Löwenheim-Skolem)**.** *If $\Phi \subseteq L^S$ is satisfiable, then Φ has a model of cardinality $\leq |L^S|$.*

Proof idea. By the proof of completeness (and the term interpretation in Henkin's theorem). □

Applications. Generating smaller models which are easier to manipulate.

Corollary 7.5.2. *If Φ is satisfiable and at most countable, then Φ has a model that is at most countable.*

Theorem 7.5.3 ("Upward" Löwenheim-Skolem)**.** *If Φ has an infinite model, then Φ has arbitrarily large infinite models.*

Proof idea. Introduce new constants and formulas stating the constants are distinct. Show finite satisfiability, and hence satisfiability by compactness. □

Theorem 7.5.4 (Löwenheim, Skolem, Tarski)**.** *If Φ has an infinite model, then Φ has a model of cardinality κ for any infinite cardinal $\kappa \geq |\Phi|$.*

Proof idea. Assume $\Phi \subseteq L^S$ with $|S| \leq \kappa$. As in the previous proof, introduce new constants and formulas stating they are distinct, with the resulting language having cardinality $\leq \kappa$. By downward Löwenheim-Skolem, the set of formulas has a model of cardinality $\leq \kappa$. But this model must have cardinality $\geq \kappa$ by construction, so it has cardinality κ. □

Applications. Generating large models. Proving first-order non-axiomatizability (e.g. of isomorphism for infinite structures).

Theorem 7.5.5.

(a) The isomorphism class of an infinite structure is not first-order axiomatizable.

(b) The elementary equivalence class of a structure is first-order axiomatizable, and is the smallest first-order axiomatizable class containing the structure.

Proof idea. For (a), use upward Löwenheim-Skolem; for (b), let the axioms be the theory of the structure. □

Corollary 7.5.3. *Any infinite structure has a non-isomorphic but elementarily equivalent structure. (In particular, isomorphism is stronger than elementary equivalence.)*

Remark. Compactness and its consequences are far less useful within *finite* model theory, since they usually involve infinite models. When working with finite models, other techniques such as the Ehrenfeucht-Fraïssé methods are required.

7.6 Syntactic Interpretations and Normal Forms

Theorems

Recall a formula φ is *term reduced* if its atomic subformulas are of the form $x \equiv y$, $x \equiv c$, $x \equiv f x_1 \cdots x_n$, and $R x_1 \cdots x_n$.

Theorem 7.6.1 (Term reduction). *For every formula φ, there is a logically equivalent term reduced formula φ^* with* $\operatorname{free}(\varphi) = \operatorname{free}(\varphi^*)$.

Proof idea. Define φ^* by recursion, then prove properties by induction. □

Applications. Simplifying inductive proofs and recursive definitions.

Theorem 7.6.2 (Relational symbol sets). *Let S be a symbol set, and S^r the corresponding relational symbol set. For an S-structure $\mathscr{A}$, let $\mathscr{A}^r$ denote the corresponding relational S^r-structure.*

(a) For all $\varphi \in L^S$, there exists $\varphi^r \in L^{S^r}$ such that for all S-interpretations $(\mathscr{A}, \beta)$,

$$(\mathscr{A}, \beta) \models \varphi \iff (\mathscr{A}^r, \beta) \models \varphi^r$$

(b) For all $\psi \in L^{S^r}$, there exists $\psi^{-r} \in L^S$ such that for all S-interpretations $(\mathscr{A}, \beta)$,

$$(\mathscr{A}, \beta) \models \psi^{-r} \iff (\mathscr{A}^r, \beta) \models \psi$$

Proof idea. Define φ^r and ψ^{-r} by recursion on term-reduced formulas, and prove equivalence by induction. □

Applications. Simplifying arguments.

Theorem 7.6.3 (Syntactic interpretations). *Let S and S' be symbol sets and I be a syntactic interpretation of S' in S.*

For every $\psi \in L^{S'}$, there exists $\psi^I \in L^S$ with $\operatorname{free}(\psi^I) \subseteq \operatorname{free}(\psi)$ *and such that for all interpreting S-structures $\mathscr{A}$ (i.e., $\mathscr{A} \models \Phi_I$) and assignments β over the interpreted structure $\mathscr{A}^{-I}$,*

$$(\mathscr{A}^{-I}, \beta) \models \psi \iff (\mathscr{A}, \beta) \models \psi^I$$

Proof idea. Define ψ^I by recursion on term-reduced formulas using the syntactic interpretation, then prove equivalence by induction. □

Applications. Talking about one structure within another structure. In particular relativization and extension by syntactic definition.

Corollary 7.6.1 (Relativization). *Let S' be a symbol set. Fix a unary relation symbol $P \notin S'$ and set $S = S' \cup \{P\}$. For all $\psi \in L_0^{S'}$, there exists $\psi^P \in L_0^S$ such that for all S-structures $(\mathscr{A}, P^A)$ where P^A is S'-closed,*

$$[P^A]^{\mathscr{A}} \models \psi \iff (\mathscr{A}, P^A) \models \psi^P$$

Proof idea. Syntactically interpret S' in S according to the identity, using P to delimit the interpreted domain. □

Applications. Talking about a substructure within a larger structure. For example, talking about the field of scalars within a (one-sorted) vector space.

Corollary 7.6.2 (Definitions)**.** *Let S be a symbol set and $\Phi \subseteq L_0^S$. Let $s \notin S$ and δ_s be an S-definition of s in Φ. Let I be the corresponding syntactic interpretation.*

(a) For all $\varphi \in L_0^S$, $\Phi \models \varphi$ iff $\Phi \cup \{\delta_s\} \models \varphi$.

(b) For all $\psi \in L_0^{S\cup\{s\}}$, $\Phi \cup \{\delta_s\} \models \psi \leftrightarrow \psi^I$.

(c) For all $\psi \in L_0^{S\cup\{s\}}$, $\Phi \cup \{\delta_s\} \models \psi$ iff $\Phi \models \psi^I$.

Applications. Allowing the introduction of defined symbols which simplify notation but preserve theory.

Recall a formula φ is in *disjunctive normal form* if it is a disjunction of conjunctions of atomic or negated atomic formulas.

Theorem 7.6.4 (Disjunctive normal form)**.** *For every quantifier free formula φ, there is a logically equivalent formula ψ in disjunctive normal form.*

Proof idea. Note φ is in the boolean closure of its finitely many atomic subformulas. Consider now all finitely many possible configurations of these subformulas over structures and tuples satisfying φ. Each such configuration can be expressed as a conjunction, and the disjunction ψ over all of them is equivalent to φ. □

Recall a formula φ is in *prenex normal form* if $\varphi = Q_1x_1\cdots Q_nx_n\psi$ where $Q_i \in \{\forall,\exists\}$ for $1 \leq i \leq n$ and ψ is quantifier free.

Theorem 7.6.5 (Prenex normal form)**.** *For every formula φ, there is a logically equivalent formula ψ in prenex normal form with* $\mathrm{free}(\varphi) = \mathrm{free}(\psi)$.

Proof idea. Induction on the number of quantifiers in φ, using basic properties of logical equivalence to move quantifiers to the left. □

Theorem 7.6.6 (Skolem normal form)**.** *For every formula φ, there is a universal formula ψ in prenex normal form such that* $\mathrm{free}(\varphi) = \mathrm{free}(\psi)$, *$\psi \models \varphi$, and φ and ψ are equivalent for satisfaction.*

Proof idea. Induction on the number of existential quantifiers in φ, introducing new function and constant symbols to construct 'witness terms' and eliminate existential quantifiers. □

7.7 Extensions of First-Order Logic

Theorems

Recall $\mathscr{L}_\mathrm{I}$ denotes first-order logic.

Theorem 7.7.1 ($\mathscr{L}_{\mathrm{II}}$)**.** *For second-order logic $\mathscr{L}_{\mathrm{II}}$ (which allows quantification over n-ary relations on the domain of a model),*

(a) $\mathscr{L}_{\mathrm{II}}$ is more expressively powerful than $\mathscr{L}_{\mathrm{I}}$.

(b) Compactness fails in $\mathscr{L}_{\mathrm{II}}$.

(c) Löwenheim-Skolem fails in $\mathscr{L}_{\mathrm{II}}$.

(d) Completeness fails in $\mathscr{L}_{\mathrm{II}}$.

Proof idea. For (a), note for example $\mathscr{L}_{\mathrm{II}}$ can axiomatize the Peano structure on $\mathbb{N}$ up to isomorphism (by expressing the induction axiom), while $\mathscr{L}_{\mathrm{I}}$ cannot.

For (b), construct a sentence φ_{fin} which states 'every injective function (on the domain of the model) is surjective'. Then a model satisfies φ_{fin} iff the model is finite. Therefore the set

$$\Phi = \{\varphi_{\mathrm{fin}}\} \cup \{\varphi_{\geq n} \mid n \geq 2\}$$

is finitely satisfiable but not satisfiable.

For (c), construct a sentence φ_{unc} which states 'there exists a linear ordering (on the domain of the model) under which not every element has only finitely many predecessors'. Then a model satisfies φ_{unc} iff the model is uncountable.

Now (d) follows from (b) and (c) since if completeness holds (in the usual sense), then compactness and Löwenheim-Skolem follow. □

Theorem 7.7.2 ($\mathscr{L}_{\omega_1\omega}$)**.** *For the infinitary logic $\mathscr{L}_{\omega_1\omega}$ (which allows infinite disjunctions and conjunctions),*

(a) $\mathscr{L}_{\omega_1\omega}$ is more expressively powerful than $\mathscr{L}_{\mathrm{I}}$.

(b) Compactness fails in $\mathscr{L}_{\omega_1\omega}$.

(c) Löwenheim-Skolem holds in $\mathscr{L}_{\omega_1\omega}$.

(d) Completeness holds in $\mathscr{L}_{\omega_1\omega}$ for infinitary proofs.

Proof idea. For (a), note for example $\mathscr{L}_{\mathrm{II}}$ can axiomatize the Peano structure on $\mathbb{N}$ up to isomorphism (by expressing the fact that every element is a finite successor of 0), while $\mathscr{L}_{\mathrm{I}}$ cannot.

For (b), construct a sentence $\varphi_{\mathrm{fin}} = \bigvee_{n\geq 1}\varphi_n$ where φ_n states 'there exist exactly n elements'. Then a model satisfies φ_{fin} iff the model is finite.

For (c), if φ has a model, construct an at most countable submodel one step at a time in countably many steps. More specifically, given a model of φ, start with a nonempty subset of the domain which contains all constants. Then, one step at a time, close the subset under function applications and existential quantifications over elements from the previous step. Take the union. Argue that it is an at most countable submodel of φ. □

Theorem 7.7.3 ($\mathscr{L}_Q$)**.** *For the logic $\mathscr{L}_Q$ (which allows quantification over uncountably many elements),*

(a) $\mathscr{L}_Q$ is more expressively powerful than $\mathscr{L}_{\mathrm{I}}$.

(b) Compactness holds in $\mathscr{L}_Q$ for at most countable sets of formulas.

(c) *Löwenheim-Skolem fails in $\mathcal{L}_Q$.*

(d) *Completeness holds in $\mathcal{L}_Q$ for at most countable sets of formulas.*

Proof idea. For (a) and (c), note $\varphi_{\text{unc}} = Qx\, x \equiv x$ characterizes uncountable models. Note (b) follows from (d). □

Remark. The results in this chapter suggest that compactness or Löwenheim-Skolem (or both) will fail to hold in any logical system more expressively powerful than $\mathcal{L}_I$. This is confirmed by Lindström's theorems.

7.8 Limitations of the Formal Method

Theorems

Lemma 7.8.1. *Let T be a first-order theory.*

(a) *If T is recursively axiomatizable, T is recursively enumerable.*

(b) *If T is recursively enumerable and complete, then T is decidable.*

Proof idea. For (a), write $T = \Phi^{\vdash}$ where Φ is a recursive set of axioms. Systematically generate all possible sequent derivations, and use the decision procedure for Φ to decide for each one whether it is a derivation from Φ. List only the consequences derivable from Φ.

For (b), to decide whether $\varphi \in T$, enumerate T until either φ or $\neg\varphi$ is found. □

Theorem 7.8.1 (Halting problem)**.** *The set of programs which halt on an empty input is undecidable.*

Proof idea. Diagonalization.

First, choose an effective coding of programs as strings over a finite alphabet. Then argue that the set of programs which halt on their own coding is undecidable, lest we could diagonalize and construct a program different from every program. Now reduce this case to the desired case. □

Applications. Reducing to other problems to establish undecidability.

Theorem 7.8.2 (Undecidability of first-order logic (Church))**.** *The first-order validities are undecidable.*

Proof idea. Reduce the halting problem to this problem.

For a given program, choose an appropriate (effectively given) symbol set and effectively construct a sentence which describes the operation and halting of the program (on the empty input). Then the program halts iff the sentence is valid, so the validities must be undecidable.

In more detail, given a program P with instructions $\alpha_0, \ldots, \alpha_k$ and using at most r registers, choose a symbol set which is suitable for representing ordered numerals and contains an $(r+2)$-ary relation symbol R. Use R to describe the states reached by P under its computation on the empty input (where a *state* consists of a step, instruction pointer, and register dump, all represented as numerals). Construct a sentence

$$\varphi_P = (\psi_E \wedge \psi_I \wedge \psi_{\alpha_0} \wedge \cdots \wedge \psi_{\alpha_k}) \to \psi_H$$

where ψ_E asserts basic properties of the ordering, ψ_I describes the initial state, ψ_{α_i} describes the state transitions induced by instruction α_i, and ψ_H asserts the existence of a halting state. Then φ_P is valid iff P halts on the empty input. □

Theorem 7.8.3 (Trahtenbrot)**.** *The first-order finite validities are not recursively enumerable.*

Proof idea. Reduce the halting problem to this problem.

The set Φ_{fs} of finitely *satisfiable* sentences is recursively enumerable, since it is possible to effectively generate all first-order sentences and all finite models. But Φ_{fs} is not decidable, lest it is possible to decide the halting problem by making use of the antecedent of φ_P from the last proof. Now $\varphi \notin \Phi_{\text{fs}}$ iff $\neg\varphi \in \Phi_{\text{fv}}$, so Φ_{fv} cannot be recursively enumerable. □

Applications. Reducing to other problems, for example to show that the validities in stronger logical systems (where finiteness can be characterized somehow) are not recursively enumerable. See the incompleteness of second-order logic (below), Lindström's Second Theorem.

Theorem 7.8.4 (Incompleteness of second-order logic (Gödel))**.** *The second-order validities are not recursively enumerable.*

Proof idea. Reduce the first-order finite validities to the second-order validities.

Let φ_{fin} be a second-order sentence characterizing the finite models. Then for a first-order sentence ψ, $\varphi_{\text{fin}} \to \psi$ is a second-order validity iff ψ is a first-order finite validity. The result follows from Trahtenbrot's theorem. □

Applications. We know that there is no correct and complete sequent calculus for second-order logic in general. This result shows that there is no such calculus for even the validities.

Theorem 7.8.5 (Undecidability of arithmetic)**.** *The first-order theory of ordinary arithmetic is undecidable.*

Proof idea. Reduce the halting problem to this problem, as with the proof of the undecidability of first-order logic. The only new difficulty is talking about sequences of numbers within arithmetic, which can be overcome using an effective coding (e.g. a prime coding, p-adic coding, etc.). □

Remark. This proof demonstrates that ordinary arithmetic *allows representations* of decidable predicates. This is also true for the axiomatic system of Peano arithmetic.

Corollary 7.8.1. *Ordinary arithmetic is not recursively enumerable, and so not recursively axiomatizable. Equivalently, for any recursive axiomatization of ordinary arithmetic, there exist true statements which cannot be proved.*

Proof idea. Since the theory is complete, this follows from the lemma above. □

For the following theorems, we work in the language $L^{S_{\mathrm{ar}}}$ of first-order arithmetic. Fix an effective, bijective coding (Gödel numbering) of formulas by natural numbers, and let $\mathbf{n}^{\varphi}$ denote the code of φ.

Theorem 7.8.6 (Fixed point theorem (Carnap))**.** *Let Φ be a set of sentences which allows representations. For every formula $\psi(x)$, there exists a sentence φ such that*

$$\Phi \vdash \varphi \leftrightarrow \psi(\mathbf{n}^{\varphi})$$

Proof idea. Diagonalization.

For notational convenience, identify formulas with their code numbers. Now let $\varphi_0(x), \varphi_1(x), \dots$ be an enumeration of formulas with one free variable x. For each k, consider the enumeration

$$\mathbf{E}_k: \quad \varphi_k(\varphi_0) \quad \varphi_k(\varphi_1) \quad \cdots \quad \varphi_k(\varphi_k) \quad \cdots$$

of sentences obtained by syntactically substituting formulas (code numbers) into φ_k. Then the diagonal enumeration is given by

$$\mathbf{D}: \quad \varphi_0(\varphi_0) \quad \varphi_1(\varphi_1) \quad \cdots \quad \varphi_k(\varphi_k) \quad \cdots$$

Substitute the sentences (code numbers) on $\mathbf{D}$ into ψ to obtain the enumeration

$$\mathbf{D}^*: \quad \psi(\varphi_0(\varphi_0)) \quad \psi(\varphi_1(\varphi_1)) \quad \cdots \quad \psi(\varphi_k(\varphi_k)) \quad \cdots$$

Since sentences on $\mathbf{D}^*$ can be obtained by effective syntactic operation, and Φ allows representations, it can be argued that $\mathbf{D}^*$ is 'equivalent' to some $\mathbf{E}_k$. In other words, for some k, for all x, $\Phi \vdash \varphi_k(\varphi_x) \leftrightarrow \psi(\varphi_k(\varphi_x))$. Now $\mathbf{D}$ and $\mathbf{E}_k$ have their $(k+1)$-th entry in common, hence we obtain a fixed point at $x = k$ by letting $\varphi = \varphi_k(\varphi_k)$. Then $\Phi \vdash \varphi \leftrightarrow \psi(\varphi)$ as desired. □

Applications. Self-referential statements. Tarski's theorem. Gödel's theorems.

Theorem 7.8.7 (Tarski)**.** *Let Φ be a consistent set of sentences which allows representations. Then Φ cannot represent its own theory.*

Proof idea. Suppose χ represents the theory, so for all sentences φ, if $\Phi \models \varphi$ then $\Phi \vdash \chi(\mathbf{n}^{\varphi})$, and if not $\Phi \models \varphi$ then $\Phi \vdash \neg\chi(\mathbf{n}^{\varphi})$. Let φ be a fixed point of $\psi = \neg\chi$ (so intuitively φ states 'I am not true'). Then

$$\Phi \models \varphi \iff \Phi \vdash \varphi \iff \Phi \vdash \neg\chi(\mathbf{n}^{\varphi}) \iff \text{not } \Phi \models \varphi$$

—a contradiction. □

Corollary 7.8.2. *Ordinary arithmetic cannot represent its own theory.*

Theorem 7.8.8 (First incompleteness theorem (Gödel))**.** *Let Φ be a consistent, recursive set of sentences which allows representations. Then the theory of Φ is incomplete, that is, there exists a sentence φ such that not $\Phi \vdash \varphi$ and not $\Phi \vdash \neg\varphi$.*

Proof idea. If the theory of Φ is complete, then it is decidable (by the lemma above). But then since Φ allows representations, Φ can represent its theory, contradicting Tarski's theorem (in syntactic form). □

Theorem 7.8.9 (Second incompleteness theorem (Gödel))**.** *Let Φ be a consistent, recursive set of sentences extending Peano arithmetic. Then Φ cannot prove its own consistency.*

Proof idea. Since Φ is recursive and allows representations (since Peano arithmetic does), Φ can represent derivability from Φ. That is, there exists a formula $\delta(x)$ such that for all sentences φ, if $\Phi \vdash \varphi$, then $\Phi \vdash \delta(\mathbf{n}^{\varphi})$. (This does *not* mean that Φ can represent its theory!) Note Φ is consistent iff not $\Phi \vdash \neg 0 \equiv 0$, hence the sentence $\varphi_{\text{con}} = \neg\delta(\mathbf{n}^{\neg 0 \equiv 0})$ expresses the consistency of Φ.

Now let φ be a fixed point for $\psi = \neg\delta$ (so intuitively φ states 'I am not provable'). Since Φ is consistent, not $\Phi \vdash \varphi$. It can be shown that

$$\Phi \vdash \varphi_{\text{con}} \to \neg\delta(\mathbf{n}^{\varphi})$$

Therefore if $\Phi \vdash \varphi_{\text{con}}$, then $\Phi \vdash \neg\delta(\mathbf{n}^{\varphi})$. But then since φ is a fixed point, $\Phi \vdash \varphi$—a contradiction. □

Applications. Derailing Hilbert's program.

7.9 Elementary Equivalence

Theorems

Lemma 7.9.1 (Partial isomorphisms)**.**

(a) *If S is relational and $\mathscr{A}$ and $\mathscr{B}$ are S-structures, then for all r-tuples $\boldsymbol{a} \in A$ and $\boldsymbol{b} \in B$, the map $\boldsymbol{a} \mapsto \boldsymbol{b}$ determines a partial isomorphism iff for all atomic φ, $\mathscr{A} \models \varphi[\boldsymbol{a}]$ iff $\mathscr{B} \models \varphi[\boldsymbol{b}]$.*

(b) *If $\mathscr{A} \cong \mathscr{B}$, then $\mathscr{A} \cong_p \mathscr{B}$.*

(c) *If $\mathscr{A} \cong_p \mathscr{B}$, then $\mathscr{A} \cong_f \mathscr{B}$.*

(d) *If $\mathscr{A} \cong_f \mathscr{B}$ and A is finite, then $\mathscr{A} \cong \mathscr{B}$.*

(e) *If $\mathscr{A} \cong_p \mathscr{B}$ and A and B are at most countable, then $\mathscr{A} \cong \mathscr{B}$.*

Proof idea. (a) is immediate since atomic formulas describe all relations (including equality) satisfied by elements. (b)-(d) are trivial.

For (e), proceed one step at a time in countably many steps using the back and forth properties. In detail, write $A = \{a_0, a_1, \ldots\}$ and $B = \{b_0, b_1, \ldots\}$. Then starting from the empty map, recursively build an ascending chain of partial isomorphisms using the back and forth properties with elements $a_0, b_0, a_1, b_1, \ldots$. Take the union. Argue that it is an isomorphism. □

Remark. Note that (a) does not hold in general for non-relational symbol sets or non-atomic formulas, both of which (either explicitly with existential quantifiers or implicitly with 'witness terms') may refer to existence of other elements not involved in the partial isomorphism.

Corollary 7.9.1 (Cantor)**.** *Any two countable dense linear orderings without endpoins are isomorphic.*

Proof idea. Argue that any two arbitrary dense linear orderings without endpoints are partially isomorphic, then appeal to (e) above. □

Theorem 7.9.1 (Fraïsse, Hintikka, Ehrenfeucht)**.** *Let S be a finite symbol set and $\mathscr{A}$ and $\mathscr{B}$ S-structures. Then the following are equivalent:*

(a) $\mathscr{A} \equiv \mathscr{B}$

(b) $\mathscr{A} \cong_f \mathscr{B}$

(c) $\mathscr{A} \models \varphi^n_{\mathscr{B}}$ *for all* $n \geq 1$

(d) The responding player has a winning strategy in the Ehrenfeucht game for $\mathscr{A}$ *and* $\mathscr{B}$

Proof idea. (b) $\Longleftrightarrow$ (d) is immediate from the definition of the Ehrenfeucht game.

(a) $\Longrightarrow$ (c) is immediate since $\mathscr{B} \models \varphi^n_{\mathscr{B}}$ for all $n \geq 1$. For (c) $\Longrightarrow$ (b), argue by induction on n that if $\mathscr{A} \models \varphi^n_{\mathscr{B},\boldsymbol{b}}[\boldsymbol{a}]$, then by construction of the formula the map $\boldsymbol{a} \mapsto \boldsymbol{b}$ is a partial isomorphism admitting of extension by the back and forth properties n times. The result follows for $\boldsymbol{a} = \boldsymbol{b} = \emptyset$.

Finally, for (b) $\Longrightarrow$ (a), argue by induction on formulas that if $\boldsymbol{a} \mapsto \boldsymbol{b}$ is a partial isomorphism admitting of extension by the back and forth properties n times, then for all formulas φ with quantifier rank $\leq n$, $\mathscr{A} \models \varphi[\boldsymbol{a}] \iff \mathscr{B} \models \varphi[\boldsymbol{b}]$. Use the back and forth property to handle the quantifier case. Again take $\boldsymbol{a} = \boldsymbol{b} = \emptyset$. □

The proof yields the corollary:

Corollary 7.9.2. *Let S be a finite symbol set and $\mathscr{A}$ and $\mathscr{B}$ S-structures. Then the following are equivalent:*

(a) $\mathscr{A} \equiv_m \mathscr{B}$

(b) $\mathscr{A} \cong_m \mathscr{B}$

(c) $\mathscr{A} \models \varphi^n_{\mathscr{B}}$ *for all* $1 \leq n \leq m$

(d) The responding player has a winning strategy in the m-Ehrenfeucht game for $\mathscr{A}$ *and* $\mathscr{B}$

Applications. Proving elementary equivalence. Establishing non-axiomatizability results (even in finite model theory). Proving the completeness of theories (since a theory T is complete iff any two models of T are elementarily equivalent).

Remark. Fix finite S and let $\mathfrak{K}$ be a class of S-structures. If for every m there exist S-structures $\mathscr{A}$ and $\mathscr{B}$ with $\mathscr{A} \cong_m \mathscr{B}$, $\mathscr{A} \in \mathfrak{K}$ and $\mathscr{B} \notin \mathfrak{K}$, then $\mathfrak{K}$ is not axiomatizable by a first-order sentence (i.e., $\mathfrak{K}$ is not elementary). Indeed, by the corollary, if φ is any sentence, then for m the quantifier rank of φ, the assumption implies $\mathfrak{K} \neq \mathrm{Mod}^S \varphi$.

This approach can be used in finite model theory, if the assumption holds among finite structures.

When working with finite structures, the intuition is that if structures are 'locally' similar and sufficiently large, they should be finitely isomorphic up to a point. For a partial isomorphism to be extensible, elements that are sufficiently 'close' to each other (under the relations of the structure) need to be mapped to elements having the same 'close'-ness, and elements that are 'far' apart need to be kept 'far' apart. This prevents any conflicts from occurring when an extension is made. Of course, this gets harder to maintain as the partial isomorphism grows. A 'strategy' based on these observations is often formalized using truncated distance functions.

7.10 Uniqueness of First-Order Logic

Theorems

Lemma 7.10.1 (Expressive power)**.** *Let $\mathscr{L}$ be a regular logical system with $\mathscr{L}_\mathrm{I} \le \mathscr{L}$. Let S be relational and suppose $\psi \in L(S)$ is not logically equivalent to a first-order sentence. Then for every finite $S_0 \subseteq S$ and $m \in \mathbb{N}$, there exist S-structures $\mathscr{A}$ and $\mathscr{B}$ with*

$$\mathscr{A}|_{S_0} \cong_m \mathscr{B}|_{S_0} \quad \textit{and} \quad \mathscr{A} \models_{\mathscr{L}} \psi \textit{ and } \mathscr{B} \models_{\mathscr{L}} \neg\psi$$

Proof idea. Fix finite $S_0 \subseteq S$ and $m \in \mathbb{N}$. Define a first-order sentence by taking a disjunction over the m-isomorphism types of (S_0-reducts of) models of ψ:

$$\varphi = \bigvee\{\, \varphi^m_{\mathscr{A}|_{S_0}} \mid \mathscr{A} \text{ an } S\text{-structure and } \mathscr{A} \models_{\mathscr{L}} \psi \,\}$$

Let $\varphi^* \in L(S)$ be the corresponding sentence in $\mathscr{L}$. Clearly ψ implies φ, and so by assumption φ cannot imply ψ. Therefore there must exist S-structures $\mathscr{A}$ and $\mathscr{B}$ satisfying φ with $\mathscr{A} \models \psi$ and $\mathscr{B} \models \neg\psi$. Now $\mathscr{B} \models \varphi^m_{\mathscr{A}|_{S_0}}$, so $\mathscr{A}|_{S_0} \cong_m \mathscr{B}|_{S_0}$. □

Remark. Intuitively, this result shows that if ψ is more powerful than a first-order sentence, we can find structures disagreeing on ψ which agree up to any desired amount on first-order theory.

Note the similarity between this result and the Fraïsse method for establishing first-order non-axiomatizability. Indeed, by assumption, the class of models of ψ is not elementary.

Lemma 7.10.2 (Compact systems)**.** *Let $\mathscr{L}$ be a compact regular logical system with $\mathscr{L}_\mathrm{I} \le \mathscr{L}$. Let S be relational and suppose $\psi \in L(S)$. Then there exists finite $S_0 \subseteq S$ such that for all S-structures $\mathscr{A}$ and $\mathscr{B}$,*

$$\mathscr{A}|_{S_0} \cong \mathscr{B}|_{S_0} \quad \Longrightarrow \quad (\mathscr{A} \models_{\mathscr{L}} \psi \iff \mathscr{B} \models_{\mathscr{L}} \psi)$$

Proof idea. Model theory within models.

In $\mathscr{L}_I$, so also in $\mathscr{L}$, we can describe isomorphism between two substructures of a model. Using relativization in $\mathscr{L}$, we can also express that a given statement is true in (relative to) a substructure. Therefore, if U and V represent substructures, the fact that isomorphism between U and V implies agreement on ψ by U and V can be captured by a logical implication in $\mathscr{L}$ of the form

$$\Phi^* \models \psi^U \leftrightarrow \psi^V$$

where $\Phi \subseteq L_I(S)$ describes isomorphism between U and V in $\mathscr{L}_I$, and $\Phi^* \subseteq L(S)$ is the corresponding description in $\mathscr{L}$.

Now by compactness of $\mathscr{L}$, there exists finite $\Phi_0 \subseteq \Phi$ with $\Phi_0^* \models \psi^U \leftrightarrow \psi^V$. For a finite $S_0 \subseteq S$, $\Phi_0 \subseteq L_I(S_0)$. Now if $\mathscr{A}$ and $\mathscr{B}$ are S-structures and $\mathscr{A}|_{S_0} \cong \mathscr{B}|_{S_0}$, we can construct a structure $\mathscr{C}$ containing $\mathscr{A}$ and $\mathscr{B}$ and describing this isomorphism as an isomorphism between substructures with $U = A$ and $V = B$. Then $\mathscr{C} \models \Phi^*$, so $\mathscr{C} \models \psi^A \leftrightarrow \psi^B$, so $\mathscr{A} \models \psi$ iff $\mathscr{B} \models \psi$ by relativization. □

Remark. Intuitively, this result shows that the meaning of any sentence in a compact regular logical system depends upon only finitely many symbols.

Lemma 7.10.3 (LöSko systems)**.** *Let $\mathscr{L}$ be a regular logical system where $\mathscr{L}_{\mathrm{I}} \leq \mathscr{L}$ and where Löwenheim-Skolem holds. Let S be relational and suppose $\psi \in L(S)$ is not logically equivalent to a first-order sentence. Then one of the following holds:*

(a) For all finite $S_0 \subseteq S$, there exist S-structures $\mathscr{A}$ and $\mathscr{B}$ with

$$\mathscr{A}|_{S_0} \cong \mathscr{B}|_{S_0} \quad \textit{and} \quad \mathscr{A} \models \psi \textit{ and } \mathscr{B} \models \neg\psi$$

(b) For suitable S^* with $S \subseteq S^*$ and a unary relation symbol $W \in S^*$, there exists $\chi^* \in L(S^*)$ having models with arbitrarily large sets W but having no model with infinite set W.

Proof idea. Model theory within models, again.

Let χ be a first-order sentence describing the situation in the first lemma above. That is, let χ describe that the reducts of two substructures are finitely (or partially) isomorphic, and that the substructures disagree on ψ. Use the symbol W in χ to index the sets of partial isomorphisms. Let χ^* correspond to χ.

If (a) does not hold, choose a witness S_0 and construct χ^* for S_0. By the first lemma above, we know there exist models of χ^* with arbitrarily large sets W. Now suppose χ^* has a model with infinite W. It can be shown that the substructure reducts described by χ^* are actually partially isomorphic in this model. Now by Löwenheim-Skolem, we may assume them to be at most countable. But then they are actually isomorphic, contradicting choice of S_0. Therefore χ^* has no model with infinite W. □

Theorem 7.10.1 (First theorem (Lindström))**.** *Let $\mathscr{L}$ be a compact regular logical system where $\mathscr{L}_{\mathrm{I}} \leq \mathscr{L}$ and Löwenheim-Skolem holds. Then $\mathscr{L}_{\mathrm{I}} \sim \mathscr{L}$.*

Proof idea. If ψ is not logically equivalent to a first-order formula, then by the third lemma either the meaning of ψ does not depend upon only finitely many symbols, so compactness does not hold (by the second lemma), or else we can find sets of sentences with only finite models, so compactness does not hold again. □

Applications. Uniqueness of first-order logic.

Theorem 7.10.2 (Second theorem (Lindström))**.** *Let $\mathscr{L}$ be an effectively regular logical system where $\mathscr{L}_{\mathrm{I}} \leq_{\mathrm{eff}} \mathscr{L}$, Löwenheim-Skolem holds, and the validities are recursively enumerable. Then $\mathscr{L}_{\mathrm{I}} \sim_{\mathrm{eff}} \mathscr{L}$.*

Proof idea. Reduce Trahtenbrot's theorem to this problem.

In more detail, first argue $\mathscr{L} \leq \mathscr{L}_{\mathrm{I}}$. Suppose towards a contradiction $\psi \in L(S)$ is not logically equivalent to a first-order sentence. By effectiveness, assume S is finite (and hence recursive) and relational. By the previous lemma, there must exist recursive $S^* \supseteq S$ with $W \in S^*$ and $\chi^* \in L(S^*)$ such that χ^* has models with arbitrarily large finite W. As these models vary, W varies over all finite sets. By Trahtenbrot's theorem, there exists recursive S^{**} disjoint from S^* such that the finite S^{**}-validities are not recursively enumerable. But for $\varphi \in L_I(S^{**})$,

$$\varphi \text{ is a finite } S^{**}\text{-validity} \quad \Longleftrightarrow \quad \models_{\mathscr{L}} \chi^* \to (\varphi^*)^W$$

and the validities on the right are recursively enumerable in $\mathscr{L}$—a contradiction. Therefore $\mathscr{L} \leq \mathscr{L}_{\mathrm{I}}$.

Now $\mathscr{L} \leq_{\mathrm{eff}} \mathscr{L}_{\mathrm{I}}$ is witnessed by the following search procedure: given $\psi \in L(S)$, enumerate validities in $\mathscr{L}$ until one of the form $\psi \leftrightarrow (\varphi^*)$ is found where $\varphi \in L_I(S)$, then return φ. □

Applications. Uniqueness of first-order logic.

Chapter 8

Computability

This chapter covers computability theory from [3].

8.1 Computable Functions

Theorem 8.1.1. *The class $\mathscr{C}$ of computable functions contains the basic functions and is closed under substitution, primitive recursion, and minimalization. (In particular, $\mathscr{R} \subseteq \mathscr{C}$ where $\mathscr{R}$ is the class of partial recursive functions.)*

Proof concept. Exhibit URM programs for each of the basic functions, and each of the operations. □

Applications. Establishing computability of a large class of common functions.

8.2 Fundamental Theorem of Computability

Theorems

Theorem 8.2.1 (Fundamental theorem of computability)**.** *The standard formalizations of the notion of effective computability each yield the same class of functions.*

Proof concept. On a case by case basis, establish that one formalization is equivalent to another, often by 'implementing' or 'simulating' each in terms of the other. □

Applications. Establishing computability of a larger class of functions more easily. Justifying Church's Thesis.

Thesis (Church)**.** *The standard formalizations of effective computability correctly capture the informal notion.*

Idea. This is not a theorem susceptible to proof. It is an informal claim supported by various pieces of evidence including (i) the character of the formalizations, (ii) the fundamental theorem, (iii) the observed extensiveness of the class of functions obtained, etc. □

Applications. Establishing computability of functions with informal arguments.

Remark. We have the following relationships:

$$\begin{aligned}&\text{primitive recursive functions } (\mathscr{PR})\\ &\qquad\qquad \subset \text{general recursive functions} \;=\; \mu\text{-recursive functions } (\mathscr{R}_0)\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \subset \text{partial recursive functions } (\mathscr{R})\end{aligned}$$

Recall also the term *recursive function* generally means μ-recursive function (p. 50), and all functions of this type are total. The first inclusion above is strict since some total functions make essential use of the μ operator (like the Ackermann function).

Remark. Let $\Sigma = \{a_1, \dots, a_k\}$ be a finite alphabet. Define a k-adic encoding $\hat{}: \Sigma^* \to \mathbb{N}$ as follows: $\hat{\Lambda} = 0$, and

$$\widehat{a_{r_0} \cdots a_{r_m}} = r_0 + r_1 k + \cdots + r_m k^m$$

We claim this encoding is bijective.

Proof. Note that to establish injectivity, it is sufficient to show that

$$d_0 + d_1 k + \cdots + d_m k^m = 0 \quad \Longrightarrow \quad d_0 = \cdots = d_m = 0$$

whenever $1 - k \le d_j \le k - 1$ for all $0 \le j \le m$. We first prove a lemma:

Lemma 8.2.1. *For $k \ge 1$, $m \ge 0$, and $d_j \le k - 1$ for $0 \le j \le m$,*

$$d_0 + d_1 k + \cdots + d_m k^m < k^{m+1}$$

Proof. By induction on m. For $m = 0$, the result holds since $d_0 \le k - 1 < k$. If $m > 0$ and the result holds for $m - 1$, then we have

$$\begin{aligned} d_0 + d_1 k + \cdots + d_m k^m &= \left(d_0 + d_1 k + \cdots + d_{m-1} k^{m-1}\right) + d_m k^m \\ &< k^m + (k-1) k^m \\ &= k^{m+1} \end{aligned}$$

Hence the result holds for m. □

We now return to the above claim. If $m = 0$, the result is trivial. Suppose $m > 0$, the result holds for $m - 1$, and

$$d_0 + d_1 k + \cdots + d_m k^m = 0$$

If $d_m = 0$, we are done by induction. Suppose towards a contradiction, and without loss of generality, that $d_m < 0$. Then $-d_m \ge 1$, and we have

$$d_0 + d_1 k + \cdots + d_{m-1} k^{m-1} = (-d_m) k^m \ge k^m$$

—contradicting the above lemma. Hence $d_m = 0$, and the result holds for m.

To establish surjectivity, we first prove the following lemma:

Lemma 8.2.2. *For $k \geq 1$ and $m \geq 0$, let S_m denote the set of all n for which there exist $r_0, \ldots, r_m$ with $1 \leq r_j \leq k$ for $0 \leq j \leq m$ and $n = r_0 + r_1 k + \cdots + r_m k^m$. Then*

$$S_m = \{ n \in \mathbb{N} \mid 1 + k + \cdots + k^m \leq n \leq k + k^2 + \cdots + k^{m+1} \}$$

Proof. Let S_m^* denote the set on the right. Clearly $S_m \subseteq S_m^*$. Note that $|S_m^*| = k^{m+1}$. But by the injectivity of the encoding, $|S_m| = k^{m+1}$. It follows that $S_m = S_m^*$. □

It is then immediate that $\mathbb{N} - \{0\} = \bigcup_{m \in \mathbb{N}} S_m$ (and, in fact, this union is disjoint). Thus the encoding is surjective. □

Remark. There are numerous ways of effectively coding objects as natural numbers. Examples include *prime coding* ([3, p. 41]), *binary coding* ([3, ex. 2.4.16(5)]), *k-adic coding* ([3, p. 61]), and various ad hoc codings (like simpler prime codings for fixed-length sequences and even-odd coding [3, ex. 2.4.16(2)]).

Every coding serves a similar purpose, though there are differences between them. For example, a k-adic coding can accommodate arbitrary-length sequences, but the values in a sequence must come from a finite set (of cardinality $\leq k$). On the other hand, certain ad hoc codings accommodate only fixed-length sequences, but with values from a (countably) infinite set. The prime and binary codings cited allow coding of arbitrary-length sequences with values from a (countably) infinite set, so in this sense they are more powerful than some others. However, simpler coding schemes may also be combined to produce more powerful ones (for example, see the proof of the computability of the Ackermann function in [3, p. 46–47]).

8.3 Gödel Numbering and Parametrization

Theorems

Theorem 8.3.1 (Gödel numbering)**.** *The set $\mathscr{P}$ of programs is effectively denumerable.*

Proof idea. Show that URM programs can be effectively coded as finite sequences of numbers, which can in turn be effectively coded as numbers. □

Applications. A fundamental result supporting the rest of the theory. It allows us to treat code as data, opening the door to effective operations on code (and hence on computable functions) as seen in the s-m-n theorem, universal programs, and much other theory. It also provides us an indexing of computable functions.

Corollary 8.3.1. *For each n, $\mathscr{C}_n$ is denumerable, and $\mathscr{C}$ is denumerable.*

Applications. Diagonalizing on computable functions (etc.).

Theorem 8.3.2 (Kleene's s-m-n theorem)**.** *For all $m \geq 1$, there exists a primitive recursive function $s^m(e, \boldsymbol{x})$ such that for all n and e,*

$$\phi^{(n)}_{s^m(e,\boldsymbol{x})}(\boldsymbol{y}) \simeq \phi^{(m+n)}_e(\boldsymbol{x}, \boldsymbol{y})$$

for all m-tuples $\boldsymbol{x}$ and n-tuples $\boldsymbol{y}$.

In particular, if $f(x, y)$ is computable, there exists a primitive recursive function $s(x)$ such that $\phi_{s(x)}(y) \simeq f(x, y)$ for all y.

Proof idea. Program hacking. Have s^m decode e to obtain P_e, hack P_e by prepending code which first shifts register contents and loads $\boldsymbol{x}$ into the lowest registers, and return the code of the hacked program. □

Applications. Effectively currying computable functions and indexing the resulting functions. Used in conjunction with with universal functions to perform effective operations on computable functions, and in conjunction with the second recursion theorem to define computable functions in terms of their own source code.

8.4 Universal Programs

Theorems

Theorem 8.4.1 (Universal programs)**.** *For each $n \geq 1$, the universal function $\psi_U^{(n)}(e, \boldsymbol{x})$ for n-ary computable functions is computable.*

Proof idea. Program simulation. Construct a program for $\psi_U^{(n)}$ which decodes e to obtain P_e and recursively simulates the computation $P_e(\boldsymbol{x})$.

In more detail, first prove that the state of a URM computation at any point in time (determined by the register contents and instruction pointer at that time) can be effectively coded as a number. Then prove that the function $\sigma_n(e, \boldsymbol{x}, t)$ giving the (coded) state of the computation $P_e(\boldsymbol{x})$ after t steps is primitive recursive since (i) the initial state at $t = 0$ is known and (ii) the state at $t + 1$ can be determined from the state at t by examining P_e. Any value of $\psi_U^{(n)}$ can then be obtained from σ_n using minimalization. □

Applications. Simulating programs and computations (including multitasking or multithreading). Effectively diagonalizing on computable functions (etc.). Used in conjunction with the *s-m-n* theorem to perform effective operations on computable functions in various ways.

Corollary 8.4.1. *For each $n \geq 1$, the following predicates are primitive recursive:*

(a) $S_n(e, \boldsymbol{x}, y, t) \equiv$ '$P_e(\boldsymbol{x})$ converges to y in $\leq t$ steps'

(b) $H_n(e, \boldsymbol{x}, t) \equiv$ '$P_e(\boldsymbol{x})$ converges in $\leq t$ steps'

Corollary 8.4.2 (Kleene's normal form theorem)**.** *There is a primitive recursive function $U(x)$ and for each $n \geq 1$ a primitive recursive predicate $T_n(e, \boldsymbol{x}, z)$ such that for all e and $\boldsymbol{x}$,*

$$\phi_e^{(n)}(\boldsymbol{x}) \simeq U(\mu z\, T_n(e, \boldsymbol{x}, z))$$

Proof idea. Let z code a pair (y, t). Set $T_n(e, \boldsymbol{x}, z) \equiv S_n(e, \boldsymbol{x}, y, t)$ and $U(z) = y$. □

Applications. Obtaining any computable function with at most one μ operation.

Remark. In the proof of the normal form theorem, a search for a pair of numbers is facilitated by a single μ operation which makes use of an effective coding of pairs. In general, searches for complex objects (pairs, triples, arbitrary sequences, program instructions, programs, etc.) may be facilitated with a single μ operation provided that the objects can be effectively coded.

In many cases, this technique allows us to avoid the use of nested searches (for example, $\mu x \mu y$). This becomes particularly important in contexts where a direct nested search does not work.

Remark. Universal functions are in a certain sense inversely related to the functions of the s-m-n theorem. To illustrate this, recall that for any computable function $f(x,y)$, the s-m-n theorem establishes the existence of a recursive function $s(x)$ such that for all x and y,

$$\phi_{s(x)}(y) \simeq f(x,y)$$

Thus the s-m-n theorem allows us to obtain from a given computable function f an effectively indexed family of derived computable functions. Each of the derived functions is just a curried version of f.

Conversely, given any recursive function $s(x)$, the universal function ψ_U allows us to define a computable function $f(x,y)$ by

$$f(x,y) \simeq \psi_U(s(x),y) \simeq \phi_{s(x)}(y)$$

Thus from an indexed family of computable functions (here the $\phi_{s(x)}$) we can build a single computable function f. The indexed functions can be obtained from f through currying. In this way then, the application of a universal function can be seen as an inverse to the currying process. Note this is similar to differentiation and integration.

8.5 Decidability and Partial Decidability

Theorems

Theorem 8.5.1 (Algebra of decidability)**.** *The class of decidable predicates is closed under negation, conjunction, disjunction, and bounded quantification; it is not closed under unbounded quantification.*

Theorem 8.5.2 (Halting problem)**.** *The predicate '$y \in W_x$' is undecidable.*

Proof idea. Use diagonalization to prove '$x \in W_x$' undecidable, then reduce. □

Applications. The halting problem shows that there is no general effective method to decide whether a program halts on a given input. It (or, more commonly, the problem '$x \in W_x$') is often used to prove undecidability of other problems through reduction.

Theorem 8.5.3 (Rice)**.** *If $\emptyset \subset \mathscr{A} \subset \mathscr{C}_1$, the problem '$\phi_x \in \mathscr{A}$' is undecidable.*

Proof idea. Use the s-m-n theorem to reduce $x \in W_x$ to $\phi_x \in \mathscr{A}$. □

Applications. Establishing undecidability of many classes of computable functions.

Theorem 8.5.4 (Partial decidability)**.** *The following are equivalent:*

(a) $P(\boldsymbol{x})$ *is partially decidable*

(b) *There exists a computable function* $f(\boldsymbol{x})$ *such that* $P(\boldsymbol{x}) \iff \boldsymbol{x} \in \mathrm{domain}(f)$

(c) *There exists a primitive recursive predicate* $R(\boldsymbol{x}, y)$ *such that* $P(\boldsymbol{x}) \iff \exists y R(\boldsymbol{x}, y)$

(d) *(Matiyasevich's theorem)* $P(\boldsymbol{x})$ *is diophantine*

Proof idea. Trivially (a) $\iff$ (b). For (b) $\implies$ (c), set $R(\boldsymbol{x}, y) \equiv H_n(e, \boldsymbol{x}, y)$ where $f = \phi_e$, and for (c) $\implies$ (b) set $f(\boldsymbol{x}) = \mu y\, R(\boldsymbol{x}, y)$. □

Applications. Matiyasevich's theorem solves Hilbert's tenth problem (answering in the negative by establishing the existence of undecidable diophantine predicates).

Theorem 8.5.5 (Algebra of partial decidability)**.**

(a) *The class of partially decidable predicates is closed under conjunction, disjunction, existential quantification, and bounded universal quantification; it is not closed under negation or unbounded universal quantification.*

(b) $P(\boldsymbol{x})$ *is decidable iff both* $P(\boldsymbol{x})$ *and* $\neg P(\boldsymbol{x})$ *are partially decidable.*

Proof idea. For (a), closure under existential quantification follows from part (c) of the previous theorem and searching for pairs; the rest is straightforward (consult exercises below).

For (b), use multitasking. □

Theorem 8.5.6. *The function* $f(\boldsymbol{x})$ *is computable iff* $f(\boldsymbol{x}) \simeq y$ *is partially decidable.*

8.6 Recursive and Recursively Enumerable Sets

Theorems

Theorem 8.6.1 (Recursive enumerability)**.** *For a set* A*, the following are equivalent:*

(a) A *is recursively enumerable (r.e.)*

(b) *The predicate '*$x \in A$*' is partially decidable*

(c) *There exists a primitive recursive predicate* $R(x, y)$ *such that* $x \in A \iff \exists y\, R(x, y)$

(d) *There exists a partially decidable predicate* $M(x, \boldsymbol{y})$ *such that* $x \in A \iff \exists \boldsymbol{y}\, M(x, \boldsymbol{y})$

(e) A *is the domain of a unary computable function (*$A = W_k$ *for some* k*)*

(f) A *is the range of a computable function (*$A = E_k^{(n)}$ *for some* k*)*

(g) *If* $A \neq \emptyset$*,* A *is the range of a unary primitive recursive function*

(h) *(Matiyasevich)* A *is diophantine*

(i) *(Matiyasevich)* A *is the set of nonnegative values of a diophantine polynomial*

Applications. See the notes on Matiyasevich's theorem above.

Theorem 8.6.2. *Let A be a set.*

(a) *A is recursive iff both A and $\overline{A}$ are r.e.*

(b) *If A is infinite, A is recursive iff A can be recursively enumerated in increasing order.*

Proof idea. For (a), use multitasking. For (b), for the forward implication, recursively search for the elements in A in order; for the other implication, decide $x \in A$ by searching the ordered list and halting once you have found x or something larger. □

Corollary 8.6.1. *Every infinite r.e. set contains an infinite recursive subset.*

Proof idea. Recursively enumerate an infinite subset in increasing order. □

Recall $K = \{x \mid x \in W_x\}$ is a creative set.

Theorem 8.6.3 (Rice-Shapiro)**.** *Suppose $\mathscr{A} \subseteq \mathscr{C}_1$ and $A = \{x \mid \phi_x \in \mathscr{A}\}$ is r.e. For all $f \in \mathscr{C}_1$,*

$$f \in \mathscr{A} \iff \textit{there exists a finite function } \theta \subseteq f \textit{ with } \theta \in \mathscr{A}$$

Proof idea. Show if either direction of the implication fails, then the s-m-n theorem (with the help of universal programs) can be used to reduce $\overline{K}$ to A, so A is not r.e. □

Applications. Establishing non-recursive enumerability of classes of functions. Rice's theorem. The Myhill-Sheperdson theorem.

Theorem 8.6.4. *Suppose $\mathscr{A} \subset \mathscr{C}_1$ and $f_\emptyset \in \mathscr{A}$. Then $A = \{x \mid \phi_x \in \mathscr{A}\}$ is productive.*

Proof idea. Proceed as in Rice's theorem to reduce $\overline{K}$ to A. □

Corollary 8.6.2. *Suppose $\mathscr{A} \subseteq \mathscr{C}_1$ and $A = \{x \mid \phi_x \in \mathscr{A}\}$ is r.e. with $\emptyset \subset A \subset \mathbb{N}$. Then A is creative.*

Theorem 8.6.5. *A productive set contains an infinite r.e. subset.*

Proof idea. Starting with an index for the empty set, repeatedly apply the productive function for the set to construct an ascending chain of finite subsets. The elements adjoined at each step form an infinite r.e. subset. □

Theorem 8.6.6 (Post)**.** *There exists a simple set.*

Proof idea. Construct a computable function whose range contains an element from every infinite r.e. set, picking elements carefully so that the complement of the range remains infinite. □

8.7 Incompleteness

Theorems

Recall $\mathscr{T}$ denotes the set of true formal statements in ordinary arithmetic on $\mathbb{N}$.

Lemma 8.7.1 (Representations)**.** *For any decidable predicate $M(\boldsymbol{x})$, there is a statement $\sigma_M(\boldsymbol{x})$ of formal arithmetic such that, for any $\boldsymbol{a} \in \mathbb{N}^n$,*

$$M(\boldsymbol{a}) \iff \sigma_M(\mathbf{a}) \in \mathscr{T}$$

Applications. This result establishes that formal arithmetic *allows representations* of decidable predicates. Intuitively, this means that formal arithmetic is rich enough to describe machine computations. On the other hand, formal arithmetic is finitary enough that machines can compute with its objects (statements, proofs, etc.). This relationship makes computability a useful tool when studying formal arithmetic, and opens the door to self-reference as it (explicitly or implicitly) occurs in results like Carnap's fixed-point theorem, Tarski's theorem, Gödel's theorems, etc.

Theorem 8.7.1. *$\mathscr{T}$ is productive.*

Proof idea. Talk about K in arithmetic. More specifically, using the above lemma, construct formal statements corresponding to $n \in K$ and $n \notin K$, and use the latter type to reduce $\overline{K}$ to $\mathscr{T}$. □

Lemma 8.7.2. *In any recursively axiomatized formal system of arithmetic, the set of all provable statements is recursively enumerable.*

Proof idea. Since the set of axioms is recursive and proofs and statements are finite,

$$P(\sigma, p) \equiv \text{`}p \text{ is a proof of } \sigma \text{ from the axioms'}$$

is a decidable predicate. Then σ is provable iff there exists p with $P(\sigma, p)$, which is partially decidable, so provable statements are r.e. □

Theorem 8.7.2 (Gödel incompleteness)**.** *In any recursively axiomatized formal system of arithmetic in which provable statements are true, there exists a statement σ that is true but not provable (and hence $\neg\sigma$ is not provable either).*

Proof idea. Let $\mathscr{P}$ denote the set of provable statements. By assumption $\mathscr{P} \subseteq \mathscr{T}$. By the above, $\mathscr{P}$ is r.e. and $\mathscr{T}$ is productive, hence there exists $\sigma \in \mathscr{T} - \mathscr{P}$. □

Recall Peano arithmetic is a particular recursive axiomatization of arithmetic.

Lemma 8.7.3 (Representations)**.** *For any decidable predicate $M(\boldsymbol{x})$, there is a statement $\sigma_M(\boldsymbol{x})$ of formal arithmetic such that, for all $\boldsymbol{a} \in \mathbb{N}^n$,*

(a) If $M(\boldsymbol{a})$ holds, then $\sigma_M(\mathbf{a})$ is provable in Peano arithmetic.

(b) If $M(\boldsymbol{a})$ does not hold, then $\neg\sigma_M(\mathbf{a})$ is provable in Peano arithmetic.

Applications. As above, but for Peano arithmetic.

Theorem 8.7.3 (Gödel incompleteness)**.** *There exists a statement σ of formal arithmetic such that*

(a) If Peano arithmetic is consistent, σ is not provable.

(b) If Peano arithmetic is ω-consistent, $\neg\sigma$ is not provable.

Proof idea. Again, talk about K in arithmetic and use the productivity of $\overline{K}$.

In more detail, using the above lemma construct a formal statement $\mathbf{n} \in \mathbf{K}$ that corresponds to $n \in K$. Consider the sets

$$P = \{\, n \mid \mathbf{n} \in \mathbf{K} \text{ is provable} \,\}$$
$$R = \{\, n \mid \mathbf{n} \in \mathbf{K} \text{ is refutable} \,\}$$

Now $K \subseteq P$, and consistency implies P and R are disjoint, so $R \subseteq \overline{K}$. Since R is r.e. and $\overline{K}$ is productive, there exists $n \in \overline{K} - R$. It follows that $\neg(\mathbf{n} \in \mathbf{K})$ is not provable. If ω-consistency holds, $K = P$, so $n \notin P$ and $\mathbf{n} \in \mathbf{K}$ is not provable either. □

Theorem 8.7.4 (Rosser incompleteness)**.** *There exists a statement σ of formal arithmetic such that if Peano arithmetic is consistent, neither σ nor $\neg\sigma$ is provable.*

Proof idea. Talk about $K_0 = \{x \mid \phi_x(x) = 0\}$ and $K_1 = \{x \mid \phi_x(x) = 1\}$ in arithmetic, and appeal to their recursive inseparability. □

Remark. Although not clear from these notes, the Gödel and Rosser theorems have constructivist and finitist proofs. That is, given an explicit specification of Peano arithmetic, the proofs (in full detail) show how to explicitly construct the undecided statements σ, and from a proof of σ how to explicitly construct a proof of $\neg\sigma$.

Remark. The Gödel and Rosser theorems apply to all sufficiently strong recursively axiomatized formal systems. In particular, adding any recursive number of axioms does not eliminate incompleteness.

Theorem 8.7.5. *In any recursively axiomatized, ω-consistent formal system of arithmetic in which all decidable predicates are representable, the set of provable statements is creative.*

Proof idea. Again, talk about K, then reduce K to the set of provable statements. □

Theorems

In our notes on many-one reducibility, we mostly ignore the trivial sets $\emptyset$ and $\mathbb{N}$ and their corresponding m-degrees $\mathbf{o}$ and $\mathbf{n}$.

Theorem 8.7.6 (Many-one reducibility)**.** *Let A and B be sets.*

(a) $\leq_{\mathrm{m}}$ is reflexive and transitive; $\equiv_{\mathrm{m}}$ is an equivalence relation.

(b) $A \leq_{\mathrm{m}} B$ iff $\overline{A} \leq_{\mathrm{m}} \overline{B}$.

(c) If A is recursive and $B \leq_{\mathrm{m}} A$, then B is recursive.

(d) If A is recursive, $A \leq_{\mathrm{m}} B$ (if B is nontrivial).

(e) *If A is r.e. and $B \leq_{\mathrm{m}} A$, then B is r.e.*

(f) *If A is r.e. but not recursive, then $A \not\leq_{\mathrm{m}} \overline{A}$ and $\overline{A} \not\leq_{\mathrm{m}} A$.*

(g) *A is r.e. iff $A \leq_{\mathrm{m}} K$ (K is m-complete).*

Theorem 8.7.7 (Many-one degrees)**.**

(a) *There exists one (nontrivial) recursive m-degree, denoted $\mathbf{0}_m$, which consists of all (nontrivial) recursive sets; and $\mathbf{0}_m \leq_{\mathrm{m}} \mathbf{a}$ for all (nontrivial) m-degrees $\mathbf{a}$.*

(b) *Any r.e. m-degree consists only of r.e. sets.*

(c) *If $\mathbf{a} \leq_{\mathrm{m}} \mathbf{b}$ and $\mathbf{b}$ is r.e., then $\mathbf{a}$ is r.e.*

(d) *There exists a maximum r.e. m-degree, denoted $\mathbf{0}'_m$ (the m-degree of K), and $\mathbf{0}_m <_{\mathrm{m}} \mathbf{0}'_m$.*

(e) *Any two m-degrees have a unique least upper bound.*

Remark. Note the r.e. m-degrees form an initial segment of the m-degrees.

Theorem 8.7.8 (Myhill)**.** *The m-complete sets (those in $\mathbf{0}'_m$) are precisely the creative sets.*

Proof idea. The forward implication is trivial. The reverse implication requires the second recursion theorem and is discussed in the notes on Chapter 11 below. □

Corollary 8.7.1. *If $\mathbf{a}$ is the m-degree of a simple set, then $\mathbf{0}_m <_{\mathrm{m}} \mathbf{a} <_{\mathrm{m}} \mathbf{0}'_m$.*

Theorem 8.7.9 (Turing reducibility)**.** *Let A and B be sets.*

(a) *$\leq_{\mathrm{T}}$ is reflexive and transitive; $\equiv_{\mathrm{T}}$ is an equivalence relation.*

(b) *If $A \leq_{\mathrm{m}} B$, then $A \leq_{\mathrm{T}} B$.*

(c) *$A \equiv_{\mathrm{T}} \overline{A}$*

(d) *If A is recursive, $A \leq_{\mathrm{T}} B$.*

(e) *If A is recursive and $B \leq_{\mathrm{T}} A$, then B is recursive.*

(f) *If A is r.e., then $A \leq_{\mathrm{T}} K$ (K is T-complete).*

Remark. Note it is *not* true in general that if A is r.e. and $B \leq_{\mathrm{T}} A$, then B is r.e.

Theorem 8.7.10 (Turing degrees)**.**

(a) *There exists one recursive degree, denoted $\mathbf{0}$, which consists of all recursive sets and is the unique minimum degree.*

(b) *There exists a maximum r.e. degree, denoted $\mathbf{0}'$ (the degree of K), and $\mathbf{0} < \mathbf{0}'$.*

(c) *Any two degrees have a unique least upper bound.*

(d) *(Friedberg-Muchnik) There exist incomparable r.e. degrees.*

(e) *For any r.e. degree $\mathbf{0} < \mathbf{a} < \mathbf{0}'$, there exists an incomparable r.e. degree $\mathbf{b}$.*

(f) *(Sacks density theorem) The r.e. degrees are dense.*

(g) *(Sacks splitting theorem) For any r.e. degree* $\mathbf{a} > \mathbf{0}$*, there exist r.e. degrees* $\mathbf{b}, \mathbf{c}$ *less than* $\mathbf{a}$ *with least upper bound* $\mathbf{a}$.

(h) *(Lachlan and Yates) There exist r.e. degrees* $\mathbf{a}, \mathbf{b} > \mathbf{0}$ *with greatest lower bound* $\mathbf{0}$.

(i) *(Lachlan and Yates) There exist r.e. degrees* $\mathbf{a}, \mathbf{b}$ *having no greatest lower bound (either among r.e. degrees or all degrees).*

(j) *(Schoenfield) There exists a non-r.e. degree* $\mathbf{a} < \mathbf{0}'$.

(k) *(Spector) There exists a minimal degree.*

Recall $K^A = \{x \mid x \in W_x^A\}$ is the A-relativized version of K.

Theorem 8.7.11 (Turing jump)**.** *Let A and B be sets.*

(a) *K^A is A-r.e. and if B is A-r.e. then $B \leq_{\mathrm{T}} K^A$ (K^A is T-complete among A-r.e. sets).*

(b) $A <_{\mathrm{T}} K^A$.

Proof idea. Relativized versions of the standard arguments. □

Corollary 8.7.2. *For any degree* $\mathbf{a}$, $\mathbf{a} < \mathbf{a}'$ *(where* $\mathbf{a}'$ *is the jump of* $\mathbf{a}$*, that is, the degree of K^A for any $A \in$* $\mathbf{a}$*).*

8.8 Effective Operations on Partial Functions

Theorems

Theorem 8.8.1. *Let $\Phi : \mathscr{F}_m \to \mathscr{F}_n$ be an operator. Then Φ is recursive iff*

(a) *Φ is continuous (and hence monotone), that is, for all $f \in \mathscr{F}_m$ and $\boldsymbol{x}, y$,*

$$\Phi(f)(\boldsymbol{x}) \simeq y \iff \text{there exists a finite } \theta \subseteq f \text{ with } \Phi(\theta)(\boldsymbol{x}) \simeq y$$

(b) *the function $\phi(z, \boldsymbol{x})$ given by*

$$\phi(z, \boldsymbol{x}) \simeq \begin{cases} \Phi(\theta)(\boldsymbol{x}) & \text{if } z = \widetilde{\theta} \text{ for finite } \theta \in \mathscr{F}_m \\ \text{undefined} & \text{otherwise} \end{cases}$$

is computable.

Applications. Easily prove operators to be recursive.

Theorem 8.8.2 (Myhill-Sheperdson I)**.** *Let $\Psi : \mathscr{F}_m \to \mathscr{F}_n$ be a recursive operator. Then there exists a total computable function h such that for all e,*

$$\Psi(\phi_e^{(m)}) = \phi_{h(e)}^{(n)}$$

Proof idea. Proceed by search.

Let $\psi(z, \boldsymbol{x})$ witness recursiveness of Ψ. Using the s-m-n theorem together with universal programs, construct $\phi_{h(e)}^{(n)}$ to search for finite $\theta \subseteq \phi_e^{(m)}$ with $\psi(\widetilde{\theta}, \boldsymbol{x})$ defined, and return such a value if found. □

Call a total function h a *extensional* (for m and n) if $\phi_a^{(m)} = \phi_b^{(m)}$ implies $\phi_{h(a)}^{(n)} = \phi_{h(b)}^{(n)}$.

Theorem 8.8.3 (Myhill-Sheperdson II)**.** *Let h be an extensional total computable function. Then there exists a unique continuous operator $\Psi : \mathscr{F}_m \to \mathscr{F}_n$ such that, for all e,*

$$\Psi(\phi_e^{(m)}) = \phi_{h(e)}^{(n)}$$

Moreover, Ψ is recursive.

Proof idea. Proceed by extension.

First note that h naturally induces an operation on computable functions, which includes in particular the finite functions. Any continuous operator Ψ extending this operation must be defined in the obvious way in terms of the finite functions, and hence is unique (if it exists). By the Rice-Shapiro theorem, the operation on computable functions is continuous, and it follows that Ψ is well defined. Now Ψ is trivially continuous by definition, and its operation on finite functions can easily be computed using h, so Ψ is recursive. □

Applications. The Myhill-Sheperdson theorem establishes the equivalence of two formalizations of the notion of an effective operation on computable functions. It thereby allows one to conveniently switch between the two as needed.

Theorem 8.8.4 (Kleene's first recursion theorem)**.** *Let $\Phi : \mathscr{F}_m \to \mathscr{F}_n$ be a continuous operator. Then there exists a function f_Φ which is a least fixed point for Φ, that is,*

(a) $\Phi(f_\Phi) = f_\Phi$

(b) $\Phi(g) = g$ *implies* $f_\Phi \subseteq g$

Moreover, if Φ is recursive, then f_Φ is computable.

Proof idea. Construct the fixed point by iterating Φ and taking a limit. Specifically, recursively (in the set theoretic sense) define the sequence

$$\begin{aligned} f_0 &= f_\emptyset \\ f_{n+1} &= \Phi(f_n) \\ f_\Phi &= \bigcup_{n \geq 0} f_n \end{aligned}$$

If Φ is recursive, then by the Myhill-Sheperdson theorem (part I) we can effectively index the functions in this sequence, so values of f_Φ are computable. □

Applications. Proving existence and computability of a very broad class of recursive functions. In the semantics of formal programming languages, giving meaning to recursively defined programs.

Remark. Note that the proof of the recursion theorem relies on a basic set-theoretic recursion principle to ensure the existence of the sequence (f_n), and so ultimately the existence of f_Φ.

Note also the method used to construct the fixed point works in any context where there is a monotone operator behaving appropriately at limit points, and is used in other areas of mathematics.

8.9 Recursion Theorem

Theorems

Theorem 8.9.1 (Kleene's second recursion theorem)**.** *Let f be a total unary computable function. Then for any $m \geq 1$, there exists n such that $\phi^{(m)}_{f(n)} = \phi^{(m)}_n$.*

Proof. Diagonalize effective enumerations of computable functions.

In more detail, consider effective enumerations of the form

$$\mathbf{E}^m_k: \quad \phi^{(m)}_{\phi_k(0)} \quad \phi^{(m)}_{\phi_k(1)} \quad \cdots \quad \phi^{(m)}_{\phi_k(k)} \quad \cdots$$

and consider the diagonal enumeration

$$\mathbf{D}: \quad \phi^{(m)}_{\phi_0(0)} \quad \phi^{(m)}_{\phi_1(1)} \quad \cdots \quad \phi^{(m)}_{\phi_k(k)} \quad \cdots$$

Note $\mathbf{D}$ and $\mathbf{E}_k$ have their $(k+1)$-th functions in common. Transform the diagonal with f to obtain a new effective enumeration

$$\mathbf{D}^*: \quad \phi^{(m)}_{f(\phi_0(0))} \quad \phi^{(m)}_{f(\phi_1(1))} \quad \cdots \quad \phi^{(m)}_{f(\phi_k(k))} \quad \cdots$$

By the s-m-n theorem, $\mathbf{D}^* = \mathbf{E}^m_k$ for some total computable function ϕ_k. But then $\mathbf{D}$ and $\mathbf{E}_k$ have their $(k+1)$-th functions in common, that is, $\phi^{(m)}_{f(\phi_k(k))} = \phi^{(m)}_{\phi_k(k)}$, so $n = \phi_k(k)$ is a fixed point for f as desired. □

Applications. Proving existence and computability of a very broad class of recursive functions. Used in conjunction with the s-m-n theorem to establish the existence of programs defined in terms of their own source code.

Remark. Note that the second recursion theorem is more general than the first in that it applies not only to extensional functions, but it does not establish existence or computability of a *least* fixed point.

Corollary 8.9.1. *Let $f(x, y)$ be any computable function. Then there exists e such that*

$$\phi_e(y) \simeq f(e, y)$$

Applications. Programs which output their own source code, initial segments of their own computations, etc.

Theorem 8.9.2 (Myhill)**.** *Any creative set is m-complete.*

Proof idea. Given a creative set A and an r.e. set B, reduce B to A by making use of the productive function for $\overline{A}$. This requires carefully crafting 'intermediate' r.e. sets.

Let p be the productive function for $\overline{A}$. Using the s-m-n theorem and the second recursion theorem, construct a total computable function $n(y)$ such that

$$W_{n(y)} = \begin{cases} \{p(n(y))\} & \text{if } y \in B \\ \emptyset & \text{otherwise} \end{cases}$$

Then argue $y \in B$ iff $p(n(y)) \in A$. □

Applications. Structure of the m-degrees (see notes on Chapter 9 above).

8.10 Computational Complexity

Theorems

In what follows, Φ denotes an arbitrary computational complexity measure.

Theorem 8.10.1 (Rabin)**.** *Let b be a total computable function. Then there exists a total computable function f taking only values in $\{0,1\}$ such that, if e is any index of f, $\Phi_e(n) > b(n)$ for almost all n.*

Proof idea. Diagonalize against all programs not satisfying the desired property.

In more detail, define f in stages. At stage n, let i_n be the least new index $i \le n$ such that $\Phi_i(n) \le b(n)$, or be undefined if no such index exists. Make $f(n)$ differ from $\phi_{i_n}(n)$ if i_n is defined. Then if $f = \phi_e$, it must be that $\Phi_e(n) > b(n)$ for almost all n, or else $e = i_n$ for some n. □

Applications. Establishing the existence of arbitrarily hard problems.

Theorem 8.10.2 (Blum's speed-up theorem)**.** *Let r be a total computable function. Then there exists a total computable function f such that, if j is any index for f, there exists another index k with $r(\Phi_k(n)) < \Phi_j(n)$ for almost all n.*

Proof idea. Define an infinite sequence of programs $\dots, F_2, F_1, F_0$ leading up to a program F_0 for f where for all n, $f_{F_{n+1}} \subseteq f_{F_n}$ almost everywhere. At each stage n, for input x, check that the computation $F_{i+1}(x)$ is sufficiently simpler (relative to Φ) than the computation $P_i(x)$ for all indices $n \le i < x$; wherever this fails, ensure $f_{F_n}(x) \neq \phi_i(x)$, and hence $f(x) \neq \phi_i(x)$. (Note that the second recursion theorem is required to make sense of this definition.)

It can be shown that f is a total. If P_j computes f, then F_{j+1} also computes f almost everywhere, and sufficiently simply for all $x > j$ (otherwise the definition of our sequence would have ensured $f(x) \neq \phi_j(x)$). Now the desired program for f can be obtained by patching F_{j+1} with a table lookup for finitely many values. □

Applications. Establishing that no matter which computational complexity measure we choose, there will always be functions with no 'best' implementation under this measure.

Recall $\mathscr{C}_b = \{f \mid \exists e[f = \phi_e \wedge \Phi_e(n) \le b(n)$ for almost all $n]\}$.

Theorem 8.10.3 (Borodin's gap theorem)**.** *Let r be a total computable function such that $r(x) \ge x$ for all x. Then there exists a total computable function b with $\mathscr{C}_b = \mathscr{C}_{r\circ b}$.*

Proof idea. Use cardinality when constructing b.

To calculate $b(x)$, apply r to obtain a sequence $k_0 = 0$, $k_{i+1} = r(k_i) + 1$ for $i \le x$, and consider the induced disjoint sequences

$$[k_0, r(k_0)] \quad [k_1, r(k_1)] \quad \cdots \quad [k_x, r(k_x)]$$

Now let $b(x)$ be the least k_j such that $\Phi_i(x) \not\in [k_j, r(k_j)]$ for all $i < x$ (such a k_j exists since there are at most x values $\Phi_i(x)$ for $i < x$, but $x+1$ disjoint intervals). It is then immediate that for any e, $\Phi_e(x) \not\in [b(x), r(b(x))]$ for any $x > e$. □

Applications. Tells against our intuition that with increased resources we should always be able to compute more functions.

Remark. The preceding three proofs all share a similar underlying approach. During a construction, at a parameter x, all programs P_i with $i \le x$ are considered, and any such programs *not* satisfying a desired property are 'ruled out' or 'cancelled' in some manner. Then, relative to the construction, it can be proved that any P_i satisfies the property almost everywhere (often for all $x > i$).

Theorem 8.10.4 (Elementary functions)**.**

(a) The class $\mathscr{E}$ of elementary functions contains the basic functions and is closed under substitution, limited recursion, and bounded minimalization.

(c) The class of elementary predicates is closed under negation, conjunction, disjunction, and bounded quantification.

Corollary 8.10.1. *The computation state functions σ_n are elementary.*

Define $B_0(z) = z$ and $B_k(z) = 2^{B_{k-1}(z)}$ for all $z > 0$.

Theorem 8.10.5. *If f is elementary, there exists k such that $f(\boldsymbol{x}) \le B_k(\max(\boldsymbol{x}))$ for all $\boldsymbol{x}$.*

Proof idea. Induction on $\mathscr{E}$. □

Theorem 8.10.6. *The following are equivalent:*

(a) f is elementary

(b) f is computable in elementary time

(c) There exists k such that $f(\boldsymbol{x})$ is computable in time $\le B_k(\max(\boldsymbol{x}))$ for all $\boldsymbol{x}$.

Proof idea. (a) $\Longrightarrow$ (b) by induction on $\mathscr{E}$, (b) $\Longrightarrow$ (c) by the previous theorem, and (c) $\Longrightarrow$ (a) by the previous corollary and the fact that $B_k(\max(\boldsymbol{x}))$ is elementary. □

Corollary 8.10.2. $\mathscr{E} \subset \mathscr{PR}$

Proof idea. Trivially $\mathscr{E} \subseteq \mathscr{PR}$, and $B_x(x) \in \mathscr{PR} - \mathscr{B}$. □

Chapter 9

Sets (Lectures)

This chapter contains notes from Leo Harrington's set theory course at UC Berkeley in Spring 2007, with supplementary information from [6].

We draw a distinction between *naive set theory* and *axiomatic set theory*. In naive set theory, we intuitively understand a *set* to be a collection of objects, and we do things with sets that seem obviously to be justified. We also use freely objects that are familiar in mathematics (such as numbers, relations, functions, etc.) without explicitly specifying their ontology.

In axiomatic set theory, on the other hand, we explicitly specify all of the axioms (or assumptions) of our theory. We also reduce all of our mathematical objects to sets by constructing appropriate sets to serve as these objects. Interestingly we do not define what a 'set' is in our axiomatic treatment; any objects that satisfy the axioms of our theory are acceptable. We also do not define the relation '$\in$' of membership among sets. These constitute the two primitive notions of our theory, and we use them to construct other definitions:

Definition 9.0.1. Let a, b be sets. Then we write $a \subseteq b$ and say that *a is a subset of b* iff for all sets x, if $x \in a$, then $x \in b$. In other words, $a \subseteq b$ iff

$$\forall x (x \in a \implies x \in b)$$

We assume as understood the notion of a *property*. Given a property $P(x)$, it will be convenient to be able to refer to the collection of objects in our theory satisfying P.

Definition 9.0.2. Let $P(x)$ be a property. We define the *class*

$$\mathscr{C} = \{x \mid P(x)\}$$

of all objects (sets) satisfying P. For an object a, we write $a \in \mathscr{C}$ iff $P(a)$ holds. If $\mathscr{D}$ is another class, we write $\mathscr{C} \subseteq \mathscr{D}$ iff for all objects a, if $a \in \mathscr{C}$, then $a \in \mathscr{D}$. We write $\mathscr{C} = \mathscr{D}$ iff $\mathscr{C} \subseteq \mathscr{D}$ and $\mathscr{D} \subseteq \mathscr{C}$.

Intuitively, a class $\mathscr{C}$ itself forms a set, but we do not require this in our definition. In fact, we cannot, for this would lead us quickly to contradiction.

Frege had an ‘axiom’ to this effect: for all properties $P(x)$, the class $\mathscr{C} = \{x \mid P(x)\}$ is a set. On this theory, sets are simply a ‘manner of speaking’ about properties; all statements about sets can be translated into statements about properties, and vice versa. But this assumption is problematic, as Russell demonstrated. For let

$$R = \{x \mid x \notin x\}$$

By definition R is the class of all sets a such that $a \notin a$. In other words, for all sets a, $a \in R$ iff $a \notin a$. Now if R is a set, then $R \in R$ iff $R \notin R$—a contradiction. Hence R cannot be a set. In other words, Frege was wrong.

This result is called the *Russell paradox*, and it motivates the development of a more subtle axiomatic set theory. We must specify axioms that allow us to work with certain sets, while at the same time preventing us from encountering contradictions.

It should be noted that *classes* in our treatment do remain simply a ‘manner of speaking’. Every statement about classes can be translated into a statement about properties, and vice versa. We do not encounter the Russell paradox at the level of classes, however, because the properties used in defining classes can only refer to sets. That is, when we form the class

$$\mathscr{C} = \{x \mid P(x)\}$$

the objects x satisfying $P(x)$ must be sets. Note also that the notation ‘$\in$’ for set membership is being abused slightly for use with classes. The set membership relation is undefined, while we have defined the ‘$\in$’ notation for classes above.

Most of our axioms will state that certain classes do indeed form sets. These axioms can be seen as telling us how we may legitimately construct objects in our model. But what does it mean for a class to ‘form’ a set? Intuitively, this just means that there exists some set whose elements are precisely the objects satisfying the property associated with the class. We formalize this:

Definition 9.0.3. Let a be a set. Then a determines a class

$$\mathscr{C}(a) = \{x \mid x \in a\}$$

which we call the *class determined by a*. For a class $\mathscr{D}$, we write $\mathscr{D} \subseteq a$ iff $\mathscr{D} \subseteq \mathscr{C}(a)$, we write $a \subseteq \mathscr{D}$ iff $\mathscr{C}(a) \subseteq \mathscr{D}$, and we write $\mathscr{D} \equiv a$ iff $\mathscr{D} \subseteq a$ and $a \subseteq \mathscr{D}$.

Now let $\mathscr{C}$ be a class. Then we say that $\mathscr{C}$ *forms a set* (or $\mathscr{C}$ *is a set*) iff there exists a set a such that $\mathscr{C} \equiv a$.

Note that for a class $\mathscr{C}$ and a set a, even if $\mathscr{C} \equiv a$, technically we cannot write $\mathscr{C} = a$, for classes are not objects in our theory. But since $\mathscr{C}(a)$ intuitively ‘captures’ a completely, we adopt the following axiom:

Axiom (Extension)**.** *Let a, b be sets. If $\mathscr{C}(a) = \mathscr{C}(b)$, then $a = b$. Equivalently, if*

$$\forall x(x \in a \iff x \in b)$$

then $a = b$.

We obtain the converse of the axiom of extensionality from logic. That is, if $a = b$, then $\mathscr{C}(a) = \mathscr{C}(b)$. Hence $a = b$ iff $\mathscr{C}(a) = \mathscr{C}(b)$. *Therefore we adopt the convention of identifying a set a with the class $\mathscr{C}(a)$ that it determines.* In particular we may now write $\mathscr{C}(a) = a$, understanding this as a convention.

We obtain our first theorem:

Theorem 9.0.7. *Let $\mathscr{C}$ be a class. Then $\mathscr{C}$ forms at most one set.*

Proof. Suppose $\mathscr{C} \equiv a$ and $\mathscr{C} \equiv b$. Then $\mathscr{C}(a) = \mathscr{C} = \mathscr{C}(b)$. Hence $a = b$ by the axiom of extensionality. □

Based on this theorem, we adopt the convention that *if a class $\mathscr{C}$ forms a set, then we identify $\mathscr{C}$ with the set it forms.* In other words, if $\mathscr{C} \equiv a$, then we simply write $\mathscr{C} = a$, again understanding this as a convention. It is important to emphasize that this convention only applies to classes forming sets. If a class does not form a set, we do not identify it with any set. We reiterate that most of our axioms will assert that a given class $\mathscr{C}$ forms a set. Our guiding intuition here is that if we can somehow 'see' all of the elements of a class $\mathscr{C}$ (for example in a natural way using sets that already exist), then $\mathscr{C}$ should form a set.

Axiom (Subset (schema))**.** *Let $\mathscr{C}$ be a class and let a be a set. If $\mathscr{C} \subseteq a$, then $\mathscr{C}$ is a set. Equivalently, if $P(x)$ is a property and a is a set, then the class*

$$\mathscr{D} = \{x \mid x \in a \wedge P(x)\}$$

is a set. This set is unique by the axiom of extensionality, and we refer to it as the subset of a determined by P.

Remark. Formally we define properties $P(x)$ using first-order logic. Let $S = \{\in\}$ be our symbol set and consider the first order language $\mathscr{L}_S$. Then a property is by definition a formula $\varphi(x, y_1, \dots, y_n) \in \mathscr{L}_S$ (where free variables $y_1, \dots, y_n$ are used for parameters). With this formalization, the subset axiom schema reads as follows:

> For all formulas $\varphi(x, y_1, \dots, y_n) \in \mathscr{L}_S$, the sentence
>
> $$\forall a_1 \cdots \forall a_n \forall a \exists b \forall x (x \in b \iff (x \in a \wedge \varphi(x, a_1, \dots, a_n)))$$
>
> is an axiom in our theory.

Note that this defines infinitely many axioms, hence it is technically an axiom schema. We will, however, often be imprecise and refer to this simply as the 'subset axiom'.

Corollary 9.0.3. *For all classes $\mathscr{C}$, $\mathscr{C}$ is a set iff $\mathscr{C} \subseteq a$ for some set a.*

Proof. Immediate from definitions and the subset axiom. □

We now introduce some axioms for forming sets. The following is the simplest:

Axiom (Pairing)**.** *Let a, b be sets. Then the class*

$$\mathscr{C} = \{x \mid x = a \vee x = b\}$$

is a set. This set is unique by the axiom of extensionality. We denote it by $\{a, b\}$*, and refer to it as* the (unordered) pair of a, b.

Note that $\{a, b\} = \{b, a\}$, which conforms to our intuitive expectation that sets, at least intrinsically, are not distingiushed by any ordering of their elements.

Given sets a, b, we intuitively want to be able to talk about the collection of all objects that are in a or b (here and in the future we use the *inclusive* sense of 'or', to refer to objects in either a or b or both). In fact, for any finite number of sets $a_1, \ldots, a_n$, we want to be able to talk about the collection of all objects that reside in any one of the sets a_i—that is, the collection of all objects x such that there exists some $1 \leq i \leq n$ with $x \in a_i$. But we do not want to restrict ourselves to even finitely many sets. We adopt the following axiom:

Axiom (Union)**.** *Let b be a set. Then the class*

$$\mathscr{C} = \{x \mid \exists a(a \in b \wedge x \in a)\}$$

is a set. This set is unique by the axiom of extensionality. We denote it by $\bigcup b$ *and refer to it as* the union over b.

Using this and the pairing axiom, we define for sets a, b the (unique) set

$$a \cup b = \bigcup\{a, b\}$$

to be the union of a and b. A similar procedure could be done for any finite number of sets.

Intuitively, if we can 'see' a set a, then we should be able to 'see' all of the subsets of a. It is useful to have an axiom that guarantees ahead of time that there exists a set containing as elements all such subsets in which we might be interested.

Axiom (Powerset)**.** *Let a be a set. Then the class*

$$\mathscr{C} = \{x \mid x \subseteq a\}$$

is a set. This set is unique by the axiom of extensionality. We denote it by $\mathscr{P}(a)$ *and refer to it as* the powerset of a.

Note that for a set a, $x \subseteq a$ iff $x \in \mathscr{P}(a)$ by definition.

Next we motivate the axiom of replacement. For any property $P(x, y)$, note that we have the following potential operation: for each set a, construct the class

$$\mathscr{O}(a) = \{x \mid P(x, a)\}$$

We make a definition:

Definition 9.0.4. Let b be a set and $P(x, y)$ a property. If for all $a \in b$, the class

$$\mathscr{O}(a) = \{x \mid P(x, a)\}$$

forms a set, then we say that $\mathscr{O}$ *is a definable operation on* b.

Intuitively, we want to guarantee the existence of the 'image set' of any set under a definable operation. We adopt the following axiom:

Axiom (Replacement (schema))**.** *Let b be a set and $\mathscr{O}$ be a definable operation on b. Then the class*

$$\mathscr{C} = \{x \mid (\exists a \in b)(x = \mathscr{O}(a))\}$$

is a set.

We note finally a rough form of the axiom of infinity, which asserts the existence of the set of natural numbers, and hence in particular the existence of an infinite set:

Axiom (Infinity)**.** *The class ω of natural numbers is a set.*

We will return to this axiom in greater detail later on. Let A, B be sets. Define

$$C = \left(\bigcup A\right) \cup \left(\bigcup B\right)$$

Intuitively, C can be viewed as the set containing the 'background' objects out of which A and B are built. Note that for all $a \in A$ and $b \in B$,

$$a \subseteq \bigcup A \subseteq C \quad \text{and} \quad b \subseteq \bigcup B \subseteq C$$

or equivalently $a, b \in \mathscr{P}(C)$. Hence $A \subseteq \mathscr{P}(C)$ and $B \subseteq \mathscr{P}(C)$, or $A, B \in \mathscr{P}(\mathscr{P}(C))$. Therefore we have recovered a hierarchy of sets using the pairing, union, and powerset axioms. This type of procedure can be carried out to prove that a given class is a set. More specifically, if we can prove that all of the elements in a given class belong to some set, then the class must be a set (by the subset axiom). We illustrate this in the following example:

Example. Let A be a set and consider the class

$$\mathscr{C} = \{\mathscr{P}(a) \mid a \in A\} = \{x \mid (\exists a \in A)(x = \mathscr{P}(a))\}$$

Note that $\mathscr{C}$ is guaranteed to be a set by the axiom of replacement. Nevertheless, we can prove directly that $\mathscr{C}$ is a set by finding a set in which all of the elements of $\mathscr{C}$ live, and then applying the subset axiom.

If $x \in \mathscr{C}$, then $x = \mathscr{P}(a)$ for some $a \in A$. That is, x is a set of (all) subsets of the element a. Recall however that $a \subseteq \bigcup A$, so x is a set of subsets of $\bigcup A$. This implies that $x \subseteq \mathscr{P}(\bigcup A)$, or $x \in \mathscr{P}(\mathscr{P}(\bigcup A))$. Since x was arbitrary, we have $\mathscr{C} \subseteq \mathscr{P}(\mathscr{P}(\bigcup A))$. Hence by the subset axiom, $\mathscr{C}$ is a set. In fact,

$$\mathscr{C} = \{x \mid x \in \mathscr{P}(\mathscr{P}(\bigcup A)) \land (\exists a \in A)(x = \mathscr{P}(a))\}$$

We will examine in greater detail the hierarchical nature of set theory later on. As a rough preview, we will define the 'hierarchy of sets' by iterating the powerset operation:

$$\begin{aligned}
V_0 &= \emptyset \\
V_1 &= \mathscr{P}(\emptyset) = \{\emptyset\} \\
&\vdots \\
V_{n+1} &= \mathscr{P}(V_n) \\
&\vdots \\
V_\omega &= \bigcup\{V_n \mid n \in \omega\} \\
V_{\omega+1} &= \mathscr{P}(V_\omega) \\
&\vdots
\end{aligned}$$

The objects of our theory will live in the hierarchy.

For now we return to simpler things. First let us note that we need the existence of the empty set. In fact, this follows from the existence of *any* set by the subset axiom, for if a is a set, then we define

$$\emptyset = \{x \mid x \in a \wedge x \neq x\}$$

In any case, we simply state the empty set axiom:

Axiom (Empty set)**.** *The class*

$$\mathscr{C} = \{x \mid x \neq x\}$$

is a set. This set is unique by the axiom of extensionality. We denote it by $\emptyset$*, and call it* the empty set. *A set* a *is said to be* empty *iff* $a = \emptyset$*, and* nonempty *iff it is not empty.*

We now define the intersection:

Theorem 9.0.8. *Let* b *be a nonempty set. Then the class*

$$\mathscr{C} = \{x \mid (\forall a \in b)(x \in a)\}$$

is a set. This set is unique by the axiom of extensionality. We denote it by $\bigcap b$*, and refer to it as* the intersection over b.

Proof. Since b is nonempty, we may choose $a \in b$. Now by the subset axiom, the class

$$\begin{aligned}
\mathscr{D} &= \{x \mid x \in a \wedge (\forall c \in b)(c \neq a \implies x \in c)\} \\
&= \{x \mid (\forall c \in b)(x \in c)\} = \mathscr{C}
\end{aligned}$$

is a set as desired. □

The hypothesis that b be nonempty is necessary. If $b = \emptyset$, then

$$\mathscr{C} = \{ x \mid (\forall a)(a \in \emptyset \implies x \in a) \}$$

But $a \in \emptyset$ is always false, so the implication $a \in \emptyset \implies x \in a$ is vacuously true for all a, no matter what x is. Hence $\mathscr{C}$ is the class of all sets in this case, which is not a set (if it were, we could form the subset of all sets that are not members of themselves, and arrive at the Russell paradox).

In light of this 'glitch', we introduce the notation

$$\bigcap_A B = A \cap \left(\bigcap B\right) = \{ x \in A \mid (\forall b \in B)(x \in b) \}$$

which is always guaranteed to be a set by the subset axiom.

It is instructive to note at this point that we have a decent amount of expressive power already. To demonstrate this, we introduce a definition:

Definition 9.0.5. Let a be a set and $A \subseteq \mathscr{P}(a)$. Then a set $m \in A$ is called *the smallest element in A* iff for all $x \in A$, $m \subseteq x$.

Note that a smallest element need not exist, but if it does exist it is unique, for if m, m' are both smallest elements in A, then $m \subseteq m'$ and $m' \subseteq m$ by the definition, hence $m = m'$ by the axiom of extensionality. We obtain a characterization:

Proposition 9.0.1. *Let a be a set and $A \subseteq \mathscr{P}(a)$. If A has a smallest element m, then*

$$m = \bigcap_a A = \{ x \in a \mid (\forall b \in A)(x \in b) \}$$

Conversely, if $m = \bigcap_a A \in A$, then m is the smallest element of A.

Proof. Suppose m is the smallest element in A. We claim $m = \bigcap_a A$. First note $m \subseteq a$. Now note that $m \subseteq \bigcap_a A$, for if $x \in m$ and $y \in A$, then since m is the smallest element in A, $m \subseteq y$ and thus $x \in y$. Since y was arbitrary, $x \in \bigcap_a A$. Since $m \in A$, trivially $\bigcap_a A \subseteq m$. Hence $m = \bigcap_a A$ by the axiom of extensionality.

The second claim of the proposition is trivial. □

We introduce additional termonology:

Definition 9.0.6. Let a be a set and $A \subseteq \mathscr{P}(a)$. We say that *$A$ is closed under arbitrary intersections (over a)* iff for all $B \subseteq A$, $\bigcap_a B \in A$.

The following is then a corollary of the previous proposition:

Corollary 9.0.4. *Let a be a set, $A \subseteq \mathscr{P}(a)$, and suppose A is closed under arbitrary intersections. Then there exists a smallest element in A, namely*

$$m = \bigcap_a A$$

More generally, for any $x \subseteq a$, there exists a smallest $m_x \subseteq a$ with the properties that $x \subseteq m_x$ and $m_x \in A$, namely

$$m_x = \bigcap_a \{ y \in A \mid x \subseteq y \}$$

We can generalize the notion of closure to arbitrary definable operations:

Definition 9.0.7. Let A be a set and $\mathscr{O}$ be a definable operation such that for all $X \subseteq A$, $\mathscr{O}(X) \subseteq A$. Then for $B \subseteq A$, we say that B *is closed under* $\mathscr{O}$ iff for all $X \subseteq B$, $\mathscr{O}(X) \subseteq B$.

Example. In this example we use script letters $\mathscr{A}, \mathscr{B}$ to refer to sets so that the notation is easier to follow. Let A and $\mathscr{O}$ be as in the above definition and define

$$\mathscr{A} = \{S \subseteq A \mid S \text{ is closed under } \mathscr{O}\}$$

Note that $\mathscr{A} \subseteq \mathscr{P}(A)$. We claim that $\mathscr{A}$ is closed under arbitrary intersections (over A). To verify this, let $\mathscr{B} \subseteq \mathscr{A}$. We must show that $\bigcap_A \mathscr{B} \in \mathscr{A}$, or equivalently that $\bigcap_A \mathscr{B}$ is closed under $\mathscr{O}$.

To verify the latter fact, let $X \subseteq \bigcap_A \mathscr{B}$. This implies that for all $S \in \mathscr{B}$, $X \subseteq S$, and hence since S is closed under $\mathscr{O}$ by hypothesis, $\mathscr{O}(X) \subseteq S$. Hence $\mathscr{O}(X) \subseteq \bigcap_A \mathscr{B}$. But since X was arbitrary, this just means that $\bigcap_A \mathscr{B}$ is closed under $\mathscr{O}$ as desired.

Hence $\bigcap_A \mathscr{B} \in \mathscr{A}$ and, since $\mathscr{B}$ was arbitrary, $\mathscr{A}$ is closed under arbitrary intersections.

We can apply the more general notion of a closure to define the concept of a *finite* set. Intuitively we know what a finite set is, but we must define it formally from the axioms of set theory. Later on we will do this in a different way after we construct the natural numbers, but for now we take another route. Our guiding intuition is as follows: any finite set should be obtainable by starting with the empty set and repeatedly adjoining one element.

Let A be a set. For $x \subseteq A$ and $a \in A$, we produce the set

$$x \cup \{a\}$$

which is intuitively obtained from x by adjoining the element a. Now for a given set $X \subseteq \mathscr{P}(A)$, we consider the definable operation

$$\mathscr{O}_A(X) = \{x \cup \{a\} \mid x \in X \wedge a \in A\}$$

where $\mathscr{O}_A(X)$ consists of all subsets of A obtainable by adjoining a single element $a \in A$ to a subset of A in X. Define

$$\mathrm{FIN}_A = \bigcap\nolimits_{\mathscr{P}(A)} \{X \subseteq \mathscr{P}(A) \mid \emptyset \in X \wedge X \text{ closed under } \mathscr{O}_A\}$$

to be the smallest $M \subseteq \mathscr{P}(A)$ such that $\emptyset \in M$ and M is closed under $\mathscr{O}_A$. Intuitively, it is clear that FIN_A consists of all of the finite subsets of A. In fact, we will temporarily say that a set A is *finite* iff $A \in \mathrm{FIN}_A$, and a set A is infinite iff A is not finite. The reader may verify in the following exercise that this definition yields expected results.

Exercise. We assume the preceding definitions.

(a) If $A \subseteq A'$, then A is finite iff $A \in \mathrm{FIN}_{A'}$.

(b) If $A \subseteq A'$ and A' is finite, then A is finite.

(c) If A is finite, and for all $a \in A$, a is finite, then $\bigcup A$ is finite.

Intuitively, these results state respectively: a subset is finite iff it is a finite subset, any subset of a finite set is finite, and a finite union of finite sets is finite.

It should be noted that without the axiom of infinity, there is no guarantee that there exist any infinite sets. In fact, the other axioms of set theory are consistent with the existence of only finite sets. We desire to implement a pairing operation

that will capture the ordering of the two elements in the pair:

Definition 9.0.8. We call $\mathscr{O}$ a *pairing operation* iff for all sets x, y, $\mathscr{O}(x, y)$ is a set and

$$\mathscr{O}(x, y) = \mathscr{O}(u, v) \implies x = u \text{ and } y = v$$

The standard pairing operation is the simplest. Note that

$$\mathscr{P}(\{x, y\}) = \{\emptyset, \{x\}, \{y\}, \{x, y\}\}$$

There are four definable operations determining this set, none of which are pairing operations. On the other hand, there are sixteen operations determining the set $\mathscr{P}(\mathscr{P}(\{x, y\}))$, the following four of which are pairing operations:

$$\begin{aligned} &\{\{x\}, \{x, y\}\} \\ &\{\{y\}, \{x, y\}\} \\ &\{\emptyset, \{x\}, \{x, y\}\} \\ &\{\emptyset, \{y\}, \{x, y\}\} \end{aligned}$$

We choose the first one, known as the *Kuratowski pair*:

Definition 9.0.9. Let x, y be sets. Then the *ordered pair of x and y* is defined to be

$$\langle x, y\rangle = \mathscr{O}_K(x, y) = \{\{x\}, \{x, y\}\}$$

More generally, z is an *ordered pair* iff there exist sets x, y with $z = \langle x, y\rangle$.

Note that $\langle x, y\rangle$ is a (unique) set by the subset axiom since $\langle x, y\rangle \in \mathscr{P}(\mathscr{P}(\{x, y\}))$. We must verify that $\langle x, y\rangle$ is indeed a pairing operation. In doing so, we first construct operations with which we may recover x and y from $\langle x, y\rangle$.

Let $z = \langle x, y\rangle$. Note that $\bigcup z = \{x, y\}$ and $\bigcap z = \{x\}$. It is immediate then that $x = \bigcup\bigcap z$. Now if $\bigcup z = \bigcap z$, then $y = x$. On the other hand, if $\bigcup z \neq \bigcap z$, then $y = \bigcup(\bigcup z - \bigcap z)$. In either case we have recovered x and y from z using the following operations:

$$\mathscr{X}(z) = \bigcup\bigcap z$$
$$\mathscr{Y}(z) = \begin{cases} \bigcup\bigcap z & \text{if } \bigcup z = \bigcap z \\ \bigcup(\bigcup z - \bigcap z) & \text{otherwise} \end{cases}$$

This allows us to prove that the ordered pair is indeed a pairing operation. For if $z = \langle x, y\rangle = \langle u, v\rangle$, then $x = \mathscr{X}(z) = u$ and $y = \mathscr{Y}(z) = v$.

Definition 9.0.10. Let A and B be sets. Then we define the *cartesian product of A and B* as

$$A \times B = \{ \langle x, y \rangle \mid x \in A \wedge y \in B \}$$

Note that $A \times B$ is indeed a set since if $\langle x, y \rangle \in A \times B$, we have $\{x, y\} \subseteq A \cup B$, hence

$$\langle x, y \rangle \in \mathscr{P}(\mathscr{P}(\{x, y\})) \subseteq \mathscr{P}(\mathscr{P}(A \cup B))$$

Therefore $A \times B \subseteq \mathscr{P}(\mathscr{P}(A \cup B))$, so $A \times B$ is a set by the subset axiom. We may also verify that $A \times B$ is a set using the replacement axiom a few times. To this end, fix $a \in A$ and construct

$$\{a\} \times B = \{ \langle a, y \rangle \mid y \in B \}$$

This is a set by the replacement axiom with $\mathscr{O}_1(y) = \langle a, y \rangle$. Now define for $x \in A$ the operation $\mathscr{O}_2$ given by $\mathscr{O}_2(x) = \{x\} \times B$. Applying the replacement axiom again, we obtain that

$$\{ \{x\} \times B \mid x \in A \}$$

is a set. Then we note that

$$A \times B = \bigcup \{ \{x\} \times B \mid x \in A \}$$

Definition 9.0.11. A *relation* R is a set of ordered pairs. We write xRy iff $\langle x, y \rangle \in R$. In addition we define the classes

$$\begin{aligned} \text{domain}(R) &= \{ x \mid (\exists y)(\langle x, y \rangle \in R) \} \\ \text{range}(R) &= \{ y \mid (\exists x)(\langle x, y \rangle \in R) \} \\ \text{field}(R) &= \text{domain}(R) \cup \text{range}(R) \end{aligned}$$

called the *domain*, *range*, and *field* of R, respectively.

Note that the domain and range of a relation are sets. This can be seen using the replacement axioms and the operations $\mathscr{X}$ and $\mathscr{Y}$ introduced earlier, or using the subset axioms. Hence the field is also a set by the union axiom.

Relations can be used to define equivalence relations, graphs, functions, and other mathematical objects. We first make some preliminary definitions:

Definition 9.0.12. Let A be a set. Then we define the *relation R_A induced by A* to be

$$R_A = \{ z \mid z \in A \wedge (\exists x)(\exists y)(z = \langle x, y \rangle) \}$$

We may then extend our previous definitions of domain and range to arbitrary sets by writing $\text{domain}(A) = \text{domain}(R_A)$ and $\text{range}(A) = \text{range}(R_A)$.

Definition 9.0.13. Let R, S be relations. We define the *composite of R and S* by

$$S \circ R = \{ \langle x, z \rangle \mid (\exists y)(\langle x, y \rangle \in R \wedge \langle y, z \rangle \in S) \}$$

We note that composition of relations satisfies associativity:

Proposition 9.0.2. *Let R, S, T be relations. Then*

$$(R \circ S) \circ T = R \circ (S \circ T)$$

Proof. Suppose $\langle x, z\rangle \in (R \circ S) \circ T$. This means that there exists y such that $\langle x, y\rangle \in T$ and $\langle y, z\rangle \in R \circ S$. But the last statement means that there exists y' such that $\langle y, y'\rangle \in S$ and $\langle y', z\rangle \in R$. It follows from $\langle x, y\rangle \in T$ and $\langle y, y'\rangle \in S$ that $\langle x, y'\rangle \in S \circ T$. Finally it follows from this result and $\langle y', z\rangle \in R$ that $\langle x, z\rangle \in R \circ (S \circ T)$. Hence

$$(R \circ S) \circ T \subseteq R \circ (S \circ T)$$

The reverse inclusion is similar. □

We now define a function:

Definition 9.0.14. We say that a relation R is a *function* iff

$$(\langle x, y\rangle \in R \wedge \langle x, z\rangle \in R) \implies y = z$$

For $x \in \mathrm{domain}(R)$, we denote by $R(x)$ the unique $y \in \mathrm{range}(R)$ with $\langle x, y\rangle \in R$.

We introduce some additional notation. Our definitions are as general as possible, though we are primarily interested in using them with relations.

Definition 9.0.15. Let R and A be sets. Then we define *the image of A under R* to be

$$R[A] = \{\, y \mid (\exists x \in A)(\langle x, y\rangle \in R)\,\}$$

Definition 9.0.16. Let R be a set. Then we define the *inverse of R* to be

$$R^{-1} = \{\, \langle y, x\rangle \mid \langle x, y\rangle \in R\,\}$$

Note that R^{-1} is always a relation, even if R is not a relation. Note also that for any set A, $R = A \times A$ is a relation and $R^{-1} = R$. If A has at least two distinct elements, then neither R nor R^{-1} is a function.

Let R be a set. We examine $R \circ R^{-1}$. Note that $\langle x, z\rangle \in R \circ R^{-1}$ iff there exists y such that $\langle x, y\rangle \in R^{-1}$ and $\langle y, z\rangle \in R$, or equivalently $\langle y, x\rangle \in R$ and $\langle y, z\rangle \in R$. Now the latter holds iff $x, z \in R[\{y\}]$, or equivalently $\langle x, z\rangle \in R[\{y\}] \times R[\{y\}]$. Hence we obtain the following identity:

Proposition 9.0.3. *For any set R,*

$$R \circ R^{-1} = \bigcup \{\, R[\{x\}] \times R[\{x\}] \mid x \in \mathrm{domain}(R)\,\}$$

If R is a function, then $R \circ R^{-1} = I_{\mathrm{range}(R)}$, where $I_{\mathrm{range}(R)}$ denotes the identity function on $\mathrm{range}(R)$.

Proof. To prove the second claim, suppose R is a function. Then for any $x \in \text{domain}(R)$, $R[\{x\}] = \{R(x)\}$. Hence we have

$$\begin{aligned} R \circ R^{-1} &= \bigcup \{ \{ \langle R(x), R(x) \rangle \} \mid x \in \text{domain}(R) \} \\ &= \{ \langle R(x), R(x) \rangle \mid x \in \text{domain}(R) \} \\ &= \{ \langle y, y \rangle \mid y \in \text{range}(R) \} \\ &= I_{\text{range}(R)} \end{aligned}$$

□

Note that the second part of the above proposition is not in general true for arbitrary relations.

We wish to examine properties of the image set operator $R[-]$. Before doing so, we introduce a preliminary definition:

Definition 9.0.17. A set R is said to be *single rooted* iff

$$(\langle x, y \rangle \in R \wedge \langle x', y \rangle \in R) \implies x = x'$$

Proposition 9.0.4. *A set R is single rooted iff R^{-1} is a function.*

Now we obtain a theorem:

Theorem 9.0.9. *Let R be a set. Let $\mathscr{A}$ be a set (of sets) and A and B be sets. Then*

(a) $R[\bigcup \mathscr{A}] = \bigcup \{ R[X] \mid X \in \mathscr{A} \}$

(b) $R[\bigcap \mathscr{A}] \subseteq \bigcap \{ R[X] \mid X \in \mathscr{A} \}$ *(where $\mathscr{A} \neq \emptyset$)*

(c) $R[A - B] \supseteq R[A] - R[B]$

If R is single rooted, then equality holds in (b) and (c).

Proof. We prove (b). Suppose that $y \in R[\bigcap \mathscr{A}]$. Then there exists $x \in \bigcap \mathscr{A}$ such that $\langle x, y \rangle \in R$. Now if $X \in \mathscr{A}$, then $x \in X$, hence $y \in R[X]$ since $\langle x, y \rangle \in R$. But since X was arbitrary, this means $y \in \bigcap \{ R[X] \mid X \in \mathscr{A} \}$. Hence (b) holds.

Suppose R is single rooted and $y \in \bigcap \{ R[X] \mid X \in \mathscr{A} \}$. Then since $\mathscr{A} \neq \emptyset$, we may choose $X \in \mathscr{A}$ with $x \in X$ such that $\langle x, y \rangle \in R$. Now if $X' \in \mathscr{A}$, then there exists $x' \in X'$ such that $\langle x', y \rangle \in R$. But $x = x'$ since R is single rooted, hence $x \in X'$. Since X' was arbitrary, we have $x \in \bigcap \mathscr{A}$, so $y \in R[\bigcap \mathscr{A}]$ and the reverse inclusion holds. □

We introduce some notation and termonology to work with functions:

Definition 9.0.18. If F is a function, we write $F : A \to B$ iff $\text{domain}(F) = A$ and $\text{range}(F) \subseteq B$. We write $a \mapsto b$ iff $a \in \text{domain}(F)$ and $F(a) = b$. We also write $F = \langle F(a) \mid a \in A \rangle$.

We say that F is *injective* (or *one-to-one*) iff F is single rooted. If $F : A \to B$, we say that F is *surjective* (or *onto*) iff $\text{range}(F) = B$. We say that F is *bijective* (or that F is a *one-to-one correspondence*) iff F is injective and surjective.

Remark. Note that 'surjective' is not really well-defined here. Surjectivity depends on the set into which we regard a function as mapping, and it is not determined in advance just which set this is. If $F : A \to B$ and $B \subseteq B'$, then $F : A \to B'$ also. More generally, if $\mathrm{range}(F) \subseteq B'$, then $F : A \to B'$. Hence when we say that a function is surjective, we must understand this *relative to some specified target set.* (For example, a function is always surjective on its range.) When we write '$F : A \to B$', we will generally understand B to be our working target set.

We may obtain a function from a relation in a natural way:

Definition 9.0.19. Let $R \subseteq A \times B$ be a relation. Define the function

$$\begin{aligned} F_R : A &\to \mathscr{P}(B) \\ a &\mapsto R[\{a\}] \end{aligned}$$

mapping $a \in A$ to its image set under R. We call F_R *the function induced by R.*

In the following theorem, we see that this definition actually provides us with an alternate way to speak about relations. More specifically, for given sets A and B, we may naturally identify relations from A to B and functions from A to $\mathscr{P}(B)$.

Theorem 9.0.10. *Let A and B be sets. Define*

$$\mathscr{A} = \{R \mid R \subseteq A \times B\} \qquad \mathscr{B} = \{F \mid F : A \to \mathscr{P}(B)\}$$

Then the function

$$\begin{aligned} \mathscr{F} : \mathscr{A} &\to \mathscr{B} \\ R &\mapsto F_R \end{aligned}$$

is a bijection from $\mathscr{A}$ to $\mathscr{B}$.

Proof. We construct an inverse function $\mathscr{G} : \mathscr{B} \to \mathscr{A}$ directly. For $F \in \mathscr{B}$, define

$$\mathscr{G}(F) = R_F = \{\langle a, b\rangle \mid a \in A \wedge b \in F(a)\}$$

We claim that for $R \in \mathscr{A}$, $R_{(F_R)} = R$ and for $F \in \mathscr{B}$, $F_{(R_F)} = F$. From this injectivity and surjectivity of $\mathscr{F}$ follow, for if $F_R = F_{R'}$, then

$$R = R_{(F_R)} = R_{(F_{R'})} = R'$$

and if $F \in \mathscr{B}$, then for $R = R_F$ we have $F_R = F$.

To prove that $R_{(F_R)} = R$ for $R \in \mathscr{A}$, note that $\langle a, b\rangle \in R_{(F_R)}$ iff $a \in A$ and $b \in F_R(a)$, that is, iff $b \in R[\{a\}]$, or equivalently $\langle a, b\rangle \in R$. Similarly for $F \in \mathscr{B}$, $F_{(R_F)}(a) = X$ iff $R_F[\{a\}] = X$, or equivalently $F(a) = X$. □

Previously we constructed a natural bijection

$$\mathscr{F} : \{R \mid R \subseteq A \times B\} \to \{F \mid F : A \to \mathscr{P}(B)\}$$

for fixed sets A and B. If $f : A \to B$ is a bijection, we can construct natural bijections between other sets constructed from A and B. For example, define

$$\begin{aligned} F : \mathscr{P}(A) &\to \mathscr{P}(B) \\ X &\mapsto f[X] = \{ f(x) \mid x \in X \} \end{aligned}$$

This is easily seen to be a bijection. To verify injectivity, note that if $f[X] = f[X']$ and $x \in X$, then $f(x) \in f[X] \subseteq f[X']$. Hence there exists $x' \in X'$ with $f(x) = f(x')$. But since f is injective, $x = x'$, so $x \in X'$. Hence $X \subseteq X'$, and similarly $X' \subseteq X$, so $X = X'$. To verify surjectivity, note that if $Y \subseteq B$, then we may set

$$X = f^{-1}[Y] = \{ x \in A \mid f(x) \in Y \} \subseteq A$$

and trivially $f[X] \subseteq Y$. But $Y \subseteq f[X]$ by the surjectivity of f, hence $F(X) = f[X] = Y$.

Similarly we may define

$$\begin{aligned} F : A \times A &\to B \times B \\ (a, a') &\mapsto (f(a), f(a')) \end{aligned}$$

We leave it to the reader to verify that F is a bijection.

We now introduce some termonology and results for working with bijections.

Definition 9.0.20. Let $f : A \to B$. Then $g : B \to A$ is called a *left inverse of f* iff $g \circ f = I_A$.

Note that for $f : A \to B$ and $b \in B$, if $g : B \to A$ is a left inverse of f, then we must have, for each $a \in A$ with $f(a) = b$, $g(b) = g(f(a)) = a$. But g is a function, hence there is at most one $a \in A$ with $f(a) = b$. But this just means that f is injective, or equivalently f^{-1} is a function. We have, for $b \in \text{range}(f)$, $g(b) = f^{-1}(b)$. More generally, set $C = B - \text{range}(f)$. Then any left inverse g of f will be of the form $g = f^{-1} \cup h$ where $h : C \to A$ is arbitrary.

We have shown that if a function $f : A \to B$ has a left inverse, then it is injective. Does the converse hold? Almost. We encounter a small glitch if $A = \emptyset$ but $B \neq \emptyset$, in which case $f = \emptyset$ is injective, but f has no left inverses (since there are no functions from B into A). If $A \neq \emptyset$ and f is injective, however, then we may choose $a \in A$ and define $g : B \to A$ by

$$g = f^{-1} \cup ((B - \text{range}(f)) \times \{a\})$$

It is immediate that $g \circ f = I_A$. Hence we have a theorem:

Theorem 9.0.11. *Let $f : A \to B$. If f has a left inverse, then f is injective. If $A \neq \emptyset$, then the converse holds too.*

We may similarly characterize surjectivity:

Definition 9.0.21. Let $f : A \to B$. Then $g : B \to A$ is called a *right inverse of f* iff $f \circ g = I_B$.

It is immediate that the existence of a right inverse implies surjectivity, for if $f : A \to B$, and $g : B \to A$ is a right inverse of f, then for all $b \in B$, $f(g(b)) = b$.

The converse claim is less trivial. If $f : A \to B$ is surjective, then for each $b \in B$, the preimage set

$$f^{-1}[\{b\}] = \{x \in A \mid f(x) = b\}$$

is nonempty. We can define a relation $R \subseteq B \times A$ by

$$R = \{\langle y, x\rangle \mid y \in B \wedge x \in A \wedge f(x) = y\}$$

relating each $b \in B = \mathrm{domain}(R)$ to its preimages in A under f. What we desire is a function $g : B \to A$ such that $g \subseteq R$. For then we have, for all $b \in B$, $g(b) \in f^{-1}[\{b\}]$, so $f(g(b)) = b$. Hence in this case g is a right inverse for f.

Unfortunately *the existence of such a function g does not follow from our current axioms*. We must adopt a version of the *axiom of choice*:

Axiom (Choice)**.** *Let C, D be sets and $R \subseteq C \times D$ with* $\mathrm{domain}(R) = C$*. Then there exists a function $g : C \to D$ such that $g \subseteq R$.*

This axiom is called the 'axiom of choice' because *it allows us to make, possibly infinitely many, arbitrary choices*. In our present case, we know that there exists for each $b \in B$ at least one preimage $a \in A$ under f. The axiom of choice allows us then to actually choose, for each $b \in B$, an arbitrary preimage $a_b \in A$. Formally, this means the axiom gives us a function $g : B \to A$ mapping $b \mapsto a_b$ where $f(a_b) = b$—in other words, a right inverse for f. Hence we can complete our desired characterization of surjectivity:

Theorem 9.0.12. *Let $f : A \to B$. Then f has a right inverse iff f is surjective.*

Remark. It is important to note that the axiom of choice is *not* required to choose an element from a single nonempty set. Nor is it required to choose an element from each of an arbitrary finite number of nonempty sets. Nor is it required to choose an element from each of possibly infinitely many nonempty sets when it is possible to *define* (using a first-order formula) the element we are choosing from each nonempty set. For example, we will see later that the set ω of natural numbers is *well ordered*, meaning that every nonempty subset has a least element. We can define a choice function for ω, without appealing to the axiom of choice, by selecting the least element from each nonempty subset. The axiom of choice is really necessary only when we must make infinitely many arbitrary choices.

The following proposition uses the axiom of choice:

Proposition 9.0.5. *Let $\mathscr{A}$ be a nonempty set (of sets). Then*

$$\bigcap_{A \in \mathscr{A}} \Big(\bigcup A\Big) = \bigcup \{ \bigcap_{A \in \mathscr{A}} G(A) \mid G \textit{ a function } \wedge \mathrm{domain}(G) = \mathscr{A} \wedge (\forall A \in \mathscr{A})(G(A) \in A)\}$$

Proof. One direction is easy. If x is in the set on the right hand side, then there exists some function G satisfying the specified properties such that $x \in \bigcap_{A \in \mathscr{A}} G(A)$. Fix $A \in \mathscr{A}$. Then we know $x \in G(A)$. But $G(A) \in A$, hence $x \in \bigcup A$. Since A was arbitrary, we have $x \in \bigcap_{A \in \mathscr{A}} (\bigcup A)$.

To prove the converse inclusion, suppose $x \in \bigcap_{A \in \mathscr{A}} (\bigcup A)$. Then for each $A \in \mathscr{A}$, there exists at least one set $a \in A$ with $x \in a$. Define the relation

$$R = \{\langle A, a\rangle \mid A \in \mathscr{A} \wedge a \in A \wedge x \in a\}$$

Note that $R \subseteq \mathscr{A} \times \bigcup \mathscr{A}$ and $\operatorname{domain}(R) = \mathscr{A}$. Hence by the axiom of choice there exists a function $G : \mathscr{A} \to \bigcup \mathscr{A}$ such that, for all $A \in \mathscr{A}$, $G(A) \in A$ and $x \in G(A)$. Hence $x \in \bigcap_{A \in \mathscr{A}} G(A)$. But then x is in the set on the right hand side above, so the reverse inclusion holds as desired. □

This result is known as the *general distributivity law for union and intersection.*

Let A be a set. It is sometimes useful to 'index' the elements of A by constructing some 'indexing function' $f : I \to A$ such that

$$A = \{f(i) \mid i \in I\} = \operatorname{range}\langle f(i) \mid i \in I\rangle$$

A trivial way to do this, of course, is to set $I = A$ and $f = I_A$. But other indexing sets will be more useful in most circumstances.

We introduce a generalized cartesian product:

Definition 9.0.22. Let $\mathscr{F} = \langle A_i \mid i \in I\rangle$ be given. Then the *cartesian product (over $\mathscr{F}$)* is

$$\prod \mathscr{F} = \prod_{i \in I} A_i = \{f \mid f \text{ a function} \wedge \operatorname{domain}(f) = I \wedge (\forall i \in I)(f(i) \in A_i)\}$$

Note that if $f \in \prod_{i \in I} A_i$, then $f : I \to \bigcup_{i \in I} A_i$, so $f \in \mathscr{P}(I \times \bigcup_{i \in I} A_i)$. Hence $\prod_{i \in I} A_i \subseteq \mathscr{P}(I \times \bigcup_{i \in I} A_i)$, and so $\prod_{i \in I} A_i$ is a set by the subset axiom.

Intuitively, $f \in \prod_{i \in I} A_i$ may be viewed as an 'I-tuple' of values over $\bigcup_{i \in I} A_i$, where for $i \in I$, the 'i-th coordinate' $f(i)$ of f is an element of A_i.

Note that we encounter a slight 'glitch' with the above definition: if there exists some $i \in I$ with $A_i = \emptyset$, then $\prod_{i \in I} A_i = \emptyset$. Hence we typically assume that $A_i \neq \emptyset$ for all $i \in I$. By the axiom of choice, we are then guaranteed that $\prod_{i \in I} A_i \neq \emptyset$. For we may construct the relation

$$R = \{\langle i, a\rangle \mid i \in I \wedge a \in A_i\}$$

Now $\operatorname{domain}(R) = I$ by hypothesis. Hence by the axiom of choice, there exists a function f with $\operatorname{domain}(f) = I$ and $f(i) \in A_i$ for all $i \in I$. But then $f \in \prod_{i \in I} A_i$.

We now define some properties that may be satisfied by relations. Again, we make our definitions here as general as possible, though we expect them to be used primarily with relations.

Definition 9.0.23. Let R be a set and A be a set.

1. R is *reflexive on* A iff $\langle a, a\rangle \in R$ for all $a \in A$.
2. R is *irreflexive on* A iff $\langle a, a\rangle \notin R$ for all $a \in A$.
3. R is *symmetric* iff $\langle a, b\rangle \in R$ implies $\langle b, a\rangle \in R$.

4. R is *asymmetric* iff $\langle a,b\rangle \in R$ implies $\langle b,a\rangle \notin R$.
5. R is *antisymmetric* iff for $a \neq b$, $\langle a,b\rangle \in R$ implies $\langle b,a\rangle \notin R$.
6. R is *transitive* iff ($\langle a,b\rangle \in R$ and $\langle b,c\rangle \in R$) implies $\langle a,c\rangle \in R$.
7. R is *connected on A* iff for $a,b \in A$, $\langle a,b\rangle \in R$ or $\langle b,a\rangle \in R$.

These properties may be characterized set theoretically. For a set A, define

$$=_A\ = \{\langle x,y\rangle \in A \times A \mid x = y\} = \{\langle x,x\rangle \mid x \in A\}$$

Theorem 9.0.13. *Let R be a set and A be a set.*

1. *R is reflexive on A iff $R \supseteq =_A$. For any S, $S \cup =_A$ is reflexive on A.*
2. *R is irreflexive on A iff $R \cap =_A = \emptyset$. For any S, $S - =_A$ is irreflexive on A.*
3. *R is symmetric iff $R^{-1} \subseteq R$. For any S, $S \cup S^{-1}$ is symmetric.*
4. *R is asymmetric iff $R^{-1} \cap R = \emptyset$. For any S, $S - S^{-1}$ is asymmetric.*
5. *R is transitive iff $R \circ R \subseteq R$. For any S, the set*

$$S \cup (S \circ S) \cup (S \circ S \circ S) \cup \cdots$$

is transitive.

We may now define equivalence relations:

Definition 9.0.24. Let R be a relation on A. We say that R is an *equivalence relation on A* iff R is reflexive on A, symmetric, and transitive. For $a \in A$, we call the set

$$[a]_R = R[\{a\}] = \{b \in A \mid \langle a,b\rangle \in R\}$$

the *equivalence class of a under R*.

The following proposition is immediate from definitions:

Proposition 9.0.6. *Let R be an equivalence relation on A and $a,b \in A$. Then*

$$[a]_R \cap [b]_R \neq \emptyset \implies [a]_R = [b]_R$$

Note that for $a \in A$, $a \in [a]_R$. For $C = [a]_R$, we see that $C \times C \subseteq R$. Also, if $\langle a,b\rangle \in R$, then $\langle a,b\rangle \in C \times C$. Hence we have a set $\mathscr{C} = \{[a]_R \mid a \in A\}$ of nonempty pairwise disjoint sets with $A = \bigcup \mathscr{C}$ and $R = \bigcup\{C \times C \mid C \in \mathscr{C}\}$.

Definition 9.0.25. Let R be an equivalence relation on A. We define $A/R = \{[a]_R \mid a \in A\}$ and define the *projection* map

$$\begin{aligned} \pi_R : A &\to A/R \\ a &\mapsto [a]_R \end{aligned}$$

Example. As an example of an equivalence relation, let $f : A \to B$ and define

$$E_f = \{\langle a_1, a_2\rangle \in A \times A \mid f(a_1) = f(a_2)\}$$

It is trivially verified that E_f is an equivalence relation. For $b \in \operatorname{range}(f)$, define

$$C_b = \{a \in A \mid f(a) = b\}$$

Since $b \in \operatorname{range}(f)$, there exists $a_b \in A$ with $f(a_b) = b$. We see that $C_b = [a_b]_{E_f}$. Hence we have $A/E_f = \{C_b \mid b \in \operatorname{range}(f)\}$.

The previous example leads us to a general theorem:

Theorem 9.0.14. *Let $f : A \to B$. Then there exists an equivalence relation $E \subseteq A \times A$, a set $\widehat{B} \subseteq B$, and a bijection $\widehat{f} : A/E \to \widehat{B}$ mapping $[a]_E \mapsto f(a)$ and satisfying $f = \widehat{f} \circ \pi_E$.*

We now introduce the notions of partial and linear orderings:

Definition 9.0.26. Let R be a relation on A. We say that R is a *partial ordering on A* iff R is reflexive on A, antisymmetric, and transitive. We say that R is a *linear ordering on A* iff R is a partial ordering on A and R is connected on A.

Example. Let A be a set. Then $\mathscr{P}(A)$ is partially ordered under the relation

$$\subseteq_{\mathscr{P}(A)} = \{\langle X, Y\rangle \in \mathscr{P}(A) \times \mathscr{P}(A) \mid X \subseteq Y\}$$

This example is actually totally typical in the sense that any arbitrary partial ordering naturally induces a partial ordering under set inclusion. To see this, let $\langle A, \preceq\rangle$ be an arbitrary partial ordering. For $a \in A$, define the *segment of a* by

$$S_a = \{x \in A \mid x \preceq a\}$$

and set $\mathscr{A} = \{S_a \mid a \in A\}$. Define $\mathscr{F} : A \to \mathscr{A}$ by $a \mapsto S_a$. Then $\mathscr{F}$ is a bijection and $a \preceq b$ iff $\mathscr{F}(a) = S_a \subseteq S_b = \mathscr{F}(b)$.

Example. As another example of a partial ordering, let R be an arbitrary transitive relation on A. Define the equivalence relation

$$E_R = \{\langle a, b\rangle \mid a = b \vee (\langle a, b\rangle \in R \wedge \langle b, a\rangle \in R)\}$$

Then R induces a relation $\widehat{R}$ on A/E_R defined by

$$\langle [a]_R, [b]_R\rangle \in \widehat{R} \iff \langle a, b\rangle \in R$$

The reader may verify that $\widehat{R}$ is a well defined partial ordering on A/E_R.

This example actually illustrates a very general principle regarding equivalence relations. Defining an equivalence relation and then moving into the set of equivalence classes allows us in effect to 'identify' any two objects related by the relation. The resulting objects (the equivalence classes) will in general satisfy new, desired, properties as a result of this identification process. Above, E_R can be seen as identifying objects that are already equal or else related to each other in both directions under R. This has the effect of imposing reflexivity and antisymmetry among the resulting objects under (an appropriate redefinition of) R. Since R was already transitive, we obtain a partial ordering.

In our definitions of orderings above, we have required reflexivity. Intuitively our definitions model a 'less than or equal to' relation $\leq$. One may instead model a 'strictly less than' relation $<$ and require irreflexivity. (Of course, one may alternately require neither.) We make this distinction formally:

Definition 9.0.27. Let R be an ordering relation on A. Then the relation $R - =_A$ is called the *strict ordering induced by* R. The relation $R \cup =_A$ is called the *nonstrict ordering induced by* R.

Note that in our treatment, an ordering R is always nonstrict.

We now work more with partial orderings.

Definition 9.0.28. Let P be partially ordered under R. Then $C \subseteq P$ is called a *chain* iff R induces a linear ordering on C, that is, iff C is linearly ordered under $R \cap (C \times C)$.

Note that for any partial ordering P, $\emptyset \subseteq P$, and $\emptyset$ is a chain by the above definition. Partial orderings and chains arise in many areas of mathematics. We give two simple examples from algebra:

Example. Let G be a group and set

$$\mathscr{S} = \{H \mid H \text{ is a subgroup of } G\}$$

Then P is partially ordered under inclusion. Note that if $\mathscr{C} \subseteq \mathscr{S}$ is a nonempty chain, then $\bigcup \mathscr{C} \in \mathscr{S}$. Indeed, if $x, y \in \bigcup \mathscr{C}$, then there exist $H_1, H_2 \in \mathscr{C}$ with $x \in H_1$ and $y \in H_2$. But $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$ since $\mathscr{C}$ is a chain. If $H_1 \subseteq H_2$, then $x, y \in H_2$, hence $xy \in H_2 \subseteq \bigcup \mathscr{C}$ since H_2 is a subgroup. Similarly if $H_2 \subseteq H_1$. Hence $\bigcup \mathscr{C}$ is closed under the group operation. Now $e \in \bigcup \mathscr{C}$ since $e \in H$ for some $H \in \mathscr{C}$ ($\mathscr{C}$ nonempty). Finally, if $x \in \bigcup \mathscr{C}$, then $x \in H$ for some $H \in \mathscr{C}$, hence $x^{-1} \in H \subseteq \bigcup \mathscr{C}$. Thus $\bigcup \mathscr{C}$ is closed under inverses. We have established that $\bigcup \mathscr{C}$ is a subgroup of G, that is, $\bigcup \mathscr{C} \in \mathscr{S}$. In the language of closures, we may say that $\mathscr{S}$ is closed under the taking of unions over nonempty chains.

Example. Let R be a commutative ring and set

$$\mathscr{S} = \{I \mid I \text{ is a proper ideal of } R\}$$

Then the reader may verify that $\mathscr{S}$ is closed under unions over nonempty chains.

Definition 9.0.29. Let P be partially ordered under R. Then $M \in P$ is called a *maximal element* iff there does not exist $X \in P$ with $M \neq X$ such that $\langle M, X \rangle \in R$. The element $M \in P$ is called a *maximum element* iff for all $X \in P$, $\langle X, M \rangle \in R$.

We similarly define *minimal* and *minimum* elements.

Note that there may exist multiple maximal [minimal] elements, but a maximum [minimum] element is unique whenever it exists (by antisymmetry). Also the concepts of maximal [minimal] and maximum [minimum] concide for linear orderings.

We may now state a key result for partial orderings:

Theorem 9.0.15 (Zorn)**.** *Let P be partially ordered under inclusion and suppose that for all chains $C \subseteq P$, $\bigcup C \in P$ (in other words, P is closed under unions over arbitrary chains). Then there exists a maximal element $M \in P$.*

This result is called 'Zorn's lemma' for historical reasons. It is extremely useful in many areas in mathematics. For example, in linear algebra it is used to prove that every vector space has a basis (a basis is simply a maximal linearly independent set of vectors). We give a simpler example:

Example. Let G be a group, $g \in G$, $g \neq e$. Then there exists a maximal subgroup H of G such that $g \notin H$. In other words, there exists a maximal element in

$$\mathscr{S} = \{H \mid H = \emptyset \vee (H \text{ is a subgroup of } G \wedge g \notin H)\}$$

Indeed, this follows quickly from the results of the first example above and Zorn's lemma. Note that we include $\emptyset \in \mathscr{S}$ in order to account for the empty chain.

Often we are interested in working with the smallest set with a given property:

Definition 9.0.30. Let A be a set, $\mathscr{A} \subseteq \mathscr{P}(A)$, and $P(x)$ be a property. (Note that $\mathscr{A}$ is partially ordered under inclusion.) We say that $S \in \mathscr{A}$ is *the smallest set satisfying P* iff S is the minimum element under inclusion in

$$\mathscr{B} = \{X \mid X \in \mathscr{A} \wedge P(X)\}$$

or equivalently iff for all $X \in \mathscr{A}$, $P(X)$ implies $S \subseteq X$.

Note that in order to show that $\mathscr{A}$ has a smallest element satisfying P, it is sufficient to show that $\bigcap \mathscr{B} \in \mathscr{B}$. We give an example from real analysis:

Example. Recall that a sequence $\langle a_n \mid n \in \omega \rangle$ in $\mathbb{R}$ is said to *converge* to $a \in \mathbb{R}$ iff for all $\epsilon > 0$, there exists $N \geq 0$ such that for all $n \geq N$, $|a_n - a| < \epsilon$. A subset $X \subseteq \mathbb{R}$ is *closed* iff all convergent sequences in X converge to values in X—that is, iff for all $\langle a_n \mid n \in \omega \rangle$ in X, if $\langle a_n \mid n \in \omega \rangle$ converges to $a \in \mathbb{R}$, then $a \in X$.

Let G be a group over $\mathbb{R}$. For $X \subseteq \mathbb{R}$, there exists a smallest subgroup B of G such that $X \subseteq B$ and B is closed. Simply set

$$\mathscr{B} = \{H \mid H \text{ a subgroup of } G, X \subseteq H, H \text{ closed}\}$$

and take $B = \bigcap \mathscr{B}$.

Finally, we prove a result about finite linear orderings using our alternative definition of finite given in a previous lecture:

Theorem 9.0.16. *Let $\langle A, \preceq \rangle$ be a linear ordering with A finite. Then for any nonempty $X \subseteq A$, X has a least element and a greatest element.*

Proof. We prove the theorem for least elements.

Note that $\mathscr{P}(A) = \mathrm{FIN}_A$ by a previous exercise. Hence it suffices to show that the claim holds for nonempty sets in FIN_A. Define

$$\mathscr{A} = \{X \in \mathrm{FIN}_A \mid X \neq \emptyset \implies X \text{ has a least element}\}$$

We claim that $\emptyset \in \mathscr{A}$ and that $\mathscr{A}$ is closed under the operation of adjoining a single element of A to any one of its members. Then, since FIN_A is the smallest set satisfying these conditions by construction, we have $\mathrm{FIN}_A \subseteq \mathscr{A}$ and hence $\mathscr{A} = \mathrm{FIN}_A$. This establishes our desired claim.

Vacuously, $\emptyset \in \mathscr{A}$. Suppose now $X \in \mathscr{A}$ and $a \in A$. Consider the set $X \cup \{a\}$. If $a \le x$ for all $x \in X$, then a is the least element in $X \cup \{a\}$ and so in particular $X \cup \{a\}$ has a least element. If a is not the least element, there exists $x \in X$ such that $x < a$. In particular, $X \neq \emptyset$, hence there exists a least element $x' \in X$ since $X \in \mathscr{A}$. Now $x' \le x < a$, hence x' is also the least element of $X \cup \{a\}$, and so again $X \cup \{a\}$ has a least element. This shows $X \cup \{a\} \in \mathscr{A}$, so $\mathscr{A}$ is closed as desired. □

We continue our treatment of finite linear orderings. Previously we proved that every nonempty subset of a finite linear ordering has a least and greatest element. In fact this property characterizes finite linear orderings.

To prove this, we need a version of induction. Let $\langle A, \preceq \rangle$ be a nonempty linear ordering such that every nonempty subset of A has a least and greatest element. Then in particular A has a least (or *bottom*) element β and a greatest (or *top*) element τ. In addition, if $a \in A$ and $a \prec \tau$, then $A - S_a$ is a nonempty subset of A and hence it has a (unique) least element. (Recall that for $a \in A$, $S_a = \{x \in A \mid x \preceq a\}$.) For $a \prec \tau$, we denote the least element of $A - S_a$ by $\sigma(a)$ and refer to it as the *successor of a*. Then $\sigma : (A - \{\tau\}) \to A$ is called the *successor function on A*.

Theorem 9.0.17 (Induction on finite linear orderings)**.** *Let $\langle A, \preceq \rangle$ be a nonempty linear ordering such that every nonempty subset of A has a least and greatest element. Let β and τ be the least and greatest elements of A, respectively, and let σ be the successor function on A. Suppose $S \subseteq A$ satisfies*

(i) $\beta \in S$

(ii) If $a \in S$ and $a \prec \tau$, then $\sigma(a) \in S$

Then $\tau \in S$, or equivalently $S = A$.

Proof. Suppose $\tau \notin S$. Then $A - S$ is nonempty, so it has a least element λ. Now $\lambda \neq \beta$ since $\beta \in S$ by (i). Hence there exists $\gamma \in A$ such that $\sigma(\gamma) = \lambda$ (let γ be the greatest element of the nonempty subset $S_\lambda - \{\lambda\}$). Now $\gamma \prec \lambda \prec \tau$, hence $\gamma \notin S$ by (ii). But then $\gamma \in A - S$, contradicting that λ is least in $A - S$. Hence our original supposition is false and $\tau \in S$.

To see that this result implies $S = A$, set $S' = \{a \in A \mid S_a \subseteq S\}$. Then it is immediate that S' satisfies (i) and (ii), hence $\tau \in S'$. But $A = S_\tau$, so $S = A$. The converse is trivial (if $S = A$, then $\tau \in S$). □

We can now state our theorem:

Theorem 9.0.18. *Let $\langle A, \preceq \rangle$ be a linear ordering such that every nonempty subset has a least and greatest element. Then A is finite.*

Proof. If A is empty, then we are done. Otherwise let β, τ, σ be as above. Define

$$S = \{a \in A \mid S_a \text{ is finite}\}$$

Trivially $\beta \in S$. If $a \in S$ and $a \prec \tau$, then $\sigma(a) \in S$ since $S_{\sigma(a)} = S_a \cup \{\sigma(a)\}$ is finite. Hence $\tau \in A$ by induction, so $A = S_\tau$ is finite. □

We now introduce the idea of a (finite) recursion. Intuitively, it is clear that a (finite) sequence of values can be constructed by specifying the initial element in the sequence in conjunction with an operation used to obtain the next element in the sequence from any given element. For example, to define the sequence 2^n of powers of 2, we can specify

$$2^0 = 1 \qquad \text{and} \qquad 2^{n+1} = 2^n \cdot 2 \quad (n \geq 0)$$

where our initial element is $m = 1$ and our successor operation is $m \mapsto m \cdot 2$.

While this treatment is acceptable as it stands in naive set theory, in axiomatic set theory we must specify what a recursion is and prove that recursions exist—that is, prove that the sequence intuitively determined by a given recursive setup actually forms a set (a recursive function). We do this now for finite recursions.

Definition 9.0.31. Let $\langle A, \preceq \rangle$ be a nonempty finite linear ordering and let β, τ, σ be as above. Let b be a set and let $\mathcal{O}$ be a definable operation (on the class of all sets). Then if f is a function on A such that $f(\beta) = b$ and for all $a \in A$, $a \prec \tau$ implies $f(\sigma(a)) = \mathcal{O}(f(a))$, f is called a *finite recursion (on A)*.

Theorem 9.0.19 (Recursion on finite linear orderings (schema))**.** *Let $\langle A, \preceq \rangle$ be a nonempty finite linear ordering with β, τ, σ as above. Let b be a set and $\mathcal{O}$ be a definable operation. Then there exists a unique finite recursion on A determined by b and $\mathcal{O}$.*

Proof. We proceed by induction on A. Define

$$S = \{a \in A \mid (\exists f)[f \text{ a function } \wedge \text{domain}(f) = S_a \wedge f(\beta) = b$$
$$\wedge (\forall x \in A)(x \prec a \implies f(\sigma(x)) = \mathcal{O}(f(x)))]\}$$

Note that $\beta \in S$ since $f = \{\langle \beta, b \rangle\}$ is an appropriate recursion on S_β. Suppose $a \in S$ and $a \prec \tau$. Choose f an appropriate recursion on S_a. Set

$$f' = f \cup \{\langle \sigma(a), \mathcal{O}(f(a)) \rangle\}$$

Then it is immediate that f' is a well defined appropriate recursion on $S_{\sigma(a)}$, so $\sigma(a) \in S$. By induction it follows that $\tau \in S$, and since $A = S_\tau$, this establishes the existence of the desired recursion on A.

Uniqueness of the recursion also follows by induction. Suppose f and g are two recursions on A determined by b and $\mathcal{O}$. Set

$$S' = \{a \in A \mid f(a) = g(a)\}$$

Then $f(\beta) = b = g(\beta)$, so $\beta \in S'$. If $a \in S'$ and $a \prec \tau$, then $f(a) = g(a)$ and

$$f(\sigma(a)) = \mathcal{O}(f(a)) = \mathcal{O}(g(a)) = g(\sigma(a))$$

so $\sigma(a) \in S'$. By induction, $S' = A$, so $f = g$. □

We use the notion of a finite recursion to construct the set of natural numbers in axiomatic set theory. In naive mathematics, of course, we have an intuitive notion of a set $\mathbb{N} = \{0, 1, 2, \ldots\}$ of natural numbers.[1] We can implement, for each $n \in \mathbb{N}$, a set S_n such that

$$m \neq n \implies S_m \neq S_n$$

Here we wish to define S_n to contain exactly n elements. This forces us to set $S_0 = \emptyset$. Given $S_0, \ldots, S_n$, a natural definition for S_{n+1} is

$$S_{n+1} = \{S_0, \ldots, S_n\}$$

Note that S_n is defined for all $n \in \mathbb{N}$ by naive recursion on $\mathbb{N}$, and $S_{n+1} = S_n \cup \{S_n\}$ by naive induction on $\mathbb{N}$. Intuitively then, we can simply define $\omega = \{S_n \mid n \in \mathbb{N}\}$.

Of course, this is unsatisfactory for a rigorous axiomatic development, since it relies on the set of naturals used in naive mathematics! Nevertheless, we can use it as a guide for our set theoretic construction of the naturals.

Definition 9.0.32. Let x be a set. Then the *(ordinal) successor of* x is defined to be the set

$$x^+ = x \cup \{x\}$$

We define ω to consist of of all of the terminal values of appropriately constructed recursions on finite linear orderings:

Definition 9.0.33. The class ω of *natural numbers* is defined as

$$\begin{aligned}\omega = \{ f(\tau) \mid f \text{ a recursion on a nonempty finite linear ordering } \langle A, \preceq \rangle \text{ with } \beta, \tau, \sigma \\ \wedge f(\beta) = \emptyset \wedge (\forall a \in A)(a \prec \tau \implies f(\sigma(a)) = f(a)^+) \}\end{aligned}$$

Recall that the axiom of infinity states that ω is a set. We prove a simple fact about ω that will be required later on:

Theorem 9.0.20. *The empty set is in ω. If $n \in \omega$, then $n^+ \in \omega$.*

Proof. The empty set is the terminal value of any finite recursion constructed on a linear ordering with only one element. If $n \in \omega$, then there exists a nonempty finite linear ordering $\langle A, \preceq \rangle$ with β, τ, σ and a recursion f constructed on $\langle A, \preceq \rangle$ starting at $\emptyset$ using the ordinal successor operation such that $f(\tau) = n$. Let τ' be any set not in A. Define $A' = A \cup \{\tau'\}$ and set

$$\preceq' = \preceq \cup \{ \langle x, \tau' \rangle \mid x \in A \} \cup \{\langle \tau', \tau' \rangle\}$$

Then $\langle A', \preceq' \rangle$ is a finite linear ordering extending $\langle A, \preceq \rangle$ by one element, with greatest element τ'. Define $f' = f \cup \{\langle \tau', n^+ \rangle\}$. Then it is clear that f' is a function on A' satisfying the appropriate recursive properties for $\langle A', \preceq' \rangle$. For suppose $a' \in A'$. If

[1] We denote by $\mathbb{N}$ the set of natural numbers used in naive mathematics. In our axiomatic treatment, we denote by ω our constructed set of naturals.

$a' = \beta$, then $f'(a') = f(\beta) = \emptyset$. If $a' \prec \tau \prec \tau'$, then $f'(\sigma'(a')) = f(\sigma(a')) = f(a')^+ = f'(a')^+$. Finally, if $\tau \preceq a' \prec \tau'$, then $a' = \tau$, so $\sigma'(a') = \tau'$ and hence

$$f'(\sigma'(a')) = f'(\tau') = n^+ = f(\tau)^+ = f'(a')^+$$

Hence f' is the unique appropriate recursion constructed on $\langle A', \preceq' \rangle$. Since $n^+ = f'(\tau')$, it follows that $n^+ \in \omega$ by construction of ω. □

In the language of closures, this theorem tells us that $\emptyset \in \omega$ and ω is closed under the successor operation. For a partial ordering $\langle A, \preceq \rangle$, we define

$$A_a = \{x \in A \mid x \preceq a\}$$

to be the segment of a in A. (Previously we used the notation S_a, but when working with multiple partial orderings we will need to distinguish the base sets.)

Definition 9.0.34. Let $\langle A, \preceq \rangle$ be a partial ordering. Let $a \in A$ and set $\preceq_a = \preceq \cap (A_a \times A_a)$. Then $\langle A_a, \preceq_a \rangle$ is called an *initial segment of* $\langle A, \preceq \rangle$.

Note that $\langle A_a, \preceq_a \rangle$ is a partial ordering, and it is a [finite] linear ordering if $\langle A, \preceq \rangle$ is a [finite] linear ordering.

Definition 9.0.35. Let $\langle A, \preceq \rangle$ and $\langle A', \preceq' \rangle$ be partial orderings. We write

$$\langle A, \preceq \rangle \cong \langle A', \preceq' \rangle$$

and say that $\langle A, \preceq \rangle$ is *isomorphic* to $\langle A', \preceq' \rangle$ iff there exists a bijection $\pi : A \to A'$ that is *order preserving*—that is, such that for all $a_1, a_2 \in A$

$$a_1 \preceq a_2 \iff \pi(a_1) \preceq' \pi(a_2)$$

We call π an *isomorphism* from $\langle A, \preceq \rangle$ to $\langle A', \preceq' \rangle$.

We prove a theorem for finite linear orderings:

Theorem 9.0.21. *Let $\langle A, \preceq \rangle$ and $\langle A', \preceq' \rangle$ be finite linear orderings. Then $\langle A, \preceq \rangle$ is isomorphic to an initial segment of $\langle A', \preceq' \rangle$ or vice versa (or both).*

Proof. If either linear ordering is empty, then we are done, so suppose both are nonempty. Let β, τ, σ be associated with $\langle A, \preceq \rangle$ and β', τ', σ' with $\langle A', \preceq' \rangle$.

We proceed by induction on $\langle A, \preceq \rangle$. Set

$$S = \{a \in A \mid (\exists a' \in A')(\exists g)(g : A_a \to A'_{a'} \text{ is an isomorphism})\}$$

Then $\beta \in S$ since $g = \{\langle \beta, \beta' \rangle\}$ is an isomorphism from A_β to $A'_{\beta'}$.

Now if for all $a \in S$, $a \prec \tau$ implies $\sigma(a) \in S$, then by induction $\tau \in S$ and hence there exists an isomorphism from $A_\tau = A$ onto an initial segment of $\langle A', \preceq' \rangle$. If this does not hold, then there exists $a \in S$ such that $a \prec \tau$ and $\sigma(a) \notin S$. Choose $a' \in A'$

and an isomorphism $g : A_a \to A'_{a'}$. Note $g(a) = a'$. We claim that $a' = \tau'$. For if not, then $a' \prec' \tau'$, so we can construct

$$g' = g \cup \{\langle \sigma(a), \sigma'(a'), \}\rangle$$

But then g' is an isomorphism from $A_{\sigma(a)}$ onto $A'_{\sigma'(a')}$, so $\sigma(a) \in S$—a contradiction. Hence $a' = \tau'$ and g^{-1} is an isomorphism from $A'_{\tau'} = A'$ onto an initial segment of $\langle A, \preceq \rangle$. This establishes our theorem. □

Suppose that π is an isomorphism from $\langle A, \preceq \rangle$ to $\langle A', \preceq' \rangle$, and further suppose that we have constructed a finite recursion f' on $\langle A', \preceq' \rangle$ starting with a set b' and using a definable operation $\mathscr{O}'$—that is, $f'(\beta') = b'$ and for all $a' \in A'$, $a' \prec' \tau'$ implies $f'(\sigma'(a')) = \mathscr{O}'(f'(a'))$. Then we can construct a finite recursion on $\langle A, \preceq \rangle$ by 'pullback' along π. More specifically, set $f = f' \circ \pi$. Then

$$f(\beta) = f'(\pi(\beta)) = f'(\beta') = b'$$

Moreover if $a \in A$ and $a \prec \tau$, then $\pi(a) \prec' \pi(\tau) = \tau'$, and

$$\begin{aligned} f(\sigma(a)) &= f'(\pi(\sigma(a))) \\ &= f'(\sigma'(\pi(a))) \\ &= \mathscr{O}'(f'(\pi(a))) \\ &= \mathscr{O}'(f(a)) \end{aligned}$$

Hence f must be the unique recursion on $\langle A, \preceq \rangle$ determined by b' and $\mathscr{O}'$.

We now return to our study of the natural numbers.

Theorem 9.0.22 (Induction on ω)**.** *Let $P(x)$ be a property and suppose*

(i) $P(\emptyset)$

(ii) For all $n \in \omega$, $P(n)$ implies $P(n^+)$

Then $P(n)$ holds for all $n \in \omega$.

Proof. Recall that ω consists of terminal values of specific finite recursions. Hence we can proceed by induction on finite linear orderings.

Let $\langle A, \preceq \rangle$ be an arbitrary nonempty finite linear ordering with β, τ, σ, and suppose that f is constructed by recursion on $\langle A, \preceq \rangle$ starting at $\emptyset$ using the successor operation. We claim that $P(f(\tau))$ holds. Set

$$S = \{ a \in A \mid P(f(a)) \}$$

Then $\beta \in S$ since $f(\beta) = \emptyset$ and $P(\emptyset)$ holds by (i). If $a \in S$ and $a \prec \tau$, then $P(f(a))$ holds by hypothesis and $f(\sigma(a)) = f(a)^+$ by construction of f. But then $P(f(\sigma(a)))$ holds since $P(f(a)^+)$ holds by (ii), so $\sigma(a) \in S$. By finite induction, $\tau \in S$, so $P(f(\tau))$ holds as desired.

Since $f(\tau)$ was an arbitrary element of ω, $P(n)$ holds for all $n \in \omega$. □

Previously we proved that $\emptyset \in \omega$ and ω is closed under the successor operation. We can now prove that ω is the smallest such set.

Corollary 9.0.5. *The set ω is the smallest set containing $\emptyset$ and closed under the successor operation.*

Proof. Suppose $\emptyset \in A$ and A is closed under the successor operation. We claim that $\omega \subseteq A$. Define

$$S = \{\, n \in \omega \mid n \in A\} = \omega \cap A$$

Then $\emptyset \in S$, and if $n \in S$, then $n^+ \in S$. Hence by induction on ω (with $P(x)$ defined as '$x \in S$') we have $S = \omega$, so $\omega \subseteq A$ as desired. □

Our construction of the naturals has a convenient property which we discuss presently. Note that the transitivity of a relation R on a set A can be redefined in terms the elements of A:

Definition 9.0.36. Let A be a set and $R \subseteq A \times A$. Then we say $a \in A$ is *transitive for R* iff

$$(\langle c, b\rangle \in R \wedge \langle b, a\rangle \in R) \implies \langle c, a\rangle \in R$$

for all $b, c \in A$.

It is immediate that R is transitive on A iff for all $a \in A$, a is transitive for R.

We extend this notion of transitivity to the class of all sets under the set membership relation:

Definition 9.0.37. Let a be a set. Then a is *transitive (for $\in$)* iff for all sets b and c,

$$c \in b \in a \implies c \in a$$

Note that a is transitive iff $b \in a$ implies $b \subseteq a$.

Transitive sets are convenient because we know what their elements look like *as sets*. If a is transitive and $b \in a$, then we know what the elements of b look like—they are simply other elements of a! In other words, the elements of a transitive set are simply built from other elements of the set.

The elements of ω, as well as ω itself, satisfy this property:

Theorem 9.0.23. *The set ω of natural numbers is transitive. Every $n \in \omega$ is transitive.*

Proof. Both claims are proved by induction on ω. For the first claim, define

$$S = \{\, n \in \omega \mid n \subseteq \omega\}$$

Note $\emptyset \in S$ trivially. If $n \in S$, then $n \cup \{n\} \subseteq \omega$ since $n \in \omega$ and $n \subseteq \omega$ by hypothesis. Hence $n^+ \in S$. By induction on ω, $S = \omega$, so ω is transitive.

To prove the second claim, set

$$S' = \{\, n \in \omega \mid n \text{ is transitive}\}$$

Trivially $\emptyset \in S$. We leave it to the reader to verify that if n is transitive, then n^+ is transitive. From this it follows that $S' = \omega$, so every $n \in \omega$ is transitive. □

We can define the successor function $\sigma : \omega \to \omega$ mapping $n \mapsto n^+$ (this is simply the successor operation restricted to ω). Note that σ is not surjective since $\emptyset \notin \sigma[\omega]$ (this is verified by induction). We claim however that σ is injective. Indeed, note that for any transitive set a, we have

$$\bigcup a^+ = \bigcup(a \cup \{a\}) = a$$

Hence if $m^+ = \sigma(m) = \sigma(n) = n^+$, then $m = \bigcup m^+ = \bigcup n^+ = n$.

This leads us to a generalization of the natural numbers and induction. Suppose that $S : A \to A$ is an injective function that is not surjective, and choose $e \in A - S[A]$. Now choose the smallest $B \subseteq A$ such that $e \in B$ and $S[B] \subseteq B$. (This can be done by defining

$$\mathscr{B} = \{X \subseteq A \mid e \in X \wedge S[X] \subseteq X\}$$

and noting that $\bigcap \mathscr{B} \in \mathscr{B}$. Set $B = \bigcap \mathscr{B}$.) Intuitively we can view B as a smallest set containing an initial element e and closed under a 'successor' operation S. It seems plausible that a version of induction will hold for B, and indeed this is true.

Let $P(x)$ be a property such that $P(e)$ holds and such that $P(a)$ implies $P(S(a))$ for all $a \in B$. Define

$$\mathscr{C} = \{a \in B \mid P(a)\}$$

Then $e \in \mathscr{C}$ by hypothesis. If $a \in \mathscr{C}$, then $P(a)$ holds, so $P(S(a))$ holds by hypothesis, and hence $S(a) \in \mathscr{C}$. Thus $S[\mathscr{C}] \subseteq \mathscr{C}$. Since B is the smallest subset of A satisfying these conditions, we have $B \subseteq \mathscr{C}$ and hence $\mathscr{C} = B$, so $P(a)$ holds for all $a \in B$.

We record this general phenomenon:

Definition 9.0.38. Let B be a set, $e \in B$, and $S : B \to B$ be an injective function with $e \notin S[B]$ such that the following condition holds:

> For all $C \subseteq B$, if $e \in C$ and $S[C] \subseteq C$, then $C = B$. (In other words, B is the only subset of B containing e and closed under S.)

Then we call $\langle B, e, S\rangle$ a *Peano system.*

Theorem 9.0.24 (Induction on peano systems)**.** *Let $\langle B, e, S\rangle$ be a Peano system and let $P(x)$ be a property such that $P(e)$ holds and such that $P(b)$ implies $P(S(b))$ for all $b \in B$. Then $P(b)$ holds for all $b \in B$.*

We use induction to prove an easy result:

Corollary 9.0.6. *Let $\langle B, e, S\rangle$ be a Peano system. Then $S(b) \neq b$ for all $b \in B$.*

Proof. Define

$$C = \{x \in B \mid S(x) \neq x\}$$

Then $e \in C$ since $e \notin S[B]$. If $x \in C$, then $S(x) \neq x$, hence $S(S(x)) \neq S(x)$ by injectivity of S. Thus $S(x) \in C$. But then $e \in C$ and C is closed under S, so $C = B$. □

We wish to define a natural linear ordering for Peano systems. In the following lemma, we show that we can naturally define a finite linear ordering on any 'initial segment' of a Peano system:

Lemma 9.0.1. *Let $\langle A, S, e\rangle$ be a Peano system and let $a \in A$. Then there exists a unique finite linear ordering $\preceq_a$ on a subset A' of A having bottom element e and top element a, and such that for all $x, y \in A$, $x \prec_a y$ iff $S(x) \preceq_a y$.*

Proof. We prove existence by induction on A.

Let C be the set of $a \in A$ for which there exists a finite linear ordering satisfying the required properties. Clearly $e \in C$, since the ordering $\preceq_e = \{\langle e, e\rangle\}$ works on $\{e\}$ (for the successor ordering property, recall that $e \notin S[A]$). Suppose now $a \in C$, and denote by $\preceq_a$ an appropriate finite linear ordering for a on A'. Note that $S(a) \notin A'$, for otherwise $S(a) \preceq_a a$ since a is the top element in A', so $a \prec_a a$ by the successor ordering property—a contradiction. Set $A'' = A' \cup \{S(a)\}$. Then A'' is finite since A' is finite by hypothesis. Define

$$\preceq_{S(a)} = \preceq_a \cup \{\langle x, S(a)\rangle \mid x \in A'\} \cup \{\langle S(a), S(a)\rangle\} \subseteq A'' \times A''$$

We claim $\preceq_{S(a)}$ is an appropriate ordering for $S(a)$ on A''. Indeed, reflexivity, antisymmetry, transitivity, and connectedness are easily verified. To illustrate, we prove antisymmetry: suppose $x, y \in A''$, $x \preceq_{S(a)} y$ and $y \preceq_{S(a)} x$. If $x, y \in A'$, then $x = y$ by antisymmetry of $\preceq_a$. If, say, $x \notin A'$, then $x = S(a)$. But since $S(a) \notin A'$, $S(a)$ relates forward only to itself under $\preceq_{S(a)}$, hence $y = S(a)$. Similarly if $y \notin A'$.

Thus we have that $\preceq_{S(a)}$ is a finite linear ordering on A''. It is clear that e is the bottom element and $S(a)$ is the top element under this ordering. We have only to verify the successor ordering property.

Suppose $x, y \in A$ and $x \prec_{S(a)} y$. We must have $x \in A'$. If $y \in A'$ then we are done, so suppose $y \notin A'$, that is, $y = S(a)$. If $x = a$, then $S(x) = S(a) \preceq_{S(a)} y$. Otherwise $x \prec_a a$, hence $S(x) \preceq_a a$ by the successor property on $\preceq_a$, so in this case too we have $S(x) \preceq_{S(a)} y$. Conversely, suppose $S(x) \preceq_{S(a)} y$ for arbitrary $x, y \in A$. Consider the case $S(x) \in A'$. If $y \in A'$, then we are done by the successor property on $\preceq_a$. If $y = S(a)$, then note that $S(x) \preceq_a a$, hence $x \prec_a a \preceq_{S(a)} y$, so $x \prec_{S(a)} y$ as desired. In the case $S(x) \notin A'$, we have $S(x) = S(a)$, hence $y = S(a)$ (since $S(a)$ relates forward only to itself), and $x = a$ by the injectivity of S. But $a \preceq_{S(a)} S(a)$, and $a \neq S(a)$ by a previous corollary, so $x \prec_{S(a)} y$ as desired.

This establishes the successor ordering property for $\preceq_{S(a)}$, completing the proof that $S(a) \in C$. It follows by induction on A that there exists an appropriate finite linear ordering for every $a \in A$.

We now prove uniqueness, also by induction on A. Let D be the set of $a \in A$ such that there is a unique appropriate finite linear ordering for a. Clearly $e \in D$. Suppose $a \in D$ and let $\preceq_a$ denote the unique appropriate linear ordering for a, on subset A'. Let $\preceq^1_{S(a)}$ and $\preceq^2_{S(a)}$ be appropriate linear orderings for $S(a)$, on subsets A^1 and A^2 respectively. We claim that $\preceq^1_{S(a)} = \preceq^2_{S(a)}$.

Note that $S(a) \preceq^1_{S(a)} S(a)$ by reflexivity, hence by successor ordering we have $a \prec^1_{S(a)} S(a)$. In particular, $a \in A^1$, and similarly $a \in A^2$. We thus define for $i = 1, 2$

$$A^i_a = \{x \in A^i \mid x \preceq^i_{S(a)} a\}$$

and $\preceq^i_a = \preceq^i_{S(a)} \cap (A^i_a \times A^i_a)$. We claim that $\langle A^i_a, \preceq^i_a\rangle$ are both appropriate finite linear orderings for a. Indeed, they are clearly both finite linear orderings with bottom element e and top element a. We need only verify successor ordering.

Fix $i \in \{1,2\}$. Let $x, y \in A$ and suppose $x \prec_a^i y$. Then $x \prec_{S(a)}^i y \preceq_{S(a)}^i a$, hence $S(x) \preceq_{S(a)}^i y \preceq_{S(a)}^i a$, so $S(x) \preceq_a^i y$. The converse follows similarly. Hence successor ordering holds for $\preceq_a^i$.

By the induction hypothesis, then, $A_a^i = A'$ and $\preceq_a^i = \preceq_a$ for $i = 1,2$. We claim that both $\preceq_{S(a)}^1$ and $\preceq_{S(a)}^2$ are obtained by adjoining the single element $S(a)$ to the end of $\langle A', \preceq_a \rangle$. From this it follows that $\preceq_{S(a)}^1 = \preceq_{S(a)}^2$ as desired.

Fix $i \in \{1,2\}$. We saw that $a \prec_{S(a)}^i S(a)$. But note that if $a \prec_{S(a)}^i b$ for $b \in A^i$, then by the successor ordering property we have $S(a) \preceq_{S(a)}^i b$. On the other hand $b \preceq_{S(a)}^i S(a)$ since $S(a)$ is the top element under this ordering. Hence $b = S(a)$ by antisymmetry. Thus $S(a)$ is the only element greater than a under $\preceq_{S(a)}^i$, so $A^i = A' \cup \{S(a)\}$, and $\preceq_{S(a)}^i$ is the claimed extension of $\preceq_a$.

This establishes the lemma. □

We can now define a linear ordering on an entire Peano system:

Theorem 9.0.25 (Ordering on peano systems)**.** *Let $\langle A, S, e \rangle$ be a Peano system. Then there exists a unique linear ordering $\langle A, \preceq \rangle$ such that $e \preceq a$ for all $a \in A$, and such that for all $a, b \in A$, $a \prec b \iff S(a) \preceq b$.*

Proof. For $a, b \in A$, say

$$a \preceq b \iff a \preceq_b b$$

We verify that this defines a linear ordering on A. Reflexivity is clear. To verify antisymmetry, suppose $a \preceq b$ and $b \preceq a$. Then $a \preceq_b b$ and $b \preceq_a a$. Note that $\preceq_b$ induces an appropriate finite linear ordering for a (see the uniqueness proof of the above lemma), hence this induced ordering equals $\preceq_a$ by the previous lemma. But then $b \preceq_b a \preceq_b b$, so $a = b$ as desired. The proof of transitivity is similar. If $a \preceq b$ and $b \preceq c$, then $a \preceq_b b$ and $b \preceq_c c$. Now $\preceq_c$ induces an appropriate finite linear ordering for b, which equals $\preceq_b$, so $a \preceq_c b \preceq_c c$, and hence $a \preceq_c c$. But then $a \preceq c$ as desired.

We verify connectedness by induction on A. Define

$$T = \{a \in A \mid (\forall b \in A)(a \preceq b \vee b \preceq a)\}$$

Clearly $e \in T$. In fact, $e \preceq a$ for all $a \in A$. Suppose $a \in T$, and let $b \in A$ be arbitrary. We claim that $S(a) \preceq b$ or $b \preceq S(a)$. If $S(a) \not\preceq b$, then $S(a) \not\preceq_b b$, so $a \not\prec_b b$ by the successor ordering property on $\prec_b$. If $a = b$, then $b \prec_{S(a)} S(a)$, so $b \preceq S(a)$. If $a \neq b$, then $b \prec a$ since $a \in S$. Hence $b \prec_a a$, so $b \prec_{S(a)} a \prec_{S(a)} S(a)$, and thus $b \preceq S(a)$. But this shows that $S(a) \in T$, hence by induction we have $T = A$ and connectedness holds.

The successor property is immediate from the same property for each $\preceq_b$. Hence we have proven existence of a linear ordering.

To verify uniqueness, suppose $\preceq$ and $\preceq'$ are both linear orderings on A satisfying the stated properties. Define

$$T' = \{a \in A \mid (\forall b \in A)(a \preceq b \iff a \preceq' b)\}$$

Clearly $e \in T'$. Suppose $a \in T'$ and let $b \in A$. We claim $S(a) \preceq b$ iff $S(a) \preceq' b$. Indeed, we have

$$\begin{array}{rcll} S(a) \preceq b & \iff & a \prec b & \text{by successor property of } \preceq \\ & \iff & a \prec' b & \text{since } a \in T' \\ & \iff & S(a) \preceq' b & \text{by successor property of } \preceq' \end{array}$$

Hence $S(a) \in T'$, so $T' = A$ by induction on A. But then for arbitrary $a, b \in A$, we have $a \preceq b$ iff $a \preceq' b$, so $\preceq = \preceq'$ as desired. □

Definition 9.0.39. Let $\langle A, S, e\rangle$ be a Peano system and let $\preceq$ be the unique linear ordering defined on A in the previous theorem. For $a \in A$, we denote by

$$A_a = \{b \in A \mid b \preceq a\}$$

the *initial segment* of A determined by a.

Note that by the proof of the previous theorem, the set A_a is finite for all $a \in A$. This allows us to use finite recursion on initial segments to establish recursion for an entire Peano system:

Theorem 9.0.26 (Recursion on peano systems (schema))**.** *Let $\langle A, S, e\rangle$ be a Peano system, b be an arbitrary set, and $\mathscr{O}$ be a definable operation (on the class of all sets). Then there exists a unique function H on A satisfying*

$$H(e) = b$$
$$H(S(a)) = \mathscr{O}(H(a))$$

for all $a \in A$.

Proof. For each $a \in A$, the set A_a (on the unique ordering $\preceq$ for A from the previous theorem) is finite, hence by recursion on finite linear orderings there exists a unique function H_a on A_a such that $H_a(e) = b$ and $H_a(S(x)) = \mathscr{O}(H_a(x))$ for all $x \prec a$. (Note that since H_a is uniquely defined for each $a \in A$, we do not require the axiom of choice to choose these.)

Set $H = \bigcup_{a\in A} H_a$. We claim that H satisfies the required properties.

First, we verify that H is a function. Suppose $\langle x, y\rangle, \langle x, y'\rangle \in H$. Then there exist $a, b \in A$ such that $\langle x, y\rangle \in H_a$ and $\langle x, y'\rangle \in H_b$. Now $a \preceq b$ or $b \preceq a$. Suppose the former case holds. Then $A_a \subseteq A_b$, and the restriction of H_b to A_a is equal to H_a by the uniqueness of finite recursions. But then $\langle x, y\rangle, \langle x, y'\rangle \in H_b$, and hence $y = y'$ since H_b is a function. Similar reasoning (using H_a and A_b) applies if $b \preceq a$. Thus H is a function as desired.

Clearly domain$(H) = A$, and $H(e) = b$. If $a \in A$, then $a \prec S(a)$, hence

$$H(S(a)) = H_{S(a)}(S(a)) = \mathscr{O}(H_{S(a)}(a)) = \mathscr{O}(H(a))$$

as desired. Hence H satisfies the required properties as claimed.

Uniqueness is immediate by induction on A. □

We note an easy corollary:

Corollary 9.0.7. *Any two Peano systems are isomorphic.*

Previously we saw that if $\langle A, S, e\rangle$ is a Peano system with ordering $\preceq$, and $a \in A$, then the initial segment A_a is finite. This allows us to give an alternate characterization of finite sets, namely as those sets in bijective correspondence with an initial segment of a Peano system:

Theorem 9.0.27. *Let $\langle B, S, e\rangle$ be a Peano system with ordering $\preceq$. Let A be a set. Then*

$$\mathrm{FIN}_A = \left\{X \subseteq A \mid X = \emptyset \vee (\exists b \in B)(\exists f)(f : B_b \to X \textit{ is a bijection})\right\}$$

Proof. Call the set on the right C.

First suppose $X \in C$. If $X = \emptyset$, then $X \in \mathrm{FIN}_A$. Assume X is nonempty. Choose $b \in B$ and $f : B_b \to X$ is a bijection. We verify by induction on B that if $c \preceq b$, then $f[B_c] \in \mathrm{FIN}_A$. Define

$$T = \{c \in B \mid c \preceq b \implies f[B_c] \in \mathrm{FIN}_A\}$$

Clearly $e \in T$. Suppose $c \in T$. If $c \prec b$, then $B_{S(c)} = B_c \cup \{S(c)\} \subseteq B_b$, hence

$$f[B_{S(c)}] = f[B_c] \cup \{f(S(c))\} \in \mathrm{FIN}_A$$

since $f[B_c] \in \mathrm{FIN}_A$ by hypothesis. If $c \not\prec b$, then $b \preceq c \prec S(c)$, hence $S(c) \not\preceq b$, so trivially $S(c) \in T$. In either case $S(c) \in T$. Hence $T = B$ by induction on B, so in particular $f[B_b] = X \in \mathrm{FIN}_A$. Thus $C \subseteq \mathrm{FIN}_A$.

To show $\mathrm{FIN}_A \subseteq C$, note that $\emptyset \in C$ by construction. Also for $a \in A$, $\{a\} \in C$ since $g = \{\langle e, a\rangle\}$ is a bijection on B_e. Let $X \in C$ be nonempty and choose an element $b \in B$ with $f : B_b \to X$ a bijection. If $a \in B$, $a \notin X$, then

$$g = f \cup \{\langle S(b), a\rangle\}$$

witnesses a bijection from $B_{S(b)}$ to $X \cup \{a\}$. Hence $X \cup \{a\} \in C$. But then C contains the empty set and is closed under the adjoining of single elements of B. Since FIN_A is the smallest set satisfying these conditions, $\mathrm{FIN}_A \subseteq C$. Thus $\mathrm{FIN}_A = C$. □

We return to our study of the axiom of infinity.

Theorem 9.0.28. *The following are equivalent:*

(1) *There exists an infinite set.*

(2) *There exists a Peano system.*

(3) *The class ω is a set.*

Proof. We provide a sketchy proof.

(1) $\implies$ (2): Let A be infinite and set $C = \mathrm{FIN}_A$. Note $A \notin C$ by hypothesis. Define a relation $E \subseteq C \times C$ by

$$E = \{\langle x, y\rangle \in C \times C \mid (\exists f)(f : x \to y \text{ is a bijection})\}$$

Note that E is an equivalence relation on C. Set $B = C/E$ and define $S : B \to B$ by

$$S = \{\langle [x], [x \cup \{a\}]\rangle \mid x \in C \wedge a \in A - x\}$$

where $[x]$ denotes the equivalence class of x. Note that domain$(S) = C$ since if $x \in C$, then $A \neq x$ (A infinite) hence $A - x$ is nonempty. It is easily verified that S is an injective function, and $[\emptyset] \neq S([x])$ for all $x \in C$. Hence $\langle C, S, e\rangle$ is a Peano system.

(2) $\Longrightarrow$ (3): If $\langle C, S, e\rangle$ is a Peano system, we do a recursion on C starting with $\emptyset$ and using ordinal successor. We obtain a function H with $H(e) = \emptyset$ and $H(S(c)) = S(c)^+$ for all $c \in C$. By induction on C, $H(c) \in \omega$ for all $c \in C$. Hence $H[B] \subseteq \omega$. By induction on ω, $\omega \subseteq H[B]$. Hence $H[B] = \omega$ is a set.

(3) $\Longrightarrow$ (1): If ω is a set, then ω is a Peano system, and hence has a linear ordering with no greatest element. It follows that ω is infinite. □

We implement arithmetical operations on the natural numbers using recursion. For addition, we desire for each $m \in \omega$ a function satisfying

$$\begin{aligned} A_m : \omega &\to \omega \\ n &\mapsto m + n \end{aligned}$$

where '+' here represents our naive addition operation. Under this naive definition, we have $A_m(0) = m + 0 = m$ and for all $n \in \omega$, if $A_m(n)$ is defined, then

$$A_m(n+1) = m + (n+1) = (m+n) + 1 = A_m(n) + 1$$

Using this as a motivation, we formally define A_m by recursion on ω by setting $A_m(0) = m$ and $A_m(n^+) = A_m(n)^+$ for all $n \in \omega$. Recall that A_m is unique for each $m \in \omega$, so we have a function $\langle\, A_m \mid m \in \omega\,\rangle$. We define addition on ω by

$$\begin{aligned} A : \omega \times \omega &\to \omega \\ \langle m, n\rangle &\mapsto A_m(n) \end{aligned}$$

By the uniqueness of each A_m, it follows that A is the unique function on $\omega \times \omega$ such that for all $m \in \omega$, $A(m, 0) = m$, and for all $m, n \in \omega$, $A(m, n^+) = A(m, n)^+$. From now on we will denote $A(m, n)$ by $m + n$.

In similar manner we can define multiplication on ω. Intuitively, for fixed $m \in \omega$, we want $m \cdot 0 = 0$ and, for all $n \in \omega$, we want $m \cdot (n+1) = m \cdot n + m$. Formally, for $m \in \omega$, define M_m by recurson with $M_m(0) = 0$ and $M_m(n^+) = M_m(n) + m = A(M_m(n), m)$ for all $n \in \omega$. Then define

$$\begin{aligned} M : \omega \times \omega &\to \omega \\ \langle m, n\rangle &\mapsto M_m(n) \end{aligned}$$

We will denote $M(m, n)$ by $m \cdot n$.

Previously we defined addition and multiplication on ω using recursion. We can prove using induction on ω that these operations satisfy all of the normal properties we expect. We prove a few of these for illustration:

Theorem 9.0.29. *Let $a, b, c \in \omega$. Then*

(i) [Associativity of addition] $(a + b) + c = a + (b + c)$

(ii) *[Commutativity of addition]* $a+b=b+a$

(iii) *[Associativity of multiplication]* $(a\cdot b)\cdot c=a\cdot(b\cdot c)$

(iv) *[Commutativity of multiplication]* $a\cdot b=b\cdot a$

(v) *[Distributivity of multiplication over addition]* $a\cdot(b+c)=a\cdot b+a\cdot c$

Proof. We prove (i) and (v) and leave the rest to the reader.

(i) Note that

$$(a+b)+0=a+b=a+(b+0)$$

So the result holds for $c=0$. Suppose the result holds for c. Then

$$\begin{aligned}(a+b)+c^+ &= ((a+b)+c)^+ && \text{by definition}\\ &= (a+(b+c))^+ && \text{by induction hypothesis}\\ &= a+(b+c)^+ && \text{by definition}\\ &= a+(b+c^+) && \text{by definition}\end{aligned}$$

Hence the result holds for c^+. By induction the result holds for all c.

(v) We have

$$a\cdot(b+0)=a\cdot b=a\cdot b+0=a\cdot b+a\cdot 0$$

so the result holds for $c=0$. If the result holds for c, then

$$\begin{aligned}a\cdot(b+c^+) &= a\cdot(b+c)^+ && \text{by definition of addition}\\ &= a\cdot(b+c)+a && \text{by definition of multiplication}\\ &= (a\cdot b+a\cdot c)+a && \text{by induction hypothesis}\\ &= a\cdot b+(a\cdot c+a) && \text{by associativity of addition}\\ &= a\cdot b+a\cdot c^+ && \text{by definition of multiplication}\end{aligned}$$

Thus the result holds for c^+. □

Note that for all $n\in\omega$, $n+0=n=0+n$. It is immediate that 0 is the only number satisfying this property, so we call 0 the *additive identity* element in ω. Define $1=0^+$. Then $n+1=n^+$ for all $n\in\omega$. In addition, $n\cdot 1=n=1\cdot n$ for all $n\in\omega$, and 1 is the only number satisfying this property. We call 1 the *multiplicative identity* element in ω.

Another important result is cancellation for addition and multiplication:

Theorem 9.0.30. *Let $a,b,c\in\omega$. Then*

(i) *If $a+b=a+c$, then $b=c$.*

(ii) *If $a\cdot b=a\cdot c$ and $a\neq 0$, then $b=c$.*

In what follows, we will assume as given the usual well ordering $<_\omega$ on ω, with least element 0 and successor corresponding to ordinal successor.[2] We will now briefly sketch the constructions of the integers ($\mathbb{Z}$) from the naturals, the rationals

[2] A detailed construction is given in [6].

($\mathbb{Q}$) from the integers, and the reals ($\mathbb{R}$) from the rationals. It is assumed that the reader has some familiarity with these sets from naive mathematics.

Intuitively, any integer can be represented using a pair a, b of naturals in the form $a-b$. Two integers $a-b$ and $c-d$ should be equal iff $a-b=c-d$, or equivalently, $a+d=b+c$. Using this as a motivation, we define the following relation on $\omega\times\omega$:

$$E=\{\langle\langle a,b\rangle,\langle c,d\rangle\rangle \mid a,b,c,d\in\omega\wedge a+d=b+c\}$$

Then E is an equivalence relation on $\omega\times\omega$. We set $\mathbb{Z}=(\omega\times\omega)/E$ and define a map $\omega\to\mathbb{Z}$ by $n\mapsto[\langle n,0\rangle]_E$. This is a natural embedding, hence we adopt the convention of identifying elements of ω with their images in $\mathbb{Z}$, so $\omega\subseteq\mathbb{Z}$.

Note that we can naturally extend the arithmetical operations on ω to $\mathbb{Z}$:

$$[\langle a,b\rangle]_E+[\langle c,d\rangle]_E=[\langle a+c,b+d\rangle]_E$$
$$[\langle a,b\rangle]_E\cdot[\langle c,d\rangle]_E=[\langle ac+bd,ad+bc\rangle]_E$$

The reader may verify that these are well defined extensions. For ordering, since intuitively $a-b<c-d$ iff $a+d<b+c$, we define

$$[\langle a,b\rangle]_E<_{\mathbb{Z}}[\langle c,d\rangle]_E\iff a+d<_\omega b+c$$

The reader may verify this is a well defined extension of the ordering on ω. Note that this is a linear ordering, but not a well ordering, since for example the subset

$$\mathbb{Z}^-=\{[\langle 0,a\rangle]_E \mid a\in\omega\}$$

of nonnegative integers has no least element.

A similar technique can be used to construct the set of rationals from the set of integers. Intuitively, any rational can be represented using a pair a, b of integers in the form a/b, where $b\neq 0$. We should have $a/b=c/d$ iff $ad=bc$. Using this as a motivation, we define

$$D=\{\langle a,b\rangle \mid a,b\in\mathbb{Z}\wedge b\neq 0\}$$

Now define the following relation on $D\times D$:

$$E=\{\langle\langle a,b\rangle,\langle c,d\rangle\rangle \mid \langle a,b\rangle,\langle c,d\rangle\in D\wedge ad=bc\}$$

Then E is an equivalence relation on D. We set $\mathbb{Q}=(D\times D)/E$ and define the natural embedding $\mathbb{Z}\to\mathbb{Q}$ by $a\mapsto[\langle a,1\rangle]_E$. We also adopt the convention $\mathbb{Z}\subseteq\mathbb{Q}$.

We can extend arithmetical operations from $\mathbb{Z}$ to $\mathbb{Q}$ in the obvious manner:

$$[\langle a,b\rangle]_E+[\langle c,d\rangle]_E=[\langle ad+bc,bd\rangle]_E$$
$$[\langle a,b\rangle]_E\cdot[\langle c,d\rangle]_E=[\langle ac,bd\rangle]_E$$

Ordering is given by

$$[\langle a,b\rangle]_E<_{\mathbb{Q}}[\langle c,d\rangle]_E\iff ad<_{\mathbb{Z}}bc$$

The reader may verify that these are all well defined extensions.

The construction of the reals from the rationals is not nearly as straightforward. Intuitively, we desire to implement a real number r as the set of all rational numbers less than r. Unfortunately, this will not do as a formal definition unless r is a rational number. Instead, we must characterize these sets of rationals *intrinsically*, without reference to an upper bound, and then form the set of all such sets. We require some preliminary definitions:

Definition 9.0.40. Let $\langle A, \preceq \rangle$ be a linear ordering. Then a *cut* in A is a set $C \subseteq A$ satisfying

(i) [Nontriviality] $C \neq \emptyset$ and $C \neq A$

(ii) [Downward closure] If $c \in C$, $a \in A$, and $a \prec c$, then $a \in C$.

Definition 9.0.41. A set $C \subseteq \mathbb{Q}$ is called a *Dedekind cut* iff C is a cut of $\langle \mathbb{Q}, \leq_{\mathbb{Q}} \rangle$ and C has no greatest element, that is, for all $p \in C$, there exists $q \in C$ with $p <_{\mathbb{Q}} q$.

Now define

$$\mathbb{R} = \{C \subseteq \mathbb{Q} \mid C \text{ is a Dedekind cut}\}$$

and define the natural embedding $\mathbb{Q} \to \mathbb{R}$ by $p \mapsto \{q \in \mathbb{Q} \mid q <_{\mathbb{Q}} p\}$. As before, we will adopt the convention that $\mathbb{Q} \subseteq \mathbb{R}$.

Ordering in $\mathbb{R}$ can be defined easily by

$$C \leq_{\mathbb{R}} D \iff C \subseteq D$$

The reader may verify this extends the ordering on $\mathbb{Q}$. Arithmetical operations on $\mathbb{R}$ are more complicated, and we will not detail their definitions here.[3] From now on, however, we will freely use them.

Example. We give a quick example using the axiom of choice in $\mathbb{R}$. Recall that for a set $A \subseteq \mathbb{R}$, a point $r \in \mathbb{R}$ is called a *limit point* of A iff for all $\epsilon > 0$, there exists $a \in A$ such that $|a - r| < \epsilon$.

Let $A \subseteq \mathbb{R}$ and $r \in \mathbb{R}$ be a limit point of A. By the axiom of choice, we can choose for each $n \in \omega$ a point $a_n \in A$ such that $|a_n - r| < 1/n$. The sequence $\langle a_n \mid n \in \omega \rangle$ converges to r.

We continue our study of recursion by noting that certain naively recursive functions are more complicated than those we have considered thus far. For example, naively we can define the factorial operation recursively as follows:

$$0! = 1 \qquad \text{and} \qquad (n+1)! = (n+1) \cdot n! \quad (n \in \omega)$$

Note that this involves $n+1$ in the definition of $(n+1)!$, so we must 'remember' where we are at each step in the recursive construction.

Another example is the sequence of Fibonacci numbers, where we set

$$F_0 = F_1 = 1 \qquad \text{and} \qquad F_{n+2} = F_{n+1} + F_n \quad (n \in \omega)$$

[3] See [6].

Here each element in the sequence is defined in terms of the previous two elements.

There is a trick that we can use to handle these cases. For the factorial operation, set $B = \omega \times \omega$ and define $F : B \to B$ by $\langle m, n\rangle \mapsto \langle m+1, (m+1)\cdot n\rangle$. Then by recursion on ω, we obtain a function $H : \omega \to B$ such that $H(0) = \langle 0, 1\rangle$ and $H(n+1) = F(H(n))$ for all $n \in \omega$. By induction on ω, $H(n) = \langle n, n!\rangle$ for all $n \in \omega$. Here we used pairs of naturals for the recursion, where the first coordinate in each pair keeps track of where we are in the factorial construction.

A similar technique can be applied to the Fibonacci numbers. Define $F' : B \to B$ by $\langle m, n\rangle \mapsto \langle n, m+n\rangle$. By recursion on ω, there exists $H' : \omega \to B$ with $H(0) = \langle 1, 1\rangle$ and $H(n+1) = F(H(n))$ for all $n \in \omega$. By induction on ω, $H(n) = \langle F_n, F_{n+1}\rangle$ for all $n \in \omega$. Here again we used pairs of naturals in the recursion, this time to keep track of two elements in the sequence at each step.

This trick can be generalized:

Theorem 9.0.31 (General recursion on ω)**.** *Let B be a set and define*

$$\mathscr{B} = \{\, k \mid (\exists n \in \omega)(k : n \to B)\,\}$$

Then for any $G : \mathscr{B} \to B$, there exists a unique function $H : \omega \to B$ such that

$$H(n) = G(H|_n)$$

for all $n \in \omega$.

Proof. Define $F : \mathscr{B} \to \mathscr{B}$ such that if $k : n \to B$, then

$$F(k) = k \cup \{\langle n, G(k)\rangle\}$$

Let $b = \emptyset$. By recursion on ω, there exists a function $h : \omega \to \mathscr{B}$ such that $h(0) = \emptyset$ and $h(n+1) = F(h(n))$ for all $n \in \omega$. Write $h_n = h(n)$. By induction on ω, $h_n : n \to B$ and $h_{n+1} \supseteq h_n$ for all $n \in \omega$. Define $H = \bigcup_{n\in\omega} h_n$.

It is easily verified that H is the desired recursion. □

We move now to the study of cardinal arithmetic. Intuitively, the cardinality of a set A is just the 'number of elements' in A. Unfortunately, this intuitive concept gets fuzzy when we consider infinite sets. We must give a precise treatment:

Definition 9.0.42. Let A and B be sets. We say A is *equinumerous* to B, and we write $A \approx B$, iff there exists a bijection $f : A \to B$.

Note that we have defined a property $P(x, y)$, where for sets a, b:

$$P(a, b) \iff a \approx b$$

It is easy to check that if P were a set, it would be an equivalence relation. But P is too big to be a set (its field is the class of all sets). However, it naturally induces equivalence relations. If A is a set, we can define E_A on $\mathscr{P}(A)$ as follows:

$$E_A = \{\,\langle x, y\rangle \mid x, y \subseteq A \wedge P(x, y)\,\}$$

Then E_A is an equivalence relation on $\mathscr{P}(A)$.

We can characterize finite sets in terms of equinumerosity:

Theorem 9.0.32. *Let A be a set. Then*

$$\mathrm{FIN}_A = \{X \subseteq A \mid (\exists n \in \omega)(n \approx X)\}$$

In particular, A is finite iff $A \approx n$ for some $n \in \omega$.

Proof. This follows from the fact that $\langle \omega, \sigma, 0\rangle$ is a Peano system and each $n \in \omega$ is an initial segment of ω. □

Definition 9.0.43. Let A and B be sets. We say that A is *dominated* by B, and write $A \preceq B$, iff there exists an injection $f : A \to B$.

Note $A \preceq B$ iff there exists $C \subseteq B$ with $A \approx C$.

Intuitively, any infinite set A should contain at least as many elements as ω. In fact, we can naively describe the following procedure for exhibiting a denumerable subset: first choose some initial $a_0 \in A$. Now supposing that $a_0, \ldots, a_n$ are distinct elements of A, we know

$$A - \{a_0, \ldots, a_n\} \neq \emptyset$$

since A is infinite by hypothesis. Hence we may choose $a_{n+1} \in A - \{a_0, \ldots, a_n\}$. Now $a_0, \ldots, a_n, a_{n+1}$ are distinct elements of A. We thus obtain a denumerable subset

$$\{a_0, \ldots, a_n, \ldots\} \subseteq A$$

Formally, we must use the axiom of choice and recursion to obtain such a subset:

Theorem 9.0.33. *Let A be an infinite set. Then $\omega \preceq A$.*

Proof. Define a relation $R \subseteq \mathrm{FIN}_A \times A$ by

$$R = \{\langle X, a\rangle \in \mathrm{FIN}_A \times A \mid a \in A - X\}$$

Note for all $X \in \mathrm{FIN}_A$, $A - X \neq \emptyset$ since A is infinite, hence domain$(R) = \mathrm{FIN}_A$. By the axiom of choice, there exists a function $f \subseteq R$ with domain$(f) =$ domain$(R) = \mathrm{FIN}_A$. (The function f chooses for each finite subset $X \subseteq A$ an element $a \in A$ outside of X.)

Define $F : \mathrm{FIN}_A \to \mathrm{FIN}_A$ by $X \mapsto X \cup \{f(X)\}$. Then by recursion on ω, there exists a function $H : \omega \to \mathrm{FIN}_A$ such that $H(0) = \emptyset$ and

$$H(n+1) = F(H(n)) = H(n) \cup \{f(H(n))\}$$

for all $n \in \omega$. Define $G : \omega \to A$ by $n \mapsto f(H(n))$. It is easy to verify (using induction on ω) that G witnesses the desired injection of ω into A, so $\omega \preceq A$. □

We desire to implement a class of objects in our theory which capture the notion of cardinality for arbitrary sets. This can be motivated by the special case of the natural numbers. We have seen that for finite sets, the natural numbers completely capture the notion of cardinality (if A is finite, then $A \approx n$ for some $n \in \omega$, and it turns out that this n is unique). Natural numbers are useful in large part because they can be used to describe finite cardinality relationships (without reference to specific sets), and because we can define arithmetical and other operations on them which allow us to conveniently calculate.

In the general case, we seek a definable operation $\mathscr{O}$ (over all sets) such that

$$\mathscr{O}(A) = \mathscr{O}(B) \iff A \approx B$$

for all sets A, B. Then $\mathscr{O}(A)$ becomes the cardinality of A. We will obtain such a definable operation later on using the axiom of choice. For now, however, we will use this operation:

Definition 9.0.44. For a set A, we denote by $\operatorname{card} A$ the *cardinality* of A. Our cardinality operation (to be implemented later) is such that for all sets A, B,

$$\operatorname{card} A = \operatorname{card} B \iff A \approx B$$

A set κ is called a *cardinal* iff $\kappa = \operatorname{card} A$ for some set A.

We may define arithmetical operations on cardinals.

Definition 9.0.45. Let κ, λ be cardinals. Then the *sum* $\kappa + \lambda$ is defined by

$$\kappa + \lambda = \operatorname{card}(A \cup B)$$

where $\kappa = \operatorname{card} A$ and $\lambda = \operatorname{card} B$ and $A \cap B = \emptyset$.

We must verify that this operation is well-defined.

Suppose $\kappa = \operatorname{card} A = \operatorname{card} A'$ and $\lambda = \operatorname{card} B = \operatorname{card} B'$, where $A \cap B = \emptyset$ and $A' \cap B' = \emptyset$. Then $A \approx A'$ and $B \approx B'$ by definition of cardinality, so we may choose bijections $f : A \to A'$ and $g : B \to B'$. Now it is immediate from the two disjointness assumptions that $(f \cup g) : A \cup B \to A' \cup B'$ is a bijection. Hence $A \cup B \approx A' \cup B'$, so $\operatorname{card}(A \cup B) = \operatorname{card}(A' \cup B')$. Thus the definition of cardinal addition is independent of the representative sets used, so it is well-defined. Note also that the definition of cardinal ensures that there exist sets A, B with $\operatorname{card} A = \kappa$ and $\operatorname{card} B = \lambda$. If A and B are not already disjoint, we may replace them with the respectively equinumerous disjoint sets $A \times \{0\}$ and $B \times \{1\}$.

We verify associativity and commutativity:

Proposition 9.0.7. *Let κ, λ, μ be cardinals. Then*

(i) $\kappa + \lambda = \lambda + \kappa$

(ii) $(\kappa + \lambda) + \mu = \kappa + (\lambda + \mu)$

Proof. Choose pairwise disjoint sets A, B, C such that $\kappa = \operatorname{card} A$, $\lambda = \operatorname{card} B$, and $\mu = \operatorname{card} C$. Then

$$\kappa + \lambda = \operatorname{card}(A \cup B) = \operatorname{card}(B \cup A) = \lambda + \kappa$$

Note also $\lambda + \mu = \operatorname{card}(B \cup C)$, hence

$$\begin{aligned} (\kappa + \lambda) + \mu &= \operatorname{card}((A \cup B) \cup C) && \text{since } (A \cup B) \cap C = \emptyset \\ &= \operatorname{card}(A \cup (B \cup C)) \\ &= \kappa + (\lambda + \mu) && \text{since } A \cap (B \cup C) = \emptyset \end{aligned}$$

□

Example. Consider the intervals $I_1 = (0,1)$ and $I_2 = (0,1]$ in $\mathbb{R}$. Write $\kappa = \operatorname{card} I_1$. Then by our definitions, $\kappa + 1 = \operatorname{card} I_2$. But $I_1 \approx I_2$, for we may define a bijection $f : I_2 \to I_1$ as follows:

$$f(x) = \begin{cases} 2^{-(n+1)} & \text{if } x = 2^{-n} \quad (n \in \omega) \\ x & \text{otherwise} \end{cases}$$

Hence $\kappa + 1 = \kappa$.

We can also define multiplication for cardinals:

Definition 9.0.46. Let κ, λ be cardinals. Then the *product* $\kappa \cdot \lambda$ is defined by

$$\kappa \cdot \lambda = \operatorname{card}(A \times B)$$

where $\kappa = \operatorname{card} A$ and $\lambda = \operatorname{card} B$.

It is easily verified that this operation is well-defined. Previously we defined cardinal

addition and multiplication. We may also define cardinal exponentiation, but we require a preliminary definition:

Definition 9.0.47. Let A and B be sets. Then we define

$${}^{A}B = \{f \mid f : A \to B\}$$

to be the set of all functions from A into B.

Suppose $A \approx m$ and $B \approx n$ for $m, n \in \omega$. A naive combinatorial argument shows that ${}^{A}B \approx n^m$. Indeed, to construct a function $f : A \to B$, we must make m choices of values, and for each choice we have n possible values. Hence there are

$$\underbrace{n \cdots n}_{m \text{ times}} = n^m$$

possible functions. This motivates the following definition:

Definition 9.0.48. Let κ, λ be cardinals. Then we define

$$\lambda^\kappa = \operatorname{card} {}^{A}B$$

where $\kappa = \operatorname{card} A$ and $\lambda = \operatorname{card} B$.

This is well-defined, for if $A \approx_g A'$ and $B \approx_h B'$, then the mapping $f \mapsto h \circ f \circ g^{-1}$ witnesses the equinumerosity ${}^{A}B \approx {}^{A'}B'$.

Proposition 9.0.8. *Let κ, λ, μ be cardinals. Then*

(i) $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$

(ii) $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$

Proof. We prove (i) and leave (ii) to the reader.

Let A, B, C be sets with cardinalities κ, λ, μ respectively. Then

$$(\kappa^{\lambda})^{\mu} = \operatorname{card}({}^{C}({}^{B}A)) \qquad \text{and} \qquad \kappa^{\lambda\cdot\mu} = \operatorname{card}({}^{B\times C}A)$$

Write $S = {}^{C}({}^{B}A)$ and $T = {}^{B\times C}A$. We must show that $S \approx T$. Recall by definition,

$$S = \{f \mid f : C \to {}^{B}A\} \qquad T = \{g \mid g : B\times C \to A\}$$

Note that for $f \in S$ and $c \in C$, $f(c) : B \to A$. For clarity we write $f(c) = f_c$. Now define $F : S \to T$ by mapping $f \mapsto g$ where $g(b,c) = f_c(b)$. Define $G : T \to S$ by mapping $g \mapsto f$, where $f(c)(b) = g(b,c)$. It is easily verified that $G \circ F = I_S$ and $F \circ G = I_T$, hence F is a bijection and $S \approx T$ as desired. □

We can define infinitary operations on cardinals as well.

Definition 9.0.49. Let $\langle \kappa_i \mid i \in I\rangle$ be a sequence of cardinals and $\langle A_i \mid i \in I\rangle$ be a sequence of pairwise disjoint sets such that $\kappa_i = \operatorname{card} A_i$ for all $i \in I$. Then we define the sum of the sequence of cardinals κ_i by

$$\sum_{i\in I} \kappa_i = \operatorname{card} \bigcup_{i\in I} A_i$$

Note that the existence of a sequence $\langle A_i \mid i \in I\rangle$ satisfying the required properties will follow from later results. (In particular, our usual technique for applying the axiom of choice will not work, since the classes involved are too big to be sets.) However, we can presently verify using the axiom of choice that the definition is independent of any such sequence used.

Let $\langle A_i \mid i \in I\rangle$ and $\langle B_i \mid i \in I\rangle$ be appropriate sequences. Then for all $i \in I$,

$$\operatorname{card} A_i = \kappa_i = \operatorname{card} B_i$$

Hence $A_i \approx B_i$ for all $i \in I$. By the axiom of choice, we may thus choose bijections $f_i : A_i \to B_i$ for all $i \in I$. Define $f = \bigcup_{i\in I} f_i$. Then it is immediate that

$$\bigcup_{i\in I} A_i \approx_f \bigcup_{i\in I} B_i$$

as desired.

We may similarly define infinitary multiplication:

Definition 9.0.50. Let $\langle \kappa_i \mid i \in I\rangle$ be a sequence of cardinals and $\langle A_i \mid i \in I\rangle$ be a sequence of sets such that $\kappa_i = \operatorname{card} A_i$ for all $i \in I$. Then we define the product

$$\prod_{i\in I} \kappa_i = \operatorname{card} \prod_{i\in I} A_i$$

where the product on the right is the generalized cartesian product.

The reader may verify that this is well-defined.

We define an ordering on cardinals:

Definition 9.0.51. Let κ, λ be cardinals. Then we say $\kappa \leq \lambda$ iff $A \preceq B$, where $\kappa = \operatorname{card} A$ and $\lambda = \operatorname{card} B$.

This is trivially well-defined.

A natural question is whether this ordering on cardinals is antisymmetric. That is, if $\kappa \leq \lambda$ and $\lambda \leq \kappa$, is $\kappa = \lambda$? Or equivalently, if $A \preceq B$ and $B \preceq A$, is $A \approx B$? This is answered in the affirmative by the following important theorem:

Theorem 9.0.34 (Cantor-Schröder-Bernstein)**.** *Let A and B be sets and suppose that $A \preceq B$ and $B \preceq A$. Then $A \approx B$.*

We will prove this theorem by analyzing the structure of an injection $h : A \to A$. First we need a preliminary construction.

Let $h : A \to A$ be an injection. For $a \in A$, define by recursion the function $g_a : \omega \to A$ such that $g_a(0) = a$ and $g_a(n+1) = h(g_a(n))$ for $n \in \omega$. Define

$$\begin{aligned} G : A \times \omega &\to A \\ \langle a, n \rangle &\mapsto g_a(n) \end{aligned}$$

Finally, for $n \in \omega$, define $G_n : A \to A$ by $a \mapsto G(a, n)$. Now $G_0 = I_A$, and $G_{n+1} = h \circ G_n$ for all $n \in \omega$. Hence we have

$$G_n = \underbrace{h \circ \cdots \circ h}_{n \text{ times}} = h^n$$

Proof of theorem. Choose $f : A \to B$ and $g : B \to A$ injections. Then

$$h = g \circ f : A \to A$$

is an injection. Define $A_0 = A - h[A]$, $A_n = h^n[A_0]$ for $n \geq 1$, and $A_\omega = A - \bigcup_{n \in \omega} A_n$. Note that for $n \in \omega$, $A_n \approx A_0$ since h^n is injective (by induction).

We claim that for distinct $m, n \in \omega$, $A_m \cap A_n = \varnothing$. Indeed, define

$$T = \{ n \in \omega \mid (\forall m < n)(A_m \cap A_n = \varnothing) \}$$

Trivially $0 \in T$. Suppose $n \in T$ and $m < n+1$. If $m = 0$, then $A_m \cap A_{n+1} = \varnothing$ since $A_{n+1} = h[A_n]$ and $A_0 \cap h[A] = \varnothing$ by construction. If $m > 0$, then $A_m = h[A_{m-1}]$, and $m - 1 < n$ by hypothesis, so $A_{m-1} \cap A_n$ is disjoint since $n \in T$. By the injectivity of h, it follows that $h[A_{m-1}] \cap h[A_n] = \varnothing$. Thus $A_m \cap A_{n+1} = \varnothing$ again, so $n+1 \in T$. By induction, $T = \omega$, from which our claim follows.

Note also that for $n \in \omega$, $A_n \cap A_\omega = \varnothing$. Since $A = (\bigcup_{n \in \omega} A_n) \cup A_\omega$, we have

$$\operatorname{card} A = \operatorname{card} A_0 \cdot \operatorname{card} \omega + \operatorname{card} A_\omega$$

Similarly we use $k = f \circ g : B \to B$ to define B_n $(n \in \omega)$ and B_ω for B, and obtain

$$\operatorname{card} B = \operatorname{card} B_0 \cdot \operatorname{card} \omega + \operatorname{card} B_\omega$$

We must show $\operatorname{card} A_0 = \operatorname{card} B_0$ and $\operatorname{card} A_\omega = \operatorname{card} B_\omega$.

We claim that the latter equinumerosity is witnessed by $f|_{A_\omega}$. Indeed, to see $f[A_\omega] \subseteq B_\omega$, verify the contrapositive by induction: if $f(a) \in B_n$ for some $n \in \omega$, then

$a \in A_m$ for some $m \in \omega$. For case $n = 0$, $f(a) \in B_0$ implies $a \notin g[B]$, which implies $a \in A_0$. Suppose that the result holds for n, and that

$$f(a) \in B_{n+1} = k[B_n] = f[g[B_n]]$$

Then $a = g(b_n)$ for some $b_n \in B_n$ by injectivity of f. If $b_n \notin f[A]$, then $a \notin h[A]$ by the injectivity of g, hence $a \in A_0$ again. On the other hand, if $b_n = f(a')$ for some $a' \in A$, then $a' \in A_m$ for some $m \in \omega$ by the induction hypothesis, hence

$$a = g(f(a')) = (g \circ f)(a') = h(a') \in h[A_m] = A_{m+1}$$

So $a \in A_{m+1}$, and the result holds for $n+1$ as desired.

To verify $f[A_\omega] \supseteq B_\omega$, suppose $b \in B_\omega$. Then in particular $b \notin B_0$, so $b \in k[B] = f[g[B]]$. Choose $b' \in B$ such that $b = f(g(b'))$. We claim that $g(b') \in A_\omega$, so $b \in f[A_\omega]$ as desired. Indeed, reasoning similar to that above shows that if $g(b') \in A_n$ for some $n \in \omega$, then $b' \in B_m$ for some $m \in \omega$. But in that case $b = k(b') \in k[B_m] = B_{m+1}$, contradicting that $b \in B_\omega$. Hence $g(b') \in A_\omega$ as claimed.

Since f is injective, we have established $\operatorname{card} A_\omega = \operatorname{card} B_\omega$.

To verify $\operatorname{card} A_0 = \operatorname{card} B_0$, note that for $a \in A_0$, we have $a \notin h[A]$, hence either $a \notin g[B]$ or $a = g(b)$ for some (unique) $b \in B$ but $b \notin f[A]$. Define

$$\begin{aligned} A_0^+ &= \{a \in A \mid a \notin g[B]\} \\ A_0^- &= \{a \in A \mid (\exists b \in B)(a = g(b) \wedge b \notin f[A])\} \end{aligned}$$

Then $A_0 = A_0^+ \cup A_0^-$ and $A_0^+ \cap A_0^- = \emptyset$. Similarly define

$$\begin{aligned} B_0^+ &= \{b \in B \mid b \notin f[A]\} \\ B_0^- &= \{b \in B \mid (\exists a \in A)(b = f(a) \wedge a \notin g[B])\} \end{aligned}$$

so $B_0 = B_0^+ \cup B_0^-$ and $B_0^+ \cap B_0^- = \emptyset$. Note now that $A_0^- = g[B_0^+]$ and $B_0^- = f[A_0^+]$, hence by the injectivity of f and g we have $\operatorname{card} A_0^+ = \operatorname{card} B_0^-$ and $\operatorname{card} A_0^- = \operatorname{card} B_0^+$. By the definition of cardinal addition, we have

$$\operatorname{card} A_0 = \operatorname{card} A_0^+ + \operatorname{card} A_0^- = \operatorname{card} B_0^- + \operatorname{card} B_0^+ = \operatorname{card} B_0$$

as desired. This completes the proof. □

Prevously we proved the Cantor-Schröder-Bernstein theorem. This theorem allows us to establish equinumerosities by exhibiting two injections, which is often easier than exhibiting bijections directly. We give some examples:

Example. Note $(0,1) \preceq [0,1]$ since $(0,1) \subseteq [0,1]$. But also

$$[0,1] \subseteq (-1,2) \approx (0,1)$$

where $x \mapsto 3x-1$ is a bijection from $(0,1) \to (-1,2)$. Hence $[0,1] \preceq (0,1)$, so by Cantor-Schröder-Bernstein we have $(0,1) \approx [0,1]$.

Example. We claim $[0,1] \approx \mathbb{R}$. Indeed, from the last example we know $[0,1] \approx (0,1)$, and it can be easily shown (using the inverse tangent function, for example) that $(0,1) \approx \mathbb{R}$. Hence $[0,1] \approx \mathbb{R}$ as claimed.

Note that $(0,1) \approx \mathbb{R}$ is reasonable since

$$(0,1) = \Big(\bigcup_{n\in\omega} (\tfrac{1}{2^{n+1}}, \tfrac{1}{2^n})\Big) \cup \{\tfrac{1}{2^n} \mid n \in \omega\} \qquad \mathbb{R} = \bigcup_{n\in\mathbb{Z}} (n, n+1) \cup \mathbb{Z}$$

and $\omega \approx \mathbb{Z}$ and $(n, n+1) \approx (1/2^{n+1}, 1/2^n)$.

Example. We claim $\omega \times \omega \approx \omega$. Clearly $\omega \preceq \omega \times \omega$ (take the mapping $n \mapsto \langle n, 0\rangle$). But also $\omega \times \omega \preceq \omega$, for we may define

$$\begin{aligned} \phi : \omega \times \omega &\to \omega \\ \langle m, n\rangle &\mapsto 2^m \cdot 3^n \end{aligned}$$

By the uniqueness of prime factorizations, ϕ is injective. Hence our claim follows from Cantor-Schröder-Bernstein.

Definition 9.0.52. A set A is said to be *countably infinite* iff $A \approx \omega$. A set A is said to be *countable* iff A is finite ($A \approx n$ for some $n \in \omega$) or A is countably infinite. A set A is said to be *uncountable* iff A is not countable.

We establish two useful theorems:

Theorem 9.0.35. *A subset of a countable set is countable.*

Proof. Let A be countable and $B \subseteq A$. If B is finite, we are done, so suppose B is infinite. Then A must be infinite, because any subset of a finite set is finite (this can be verified by induction on the finite cardinals).

In this case $A \approx \omega$, hence $B \preceq \omega$, so we may choose an injection $f : B \to \omega$. Define $g : \omega \to f[B]$ by recursion:

$$\begin{aligned} g(0) &= \text{the least } n \in f[B] \\ g(n+1) &= \text{the least } m > g(n) \text{ such that } m \in f[B] \end{aligned}$$

Note that g is well-defined since for any $n \in \omega$, there must exist $m > n$ such that $m \in f[B]$, lest $f[B]$ would be finite, contradicting that B is infinite and $f[B] \approx B$. It is easily verified by induction that $m < n$ implies $g(m) < g(n)$, hence g is injective. It follows that

$$f^{-1} \circ g : \omega \to B$$

is an injection, so $\omega \preceq B$. By Cantor-Schröder-Bernstein, $B \approx \omega$, so B is countable in this case as well. This completes the proof. □

Corollary 9.0.8. *A set A is countable iff $A \preceq \omega$.*

Theorem 9.0.36. *A countable union of countable sets is countable.*

Proof. Let A be a countable set of countable sets. We must prove $\bigcup A$ is countable.

By the previous corollary, $A \preceq \omega$, hence we may choose an injection $g : A \to \omega$. By the axiom of choice, we may similarly choose a sequence $\langle f_a \mid a \in A \rangle$ of injections $f_a : a \to \omega$ for all $a \in A$. Set $B = g[A] \subseteq \omega$. For $n \in B$, set $a_n = g^{-1}(n)$ and write $h_n = f_{a_n}$. Now for $x \in \bigcup A = \bigcup_{n \in \omega} a_n$, define

$$n_x = \text{the least } n \text{ such that } x \in a_n$$

Then it is immediate that

$$\begin{aligned} F : \bigcup A &\to \omega \times \omega \\ x &\mapsto \langle n_x, h_{n_x}(x) \rangle \end{aligned}$$

is an injection, so $\bigcup A \preceq \omega \times \omega$. But we have seen that $\omega \times \omega \approx \omega$. Hence $\bigcup A \preceq \omega$, so $\bigcup A$ is countable by the previous corollary as desired. □

We conclude by stating some equivalent versions of the axiom of choice:

Theorem 9.0.37 (Axiom of choice)**.** *The following are equivalent:*

(1) If $R \subseteq A \times B$ and $\text{domain}(R) = A$*, then there exists a function $f : A \to B$ such that $f \subseteq R$.*

(2) If $\langle A_i \mid i \in I \rangle$ is a sequence of nonempty sets, then the product

$$\prod_{i \in I} A_i = \{ f \mid f : I \to \bigcup_{i \in I} A_i \wedge (\forall i \in I)(f(i) \in A_i) \}$$

is nonempty.

(3) If A is a set, there exists a function F on $P = \mathscr{P}(A) - \{\emptyset\}$ such that for all $a \in P$, $F(a) \in a$.

(4) If A is a set (of sets) with $\emptyset \notin A$, then there exists a function $F : A \to \bigcup A$ such that for all $a \in A$, $F(a) \in a$.

(5) If $\mathscr{C}$ is a collection of nonempty pairwise disjoint sets, then there exists a set $D \subseteq \bigcup \mathscr{C}$ such that for all $C \in \mathscr{C}$, $C \cap D$ contains exactly one element.

Proof. Straightforward and left to the reader. □

Later on we will prove that Zorn's lemma is a consequence of the axiom of choice. Presently we will prove the converse.

Suppose Zorn's lemma holds and $R \subseteq A \times B$ with $\text{domain}(R) = A$. Define

$$P = \{ f \mid f \subseteq R \text{ a function} \}$$

Intuitively, P can be viewed as a set of approximations to a choice function for R. Note that P is partially ordered under inclusion. Let $C \subseteq P$ be a chain. We claim that $\bigcup C \in P$. Indeed, it is immediate that $\bigcup C \subseteq R$, and $\bigcup C$ is a function since C is a chain. By Zorn's lemma then, there exists a maximal $f \in P$.

We claim that domain$(f) = A$. Indeed, if this were not the case, we could choose $a \in A - \text{domain}(f)$. Now $a \in \text{domain}(R)$ by hypothesis, hence there exists $b \in B$ such that $\langle a, b\rangle \in R$. But then

$$f' = f \cup \{\langle a, b\rangle\}$$

is an element of P properly extending f—a contradiction. Hence domain$(f) = A$ and f is the desired choice function.

Since R was arbitrary, this establishes the axiom of choice from Zorn's lemma.

We now use Zorn's lemma to prove an important theorem:

Theorem 9.0.38 (Cardinal comparability)**.** *Let κ, λ be cardinals. Then $\kappa \leq \lambda$ or $\lambda \leq \kappa$. Equivalently, for arbitrary sets A, B, $A \preceq B$ or $B \preceq A$.*

Our goal will be to build an injection $f \subseteq A \times B$ such that one of domain$(f) = A$ or range$(f) = B$ holds. This will involve, ultimately, choosing an element $b \in B$ for each $a \in A$. But we must choose distinct $b \in B$ for distinct $a \in A$. This illustrates the difficulty of applying the axiom of choice directly. Zorn's lemma allows us to 'structure' our choices in such a way that we can construct an injection.

Proof of theorem. Define

$$P = \{f \mid f \subseteq A \times B \text{ an injection}\}$$

Then P is partially ordered under inclusion. If $C \subseteq P$ is a chain, then it is immediate that $C \subseteq A \times B$, and it can be verified that $\bigcup C$ is an injection since C is a chain of injections. Hence $\bigcup C \in P$. By Zorn's lemma, there exists a maximal $f \in C$.

We claim that domain$(f) = A$ or range$(f) = B$. Indeed, if neither of these are the case, then we can choose $a \in A - \text{domain}(f)$ and $b \in B - \text{range}(f)$. But then

$$f' = f \cup \{\langle a, b\rangle\}$$

is an element of P properly extending f—a contradiction. Hence the claim holds.

If domain$(f) = A$, then $f : A \to B$ is an injection, so $A \preceq B$. If range$(f) = B$, then $f^{-1} : B \to A$ is an injection, so $B \preceq A$. This completes the proof. □

The above proof illustrates a general principle in the application of Zorn's lemma. Intuitively, we 'wander' through elements f of a partially ordered set P, and if f is not already maximal, we keep going. In general, many choices will be required to construct f, and one additional choice will be used for a proper extension of f. The choices are made in an 'orderly' manner in order to produce the desired object.

We now establish some other important results in cardinal arithmetic.

Theorem 9.0.39. *Let A be a set. Then $A \prec \mathscr{P}(A)$ (that is, $A \preceq \mathscr{P}(A)$ but $A \not\approx \mathscr{P}(A)$).*

Proof. By cardinal comparability, it is sufficient to show that there is no surjection $F : A \to \mathscr{P}(A)$, so $\mathscr{P}(A) \not\preceq A$, so $A \prec \mathscr{P}(A)$.

Let $F : A \to \mathscr{P}(A)$. We wish to exhibit a set $X \subseteq A$ such that for all $a \in A$, $F(a) \neq X$. We do this by constructing X so as to disagree with $F(a)$ on the element a. Set

$$X = \{x \in A \mid x \not\in F(x)\}$$

Then $X \subseteq A$, and for all $a \in A$, we have $a \in X$ iff $a \notin F(a)$, hence $F(a) \neq X$. Thus $X \notin F[A]$, so F is not surjective. □

The above proof uses Cantor's 'diagonalization' method. To see why this is called the diagonalization method, note that any $F : A \to \mathscr{P}(A)$ can be represented informally with a table of the following sort:

F	a_1	a_2	a_3	$\cdots$
a_1	**Y**	N	N	$\cdots$
a_2	N	**N**	N	$\cdots$
a_3	N	Y	**Y**	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Here row i describes the set $F(a_i)$. Entry (i, j) is Y if $a_j \in F(a_i)$ and is N if $a_j \notin F(a_i)$. Intuitively, we construct X in the above proof by moving down the diagonal: say that $a_j \in X$ if entry (j, j) is N, and $a_j \notin X$ if entry (j, j) is Y. By construction, X is distinct from every $F(a_i)$, showing that F is not surjective.

Theorem 9.0.40. *Let A be a set. Then $\mathscr{P}(A) \approx {}^{A}2$.*

Proof. For each $B \subseteq A$, we define the *characteristic function* $\chi_B : A \to 2$ by

$$\chi_B(a) = \begin{cases} 1 & \text{if } a \in B \\ 0 & \text{otherwise} \end{cases}$$

We claim that the map $F : \mathscr{P}(A) \to {}^{A}2$ given by $B \mapsto \chi_B$ is a bijection.

Indeed, F is an injection, for if $\chi_B = \chi_{B'}$, then $a \in B$ iff $\chi_B(a) = 1$ iff $\chi_{B'}(a) = 1$ iff $a \in B'$, hence $B = B'$. In addition, F is surjective, for if $\chi \in {}^{A}2$, define

$$B = \{ a \in A \mid \chi(a) = 1 \}$$

Then it is immediate that $\chi_B = \chi$. □

Corollary 9.0.9. *For all cardinals κ, $\kappa < 2^\kappa$. In particular, there is no greatest cardinal.*

Proof. If $\kappa = \operatorname{card} A$, then $2^\kappa = \operatorname{card} \mathscr{P}(A)$ by the previous theorem, hence $\kappa < 2^\kappa$ by the theorem before that. □

The following proof also uses the diagonalization method:

Theorem 9.0.41 (König)**.** *Let $\langle \kappa_i \mid i \in I \rangle$ and $\langle \lambda_i \mid i \in I \rangle$ be sequences of cardinals such that $\kappa_i < \lambda_i$ for all $i \in I$. Then*

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i$$

Proof. Choose sequences $\langle A_i \mid i \in I \rangle$ and $\langle B_i \mid i \in I \rangle$ of pairwise disjoint sets such that $\operatorname{card} A_i = \kappa_i$ and $\operatorname{card} B_i = \lambda_i$ for all $i \in I$. Then

$$\sum_{i \in I} \kappa_i = \operatorname{card} \bigcup_{i \in I} A_i \quad \text{and} \quad \prod_{i \in I} \lambda_i = \operatorname{card} \prod_{i \in I} B_i$$

Write $A = \bigcup_{i \in I} A_i$ and $B = \prod_{i \in I} B_i$. Note that B is nonempty by the axiom of choice since each B_i is nonempty. We show that there is no surjection from A to B, so that $B \not\preceq A$. By cardinal comparability, it follows that $A \prec B$, establishing the theorem.

Suppose $F : A \to B$. Note that for each $i \in I$, $F[A_i] \subseteq B$. For $i \in I$, consider

$$C_i = \{ f(i) \mid f \in F[A_i] \}$$

Note that $C_i \preceq A_i$ by the axiom of choice. Thus we have $C_i \preceq A_i \prec B_i$, so $C_i \prec B_i$ by Cantor-Schröder-Bernstein. Since $C_i \subseteq B_i$, this implies $B_i - C_i \neq \emptyset$ for all $i \in I$.

Using the axiom of choice, we obtain a function g on I such that $g(i) \in B_i - C_i$ for all $i \in I$. Now $g \in B$, but

$$g \notin \bigcup_{i \in I} F[A_i]$$

since $g(i) \notin C_i$ for all $i \in I$. □

We now present some additional cardinality examples.

Example. Note that $\mathbb{Z} \times \mathbb{Z}$ is a countable union of countable sets:

$$\mathbb{Z} \times \mathbb{Z} = \bigcup_{m \in \mathbb{Z}} \mathbb{Z} \times \{m\}$$

Now we can easily define a surjection $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ by

$$\langle m, n \rangle \mapsto \begin{cases} m/n & \text{if } n \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Hence $\mathbb{Q}$ is equinumerous to a subset of a countable set, so is countable.

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be a nondecreasing function ($x \leq y$ implies $f(x) \leq f(y)$). For $r \in R$, define

$$f^-(r) = \sup\{ f(x) \mid x < r \}$$
$$f^+(r) = \inf\{ f(x) \mid r < x \}$$

We say that f *jumps* at r iff $f^-(r) < f^+(r)$. If f does not jump at r, f is continuous at r. We claim that the set

$$J = \{ r \in \mathbb{R} \mid f \text{ jumps at } r \}$$

is countable. Indeed, by the density of $\mathbb{Q}$ in $\mathbb{R}$, we can use the axiom of choice to choose for each $r \in J$ a rational $p_r \in \mathbb{Q}$ such that $f^-(r) < p_r < f^+(r)$. Now the map $J \to \mathbb{Q}$ defined by $r \mapsto p_r$ is an injection, since if $r, r' \in J$ and $r < r'$, then

$$p_r < f^+(r) \leq f^-(r') < p_{r'}$$

Hence J is countable as claimed.

We will write $\aleph_0 = \operatorname{card} \omega$. We have seen that $\aleph_0 \cdot \aleph_0 = \aleph_0$.

Example. Any $r \in R$ is a limit of a sequence $\langle q_n \mid n \in \omega \rangle$ of rationals $q_n \in \mathbb{Q}$. Define

$$C = \{\langle q_n \mid n \in \omega \rangle \mid (\forall n \in \omega)(q_n \in \mathbb{Q}) \wedge \lim_{n \to \infty} q_n \text{ exists}\}$$

Then $C \subseteq {}^{\omega}\mathbb{Q}$, so

$$\operatorname{card} C \leq \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$$

With the axiom of choice, this shows that $\operatorname{card}\mathbb{R} \leq 2^{\aleph_0}$.

On the other hand, $2^{\aleph_0} \leq \operatorname{card}\mathbb{R}$, since for any $S \subseteq \omega$, we can define

$$r_S = \sum_{n \in S} \frac{1}{10^{n+1}}$$

and $S \mapsto r_S$ is an injection from $\mathscr{P}(\omega)$ to $\mathbb{R}$.

By Cantor-Schröder-Bernstein, $\operatorname{card}\mathbb{R} = 2^{\aleph_0}$. In particular, $\aleph_0 < \operatorname{card}\mathbb{R}$.

We continue our study of cardinal arithmetic by examining how infinite cardinal arithmetic (specifically addition and multiplication) 'collapses' under the axiom of choice. We will see that for an infinite cardinal κ,

$$\kappa + \kappa = \kappa \cdot \kappa = \kappa$$

From this result it follows that if κ, μ are cardinals, at least one of which is infinite and neither of which is zero, then

$$\kappa + \mu = \kappa \cdot \mu = \max\{\kappa, \mu\}$$

First we prove a preliminary result as a 'warmup'. The techniques used in the proof will be generalized to prove the above results.

Theorem 9.0.42. *Let κ be infinite. Then $\kappa \cdot \omega = \kappa$.*

Proof. Clearly $\kappa \leq \kappa \cdot \omega$. We must show $\kappa \cdot \omega \leq \kappa$.

Let $\kappa = \operatorname{card} A$. We want to find a set B such that $B \approx A$ and $B \times \omega \preceq B$. We will use $B \subseteq A$. Define

$$P = \{f \mid f : B \times \omega \to B \text{ an injection } \wedge B \subseteq A\}$$

By Zorn's lemma, there exists a maximal $f \in P$, say

$$f : B \times \omega \to B \subseteq A$$

We claim that $B \approx A$. Note that if $C \subseteq A - B$ is countably infinite, then

$$C \times \omega \approx \omega \times \omega \approx \omega \approx C$$

Hence we could choose an injection $g : C \times \omega \to C$ and define a proper extension $f \cup g$ of f in P—contradicting the maximality of f. But then $A - B$ must be finite, since every infinite set has a countably infinite subset, and B must be infinite. Write $\operatorname{card}(A - B) = n$. Then we have

$$\operatorname{card} A = \operatorname{card} B + \operatorname{card}(A - B) = \operatorname{card} B + n = \operatorname{card} B$$

since B is infinite. This shows $B \approx A$ as claimed.

Thus $A \preceq A \times \omega$, so $\kappa \cdot \omega \leq \kappa$, and $\kappa \cdot \omega = \kappa$ as desired. □

Corollary 9.0.10. *Let κ be infinite. Then $\kappa + \kappa = \kappa$.*

Proof.

$$\kappa \leq \kappa + \kappa = \kappa \cdot 2 \leq \kappa \cdot \omega \leq \kappa$$

□

We now prove a more general version of the above result:

Theorem 9.0.43 (Cardinal absorption)**.** *Let κ be infinite. Then $\kappa \cdot \kappa = \kappa$.*

Proof. We clearly have $\kappa \leq \kappa \cdot \kappa$. We prove $\kappa \cdot \kappa \leq \kappa$. Define

$$P = \{ f \mid f : B \times B \to B \text{ an injection } \wedge B \subseteq A \}$$

Using Zorn's lemma, we obtain a maximal $f : B \times B \to B \subseteq A$ in P. Define $\mu = \operatorname{card} B$. Then we have $\mu \leq \mu \cdot \mu \leq \mu$, so $\mu \cdot \mu = \mu$. Also $\mu + \mu = \mu$. We claim $\kappa = \mu$, from which the result follows.

Suppose $C \subseteq A - B$ has cardinality μ. Then, since

$$(B \cup C) \times (B \cup C) - (B \times B) = (B \times C) \cup (C \times B) \cup (C \times C)$$

and since

$$\mu \cdot \mu + \mu \cdot \mu + \mu \cdot \mu \leq \mu$$

by the above, there exists an injection

$$g : [(B \cup C) \times (B \cup C) - (B \times B)] \to C$$

But then $f \cup g$ properly extends f in P—contradicting the maximality of f. Hence there exists no such subset C, so $\operatorname{card}(A - B) < \mu$. Thus

$$\mu \leq \kappa = \operatorname{card} A = \operatorname{card} B + \operatorname{card}(A - B) \leq \mu + \mu = \mu$$

so $\kappa = \mu$ as claimed. □

Note that the previous theorem is just a special case of this one, since

$$\kappa \leq \kappa \cdot \omega \leq \kappa \cdot \kappa = \kappa$$

Remark. We know from previous results that for all cardinals κ, $\kappa < 2^\kappa$. Therefore cardinal exponentiation does not collapse in the way addition and multiplication do. Attempts to prove that $2^\kappa \leq \kappa$ by a Zorn's lemma argument will, of course, fail. It is instructive to examine why this is so.

Our previous uses of Zorn's lemma hinged on the fact that the entry conditions for the partially ordered set P were 'finitary' in nature. Specifically, when checking closure under the union over a chain C, we only needed to work with finitely many elements of C. If we attempt to prove $2^\kappa \leq \kappa$, this is not so. For example, set

$$P = \{ f \mid f : \mathscr{P}(B) \to B \text{ an injection } \wedge B \subseteq A \}$$

where $\kappa = \operatorname{card} A$. Let $C \subseteq P$ be an arbitrary chain. To show $\bigcup C \in P$, we must show in particular that domain$(\bigcup C) = \mathscr{P}(B)$ for some $B \subseteq A$. We can show domain$(\bigcup C) \subseteq \mathscr{P}(B)$ for various $B \subseteq A$, but to establish the reverse inclusion for any such B, we must show that every subset of B (including B itself) is in the domain of some $f \in C$. This is not possible in general since B may be infinite.

To state this crudely: the chain in a Zorn's lemma proof is only useful when we are working with finitely many elements.

We denote by $\aleph_0$ the least infinite cardinal, that is, $\aleph_0 = \operatorname{card} \omega$. Similarly we denote by $\aleph_1$ the least uncountable cardinal. So far then we have

$$0, 1, 2, \ldots, n, \ldots, \aleph_0, \aleph_1, \ldots$$

For each cardinal μ, there exists a least cardinal greater than μ (this will be proved later on). We denote this cardinal by μ^+. Note that this does *not* correspond to ordinal successor in general. For all $n \in \omega$, we define $\aleph_{n+1} = \aleph_n^+$. We also define

$$\aleph_\omega = \aleph_0 + \aleph_1 + \cdots = \sum_{n\in\omega} \aleph_n$$

Note that if $\kappa > \aleph_n$ for all n, then

$$\kappa = \kappa \cdot \kappa \geq \kappa \cdot \omega = \sum_{n\in\omega} \kappa \geq \sum_{n\in\omega} \aleph_n = \aleph_\omega$$

Thus $\aleph_\omega$ is the least cardinal greater than every $\aleph_n$. Later on we will define $\aleph_{\omega+1} = \aleph_\omega^+$, and so on.

The *continuum hypothesis* states that $2^{\aleph_0} = \aleph_1$. *This hypothesis (first conjectured by Cantor) turns out to be independent of our axioms,* meaning that both it and its negation are consistent with the axioms, or equivalently that it can be neither proved nor disproved from the axioms.

More generally, our axioms do *not* refute that $2^{\aleph_0} = \aleph_n$ for some $n \in \omega$, nor do they refute that $2^{\aleph_0} = \aleph_{\omega+1}$. Interestingly, they *do* refute that $2^{\aleph_0} = \aleph_\omega$.

Theorem 9.0.44. $2^{\aleph_0} \neq \aleph_\omega$

Proof. Since $\aleph_n < \aleph_\omega$ for all $n \in \omega$, it follows from König's theorem that

$$\aleph_\omega = \sum_{n\in\omega} \aleph_n < \prod_{n\in\omega} \aleph_\omega = \aleph_\omega^{\aleph_0}$$

so $\aleph_\omega \neq \aleph_\omega^{\aleph_0}$. But we know

$$(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$$

Hence $\aleph_\omega \neq 2^{\aleph_0}$. □

We continue with cardinal exponentiation. The following results are easily verified:

Theorem 9.0.45. *Let κ, λ, μ be cardinals. Then* *(i) If $\kappa \leq \lambda$, then $\kappa^{\mu} \leq \lambda^{\mu}$.*

(ii) If $\kappa \leq \lambda$, then $\mu^{\kappa} \leq \mu^{\lambda}$ (except if $\kappa = \mu = 0$ and $\lambda > 0$).

Note that if $\kappa = \nu^{\delta}$, then by absorption

$$\kappa^{\lambda} = (\nu^{\delta})^{\lambda} = \nu^{\delta\cdot\lambda} = \nu^{\max\{\delta,\lambda\}}$$

provided one of δ, λ is infinite and neither is zero.

We prove a simple proposition:

Proposition 9.0.9. *Let κ be infinite. Then $\kappa^{\kappa} = 2^{\kappa}$.*

Proof. We have

$$2^{\kappa} \leq \kappa^{\kappa} \leq (2^{\kappa})^{\kappa} = 2^{\kappa\cdot\kappa} = 2^{\kappa}$$

□

An alternate proof uses sets more explicitly. Choose K with $\kappa = \operatorname{card} K$. Then

$$\begin{aligned} \kappa^{\kappa} = \operatorname{card} {}^{K}K &= \operatorname{card}\{f \mid f : K \to K\} \\ &\leq \operatorname{card}\{f \mid f \subseteq K \times K\} \\ &= \operatorname{card} \mathscr{P}(K \times K) \\ &= \operatorname{card} \mathscr{P}(K) = 2^{\kappa} \end{aligned}$$

Note that the cardinal arithmetic is describing the same thing, but it allows us to avoid getting our hands dirty working with specific sets.

If $\omega \leq \lambda < \kappa$, the above proposition does not tell us about κ^{λ} (unless we already know $\kappa = \nu^{\delta}$ for some $\delta \geq \lambda$, in which case $\kappa^{\lambda} = \kappa$).

In a previous lecture we proved the axiom of choice from Zorn's lemma. Presently we desire to prove Zorn's lemma as well as illustrate the general principle behind it. We will first sketch an *intuitive outline of a proof.*

Let P be partially ordered under inclusion and suppose P is closed under unions over arbitrary chains. We desire to prove that P contains a maximal element. Note that for each $p \in P$, either p is already maximal, or else there exists some $q \in P$ with $p \subseteq q$, $p \neq q$ (we will write $p \prec q$). Define

$$R = \{\langle p, q\rangle \mid (p \text{ maximal} \wedge q = p) \vee p \prec q\}$$

By the axiom of choice, there exists a function $F \subseteq R$ with domain$(F) = P$. Note that $p \preceq F(p)$ for all $p \in P$, and if p is not maximal, then $p \prec F(p)$.

Now do a recursion to obtain a function $\langle p_n \mid n \in \omega\rangle$ where $p_{n+1} = F(p_n)$. Either p_n is maximal for some $n \in \omega$, or else we have

$$p_0 \prec p_1 \prec p_2 \prec \cdots$$

Note that $C = \{p_n \mid n \in \omega\}$ is a chain, hence by hypothesis $\bigcup C \in P$. Write $p_{\omega} = \bigcup C$. Then $p_n \prec p_{\omega}$ for all $n \in \omega$, so we have

$$p_0 \prec p_1 \prec p_2 \prec \cdots \prec p_{\omega}$$

We can repeat the above process, starting with p_ω, to obtain a maximal element or else an extended chain

$$p_0 \prec p_1 \prec p_2 \prec \cdots \prec p_\omega \prec p_{\omega+1} \prec p_{\omega+2} \prec \cdots$$

Intuitively, we are getting closer to a maximal element in P. It seems that if we were able to continue repeating the above procedure for as long as necessary, we would eventually reach a maximal element.

In order to accomplish this, we must make precise the notion of 'continuing' the procedure. The idea is to transfer our notions of induction and recursion from ω to 'larger' orderings. We will define these orderings more generally, and in such a way that induction and recursion work for them. Notice that for these orderings (unlike ω), there is more than can be reached from an initial element using a successor operation (consider p_ω above). We must therefore define these orderings without reference to a successor operation. Instead we note the fact that every nonempty subset of ω has a least element, and generalize this notion:

Definition 9.0.53. Let $\langle A, \preceq \rangle$ be a linear ordering. Then $\langle A, \preceq \rangle$ is a *well ordering* iff every nonempty subset of A has a least element. (Equivalently, if $\emptyset \neq X \subseteq A$, then there exists $a \in X$ such that for all $b \in A$, $b \prec a$ implies $b \notin X$.)

The following version of induction is immediate:

Theorem 9.0.46 (Induction on well orderings)**.** *Let $\langle A, \preceq \rangle$ be a well ordering. Suppose $P(x)$ is a property satisfying:*

For all $a \in A$, if $P(b)$ holds for all $b \prec a$, then $P(a)$ holds.

Then $P(a)$ holds for all $a \in A$.

Proof. Suppose the claim is false. Define

$$B = \{a \in A \mid P(a) \text{ does not hold}\}$$

Then B is nonempty, so there exists a least element $a \in B$. But then $P(b)$ must hold for all $b \prec a$, so $P(a)$ holds by hypothesis—contradicting that $a \in B$. □

Note that *strong induction* on ω is just a special case of this theorem.

We also state recursion (to be proved later):

Theorem 9.0.47 (Recursion on well orderings (schema))**.** *Let $\langle A, \preceq \rangle$ be a well ordering and $\mathcal{O}$ be a definable operation (on the class of all sets). Then there exists a unique function H with* domain$(H) = A$ *such that for all $a \in A$,*

$$H(a) = \mathcal{O}(H|_{\{b \mid b \prec a\}})$$

Note that if $\langle A, \preceq \rangle$ is a well ordering and A is nonempty, then A has a least element. In addition, if $a \in A$ and there exists $b \in A$ such that $a \prec b$, then there exists a least $c \in A$ such that $a \prec c$.

Definition 9.0.54. Let $\langle A,\preceq\rangle$ be a well ordering. If A is nonempty, denote by β the least (or *bottom*) element of A. If $a \in A$ and there exists $b \in A$ such that $a \prec b$, then denote by $\sigma(a)$ the least such element, called the *(immediate) successor* of a.

In general, there are more elements in well orderings than can be reached from the bottom element using the successor operation.

Definition 9.0.55. Let $\langle A,\preceq\rangle$ be a well ordering with β,σ as above. Then $a \in A$ is called a *limit element* iff $a \neq \beta$ and for all $b \in A$, $\sigma(b) \neq a$.

Note that if a is an element of a well ordering, then the following cases are mutually exclusive and exhaustive: either a is the bottom element, a is a successor element, or a is a limit element.

If A is well ordered, $a \in A$, and there exist at least n distinct elements in A greater than a, then we recursively define

$$\sigma^n(a) = \underbrace{(\sigma \circ \cdots \circ \sigma)}_{n \text{ times}}(a) = \text{the } n\text{-th successor of } a$$

Specifically, set $\sigma^0(a) = a$ and $\sigma^{m+1}(a) = \sigma(\sigma^m(a))$ for $m < n$.

With limit elements, we can completely describe the structure of a well ordering:

Theorem 9.0.48. *Let $\langle A,\preceq\rangle$ be a well ordering with β,σ. For all $a \in A$, either $a = \sigma^n(\beta)$ for some $n \in \omega$, or $a = \sigma^n(b)$ for some $n \in \omega$ and some limit element $b \in A$.*

Proof. We use a minimal criminal argument. Define

$$X = \{x \in A \mid x \text{ satisfies neither condition}\}$$

If X is empty, then we are done, so suppose X is nonempty. Choose the least element $a \in X$. We know either $a = \beta$, a is a successor element, or a is a limit element.

Case $a = \beta$ is impossible, since $\beta = \sigma^0(\beta)$ satisfies the first condition.

If $a = \sigma(c)$ for some $c \in A$, then since $c \prec a$ we must have $c = \sigma^n(b)$ for some $n \in \omega$ and some $b \in A$, where either $b = \beta$ or b is a limit element. But then

$$a = \sigma(\sigma^n(b)) = \sigma^{n+1}(b)$$

contradicting that $a \in X$.

Finally, if a is a limit element, then $a = \sigma^0(a)$, again contradicting $a \in X$.

Thus X must be empty, establishing the theorem. □

We desire to prove the recursion theorem. We first establish a lemma:

Lemma 9.0.2. *Let $\mathscr{O}$ be a definable operation and suppose $\langle A,\preceq\rangle$ is a well ordering such that for all $a \in A$, there exists a unique function h such that (i)* $\operatorname{domain}(h) = \operatorname{seg} a$ *and (ii) for all $b \prec a$, $h(b) = \mathscr{O}(h|_{\operatorname{seg} b})$.*

Then there exists a unique function H such that $\operatorname{domain}(H) = A$ *and for all $a \in A$*

$$H(a) = \mathscr{O}(H|_{\operatorname{seg} a})$$

Proof. We first prove existence.

If $A = \emptyset$, set $H = \emptyset$. If A has a greatest element τ, then by hypothesis there exists a function h satisfying (i) and (ii) for τ. Define

$$H = h \cup \{\langle \tau, \mathscr{O}(h)\rangle\}$$

Then H is a function, domain$(H) = A$, and H is $\mathscr{O}$-constructed as desired.

If A has no greatest element, then for each $a \in A$ let h_a denote the unique function satisfying (i) and (ii) for a (note that the axiom of choice is not required here, but by replacement we have a function $\langle h_a \mid a \in A\rangle$). If $a, b \in A$ and $a \prec b$, then by uniqueness $h_b|_{\operatorname{seg} a} = h_a$. Thus $C = \{h_a \mid a \in A\}$ is a chain. Define

$$H = \bigcup C = \bigcup_{a \in A} h_a$$

Then H is a function, and domain$(H) = A$ since for $a \in A$, $a \in \text{domain}(h_{\sigma(a)}) \subseteq \text{domain}(H)$. Also, H is $\mathscr{O}$-constructed, since for all $a \in A$,

$$H(a) = h_{\sigma(a)}(a) = \mathscr{O}(h_{\sigma(a)}|_{\operatorname{seg} a}) = \mathscr{O}(H|_{\operatorname{seg} a})$$

We prove uniqueness by induction on A. Suppose that H and K both satisfy the above property. Define

$$S = \{a \in A \mid H(a) = K(a)\}$$

For $a \in A$, if $\operatorname{seg} a \subseteq S$, then $H|_{\operatorname{seg} a} = K|_{\operatorname{seg} a}$, hence

$$H(a) = \mathscr{O}(H|_{\operatorname{seg} a}) = \mathscr{O}(K|_{\operatorname{seg} a}) = K(a)$$

so $a \in S$. By induction, $S = A$, so $H = K$ as desired. □

Now we can prove the recursion theorem:

Proof of recursion theorem for well orderings. Let $\mathscr{O}$ and $\langle A, \preceq\rangle$ be given as in the statement of the theorem.

It is easily proved by induction on A using the preceding lemma that for all $a \in A$ there exists a unique $\mathscr{O}$-constructed function h on $\operatorname{seg} a$ (the lemma provides the induction step). But then the hypotheses of the lemma are satisfied for $\langle A, \preceq\rangle$. Hence there exists a unique $\mathscr{O}$-constructed function H on A as desired. □

We informally discuss operations on order types.

For two linear orderings A and B, we denote by $A + B$ a linear ordering obtained by placing a (disjoint) copy of B after a copy of A. We denote by $A \cdot B$ a linear ordering obtained by placing disjoint copies of A one after another, the copies indexed by the elements of B. If A and B are well orderings, it can be proven that $A + B$ and $A \cdot B$ will be well orderings.

If A is a well ordering, A might be finite. If A is infinite, then we can find an initial segment which looks like ω. If there is stuff left, we can keep going. We have either $A = \omega \cdot m + k$ for $m, k \in \omega$, or else $A = \omega \cdot \omega + B$ where B is a well ordering.

Note that if $\omega \cdot m + k$ and $\omega \cdot m' + k'$ are given, and $m' < m$, then

$$\omega \cdot m + k = (\omega \cdot m' + k') + \omega + \omega \cdot ((m-1) - m) + k$$

Similarly if $m' = m$ and $k' < k$, then

$$\omega \cdot m + k = (\omega \cdot m' + k') + (k - k')$$

This suggests that if one well ordering is smaller than another, it will be isomorphic to an initial segment of the other. We formalize this presently:

Definition 9.0.56. Let $\langle A, \preceq \rangle$ be a linear ordering. Then $B \subseteq A$ is an *initial segment* of A iff B is downward closed, that is, iff

$$(\forall b \in B)(\forall a \in A)(a \prec b \implies a \in B)$$

Definition 9.0.57. Let $\langle A, \preceq_A \rangle$ and $\langle B, \preceq_B \rangle$ be linear orderings. Then a map $H : A \to B$ is said to be *order preserving* iff

$$a \prec_A a' \iff H(a) \prec_B H(a')$$

We say H is *onto an initial segment* of B iff $H[A]$ is an initial segment of B.

Theorem 9.0.49 (Comparability of well orderings)**.** *Let $\langle A, \preceq_A \rangle$ and $\langle B, \preceq_B \rangle$ be given well orderings. Then at least one of the following holds:*

(i) *There exists $h : A \to B$ order preserving and onto an initial segment of B.*

(ii) *There exists $k : B \to A$ order preserving and onto an initial segment of A.*

Proof. Fix $e \notin B$. Define an operation $\mathscr{O}$ as follows: if f is a function with domain$(f) =$ seg a for $a \in A$, then set

$$\mathscr{O}(f) = \begin{cases} \text{the least } b \in B - \text{range}(f) & \text{if } B - \text{range}(f) \neq \emptyset \\ e & \text{otherwise} \end{cases}$$

Otherwise set $\mathscr{O}(f) = e$. By recursion we obtain a function H with domain$(H) = A$ such that for all $a \in A$,

$$H(a) = \mathscr{O}(H|_{\text{seg}\, a})$$

We claim that if $X \subseteq A$ is an initial segment of A and $H[X] \subseteq B$, then $H[X]$ is an initial segment of B. Indeed, suppose $b \in H[X]$ and $b' \prec_B b$. Then $b = H(a)$ for some $a \in X$, and since $b \in B$, we know that b is the least element in $B - H[\text{seg}\, a]$. But then $b' \in H[\text{seg}\, a]$. Since X is an initial segment of A, $\text{seg}\, a \subseteq X$, so $b' \in H[X]$. This establishes our claim.

Write $A' = H^{-1}[B]$. We claim that A' is an initial segment of A. Indeed, suppose $a' \in A'$ and $a \prec_A a'$. Then $H(a') = b$ for $b \in B - H[\text{seg}\, a'] \subseteq B - H[\text{seg}\, a]$. Hence $B - H[\text{seg}\, a] \neq \emptyset$, so $H(a) \in B$ by construction, and thus $a \in A'$. This shows that A' is an initial segment of A, so by the above, $H[A']$ is an initial segment of B.

Now if $a \prec_A a'$, then $H(a) \in H[\operatorname{seg} a']$ but $H(a') \notin H[\operatorname{seg} a']$ by construction. Thus if $H(a') \in B$, then since $H[\operatorname{seg} a']$ is an initial segment of B by the above, we cannot have $H(a') \preceq_B H(a)$, so $H(a) \prec_B H(a')$. This shows that H is order preserving on A'.

If $A' = A$, then we are done, since (i) holds with $h = H$.

If $A' \neq A$, let $a \in A$ be least such that $H(a) = e$. Then $A' = \operatorname{seg} a$ and $H[\operatorname{seg} a] \subseteq B$. We claim that $H[\operatorname{seg} a] = B$. Indeed, if this is not the case, then there exists $b \in B - H[\operatorname{seg} a]$. But then we would have $H(a) \in B$—a contradiction. Thus $H[\operatorname{seg} a] = B$, so (ii) holds with $k = (H|_{\operatorname{seg} a})^{-1}$. □

We continue our study of well orderings.

Theorem 9.0.50. *Let $\langle A, \preceq_A \rangle$ and $\langle B, \preceq_B \rangle$ be isomorphic well orderings. Then there exists a unique isomorphism $\pi : A \to B$.*

Proof. By hypothesis there exists at least one isomorphism. Suppose that f and g are both isomorphisms from A to B. We claim $f = g$. Define

$$X = \{x \in A \mid f(x) \neq g(x)\}$$

If $X \neq \emptyset$, choose the least $a \in X$. If $f(a) \prec_B g(a)$, then since g is an isomorphism we have $g^{-1}(f(a)) \prec_A a$. Since a is least in X, we must have

$$f(a) = g(g^{-1}(f(a))) = f(g^{-1}(f(a)))$$

But then $a = g^{-1}(f(a))$, so $g(a) = f(a)$, contradicting that $a \in X$. And similarly if $g(a) \prec_B f(a)$. Thus X must be empty, so $f = g$. □

Recall that for any given finite set A, there is only one linear ordering on A, up to isomorphism. Hence for a finite linear ordering $\langle A, \preceq \rangle$, there exists a unique $n \in \omega$ such that $\langle A, \preceq \rangle \cong \langle n, \leq \rangle$. This n can be seen as capturing the 'length' of $\langle A, \preceq \rangle$.

We desire to implement such objects for arbitrary well orderings. That is, we desire a class $\mathscr{C}$ of objects such that for all well orderings $\langle A, \preceq \rangle$, there exists a unique $\alpha \in \mathscr{C}$ such that $\langle A, \preceq \rangle \cong \alpha$. Our guiding idea is to use transfinite recursion on a given well ordering $\langle A, \preceq \rangle$ in order to construct an isomorphism whose image elements look like the elements in ω. The image of this isomorphism will be the desired object for $\langle A, \preceq \rangle$. We formalize this procedure now.

Fix a well ordering $\langle A, \preceq \rangle$ with β, σ. Define F on A by recursion using

$$F(a) = \begin{cases} \emptyset & \text{if } a = \beta \\ F(b) \cup \{F(b)\} & \text{if } a = \sigma(b) \text{ for } b \in A \\ \{F(b) \mid b \prec a\} & \text{if } a \text{ is a limit element} \end{cases}$$

(To obtain this recursion, define the following operation $\mathscr{O}$ on the class of all sets: if f is a function with $\operatorname{domain}(f) = \operatorname{seg} a$ for some $a \in A$, then set

$$\mathscr{O}(f) = \begin{cases} \emptyset & \text{if } a = \beta \\ f(b) \cup \{f(b)\} & \text{if } a = \sigma(b) \text{ for } b \in \operatorname{seg} a \\ \operatorname{range}(f) & \text{if } a \text{ is a limit element} \end{cases}$$

and set $\mathcal{O}(f) = \emptyset$ if f is not such a function. Then by recursion on A we obtain a function H with domain$(H) = A$ such that $H(a) = \mathcal{O}(H|_{\text{seg}\,a})$ for all $a \in A$. It is then immediate that $F = H$.)

In fact the above construction is unnecessarily complicated. We can verify by induction that for all $a \in A$, $F(a) = \text{range}(F|_{\text{seg}\,a}) = \{F(b) \mid b \prec a\}$. For $a = \beta$ this is trivial. If $a = \sigma(b)$, then by induction we have

$$F(a) = F(b) \cup \{F(b)\} = \{F(c) \mid c \prec b\} \cup \{F(b)\} = \{F(c) \mid c \prec a\}$$

Finally, if a is a limit element, then this holds by construction. Note that F is the unique map satisfying this condition, by the recursion theorem.

Write $\alpha = F[A]$. Note then that α is transitive, for

$$\begin{aligned} \beta \in \alpha &\implies \beta = F(a) \text{ for } a \in A \\ &\implies \beta = \{F(b) \mid b \prec a\} \\ &\implies \beta \subseteq \alpha \end{aligned}$$

We claim F is an isomorphism from $\langle A, \prec\rangle$ onto $\langle \alpha, \in_\alpha\rangle$. Indeed, if $a \prec b$, then we have $F(a) \in F(b)$. Thus F is (strictly) order preserving. It is verified by induction that $F(a) \notin F(a)$ for all $a \in A$, hence F is also injective. And F is surjective by construction. It follows that F is an isomorphism, so in particular α is (strictly) well ordered by the relation $\in_\alpha$.

We call α the *epsilon image* of $\langle A, \preceq\rangle$, and we call F the *epsilon image map*. We single out the properties that characterize α:

Definition 9.0.58. A set α is called an *ordinal* iff α is transitive (for $\in$) and *well ordered by epsilon*, that is, strictly well ordered by the relation

$$\in_\alpha = \{\langle \beta, \gamma\rangle \in \alpha \times \alpha \mid \beta \in \gamma\}$$

We define the class

$$\text{Ord} = \{\alpha \mid \alpha \text{ is an ordinal}\}$$

Thus we have established the following theorem:

Theorem 9.0.51. *Let $\langle A, \preceq\rangle$ be a well ordering. Then there exists $\alpha \in$ Ord such that*

$$\langle A, \preceq\rangle \cong \langle \alpha, \in_\alpha\rangle$$

We continue our study of ordinals by proving a uniqueness result:

Theorem 9.0.52. *Let $\langle A, \preceq_A\rangle$ and $\langle B, \preceq_B\rangle$ be well orderings and suppose $\langle A, \preceq_A\rangle \cong \langle B, \preceq_B\rangle$. Then $\langle A, \preceq_A\rangle$ and $\langle B, \preceq_B\rangle$ have the same epsilon image.*

Proof. Let $f : \langle A, \preceq_A\rangle \cong \langle B, \preceq_B\rangle$ be an isomorphism, and let g be the epsilon image map from $\langle B, \preceq_B\rangle$ onto its epsilon image β. Then

$$g \circ f : \langle A, \preceq_A\rangle \to \beta$$

is defined, and for $a \in A$,

$$\begin{aligned}(g \circ f)(a) = g(f(a)) &= \{g(b) \mid b \prec_B f(a)\} \\ &= \{g(f(a')) \mid a' \prec_A a\} \\ &= \{(g \circ f)(a') \mid a' \prec_A a\}\end{aligned}$$

since f is an isomorphism. By uniqueness of epsilon image maps, $g \circ f$ must be the epsilon image map for $\langle A, \preceq_A \rangle$, so in particular $\langle A, \preceq_A \rangle$ has epsilon image β. □

By our definition above, we know that every epsilon image is an ordinal. We now show that these two concepts are actually equivalent:

Proposition 9.0.10. *A set α is an ordinal iff it is an epsilon image.*

Proof. If α is an epsilon image, we know from above that it is an ordinal.

If α is an ordinal, we claim that it is its own epsilon image. Let f be its epsilon image map. We prove by induction that $f(\beta) = \beta$ for all $\beta \in \alpha$. Indeed, if this is true for $\operatorname{seg}\beta$, then since $\operatorname{seg}\beta = \beta$, we have

$$f(\beta) = \{f(\delta) \mid \delta \in \operatorname{seg}\beta\} = \{\delta \mid \delta \in \beta\} = \beta$$

Thus it is true for β. By induction, f is the identity map. □

This proposition and the previous two theorems give us:

Theorem 9.0.53. *Every well ordering is isomorphic to a unique ordinal.*

This theorem justifies the following definition:

Definition 9.0.59. Let $\langle A, \preceq \rangle$ be a well ordering. Then $\operatorname{type}\langle A, \preceq \rangle$ or $\operatorname{type} A(\preceq)$ denotes the unique ordinal isomorphic to $\langle A, \preceq \rangle$, and is called the *order type* (or *type*) of $\langle A, \preceq \rangle$.

Using the above proposition, it is straightforward to see that any element of an ordinal is an ordinal. Also any initial segment of an ordinal is an ordinal. More specifically, let α be an ordinal and suppose $\mathscr{C} \subseteq \alpha$ is an initial segment. If $\mathscr{C} \neq \alpha$, then $\mathscr{C} = \operatorname{seg}\beta$ for some $\beta \in \alpha$. But $\operatorname{seg}\beta = \beta$, so $\mathscr{C} \in \alpha$. Thus $\mathscr{C}$ is an ordinal in α.

Note also that for any ordinal α, $\alpha \notin \alpha$. Indeed, if $\alpha \in \alpha$ for some ordinal α, then $\{\alpha\}$ would be a nonempty subset of α with no least element (under $\in_\alpha$), contradicting that α is (strictly) well ordered by epsilon. This shows that if we view $\in$ as an ordering on the class of ordinals, then $\in$ is irreflexive. It is also transitive, since $\alpha \in \beta \in \gamma$ implies $\alpha \in \gamma$ by transitivity of γ. We now prove that it is trichotomous:

Theorem 9.0.54 (Comparability of ordinals)**.** *Let α, β be ordinals. Then exactly one of the following holds:*

$$\alpha \in \beta \quad or \quad \alpha = \beta \quad or \quad \beta \in \alpha$$

Proof. It is clear (from irreflexivity and transitivity) that at most one holds.

Define $\mathscr{C} = \alpha \cap \beta$. Then $\mathscr{C}$ is an initial segment of both α and β. We consider the following exclusive and exhaustive cases:

(i) If $\alpha = \mathscr{C} = \beta$, then $\alpha = \beta$.

(ii) If $\alpha = \mathscr{C} \neq \beta$, then $\alpha \in \beta$ by the above remarks.

(iii) If $\alpha \neq \mathscr{C} = \beta$, then $\beta \in \alpha$ by the above remarks.

(iv) If $\alpha \neq \mathscr{C} \neq \beta$, then we must have $\mathscr{C} \in \alpha \cap \beta$. But this is impossible by the above remarks since $\mathscr{C} = \alpha \cap \beta$ and $\mathscr{C}$ is an ordinal.

Thus in all cases, at least one of the above holds. This establishes the result. □

This theorem, together with the above remarks, shows that $\in$ forms a linear ordering on the class of all ordinals. We claim it forms a well ordering. Indeed, let $\mathscr{C}$ be a nonempty set of ordinals. Choose $\alpha \in \mathscr{C}$. If there is no $\beta \in \mathscr{C}$ such that $\beta \in \alpha$, then α is least in $\mathscr{C}$ under $\in$. On the other hand, if such an element exists, then since α is well ordered by $\in$, there exists a least element $\beta \in \alpha \cap \mathscr{C}$ under $\in$. This β is least in $\mathscr{C}$.

We mentioned above that any element of an ordinal is an ordinal, so the class Ord of all ordinals is transitive and (strictly) well ordered by $\in$. Thus if Ord were a set, then it would itself be an ordinal, so we would have Ord $\in$ Ord—a contradiction. Thus Ord must be a proper class. This result is known as the *Burali-Forti paradox*.

We leave the proofs of the following facts to the reader:

Proposition 9.0.11.

(i) If α is an ordinal, then α^+ is the least ordinal greater than α.

(ii) If $\mathscr{C}$ is a set of ordinals, then $\bigcup\mathscr{C}$ is the least ordinal greater than every element of $\mathscr{C}$ (in other words, $\bigcup\mathscr{C}$ is the least upper bound of $\mathscr{C}$ in Ord*).*

Let $\langle A, \preceq\rangle$ be a well ordering and $\alpha = \operatorname{type} A(\preceq)$. If $e \notin A$, then it is easy to verify that if we define a new ordering on $A' = A \cup \{e\}$ by

$$\preceq' = \preceq \cup \{\langle a, e\rangle \mid a \in A\} \cup \{\langle e, e\rangle\}$$

Then $\langle A', \preceq'\rangle$ is a well ordering and $\operatorname{type} A'(\preceq') = \alpha^+ = \alpha + 1$.

More generally, let $\mathscr{C}$ be a set of well orderings. Set

$$C = \{\langle a, A, \preceq_A\rangle \mid \langle A, \preceq_A\rangle \in \mathscr{C} \wedge a \in A\}$$

and define an equivalence relation E on C by

$$\langle a, A, \preceq_A\rangle \; E \; \langle b, B, \preceq_B\rangle \iff \operatorname{seg}_A a \cong \operatorname{seg}_B b$$

Now well order C/E by

$$[\langle a, A, \preceq_A\rangle] \preceq [\langle b, B, \preceq_B\rangle] \iff (\exists b' \preceq_B b)(\langle a, A, \preceq_A\rangle \; E \; \langle b', B, \preceq_B\rangle)$$

Then each $\langle A, \preceq_A\rangle \in \mathscr{C}$ is isomorphic to an initial segment of $\langle C/E, \preceq\rangle$, and

$$\operatorname{type} C/E(\preceq) = \bigcup\{\operatorname{type} A(\preceq_A) \mid \langle A, \preceq_A\rangle \in \mathscr{C}\}$$

We present an important theorem:

Theorem 9.0.55 (Hartogs)**.** *Let A be a set. Then there exists an ordinal α such that $\alpha \not\preceq A$.*

Proof. Define $B = \{\beta \in \mathrm{Ord} \mid \beta \preceq A\}$. We claim B is a set. Indeed, define

$$\mathscr{B} = \{\langle C, \preceq_C\rangle \mid C \subseteq A \wedge \langle C, \preceq_C\rangle \text{ a well orderering}\}$$

Then $\mathscr{B}$ is a set, and

$$B = \{\beta \mid (\exists\langle C, \preceq_C\rangle \in \mathscr{B})[\beta = \operatorname{type} C(\preceq_C)]\}$$

Therefore B is a set by replacement, as claimed.

Set $\alpha = \bigcup B$. Then α is an ordinal, and we claim that $\alpha \not\preceq A$. Indeed, if A is finite this is clear. If A is infinite, then $\beta \in B$ implies $\beta^+ \in B$. Hence if $\alpha \preceq A$, then $\alpha \in B$, so $\alpha \in \alpha^+ \in B$. But then $\alpha \in \alpha$—a contradiction. □

Hartogs' theorem is useful for recursions. Given a set A, choose an ordinal α such that $\alpha \not\preceq A$. Consider any map $h : \alpha \to A$. Then there must exist a least $\beta \in \alpha$ such that $h|_{\beta^+}$ is noninjective. Intuitively, this means that we will always run out of elements of A before we run out of elements of α. In other words, we are guaranteed at least enough elements in α to index all of the elements in A.

Theorem 9.0.56 (Numeration)**.** *Let A be a set. Then there exists an ordinal α such that $\alpha \approx A$.*

Proof. By Hartogs' theorem, choose β such that $\beta \not\preceq A$. Let F be a choice function for A. Fix $e \not\in A$ and define a recursion $H : \beta \to A \cup \{e\}$ by

$$H(\gamma) = \begin{cases} F(A - \{H(\delta) \mid \delta \in \gamma\}) & \text{if } \{H(\delta) \mid \delta \in \gamma\} \neq A \\ e & \text{otherwise} \end{cases}$$

Note that for all $\gamma \in \beta$, if $H|_\gamma$ is injective and $H[\gamma] \neq A$, then $H|_{\gamma^+}$ is injective by construction. Let γ be least such that $H|_\gamma$ is not injective. Then $\gamma \neq 0$, and also γ is not a limit ordinal by leastness. Hence $\gamma = \alpha^+$ for some $\alpha \in \beta$, and thus we must have $H|_\alpha : \alpha \to A$ a bijection, so $\alpha \approx A$. □

With this theorem, one is able to define the cardinal of a set to be the least ordinal equinumerous to that set. It is then verified that this definition satisfies the usual cardinal properties. (We do not go into details.)

Recursion from an ordinal into a set is essentially the heart of Zorn's lemma:

Theorem 9.0.57 (Zorn)**.** *Let $\langle P, \preceq\rangle$ be a partial ordering such that every chain in P has an upper bound. (That is, for all chains $C \subseteq P$, there exists $q \in P$ such that for all $p \in C$, $p \preceq q$.) Then there exists a maximal element in P.*

Proof. Let $\mathscr{C}$ be the set of all chains in P and obtain using the axiom of choice a function $F : \mathscr{C} \to P$ such that for all $C \in \mathscr{C}$, $F(C)$ is an upper bound of C. Similarly, define $G : P \to P$ such that for all $q \in P$, $q \preceq G(q)$, and $q \prec G(q)$ if q is not maximal.

By Hartogs' theorem, choose α such that $\alpha \not\preceq P$. Fix $e \notin P$ and recursively define $H : \alpha \to P \cup \{e\}$ by

$$H(\beta) = \begin{cases} G(F(\{H(\gamma) \mid \gamma \in \beta\})) & \text{if } \{H(\gamma) \mid \gamma \in \beta\} \text{ is a chain} \\ e & \text{otherwise} \end{cases}$$

By induction, $H[\alpha] \subseteq P$. Now for all $\beta \in \alpha$, if $H|_\beta$ is order preserving and if $F(H[\beta])$ is not maximal in P, then $H|_{\beta^+}$ is also order preserving.

Let τ be least such that $H|_\tau$ is not order preserving. Then $\tau \neq 0$ and τ is not a limit ordinal. Hence $\tau = \beta^+$ for some β. Then $H|_\beta$ is order preserving, and we must have $F(H[\beta])$ maximal in P. □

We define $\aleph_0$ to be the least infinite cardinal, and $\aleph_1$ to be the least uncountable cardinal. (Note that the existence of an uncountable cardinal, and hence of a least such cardinal, follows from Hartogs' theorem.) For all ordinals α, we have

$$\alpha < \aleph_0 \iff \alpha \text{ is finite}$$
$$\alpha < \aleph_1 \iff \alpha \text{ is countable}$$

Cantor used uncountable ordinals to analyze isolated points in the reals (this type of analysis is called *Cantor-Bendixson analysis*). We present this as an example of the use of countable ordinals.

Definition 9.0.60. Let $X \subseteq \mathbb{R}$. Then X is *closed* iff for all sequences $\langle x_n \mid n \in \omega \rangle$ in X, if $x_n \to r$ as $n \to \infty$, then $r \in X$.

Definition 9.0.61. Let $X \subseteq \mathbb{R}$. Then $x \in X$ is called an *isolated point* of X iff there exist $r, s \in \mathbb{R}$ such that $r < x < s$ and $(r, s) \cap X = \{x\}$.

Note that by the density of the rationals in the reals, we may assume that the r, s in the above definition are rational numbers, that is, $r, s \in \mathbb{Q}$.

Let $X \subseteq \mathbb{R}$ be a closed set. We desire to implement an operation which removes at most countably many elements from X, producing a closed subset $X' \subseteq X$ with no isolated points. We recursively define the following operation (on the class of all ordinals) which produces successive sets of isolated points:

$$\begin{aligned} A_0 &= \emptyset \\ A_{\alpha^+} &= A_\alpha \cup \{x \mid x \text{ an isolated point of } X - A_\alpha\} \\ A_\lambda &= \bigcup_{\alpha<\lambda} A_\alpha \end{aligned}$$

Note that $X - A_\alpha$ is closed for all α. Because the class of ordinals is unbounded (and X is not), there must exist a least ordinal α such that $A_\alpha = A_\beta$ for all $\beta \geq \alpha$. Then $X - A_\alpha$ has no isolated points.

We claim that α is countable. Indeed, for all $\beta < \alpha$, we have $A_{\beta^+} \neq A_\beta$ by leastness of α. Hence there exists $a_\beta \in A_{\beta^+} - A_\beta$. Now $a_\beta \in X - A_\beta$ is isolated. Choose $r_\beta, s_\beta \in \mathbb{Q}$ such that

$$(r_\beta, s_\beta) \cap (X - A_\beta) = \{a_\beta\}$$

Note that β^+ is least such that $(r_\beta, s_\beta) \cap (X - A_{\beta^+}) = \emptyset$, so β^+ and hence β is uniquely determined by $\langle r_\beta, s_\beta\rangle$. But then the map $\beta \mapsto \langle r_\beta, s_\beta\rangle$ is injective, so α is countable.

It can now be verified by induction that A_β is countable for all $\beta \leq \alpha$. Thus in particular A_α is countable, and we have established the following:

Theorem 9.0.58 (Cantor-Bendixon)**.** *Let $X \subseteq \mathbb{R}$ be closed. Then there exists a countable set $A \subseteq X$ such that $X - A$ is closed and has no isolated points.*

Another useful application of countable ordinals is the *Borel sets.* We define the set of Borel sets to be the smallest collection of subsets of $\mathbb{R}$ containing the open subsets of $\mathbb{R}$ and closed under countable union and intersection.

Formally, we can define the Borel sets from the top down as follows. Define

$$\begin{aligned}&\mathscr{A} = \{X \subseteq \mathbb{R} \mid X \text{ is open}\}\\&\mathscr{C} = \{\mathscr{X} \subseteq \mathscr{P}(\mathbb{R}) \mid \mathscr{A} \subseteq \mathscr{X} \wedge \mathscr{X} \text{ closed under } \omega\text{-union and } \omega\text{-intersection}\}\end{aligned}$$

Then it is easy to verify that $\mathscr{B} = \bigcap \mathscr{C} \in \mathscr{C}$. Thus $\mathscr{B}$ is the smallest set satisfying the desired properties, and we define $\mathscr{B}$ to be the Borel sets.

An alternate construction uses recursion on countable ordinals. Define

$$\begin{aligned}\mathscr{B}_0 &= \mathscr{A}\\\mathscr{B}_{\alpha^+} &= \{\bigcup_n X_n \mid X_n \in \mathscr{B}_\alpha\} \cup \{\bigcap_n X_n \mid X_n \in \mathscr{B}_\alpha\}\\\mathscr{B}_\lambda &= \bigcup_{\alpha<\lambda} \mathscr{B}_\alpha\end{aligned}$$

Note that $\mathscr{B}_{\aleph_1}$ is closed under countable unions and intersections. Indeed, suppose

$$X_n \in \mathscr{B}_{\aleph_1} = \bigcup_{\alpha<\aleph_1} \mathscr{B}_\alpha \qquad (n \in \omega)$$

For each $n \in \omega$, choose the least $\alpha_n < \aleph_1$ such that $X_n \in \mathscr{B}_{\alpha_n}$. Set $\alpha = \sup\{\alpha_n \mid n \in \omega\}$. Then $\alpha < \aleph_1$, being a countable union of countable ordinals, and $X_n \in \mathscr{B}_\alpha$ for all $n \in \omega$. But then

$$\{\bigcup_n X_n\,,\ \bigcap_n X_n\} \subseteq \mathscr{B}_{\alpha^+} \subseteq \mathscr{B}_{\aleph_1}$$

by construction.

We claim $\mathscr{B} = \mathscr{B}_{\aleph_1}$. Indeed, $\mathscr{B}_{\aleph_1}$ contains the open sets, and we just observed that it is closed under countable unions and intersections. Hence since $\mathscr{B}$ is the smallest set satisfying this property, $\mathscr{B} \subseteq \mathscr{B}_{\aleph_1}$. On the other hand, it is easily verified by induction that $\mathscr{B}_\alpha \subseteq \mathscr{B}$ for all α. In particular, $\mathscr{B}_{\aleph_1} \subseteq \mathscr{B}$, so $\mathscr{B} = \mathscr{B}_{\aleph_1}$.

This construction can be generalized. Given $\mathscr{C} \subseteq \mathscr{P}(\mathbb{R})$, define the operation

$$F(\mathscr{C}) = \{\bigcup_n X_n \mid X_n \in \mathscr{C}\} \cup \{\bigcap_n X_n \mid X_n \in \mathscr{C}\}$$

We can construct the smallest $\mathscr{D} \subseteq \mathscr{P}(\mathbb{R})$ containing $\mathscr{C}$ and closed under F. In the case of Borel sets, $\mathscr{B}$ is the smallest set containing the open sets and closed under F.

We calculate the cardinality of the set $\mathscr{B}$ of Borel sets. Recall that $\mathscr{B}_0$ denotes the set of open subsets of the reals. Then since $\mathscr{B}_0 \subseteq \mathscr{B}$, we have

$$2^{\aleph_0} = \operatorname{card} \mathscr{B}_0 \leq \operatorname{card} \mathscr{B}$$

We show by induction that $\operatorname{card} \mathscr{B}_\alpha \leq 2^{\aleph_0}$ for all $\alpha < \aleph_1$. Indeed, this is true for $\alpha = 0$ since $\operatorname{card} \mathscr{B}_0 = 2^{\aleph_0}$. If $\alpha = \beta^+$, then by the induction hypothesis we have

$$\operatorname{card} \mathscr{B}_\alpha \leq 2 \cdot \operatorname{card}({}^\omega \mathscr{B}_\beta) \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}$$

Finally, if α is a countable limit ordinal, then $\mathscr{B}_\alpha = \bigcup_{\beta<\alpha} B_\beta$, hence

$$\operatorname{card} \mathscr{B}_\alpha \leq \omega \cdot 2^{\aleph_0} = 2^{\aleph_0}$$

by the induction hypothesis. Now

$$\operatorname{card} \mathscr{B} = \operatorname{card} \mathscr{B}_{\aleph_1} \leq \aleph_1 \cdot 2^{\aleph_0} = 2^{\aleph_0}$$

since $\aleph_1 \leq 2^{\aleph_0}$. Thus by Cantor-Schröder-Bernstein, $\operatorname{card} \mathscr{B} = 2^{\aleph_0}$.

We now consider natural models of set theory. First we consider the *hereditarily countable* sets. Define HC by recursion on the ordinals as follows:

$$\begin{aligned} \mathrm{HC}_0 &= \emptyset \\ \mathrm{HC}_{\alpha^+} &= \{X \subseteq \mathrm{HC}_\alpha \mid X \text{ countable}\} \\ \mathrm{HC}_\lambda &= \bigcup_{\alpha<\lambda} \mathrm{HC}_\alpha \end{aligned}$$

By induction, HC_α is a transitive set of countable sets for all α, and $\alpha < \beta$ implies $\mathrm{HC}_\alpha \subseteq \mathrm{HC}_\beta$. We claim that $\mathrm{HC}_{\aleph_1} = \mathrm{HC}_{\aleph_1+1}$. Indeed, we know $\mathrm{HC}_{\aleph_1} \subseteq \mathrm{HC}_{\aleph_1+1}$. Now suppose $X \in \mathrm{HC}_{\aleph_1+1}$. Then X is countable and

$$X \subseteq \mathrm{HC}_{\aleph_1} = \bigcup_{\alpha<\aleph_1} \mathrm{HC}_\alpha$$

For each $x \in X$, choose $\alpha_x < \aleph_1$ least such that $x \in \mathrm{HC}_{\alpha_x}$ Set $\alpha = \sup\{\alpha_x \mid x \in X\}$. Then $X \subseteq \mathrm{HC}_\alpha$. But $\alpha < \aleph_1$, hence $X \in \mathrm{HC}_{\alpha+1} \subseteq \mathrm{HC}_{\aleph_1}$. This shows $\mathrm{HC}_{\aleph_1} = \mathrm{HC}_{\aleph_1+1}$.

Define $\mathrm{HC} = \mathrm{HC}_{\aleph_1}$. Then HC is hereditarily countable. Also HC is the smallest transitive set such that for all $X \subseteq \mathrm{HC}$ countable, $X \in \mathrm{HC}$. Indeed, if A is another set with this property, then $\mathrm{HC}_\alpha \subseteq A$ for all α by induction, so $\mathrm{HC} \subseteq A$.

Theorem 9.0.59. *The set* HC *is a model of all of our axioms except the powerset axiom.*

Remark. Formally, this theorem means that the *relativizations* of our axioms (except some replacement axioms) to the set HC are provable from our axioms. If ϕ is one of our axioms, then the relativization ϕ^{HC} of ϕ is obtained by replacing all quantifiers in ϕ of the form $(\forall x)$ and $(\exists x)$ with $(\forall x \in \mathrm{HC})$ and $(\exists x \in \mathrm{HC})$, respectively. We do not go into details on the notions of relativizations and models.

We define the cumulative hierarchy of sets by recursion on the ordinals:

$$\begin{aligned} V_0 &= \emptyset \\ V_{\alpha^+} &= \mathscr{P}(V_\alpha) \\ V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha \end{aligned}$$

By induction, V_α is transitive for all α, and $\alpha < \beta$ implies $V_\alpha \subseteq V_\beta$. Note that V_ω is the set of all hereditarily finite sets.

Theorem 9.0.60. *The set $V_{\omega+\omega}$ is a model of all of our axioms except the replacement axiom.*

Definition 9.0.62. Let κ be a cardinal. Then κ is *strongly inaccessible* iff

(i) κ is uncountable

(ii) For all cardinals $\lambda < \kappa$, $2^\lambda < \kappa$.

(iii) For all ordinals $\alpha < \kappa$, if $f : \alpha \to \kappa$, then $\sup\{f(\beta) \mid \beta \in \alpha\} < \kappa$.

Theorem 9.0.61. *Let κ be an inaccessible cardinal. Then V_κ is a model of all of our axioms.*

Finally, define

$$V = \bigcup_\alpha V_\alpha$$

Theorem 9.0.62. *The class V is a model of all of our axioms.*

Bibliography

[1] Sheldon Axler. *Linear Algebra Done Right.* Springer, 2nd edition, 1997.

[2] James W. Brown and Ruel V. Churchill. *Complex Variables and Applications.* McGraw Hill, 7th edition, 2004.

[3] N. J. Cutland. *Computability: An introduction to recursive function theory.* Cambridge, 1980.

[4] David S. Dummit and Richard M. Foote. *Abstract Algebra.* Wiley, 3rd edition, 2003.

[5] H.-D. Ebbinghaus, J. Flum, and W. Thomas. *Mathematical Logic.* Springer, 2nd edition, 1994.

[6] Herbert B. Enderton. *Elements of Set Theory.* Academic Press, 1977.

[7] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra.* Prentice Hall, 4th edition, 2003.

[8] Walter Rudin. *Principles of Mathematical Analysis.* McGraw Hill, 3rd edition, 1976.

[9] Donald Sarason. *Notes on Complex Function Theory.* Helson, 1994.