Learning Uni Maths

gispisquared

If only I had the theorems! Then I should find the proofs easily enough.

Bernhard Riemann

Contents

CHAPTER 1

Set Theory

AXIOM 1 (Existence). *There exists a set.*

REMARK 2. This is implied by the Axiom of Infinity; however, we include it here so that we may define the empty set which is included in the statement of that axiom.

DEFINITION 3. Let A and B be sets. If every element of A is an element of B, we say that A is a *subset* of B, denoted $A \subseteq B$.

PROPOSITION 1. *If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.*

AXIOM 4 (Extensionality). *$A = B$ iff $A \subseteq B$ and $B \subseteq A$.*

DEFINITION 5. A *sentence* is made by combining assertions of belonging (eg $x \in A$) and/or assertions of equality (eg $A = B$) using the usual logical operators: *and, or, not, implies, if and only if, there exists, for all.*

AXIOM 6 (Specification). *For every set A, every set p and every sentence $S(x,p)$ there is a set B whose elements are exactly those elements x of A for which $S(x,p)$ holds.*

DEFINITION 7. We notate this set B by $\{x \in A : S(x,p)\}$.

PROPOSITION 2. *There exists a unique set X such that for any x, the sentence $x \in X$ is false.*

DEFINITION 8. We call this set the *empty set*, notated $\emptyset$.

PROPOSITION 3. *For every set A there is a set B such that $B \notin A$.*

AXIOM 9 (Pairing). *For any two sets A and B there is a set X with $A \in X$ and $B \in X$.*

PROPOSITION 4. *There is a unique set Y such that for any a, a is in Y iff $a = A$ or $a = B$.*

DEFINITION 10. This set is called the *unordered pair* formed by A and B, denoted $\{A, B\}$.

DEFINITION 11. The set $\{A, A\}$ is denoted $\{A\}$, and called the *singleton* of A.

REMARK 12. When speaking of sets of sets, we sometimes call them *collections* — this is just another name for a set.

AXIOM 13 (Union). *For any collection X of sets there exists a set Y such that for any A in X, and any a in A, a is in Y.*

PROPOSITION 5. *For a nonempty collection X of sets there is a unique set Z such that a is in Z if and only if there exists an A in X such that a is in A.*

DEFINITION 14. This set is called the *union* of X, denoted $\bigcup X$.

For two sets A and B we define $A \cup B = \bigcup\{A, B\}$.

PROPOSITION 6. *For every nonempty collection C of sets, there is a unique set Y such that $x \in Y$ iff $x \in X$ for each X in C.*

DEFINITION 15. This set Y is called the *intersection* of C, denoted $\bigcap C$.

DEFINITION 16. Let A and B be sets. The *intersection* of A and B, notated $A \cap B$, is $\bigcap\{A, B\}$.

If $A \cap B = \emptyset$ then A and B are called *disjoint*.

AXIOM 17 (Powers). *For each set X there is a collection that contains all subsets of X.*

PROPOSITION 7. *There is a unique collection Y such that $x \in Y$ iff $x \subseteq X$.*

DEFINITION 18. This set Y is called the *power set* of X, denoted $\mathcal{P}(X)$.

DEFINITION 19. The *ordered pair* of a and b is the set defined as

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

PROPOSITION 8. *For any a, b, c, d, we have $(a, b) = (c, d)$ iff $a = c$ and $b = d$.*

PROPOSITION 9. *For any sets A and B, the set*

$$\{(x, y) : x \in A,\ y \in B\}$$

exists.

DEFINITION 20. This set is called the *Cartesian product* of A and B, denoted $A \times B$.

PROPOSITION 10. *For any set R of ordered pairs there are sets A and B such that $R \subseteq A \times B$.*

DEFINITION 21. A *binary relation* R from A to B is a subset of $A \times B$. If (a, b) is in R we write aRb.

If $A = B$ then we call it a *binary relation over* A.

DEFINITION 22. An *equivalence relation* is a binary relation $\sim$ over A such that

- $a \sim a$ (reflexive),
- $a \sim b \iff b \sim a$ (symmetric), and
- if $a \sim b$ and $b \sim c$ then $a \sim c$ (transitive).

The *equivalence class* of a under $\sim$ is

$$[a] = \{x \in A : x \sim a\}.$$

DEFINITION 23. A *partition* of a set A is a disjoint collection of nonempty subsets of A whose union is A.

A partition X of A *induces* a relation A/X, where $a\ A/X\ b$ iff a and b belong to the same element of X.

PROPOSITION 11. *The collection of equivalence classes of an equivalence relation exists and is a partition.*

DEFINITION 24. This partition is called the partition *induced* by the equivalence relation $\sim$, denoted $X/\sim$.

Proposition 12. *The equivalence relation induced by a partition induces that partition; the partition induced by an equivalence relation induces that relation.*

Definition 25. For a relation R from X to Y we define the *inverse* relation $R^{-1} : Y \to X$ by $xRy \iff yR^{-1}x$.

Definition 26. A *function* $f : A \to B$ is a relation f over A and B such that for each $a \in A$ there is exactly one $b \in B$ such that afb. We usually write this as $f(a) = b$.

Definition 27. For a set $E \subseteq A$, we define the *image* of E under f as $f(E) = \{f(x) : x \in E\}$. For a set $E \subseteq B$, we define the *inverse image* of E under F as $f^{-1}(E) = \{x \in A : f(x) \in E\}$.

Definition 28. A function f is *injective* if for each b in B, there is at most one a in A such that $f(a) = b$. It is *surjective* if for each b in B there is at least one a in A such that $f(a) = b$. A function which is both injective and surjective is *bijective*.

Definition 29. For functions $f : W \to X$ and $g : Y \to Z$, where $Y \subseteq X$, we define the *composite* $f \circ g : W \to Z$ as $(f \circ g)(x) = f(g(x))$ for all x.

Definition 30. A function x from a set I (the *index set*) to a set X is called an *indexed family* of X, and its range is an *indexed set*. We notate the indexed set by $\{x_i\}_{i \in I}$.

Definition 31. The set of families of a set X indexed by a set I is X^I.

Definition 32. For any set X we define $X^+ = X \cup \{X\}$.

Axiom 33 (Infinity). *There exists a set S containing $\emptyset$ and containing X^+ for every X in S.*

Proposition 13 (Peano Axioms). *There exists a unique set ω satisfying*

- $\emptyset \in \omega$.
- *If $n \in \omega$ then $n^+ \in \omega$.*
- *If $S \subseteq \omega$ such that $\emptyset \in S$ and $n \in S \implies n^+ \in S$ then $S = \omega$.*
- *$n^+ \neq 0$ for all $n \in \omega$.*
- *If n and m are in ω, and if $n^+ = m^+$, then $n = m$.*

Theorem 14 (Recursion). *If a is an element of a set X, and if $f : X \to X$ is a function, then there is a function $g : \omega \to X$ such that $u(0) = a$ and $u(n^+) = f(u(n))$ for all n in ω.*

Definition 34. A *partial order* is a binary relation $\leq$ on a a set A such that

- $a \leq a$ (reflexive),
- if $a \leq b$ and $b \leq a$ then $a = b$ (antisymmetric), and
- if $a \leq b$ and $b \leq c$ then $a \leq c$ (transitive).

We define $a < b$ if $a \leq b$ and $a \neq b$.

If for all a and b we have $a \leq b$ or $b \leq a$ (strongly connected), then $\leq$ is a *total order*.

A *chain* is a totally ordered subset of a partially ordered set.

Definition 35. If X is a partially ordered set, and if $a \in X$, the set $s(a) = \{x \in X : x < a\}$ is called the *initial segment* determined by a.

Definition 36. Two partially ordered sets X and Y are *similar* if there is a bijection $f : X \to Y$ such that $a \leq b \iff f(a) \leq f(b)$. This bijection is called a *similarity*.

Definition 37. Let S be a subset of a partially ordered set A, and let a be an element of A. If $s \leq a$ for every s in S, then we call a an *upper bound* of S. If $a \leq s$ for every s in S, then we call a a *lower bound* of S. If a is an upper bound of S and a lower bound of the set of upper bounds of S, then we call a a *least upper bound* of S.

Definition 38. A *well-order* on A is a total order $\leq$ on A such that every nonempty subset S of A has an element a which is a lower bound for S. The set A together with the relation $\leq$ is then called *well-ordered*.

Theorem 15 (Transfinite Induction). *Let S be a subset of a well-ordered set A such that for any $x \in A$, if $s(x) \subseteq S$ then $x \in S$. Then $S = A$.*

Definition 39. If a is an element of a well-ordered set A, and X is an arbitrary set, then a *sequence of type a* is an family of X indexed by $s(a)$.

A *sequence function* of type A is a function whose domain consists of all sequences of type a for each $a \in A$, and whose codomain is A.

Proposition 16 (Transfinite Recursion). *If A is a well-ordered set, and if f is a sequence function of type A in X, then there is a unique function $U : A \to X$ such that $U(a) = f(U|s(a))$ for each a in W.*

Proposition 17. *If two well-ordered sets are similar, then the similarity is unique.*

Theorem 18. *If X and Y are well-ordered, then either X and Y are similar, or one is similar to an initial segment of the other.*

Definition 40. An *ordinal number* is a well-ordered set α such that for any $\xi \in \alpha$ we have $s(\xi) = \xi$.

We define the ordinals $0 = \emptyset$ and $1 = 0^+$.

Proposition 19. *There is no set of all ordinal numbers.*

Proposition 20. *ω is an ordinal number.*

Proposition 21. *If α is an ordinal number then so is α^+, and so is any element of α.*

Theorem 22. *If two ordinal numbers are similar, then they are equal.*
Otherwise, one is an element of the other.

Axiom 41 (Substitution). *If p is a set and $S(a, b, p)$ is a sentence such that for each a in a set A there exists a set B_a such that $b \in B_a \iff S(a, b, p)$, then there exists a function F with domain A such that $F(a) = B_a$ for each a in A.*

Axiom 42 (Foundation). *Every set X contains a set Y such that X and Y are disjoint.*

Axiom 43 (Choice). *Let X be a collection of sets whose members are all nonempty. Then there exists a function $f : X \to \bigcup X$ such that $f(Y) \in Y$ for all $Y \in X$.*

Proposition 23. *Every relation includes a function with the same domain.*

Theorem 24 (Zorn's Lemma). *Suppose a partially ordered set P has the property that every chain in P has an upper bound in P. Then there is an element $a \in P$ such that the only upper bound for $\{a\}$ is a.*

Theorem 25 (Well-Ordering Theorem). *Every set has a well-ordering.*

Proposition 26. *Every well-ordered set is similar to a unique ordinal number.*

Proposition 27. *If a and b are ordinals, let $A = \{(x, 0) : x \in a\}$ and $B = \{(y, 1) : y \in b\}$, retaining the associated orders $\leq_A$ and $\leq_B$. Then the set $A \cup B$ is well-ordered by $\leq_A \cup \leq_B \cup (A \times B)$.*

Definition 44. The ordinal corresponding to $A \cup B$ under this well-ordering is the *ordinal sum* of a and b, denoted $a + b$.

Proposition 28. *If A and B are ordinals, the ordering on $A \times B$ where $(a, b) < (c, d)$ if either $b < d$ or $b = d$ and $a < c$ is a well-ordering on $A \times B$.*

Definition 45. The ordinal corresponding to $A \times B$ under this well-ordering is the *ordinal product* of A and B, denoted AB or $A \cdot B$.

Proposition 29. *For every pair of ordinals a, b there exists an ordinal c and a unique function $f_b : a^+ \to c$ such that such that $f_b(\emptyset) = 1$ and*

$$f_b(x) = \begin{cases} f_b(\bigcup x)x & \bigcup x \neq x \\ \bigcup_{y \in x} f_b(y) & \bigcup x = x \end{cases}.$$

Definition 46. We define $a^b = f_b(a)$.

Proposition 30. *With ordinal sums, products and exponents as defined,*

$$\begin{aligned} a + 0 = 0 + a &= a \\ a + 1 &= a^+ \\ a + (b + c) &= (a + b) + c \\ a(bc) &= (ab)c \\ a(b + c) &= ab + ac \\ a^{b+c} &= a^b a^c \\ a^{bc} &= (a^b)^c. \end{aligned}$$

However, ordinal addition and multiplication are not commutative and not right-distributive. Also, $(ab)^c$ is generally distinct from $a^c b^c$.

Definition 47. Two sets A and B are said to have the same *cardinality* (written $|A| = |B|$) if there is a bijection $f : A \to B$.

A set A has cardinality at most the cardinality of B ($|A| \leq |B|$) if there is an injection $f : A \to B$.

A set A has cardinality less than the cardinality of B ($|A| < |B|$) if $|A| \leq |B|$ and $|A| \neq |B|$.

A set A is *countable* if $|A| \leq |\omega|$, and *uncountable* otherwise.

Proposition 31. *If there exists a surjection $f : A \to B$ then $|B| \leq |A|$.*

Theorem 32 (Schröder-Bernstein). *If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.*

Theorem 33 (Cantor). *For any set A, $|\mathcal{P}(A)| > |A|$.*

Definition 48. A *cardinal number* is an ordinal number α such that for any ordinal number β with $|\alpha| = |\beta|$ we have $\alpha \subseteq \beta$.

Proposition 34. *Every element of ω, as well as ω itself, is a cardinal number.*

Proposition 35. *For any set S, there is a unique cardinal number α with $|\alpha| = |S|$.*

Definition 49. For these sets S and α we define $|S| = \alpha$.

Definition 50. A set A is said to be *finite* if $|A| \in \omega$, and *infinite* otherwise.

Proposition 36. *A set is infinite if and only if it has the same cardinality as some proper subset.*

Proposition 37. *A countable set does not have any uncountable subsets. An uncountable set has a subset with cardinality equal to ω.*

Proposition 38. *A union of countably many countable sets is countable.*

Proposition 39. *If A, B, C, D are sets such that $|A| = |B|$, $|C| = |D|$, and $A \cap C = B \cap D = \emptyset$, then $|A \cup B| = |C \cup D|$, $|A \times B| = |C \times D|$ and $|A^B| = |C^D|$.*

Definition 51. We define cardinal addition, multiplication and exponentiation as, for disjoint sets A and B,

$$|A| + |B| = |A \cup B|, \ |A| \times |B| = |A \times B|, \ |A|^{|B|} = |A^B|.$$

Proposition 40. *If a and b are ordinals, then $|a+b| = |a|+|b|$, $|ab| = |a||b|$ and $|a^b| = |a|^{|b|}$. ordinal operations are used on the left side and the cardinal operations are used on the right.*

Proposition 41. *If a and b are cardinal numbers such that $a \geq \omega$ and $a \geq b$, then $a + b = a \times b = a$. If b is finite we also have $a^b = a$.*

Definition 52. For each infinite cardinal a, consider the set $c(a)$ of all infinite cardinals strictly less than a. It is well-ordered, so it has an ordinal number α. Then $a = \aleph_\alpha$.

Remark 53. The *Continuum Hypothesis*, proven to be independent from all of the axioms of set theory we've mentioned, is that $\aleph_1 = 2^{\aleph_0}$.

The *Generalised Continuum Hypothesis* extends this to

$$\aleph_{\alpha+1} = 2^{\aleph_\alpha}$$

for all α.

Both of these statements are independent of ZFC.

References.

- *Naive Set Theory*, Halmos

CHAPTER 2

Number Systems

Definition 54. A *binary operation* on A is a function $\cdot : A \times A \to A$. We usually write $\cdot(a, b) = c$ as $a \cdot b = c$.

It is *associative* if $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for any $a,\ b,\ c$ in A.

It is *commutative* if $a \cdot b = b \cdot a$ for any $a,\ b$ in A.

Definition 55. A *monoid* is an ordered pair $(A, \cdot)$ of a set A and an associative binary operation $\cdot$ on A such that there exists an element 1, called the *identity*, such that $a \cdot 1 = 1 \cdot a = a$ for all a.

Remark 56. There are two main notations for monoid-type structures. These are

- Multiplicative notation, in which the operation is notated $a \cdot b$ or simply ab, and the identity element is 1; and
- Additive notation, in which the operation is notated $a + b$ and the identity element is 0.

Definition 57. A *group* is a monoid $(A, \cdot)$ such that for each element a of A there is an element b of A such that $ab = 1 = ba$.

A group is *abelian* if the operation is commutative.

Proposition 42. *If $ab = ba = 1$ and $ac = 1$ or $ca = 1$ then $b = c$.*

Definition 58. The element b of A such that $ab = ba = 1$ is called the *inverse* of a. In multiplicative notation, the inverse of a is notated a^{-1}. In additive notation, the inverse of a is notated $-a$.

Remark 59. We often define $\frac{a}{b} = ab^{-1}$ in multiplicative notation, and $a - b = a + (-b)$ in additive notation.

Definition 60. A *ring* is an ordered triple $(A, +, \cdot)$ such that $(A, +)$ is an abelian group, $(A \setminus \{0\}, \cdot)$ is a monoid, and the *distributive laws* hold:

$$a \cdot (b + c) = ab + ac \quad \text{and} \quad (a + b) \cdot c = ac + bc.$$

It is *commutative* if $\cdot$ is commutative.

It is *ordered* if there is a total order $\leq$ on A satisfying

- if $a \leq b$ then $a + c \leq b + c$, and
- if $0 \leq a$ and $0 \leq b$ then $0 \leq ab$.

Definition 61. A *field* is a commutative ring $(A, +, \cdot)$ such that $(A \setminus \{0\}, \cdot)$ is a group.

An *ordered field* is a field that is also an ordered ring.

Definition 62. In an ordered ring R, the *absolute value* $|a|$ of an element a of R is a if $0 \leq a$, otherwise $-a$.

PROPOSITION 43 (Triangle Inequality on ordered rings). *If a and b are in an ordered ring R, then $|a+b| \leq |a| + |b|$.*

DEFINITION 63. Let X and Y be similar well-ordered sets, and let A and B be the least elements of X and Y respectively. Assume that all other elements of X and Y are operations on A and B respectively, and let f be the similarity between A and B.

A function $\varphi : A \to B$ is said to be a *homomorphism* if for every $a, b \in A$ and every $\cdot \in X \setminus \{A\}$ we have

$$\varphi(a \cdot b) = \varphi(a) f(\cdot) \varphi(b).$$

An *isomorphism* is a bijective homorphism.

If there exists an isomorphism from A to B, then we say A and B are *isomorphic*.

PROPOSITION 44. *The property of being isomorphic is reflexive, symmetric and transitive.*

REMARK 64. We don't say that isomorphism is an equivalence relation, since it would imply there exists a set of all well-ordered sets of this type.

Such a set does not exist because if it did it would contain (S, Id_S) for each set S. This would imply the existence of a set of all sets.

THEOREM 45. *There exists a unique ordered ring $\mathbb{Z}$ (up to isomorphism) such that $\{x \in \mathbb{Z} : x \geq 0\}$ is well-ordered.*

$\mathbb{Z}$ is commutative.

DEFINITION 65. We call this set $\mathbb{Z}$ the *integers*. The *non-negative integers* $\mathbb{Z}_{\geq 0}$ are $\{n \in \mathbb{Z} : n \geq 0\}$. The *positive integers* $\mathbb{Z}^+$ are $\mathbb{Z}_{\geq 0} \setminus \{0\}$.

REMARK 66. As a byproduct of our construction, we get a canonical bijection between ω and $\mathbb{Z}_{\geq 0}$. In particular, the cardinality of a finite set is a nonnegative integer.

REMARK 67. We avoid use of the term *natural numbers*, and the symbol $\mathbb{N}$, since some use them to mean the positive integers and others use them to mean the nonnegative integers.

PROPOSITION 46. *Every ordered ring contains a unique subring isomorphic to $\mathbb{Z}$.*

DEFINITION 68. In $\mathbb{Z} \times \mathbb{Z}^+$, we define the operations

$$(a, b) + (c, d) = (ad + bc, bd), \qquad (a, b)(c, d) = (ac, bd).$$

We also define an equivalence relation $\sim$ where $(a, b) \sim (c, d) \iff ad = bc$.

We define the *rational numbers* $\mathbb{Q}$ as the partition of $\mathbb{Z} \times \mathbb{Z}^+$ induced by this equivalence relation, with $[(a, b)] + [(c, d)] = [(ad + bc, ac + bd)]$ and $[(a, b)] \cdot [(c, d)] = [(ac, bd)]$.

PROPOSITION 47. *The relation $\sim$ is an equivalence relation. Moreover, the operations $+$ and $\cdot$ are independent of the representatives of each equivalence class. With these operations, $\mathbb{Q}$ is a field.*

PROPOSITION 48. *Every ordered field contains a unique subfield isomorphic to $\mathbb{Q}$.*

Definition 69. A partially ordered set S is *complete* if every nonempty subset that has an upper bound in S has a least upper bound in S.

Proposition 49. *Let S be a complete partially ordered set. Every nonempty subset that has a lower bound in S has a greatest lower bound in S.*

Theorem 50. *There exists a unique complete ordered field, up to isomorphism.*

Definition 70. We call this field $\mathbb{R}$.

Definition 71. We define $\mathbb{Q}_{\geq 0}$, $\mathbb{Q}^+$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}^+$ in an analogous way to $\mathbb{Z}_{\geq 0}$ and $\mathbb{Z}^+$.

Definition 72. We define the *complex numbers* $\mathbb{C}$ as $\mathbb{R}^2$, with the operations

$$(a,b)+(c,d)=(a+c,b+d), \qquad (a,b)\cdot(c,d)=(ac-bd,ad+bc).$$

We usually write (a,b) as $a+bi$. We define the *conjugate* of $a+bi$ to be $\overline{a+bi}=a-bi$.

Proposition 51. $\mathbb{C}$ *is a field under these operations.*

Proposition 52. *There are unique homomorphisms $\mathbb{Z}\to\mathbb{Q}$ and $\mathbb{Q}\to\mathbb{C}$. There is also an isomorphism $\mathbb{R}\to\{x\in\mathbb{C}:x=\overline{x}\}$.*

Remark 73. Because of this, we usually take $\mathbb{Z}\subseteq\mathbb{Q}\subseteq\mathbb{R}\subseteq\mathbb{C}$.

Proposition 53. *Let $a\in\mathbb{C}$. Then, $a\overline{a}\in\mathbb{R}_{\geq 0}$.*

Proposition 54. *Let $b\in\mathbb{R}_{\geq 0}$. There exists a unique $x\in\mathbb{R}_{\geq 0}$ such that $x\cdot x=b$.*

Definition 74. We call x the *square root* of b, denoted $\sqrt{b}$.
We call $\sqrt{a\overline{a}}$ the *modulus* of a, denoted $|a|$.

Proposition 55 (Triangle Inequality over $\mathbb{C}$). *If a and b are complex numbers, then $|a+b|\leq|a|+|b|$.*

Theorem 56. $|\mathbb{Z}^+|=|\mathbb{Z}_{\geq 0}|=|\mathbb{Z}|=|\mathbb{Q}|=\omega$, *but* $|\mathbb{R}|=|\mathbb{C}|=|\mathcal{P}(\omega)|$.

Definition 75. A *polynomial* over S is an expression of the form

$$p(z)=a_0+a_1z+a_2z^2+\cdots+a_mz^m,$$

for some integer m and coefficients $a_i\in S$.

We say the *degree* of p is d, where d is the largest integer such that $a_d\neq 0$. If no such d exists, the degree is $-\infty$.

Proposition 57 (Division Algorithm). *Suppose p and s are polynomials over a field $\mathbb{F}$ with $s\neq 0$. There exist unique polynomials q,r over $\mathbb{F}$ such that $p=sq+r$ and $\deg r<\deg s$.*

Definition 76. A number $r\in\mathbb{F}$ is a root of a polynomial p over $\mathbb{F}$ if $p(r)=0$.

Proposition 58. *A polynomial over a field $\mathbb{F}$ has at most as many roots as its degree.*

Theorem 59 (Fundamental Theorem of Algebra). *Every nonconstant polynomial over $\mathbb{C}$ has a root.*

Proposition 60. *If p is a polynomial over $\mathbb{C}$ then it has a unique factorisation of the form $p(z)=c(z-r_1)\cdots(z-r_m)$, where all constants are complex numbers.*

PROPOSITION 61. *If p is a polynomial over $\mathbb{R}$ then it has a unique factorisation of the form*

$$p(x) = c(x - r_1) \cdots (x - r_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_n x + c_n),$$

where all constants are real numbers such that $b_j^2 < 4c_j$ for each j.

CHAPTER 3

Linear Algebra

Definition 77. Let $\mathbb{F}$ be a field. A *vector space over* $\mathbb{F}$ is an abelian group V (of *vectors*) together with a function $\cdot : \mathbb{F} \times V \to V$ (*scalar multiplication*) such that

- $a(bv) = (ab)v$ (compatible),
- $1v = v$ (identity), and
- $a(u + v) = au + av$ and $(a + b)v = av + bv$ (distributive).

Definition 78. Let S be a subset of V. A *linear combination* of elements of S is a vector of the form

$$\sum_{i=1}^{n} a_i s_i,$$

where each s_i is a distinct element of S.

Definition 79. A *basis* of a vector space V is a set $S \subseteq V$ such that each element of V can be uniquely represented as a linear combination of elements of S.

Remark 80. For an infinite-dimensional vector space, there are multiple different notions of a basis. This one is usually called a *Hamel basis*.

Theorem 62. *Let V be a vector space.*

- *V has a basis.*
- *Any two bases of V have the same cardinality.*

Definition 81. The *dimension* of V is the cardinality of a basis of V. If $\dim V$ is an integer, V is said to be *finite-dimensional*; otherwise, it is *infinite-dimensional*.

Definition 82. A *subspace* W of V is a nonempty subset of V which is also a vector space over $\mathbb{F}$.

Proposition 63. *A subset W of V is a subspace iff the following conditions hold:*

- *W is nonempty;*
- *$u, v \in W$ implies $u + v \in W$ (closed under addition); and*
- *if $a \in \mathbb{F}$ and $u \in W$ then $au \in W$ (closed under scalar multiplication).*

Definition 83. The *span* of a subset S of V is the set of linear combinations of elements of S.

Proposition 64. *The span of S is the intersection of all subsets of V that contain S. It is also a subspace of V.*

Definition 84. A subset S of V is *linearly independent* if any linear combination of elements of S that produces 0 has all coefficients equal to 0. Otherwise, it is *linearly dependent*.

PROPOSITION 65. *A subset S of V is a* basis *iff it is linearly independent and its span is V.*

PROPOSITION 66. *Let V be finite-dimensional with dimension d. Let S be a set of vectors in V with $|S| = d$. Then S is linearly independent iff it spans V.*

DEFINITION 85. A *linear map*, or *linear transformation*, from V to W is a group homomorphism $T : V \to W$ such that $T(\lambda v) = \lambda T(v)$ for all $\lambda \in \mathbb{F}$. A linear map from a vector space to itself is an *operator*.

The *product* of linear maps S and T is $ST = S \circ T$.

PROPOSITION 67. *The set $\mathcal{L}(V, W)$ of linear maps from V to W is a vector space. Right-multiplication by a linear map $T : U \to V$ defines a linear map from $\mathcal{L}(V, W)$ to $\mathcal{L}(U, W)$. Left-multiplication by T defines a linear map from $\mathcal{L}(W, U)$ to $\mathcal{L}(W, V)$.*

DEFINITION 86. The *null space* of a linear map T is the subset of its domain that T maps to 0.

PROPOSITION 68. *Let V be finite-dimensional, and let $T : V \to W$ be a linear transformation. Then the null space of T is a subspace of V, the image of T is a subspace of W, and the sum of the dimensions of these two subspaces equals* $\dim V$.

DEFINITION 87. A linear map $T : V \to W$ is *invertible* if there is a linear map $S : W \to V$ such that ST is the identity on V and TS is the identity on W. In this case, S is called an *inverse* of T.

PROPOSITION 69. *Two vector spaces over $\mathbb{F}$ are isomorphic iff they have the same dimension.*

PROPOSITION 70. *Suppose V and W are finite-dimensional and isomorphic, and T is a linear transformation from V to W. The following are equivalent:*

- *T is invertible.*
- *T is injective.*
- *T is surjective.*

DEFINITION 88. The *product* of vector spaces is the Cartesian product, where addition and scalar multiplication are defined componentwise.

PROPOSITION 71. *Suppose U is a subspace of V. Define the relation $a \sim b \iff b - a \in V$. Then $\sim$ is an equivalence relation, and addition and scalar multiplication are invariant under it. The partition induced by this relation is a vector space.*

DEFINITION 89. This vector space is called the *quotient space* of V over U, denoted V/U.

PROPOSITION 72. *Suppose T is a linear transformation with domain V, and let U be the null space of T. Then T is an isomorphism from V/U to the range of T.*

DEFINITION 90. A *linear functional* on V is a linear map from V to $\mathbb{F}$.

DEFINITION 91. The space of linear functionals on V is the *dual space* of V, denoted V'.

PROPOSITION 73. *If V is infinite-dimensional,* $\dim V' > \dim V$.

Proposition 74. *If V is finite-dimensional,* $\dim V' = \dim V$.

Remark 92. This construction, applied twice, yields a canonical isomorphism between V'' and V, so they are often identified.

Definition 93. For $U \subseteq V$, the *annihilator* of U, denoted U^0, is

$$\{y \in V' : y(u) = 0 \ \forall u \in U\}.$$

Definition 94. If $v_1, \dots, v_n$ is a basis of V, then there exists a basis of elements φ_j of V', where $\varphi_j v_k$ is 1 if $j = k$ and 0 otherwise.

Definition 95. This basis is called the *dual basis* of $v - 1, \dots, v_n$.

Definition 96. The *dual map* of T is the linear map $T' : W' \to V'$ defined by $T'\varphi = \varphi T$ for each $\varphi \in W'$.

Proposition 75. *T' is a linear map. The dimensions of the range of T' and the range of T coincide.*

Definition 97. Suppose V and W have finite bases $\{v_i\}_1^m$ and $\{w_i\}_1^n$ respectively. The *matrix* A of T with respect to these bases is defined by

$$Tv_k = \sum_{i=1}^{n} A_{i,k} w_i.$$

We also identify $1 \times n$ and $n \times 1$ matrices with elements of $\mathbb{F}^n$.

Proposition 76. *This defines a bijection between the space of $m \times n$ matrices and the space of linear transformations $\mathbb{F}^n \to \mathbb{F}^m$.*

Definition 98. Thus, we identify the two, and can therefore talk of the image, null space, etc of a matrix.

Definition 99. The *rank* of a matrix is the dimension of its image.

The *transpose* of a matrix is the matrix obtained by swapping rows and columns: $A^T_{j,k} = A_{k,j}$.

Proposition 77. *Let $T : V \to W$ be a linear transformation, where V and W are finite-dimensional. Pick bases $\{v_i\}$ and $\{w_i\}$ of V and W. The matrix of T' with respect to the dual bases of $\{w_i\}$ and $\{v_i\}$ is the transpose of the matrix of T with respect to $\{v_i\}$ and $\{w_i\}$.*

Corollary 78. *The rank of a matrix equals the rank of its transpose.*

Definition 100. Let $A : U \to W$ and $B : V \to W$ be linear maps. We *augment* A with B to get the linear map

$$(A|B) : U \times V \to W, \ (A|B)(x, y) = Ax + By.$$

Proposition 79. *For any $x : V \to U$ we have $Ax = B \iff (A|B)(x, -I) =$* 0.

Remark 101. Thus, to solve the linear system $Ax = B$ it suffices to find the null space of $(A|B)$. Notice also that the matrix of $(A|B)$ is simply the matrix formed by concatenating the matrices of A and B.

Proposition 80. *Let T and S be linear maps from V to W. The following are equivalent:*

- *The null spaces of T and S are the same.*
- *The images of T' and S' coincide.*
- *There is an invertible linear map $A : V \to V$ such that $AT = S$.*

Definition 102. Such linear maps are called *equivalent*.

Definition 103. A *pivot* is the first nonzero entry in a row of a matrix.

A matrix is in *row echelon form (REF)* if all rows consisting of only zeroes are at the bottom and the pivot of a nonzero row is strictly to the right of the pivot of the row above it.

A matrix is in *reduced row echelon form (RREF)* if it is in REF, all pivots are 1, and each column containing a pivot has zeroes everywhere else in the column.

Proposition 81. *Every matrix is equivalent to a unique matrix in RREF.*

Definition 104. An *elementary matrix* is a matrix that differs from the identity in exactly one entry, where that entry is nonzero in the elementary matrix.

Proposition 82. *A matrix is invertible iff it is a product of elementary matrices.*

Remark 105. The null space of a matrix in RREF is easy to find. Thus, to find the null space of a matrix, we left-multiply by elementary matrices to find an equivalent matrix in RREF. This process is known as *Gaussian elimination*. It is efficient because multiplying by an elementary matrix has simple consequences:

- An elementary matrix which has a changed entry on the main diagonal multiplies a row by a scalar.
- An elementary matrix which has a nonzero entry off the main diagonal adds a scalar multiple of one row to another.

Most authors add a third (redundant) type of row operation and elementary matrix: swapping two rows.

The next proposition shows that Gaussian elimination also helps us find bases for the span of a set of vectors.

Proposition 83. *Let T be a matrix which is equivalent to a matrix S in REF. Then,*

- *The rows of S with pivots form a basis for the span of the rows of T.*
- *Consider the columns of S with pivots. The corresponding columns of T form a basis for the span of the columns of T.*

Definition 106. Let $T : V \to V$ be a linear transformation. A subspace U of V is called *invariant* under T if $u \in U \implies Tu \in U$.

A nonzero vector $v \in V$ is called an *eigenvector* of T if there is some $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$. We call λ an *eigenvalue* of T.

Proposition 84. *λ is an eigenvalue of T if and only if $T - \lambda I$ is not invertible.*

Proposition 85. *Any set of eigenvectors of T with distinct eigenvalues is linearly independent.*

Definition 107. Suppose $T : V \to V$ is linear, and U is a subspace of V invariant under T. The *restriction* $T|_U : U \to U$ is defined by $T|_U(u) = Tu$, while the *quotient* $T/U : V/U \to V/U$ is defined by $(T/U)(v + U) = Tv + U$.

DEFINITION 108. Suppose $T : V \to V$ is a linear transformation and

$$p(z) = \sum a_i z^i,$$

where each $a_i \in \mathbb{F}$. Then $p(T) = \sum a_i T^i$.

THEOREM 86. *Every operator on a finite-dimensional nonzero complex vector space has an eigenvalue.*

DEFINITION 109. In defining the *matrix* of an operator, we choose the same basis for the domain and codomain.

PROPOSITION 87. *Suppose V is a finite-dimensional vector space and $T : V \to V$ is an operator. Then T has an upper-triangular matrix with respect to some basis of V.*

PROPOSITION 88. *Suppose $T : V \to V$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that matrix.*

DEFINITION 110. An operator is *diagonalisable* if it has a diagonal matrix with respect to some basis of the space.

PROPOSITION 89. *Let $T : V \to V$ be an operator over a finite-dimensional vector space. Then T is diagonalisable iff V has a basis consisting of eigenvectors of T.*

DEFINITION 111. An *inner product space* is a vector space V over a field $\mathbb{F}$ which is either $\mathbb{R}$ or $\mathbb{C}$, together with a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ satisfying

- $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate symmetry)
- $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$ (linearity in the first argument), and
- $\langle x, x \rangle = 0 \implies x = 0$.

PROPOSITION 90. *The dot product, defined by*

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) = \sum a_i \overline{b_i},$$

is an inner product over both $\mathbb{R}^n$ and $\mathbb{C}^n$.

PROPOSITION 91 (Cauchy-Schwarz). $|\langle u, v \rangle| \leq \|u\|\|v\|$.

DEFINITION 112. A *normed vector space* is a vector space V over $\mathbb{R}$ or $\mathbb{C}$ on which there is a *norm*: a function $\| \cdot \| : V \to \mathbb{R}$ satisfying

- $\|x\| \geq 0$, with $\|x\| = 0 \iff x = 0$,
- $\|ax\| = |a|\|x\|$, and
- $\|x + y\| \leq \|x\| + \|y\|$ (the triangle inequality).

PROPOSITION 92. *If V is an inner product space, then $\langle x, x \rangle$ is real for all x. Moreover, $\|x\| = \sqrt{\langle x, x \rangle}$ is a norm on V.*

DEFINITION 113. Two vectors x and y are *orthogonal* if $\langle x, y \rangle = 0$.

A set of vectors is *orthonormal* if each vector in the set has norm 1 and is orthogonal to all other vectors in the set.

PROPOSITION 93. *Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.*

Remark 114. Thus, we may identify a finite-dimensional inner product space over $\mathbb{F}$ with $\mathbb{F}^n$ under the usual dot product.

Theorem 94 (Schur's Theorem). *An operator over a finite-dimensional inner product space has an upper-triangular matrix with respect to an orthonormal basis of the space.*

Theorem 95 (Riesz Representation). *Any linear functional f on a finite-dimensional inner product space can be written as $f(x) = \langle x, v \rangle$ for some fixed vector v.*

Definition 115. Let U be a finite-dimensional subspace of V. The *orthogonal projection* of V onto U is the operator $P_U : V \to V$ defined by $P_U v = u$ where $u \in U$ and $\langle v - u, x \rangle = 0 \ \forall x \in U$.

Proposition 96. *The orthogonal projection is well defined, and satisfies*

$$\|P_u v\| \leq \|v\|.$$

For any $u \in U$, we have

$$\|v - P_U v\| \leq \|v - u\|.$$

Remark 116. We may use this last result to solve minimisation problems, for example least-squares regression.

Proposition 97. *Let $T : V \to W$ be linear. There exists a unique function $T^* : W \to V$ such that*

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

for every $v \in V$ and every $w \in W$. The function T^* is linear.

Definition 117. We call T^* the *adjoint* of T.

Proposition 98. *Let $T : V \to W$ be linear, where V and W are real or complex vector spaces. Let $\{v_i\}$ be an orthonormal basis for V, and let $\{w_i\}$ be an orthonormal basis for W. Then, the matrix of T^* with respect to $\{w_i\}$ and $\{v_i\}$ is the conjugate transpose of the matrix of T with respect to $\{v_i\}$ and $\{w_i\}$.*

Definition 118. Let T be an operator. If $T^* = T$, then T is *self-adjoint*. If $TT^* = T^*T$, then T is *normal*.

Proposition 99. *Every eigenvalue of a self-adjoint operator is real.*

Theorem 100 (Spectral Theorem). *Let $T : V \to V$ be normal, where V is finite-dimensional. Then T has a diagonal matrix with respect to some orthonormal basis of V.*

Remark 119. Thus, we may write $T = UBU^*$, where $UU^* = U^*U = I$ and B is diagonal.

Proposition 101. *T is positive semidefinite iff there exists an operator R such that $T = R^*R$.*

Definition 120. An operator F is a *square root* of an operator T if $R^2 = T$.

Proposition 102. *Every positive operator has a unique positive square root.*

Definition 121. If T is a positive operator, then $\sqrt{T}$ denotes the unique positive semidefinite square root of T.

DEFINITION 122. A linear transformation is an *isometry* if it preserves norms. An operator which is also an isometry is *unitary*.

PROPOSITION 103. *An linear transformation T is unitary iff $T^*T = I$.*

THEOREM 104 (Polar Decomposition). *For each operator T, there exists a unitary operator S such that $T = S\sqrt{T^*T}$.*

DEFINITION 123. The *singular values* of T are the eigenvalues of $\sqrt{T^*T}$, where each eigenvalue λ is counted the same number of times as the dimension of its eigenspace.

PROPOSITION 105. *The nonzero singular values of T and of T^* coincide.*

THEOREM 106 (Singular Value Decomposition). *Suppose $T : V \to W$ has singular values $s_1, \ldots, s_n$. Then there exist orthonormal bases $e_1, \ldots, e_n$ of V and $f_1, \ldots, f_n$ of W such that*

$$Tv = \sum_i s_i \langle v, e_i \rangle f_i$$

for all $v \in V$.

PROPOSITION 107. *Let $T : V \to W$ be a linear transformation. There exists a unique linear transformation $T^+ : W \to T$ such that*

- *$TT^+T = T$;*
- *$T^+TT^+ = T^+$;*
- *TT^+ and T^+T are self-adjoint.*

DEFINITION 124. This transformation T^+ is known as the *pseudoinverse* of T.

DEFINITION 125. A vector v is called a *generalised eigenvector* of T corresponding to an eigenvalue λ if $v \neq 0$ and $(T - \lambda I)^j v = 0$ for some positive integer j.

The *generalised eigenspace* of T corresponding to λ is the set of all generalised eigenvectors of T corresponding to λ, along with the 0 vector.

PROPOSITION 108. *For finite-dimensional V, v is a generalised eigenvector of T iff $(T - \lambda I)^{\dim V} v = 0$.*

PROPOSITION 109. *Generalised eigenvectors corresponding to distinct eigenvalues are linearly independent.*

PROPOSITION 110. *Suppose V is a finite-dimensional complex vector space, and T is an operator on V. Then there is a basis of V consisting of generalised eigenvectors of T.*

DEFINITION 126. The *multiplicity* of an eigenvalue λ of T is the dimension of the corresponding generalised eigenspace.

PROPOSITION 111. *Every operator on a nonzero finite-dimensional real vector space has an invariant subspace of dimension 1 or 2.*

DEFINITION 127. A *block diagonal matrix* is a square matrix of the form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_i is a square matrix lying along the diagonal and all other entries of the matrix are 0.

THEOREM 112 (Jordan Form). *If T is an operator on a finite-dimensional complex vector space, then there is a basis such that the matrix of T with respect to this basis is block diagonal with blocks of the form*

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix},$$

where each λ_i is a distinct eigenvalue of T.

DEFINITION 128. The *trace* of a square matrix A is the sum of the diagonal entries of A.

PROPOSITION 113. *If T is an operator over a finite-dimensional vector space V, and $\{a_i\}$ and $\{b_i\}$ are two bases for V, then the trace of the matrix of T with respect to $\{a_i\}$ equals the trace of the matrix of T with respect to $\{b_i\}$.*

DEFINITION 129. We call this quantity the *trace* of T.

PROPOSITION 114. *The trace is additive.*

PROPOSITION 115. *If V is complex, then the trace of T equals the sum of the eigenvalues of T counted according to multiplicity.*

DEFINITION 130. If π is a permutation of $\{1, 2, \dots, n\}$, the sign of π is -1^k, where $k = |\{(a, b) \in \{1, 2, \dots, n\} : a < b,\ \pi(a) > \pi(b)\}|$.

PROPOSITION 116. *Let A be $n \times n$. The determinant of the linear transformation defined by A equals*

$$\sum_{\pi} \operatorname{sign}(\pi) \prod_{i=1}^{n} A_{\pi(i),i},$$

where the sum is taken over all permutations π of $\{1, 2, \dots, n\}$.

DEFINITION 131. We call this quantity the *determinant* of T.

PROPOSITION 117. *The determinant is multiplicative.*

PROPOSITION 118. *If V is complex, then the determinant of T equals the product of the eigenvalues of T counted according to multiplicity.*

DEFINITION 132. Let T be an operator on a finite-dimensional vector space. The *characteristic polynomial* p of T is defined by

$$p(\lambda) = \det(T - \lambda I).$$

PROPOSITION 119. *If T is an operator on a finite-dimensional complex vector space, then the characteristic polynomial p of T satisfies*

$$p(z) = \prod (z - \lambda_i),$$

where λ_i are the eigenvalues of T counted according to multiplicity.

THEOREM 120 (Cayley-Hamilton). *Let p be the characteristic polynomial of T. Then $p(T) = 0$.*

References.

- *Linear Algebra Done Wrong*, Treil
- *Linear Algebra Done Right*, Axler
- *Finite-Dimensional Vector Spaces*, Halmos

CHAPTER 4

Analysis

Definition 133. A *metric space* is a nonempty set M together with a function $d : M \times M \to \mathbb{R}$ (the *metric*) such that

- $d(x, y) = 0 \iff x = y$,
- $d(x, y) = d(y, x)$ (symmetry),
- $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

Proposition 121. *In a normed vector space, the function* $d(x, y) = \|x - y\|$ *is a metric.*

Definition 134. We call this the *induced metric*.

Definition 135. In a metric space, the *open ball* $B_r(x)$ with centre x and radius r is the set of all points y with $d(x, y) < r$.

The *closed ball* $\overline{B_r(x)}$ with centre x and radius r is the set of all points y with $d(x, y) \leq r$.

Definition 136. Let E be a subset of a metric space M.

- A point p is a *limit point* of E if every open ball centred at p contains a point $q \neq p$ such that $q \in E$.
- A point p is an *interior point* of E if there is an open ball centred at p which is a subset of E.
- E is *closed* if every limit point of E is a point of E.
- E is *open* if every point of E is an interior point of E.
- E is *bounded* if it is contained in some open ball.
- The *complement* E^c of a set E is the set $M \setminus E$.
- The *interior* of E is the set of interior points of E.
- The *boundary* ∂E of E is the set of points of M that are limit points of both E and E^c.

Proposition 122. *The interior and boundary of* E *are disjoint, and their union is* E.

Proposition 123. *The following are equivalent:*

- E *is open.*
- $E \cap \partial E = \emptyset$.
- E^c *is closed.*
- $\partial E \subseteq E^c$.

Proposition 124. *Every open ball is open; every closed ball is closed.*

Proposition 125. *If* p *is a limit point of* E*, then every open ball centred around* p *contains infinitely many points of* E.

Proposition 126. *Any union of open sets is open; a finite intersection of open sets is open.*

Any intersection of closed sets is closed; a finite union of closed sets is closed.

Definition 137. The *closure* of E is the set $E \cup \partial E$.

Proposition 127. *The closure of E is closed; the interior of E is open.*

Any closed set which contains E contains the closure of E. Any open set which is contained in E is contained in the interior of E.

Proposition 128. *Suppose $X \subseteq M$ inherits the metric. A subset E of X is open relative to X iff $E = X \cap Y$ for some open set Y.*

Definition 138. An *open cover* of E is a set of open sets whose union contains E.

Proposition 129. *The following are equivalent:*

- *Every open cover of E contains a finite subset which is still an open cover of E.*
- *Every infinite subset of E contains a limit point in E.*

Definition 139. Such a set is called *compact*.

Proposition 130. *Suppose $X \subseteq M$ inherits the metric. A subset E of X is open relative to X iff E is compact relative to M.*

Proposition 131. *A compact subset of a metric space is closed and bounded; a closed subset of a compact metric space is compact.*

Proposition 132. *If S is a collection of compact subsets of a metric space such that any finite intersection of elements of S is nonempty, then $\bigcap S$ is nonempty.*

Theorem 133 (Heine-Borel). *A subset of $\mathbb{R}^n$ is compact iff it is closed and bounded.*

Theorem 134 (Weierstrass). *Every bounded infinite subset of $\mathbb{R}^n$ has a limit point.*

Definition 140. Two subsets A and B of a metric space X are *separated* if both $A \cap \overline{B}$ and $B \cap \overline{A}$.

A set E is *disconnected* if it is the union of two nonempty separated sets, and connected otherwise.

Proposition 135. *A metric space M is connected iff the only sets which are both open and closed are the empty set and M.*

Proposition 136. *A subset of $\mathbb{R}^1$ is connected iff it is an interval.*

Definition 141. A sequence $\{a_n\}$ is *convergent* if there is a point L such that for any $\varepsilon > 0$ there is an $N \in \mathbb{Z}^+$ such that $n \geq N$ implies $d(a_n, L) < \varepsilon$. We write

$$\lim_{n\to\infty} a_n = L.$$

Proposition 137. *Suppose $\{a_n\}$ and $\{b_n\}$ are sequences of complex numbers which converge to a and b respectively. Then the sequences $\{a_n+b_n\}$, $\{a_nb_n\}$, $\{\frac{a_n}{b_n}\}$ converge to $a+b$, ab, $\frac{a}{b}$ respectively (where in the last one we require $b_n \neq 0$ for each n).*

PROPOSITION 138. *A sequence in $\mathbb{R}^n$ or $\mathbb{C}^n$ converges iff it converges coordinatewise.*

DEFINITION 142. A sequence $\{p_n\}$ is *Cauchy* if for every $\varepsilon > 0$ there is an integer N such that $d(p_n, p_m) < \varepsilon$ if $m, n \geq N$.

A metric space is *complete* if every Cauchy sequence converges.

PROPOSITION 139. *Every convergent sequence is Cauchy.*

PROPOSITION 140. *Every compact metric space is complete.*

PROPOSITION 141. *$\mathbb{R}^n$ and $\mathbb{C}^n$ are complete.*

DEFINITION 143. Let $f : X \to Y$ be a function, where Y is a metric space and X is a subset of a metric space E. Let p be a limit point of X. We say that

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

if for every sequence $\{x_n\}$ in E which converges to p but does not contain p, $f(x_n)$ converges to q.

DEFINITION 144. We say that f is *continuous* at p if for every sequence $\{x_n\}$ in E which converges to p, $f(x_n)$ converges to $f(p)$.

We say that f is *continuous* on X, or simply *continuous*, if it is continuous at every point in X.

PROPOSITION 142. *A function f is continuous iff the inverse image of every open set is open.*

PROPOSITION 143. *If f is continuous, then*

- *The image of a compact set is compact.*
- *The image of a connected set is connected.*

COROLLARY 144 (Intermediate Value Theorem). *If the codomain of f is $\mathbb{R}$, then it is an interval. If the domain of f is a compact set, then the interval is closed.*

DEFINITION 145. A function f is *uniformly continuous* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $d(a, b) < \delta$ then $d(f(a), f(b)) < \varepsilon$.

THEOREM 145. *Every continuous function on a compact set is uniformly continuous.*

References.

- *Principles of Mathematical Analysis*, Rudin

CHAPTER 5

Theory of Computation

DEFINITION 146. A *deterministic finite automaton (DFA)* is a 5-tuple

$$(Q, \Sigma, \delta, q_0, F),$$

where Q is a finite set called the *states*, Σ is a finite set called the *alphabet*, $\delta : Q \times \Sigma \to Q$ is the *transition function*, $q_0 \in Q$ is the *start state*, and $F \subseteq Q$ is the set of *accept states*.

DEFINITION 147. We say that machine M *accepts* string $s = s_0 s_1 \cdots s_n$ if $\delta(\delta(\cdots\delta(\delta(q_0, s_0), s_1)\cdots, s_{n-1}), s_n)$ is an accept state.

DEFINITION 148. The set A of all strings that machine M accepts is the *language of machine M*, notated $L(M)$. We say that M *recognises* A.

DEFINITION 149. A language is a *regular language* if it is recognised by some finite automaton.

DEFINITION 150. Let A and B be languages. We define the *regular operations*

- *Union*: $A \cup B = \{x : x \in A \vee x \in B\}$.
- *Concatenation*: $A \circ B = \{xy : x \in A \wedge y \in B\}$.
- *Star*: $A^* = \{x_1 x_2 \cdots x_k : k \geq 0 \wedge \forall i,\ x_i \in A\}$.

DEFINITION 151. The *empty string* is notated ε.

REMARK 152. Note that $\varepsilon \in A^*$ for all A.

DEFINITION 153. A *nondeterministic finite automaton (NFA)* is a 5-tuple

$$(Q, \Sigma, \delta, q_0, F),$$

where Q is a finite set of states, Σ is a finite alphabet, $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \to \mathcal{P}(Q)$ is the transition function, $q_0 \in Q$ is the start state, and $F \subseteq Q$ is the set of accept states.

DEFINITION 154. Two machines are *equivalent* if they describe the same language.

PROPOSITION 146. *Every NFA has an equivalent DFA.*

COROLLARY 147. *A language is regular iff some NFA recognises it.*

DEFINITION 155. Let Σ be an alphabet. An *atomic regular expression* is one of

- a ($a \in \Sigma$),
- ε, and
- $\emptyset$.

Regular expressions are obtained by combining simpler regular expressions with the operations $\cup$, $\circ$, *.

A regular expression R *describes* a language $L(R)$ obtained by replacing each instance of a and ε with $\{a\}$ and $\{\varepsilon\}$, respectively, and then applying the regular operations.

Definition 156. A *generalised nondeterministic finite automaton* (GNFA) is a 5-tuple $(Q, \Sigma, \delta, q_s, q_a)$, where Q is a finite set of states, Σ is a finite input alphabet, $\delta : (Q - \{q_a\}) \times (Q - \{q_s\}) \to \mathcal{R}$ is the transition function, and $q_s, q_a \in Q$ are the start and accept states respectively.

A GNFA *accepts* a string w in Σ^* if $w = w_1 w_2 \cdots w_k$, where each w_i is in Σ^* and a sequence of states $q_0, q_1, \ldots, q_k$ exists such that $q_0 = q_s$, $q_k = q_a$ and for each i we have $w_i \in L(\delta(q_{i-1}, q_i))$.

Proposition 148. *A language is regular iff some GNFA recognises it.*

Theorem 149. *A language is regular iff some regular expression describes it.*

Lemma 150 (Pumping Lemma). *If A is a regular language, then there is a positive integer p such that if s is any string in A of length at least p, then s may be divided into three pieces, $s = xyz$, where y is nonempty, $|xy| \leq p$ and $x \circ y^* \circ z \subseteq A$.*

References.

- *Introduction to the Theory of Computation*, Sipser

APPENDIX A

Proofs