

Real Analysis

Muchang Bahng

Spring 2025

Contents

1 Number Systems	6
1.1 The Rationals	6
1.1.1 Field Properties	6
1.1.2 Ordered Field Properties	10
1.1.3 Norm	14
1.2 The Reals	14
1.2.1 Dedekind Completeness	15
1.2.2 Cauchy Completeness	18
1.2.3 Nested Intervals Completeness	19
1.2.4 Properties of the Real Line	20
1.2.5 Roots, Exponentials, and Logarithms	21
1.3 The Extended Reals and Hyperreals	26
1.4 Complex Numbers	27
1.5 Dual Numbers	31
1.6 Euclidean Space	31
1.7 Exercises	32
2 Euclidean Topology	62
2.1 Open Sets	62
2.2 Limit Points and Closure	64
2.3 Compactness	68
2.4 Connectedness	71
2.5 Separability	72
2.6 Perfect Sets	72
2.7 Exercises	72
3 Sequences in Euclidean Space	93
3.1 Convergence in Metric Topologies	93
3.2 Convergence in Reals	96
3.2.1 Completeness	97
3.2.2 Properties of Order	97
3.3 Arithmetic	99
3.4 Limsup and Liminf	102
3.5 Convergence Tests for Real Series	105
3.6 Euler's Number and Irrationality	109
3.7 Ratio and Root Tests	111
3.8 Exercises	113
4 Continuous Functions	121
4.1 Infinitesimal Functions	125

4.2 Asymptotic Behavior of Functions	127
4.2.1 Approximations of Functions	132
5 Continuous Functions	134
5.1 Points of Discontinuity	135
5.2 Properties of Continuous Functions	137
5.2.1 Uniform Continuity	139
5.2.2 Lipschitz Continuity	142
5.2.3 Inverse Function Theorem	143
6 Differentiation	145
6.1 Functions Differentiable at a Point	145
6.2 Tangent Line: Geometric Meaning of the Derivative, Differential	146
6.2.1 Constant Approximation	147
6.2.2 Linear Approximation	147
6.2.3 Higher Order Approximations	148
6.2.4 The Real Tangent Space	149
6.3 Rules of Differentiation over R	150
6.3.1 Basic Properties; Derivatives of Composite, Inverse Functions	150
6.3.2 Higher-Order Derivatives	154
6.4 Theorems of Differential Calculus	155
6.4.1 Fermat's Lemma, Rolle's Theorem	155
6.4.2 Mean Value Theorem, Cauchy's Finite-Increment Theorem	157
6.4.3 Taylor's Formula	159
6.5 The Study of Functions using Differential Calculus	162
6.5.1 Conditions for Monotonicity of Functions	162
6.5.2 Conditions for Extrema of Functions	163
6.5.3 Important Algebraic Inequalities	165
6.5.4 Conditions for a Function to be Convex	166
6.5.5 L'Hopital's Rule	168
6.6 Complex Analysis: An Introduction	169
6.6.1 Algebraic Extension of Field R	170
6.6.2 Geometric Interpretation of C	170
6.6.3 Sequences and Series in C	173
6.6.4 Euler's Formula	177
6.6.5 Visualizing Complex Functions	179
6.6.6 Continuity, Differentiability, Analyticity of Complex Functions	179
6.6.7 Power Series Representation of a Function	182
6.6.8 Algebraic Closedness of the Field C	182
6.7 Primitives	183
7 Integration	186
7.1 Construction of the Riemann Integral	186
7.1.1 Conditions for Integrability	188
7.1.2 The Vector Space of Riemann Integrable Functions	193
7.1.3 Lebesgue's Criterion for Riemann Integrability	193
7.2 Basic Properties of the Integral	195
7.2.1 Mean Value Theorem of the Integral	197
7.3 Connections between Integrals, Primitives, Derivatives	198
7.3.1 Integration by Parts and Taylor's Formula	202
7.3.2 Change of Variables in Integration	203
7.3.3 Additive Interval Functions and the Integral	204
7.3.4 Arc Length	205
7.4 Improper Integrals	207

7.4.1 Convergence of an Improper Integral	209
7.4.2 Improper Integrals with Multiple Singularities	211

Let's first talk about why we need analysis in general in the first place. Algebra allows us to define certain algebraic structures, which are essentially sets with operations. These operations are defined to have a finite number of arguments. For example, let's take a look at the negation $x \mapsto -x$ and the addition $x, y \mapsto x + y$ operations in a group G . We can compose these operations up to any finite length n , removing the parentheses due to associativity, but note that the "sum" below is not a single operation. It is a composition of $n - 1$ operations.

$$x_1 + x_2 + \dots + x_n \in G \quad (1)$$

$$-(-(\dots(-x))) \in G \quad (2)$$

This is still well defined due to closure, but what if we wanted to do this an infinite number of times?

$$x_1 + x_2 + \dots =? \quad (3)$$

$$\dots(-(-x)) =? \quad (4)$$

For someone who has learned about sequences and series in high school, this may not be a big jump in logic, but it is. The objects above are not even well-defined and trying to define them with algebraic tools is equivalent to the famous Zeno's paradox. So we simply need to add more tools in order to define these new mathematical objects, which we call *series*. To define series, we need to first define sequences. Can we do this with algebra? Yes, since we can simply model it as a function.

Definition 0.1 (Sequence)

A sequence is a function $f : \mathbb{N} \rightarrow X$. We usually denote a sequence by writing out the first few terms of the sequence, followed by an ellipsis.

$$a_1 = f(1), a_2 = f(2), \dots \quad (5)$$

or as an indexed set over the naturals $\{a_i\}_{i \in \mathbb{N}}$.

Therefore, we can consider series as a sequence of finite sums, each element which is well-defined.

$$x_1, x_1 + x_2, x_1 + x_2 + x_3, \dots \quad (6)$$

For any $n \in \mathbb{N}$, we can get the value of $a_n = \sum_{i=1}^n a_i$, but can we say something about the limiting behavior of a_n ? That is, maybe we can just slap a value x onto this series such that it doesn't "break" any of the rules we have in the finite sense. Unfortunately, it is not possible to define such values for all series, but it is possible for some of them, which we call *convergent series*. To rigorously determine which ones are convergent and which ones are not, we need the tools of topology and analysis. Defining the concept of sequences that model infinitely composed operations is what allows us to define differentiation and integration.

Great, we've motivated the need for analysis, but before jumping straight into real analysis, let's talk about what analysis in general works with. It studies functions of the form $f : X \rightarrow Y$, and minimally both X, Y must be *Banach spaces*, i.e. complete normed vector spaces over some field \mathbb{F} . Almost all flavors of analysis, including real (\mathbb{R}), complex (\mathbb{C}), multivariate (\mathbb{R}^n), p-adic, and functional (infinite-dimensional Banach spaces) analysis require *at least* a Banach space structure. Why are Banach spaces so great? Well if we were to define convergence in X or Y , then it only makes sense to talk about convergence with respect to a topology. So X, Y must at least be topological spaces. It would also be bad if we were to take a sequence in X and find out that it converges to some element outside of X . Therefore, we want a notion of *completeness* in the sense that all sequences that "get closer," i.e. Cauchy sequences, actually converge in X . Unfortunately, while convergence of sequences is preserved under homeomorphisms (and is thus a topological property), convergence of Cauchy sequences is not.¹ Furthermore, the notion of uniform convergence is a metric space property, not a topological one. Therefore, the concept of distances is crucial to the construction of analysis. As for the norm, I'm still not sure why we need this.²

¹Consider the sequence $a_n = 1/(n+1)$ in $(0, 1)$ and the map $f(x) = 1/x$ to the set $(1, +\infty)$. a_n is Cauchy but $f(a_n)$ is not.

²Aspinwall and Ng told me this, but I'm not sure why. The Frechet derivative seems like it can be purely defined with a metric.

But in college courses such as real and complex analysis, why do we say we work over the *fields* \mathbb{R} and \mathbb{C} rather than the Banach spaces \mathbb{R} and \mathbb{C} ? This is because of the following theorem.

Theorem 0.1 ()

Every field \mathbb{F} is a 1-dimensional vector space over itself.

Therefore, when we talk about the *field* \mathbb{R} , we are really treating it as a vector space \mathbb{R} over the field \mathbb{R} .³ Every other structure beyond this is a “bonus” property that gives us extra tools to prove stronger properties. The most notable is the total ordering on \mathbb{R} , which allows us to define upper/lower bounds and other real-analysis specific theorems like the intermediate value theorem or the mean value theorem. Other structures include the inner product or the measure.

Now that we've taken in the big picture, for each type of analysis, we should construct the underlying relevant Banach space. At the very least, we can with the tools of set theory and algebra define the rationals \mathbb{Q} as an ordered field over the quotient space $\mathbb{Z} \times \mathbb{Z}/\sim$. Furthermore, \mathbb{Q} itself is a normed vector space (over \mathbb{Q})⁴ and the only thing we need now is completeness.

1. If the norm on \mathbb{Q} is defined as the normal absolute value (Euclidean norm), completing it gives \mathbb{R} as an ordered field which also has a compatible order as that of \mathbb{Q} . We study functions mapping to and from \mathbb{R} with *single-variable real analysis*.
2. If we take the *p-adic* norm, then completing it with respect to this gives the *p-adic numbers*, which also forms a field but loses the ordering. We deal with functions over the p-adics with *p-adic analysis*.
3. We can construct \mathbb{C} by taking \mathbb{R}^2 and endowing it with a bit more structure. We get *complex analysis*.
4. We can construct \mathbb{R}^n and \mathbb{C}^n by easily defining its vector space structure and then endowing it with a norm, and showing that it is complete with respect to the norm-induced metric. This is known as *multivariate analysis*.
5. With all these defined, we can define Banach function spaces like L^p and perform analysis on operators $f : L^p \rightarrow L^q$. This is *functional analysis*.

What we have talked about so far was Cauchy completeness, but there is a different type of completeness called *Dedekind completeness*, also equivalently known as the *least-upper-bound (LUB) property*, defined only on ordered sets (with no other structure). It turns out that in an ordered field, the two forms of completeness are equivalent.⁵ Therefore, many real analysis textbooks tend to use Dedekind completeness when constructing the reals, but Cauchy completeness is in a sense more “fundamental.” We will go through both independent constructions of \mathbb{R} involving both types of completeness since both are used in future theorems.

1. *Construction from Cauchy Sequences.* We verify that \mathbb{Q} is a field and endow it with the standard Euclidean metric $d(x, y) = |x - y|$. We can then construct a new quotient space S of Cauchy sequences in \mathbb{Q} , define all the ordered field operations/relations, and finally show that S satisfies Cauchy completeness. Most would end here and claim that this is \mathbb{R} , but we must also prove the Archimedean property with this order. Once done, now we can truly claim $S = \mathbb{R}$.
2. *Construction from Dedekind Cuts.* We verify that \mathbb{Q} is a field, put an order on it, and verify that it is an ordered field. We then construct a new set D of *Dedekind cuts* from \mathbb{Q} , define the compatible ordered field operations/relations, and show that this new set D satisfies the least-upper bound property. We claim that $D = \mathbb{R}$.

³Thanks to Prof. Lenny Ng for clarifying this.

⁴Note that while we define the norm and metric to usually map to \mathbb{R}^+ , \mathbb{R} isn't even defined yet and so to avoid circular definitions, we define the norm on the rationals to have codomain \mathbb{Q} .

⁵Actually, this is not true. Dedekind completeness is equivalent to Cauchy completeness plus the Archimedean property. An example of a Cauchy-complete non Archimedean field is the field F of rational functions over \mathbb{R} , with positive cone consisting of those functions f/g such that the leading coefficients of f, g have the same algebraic sign. The Cauchy completion of this into the equivalence classes of Cauchy sequences in F results in a non-Archimedean field.

1 Number Systems

1.1 The Rationals

1.1.1 Field Properties

Definition 1.1 (Field)

A **field** is an algebraic structure $(\mathbb{F}, +, \cdot)$ where

1. \mathbb{F} is an abelian group under $+$, with 0 being the *additive identity*.
2. $\mathbb{F} \setminus \{0\}$ is an abelian group under \cdot , with 1 being the *multiplicative identity*.
3. It connects the two operations through the *distributive property*.

$$x \cdot (y + z) = x \cdot y + x \cdot z \quad (7)$$

Lemma 1.1 (Left = Right Distributivity)

Left and right distributivity are equivalent.

$$x \cdot (y + z) = (y + z) \cdot x \quad (8)$$

Proof.

$$\begin{aligned} x \cdot (y + z) &= x \cdot y + x \cdot z && \text{(Distributive)} \\ &= y \cdot x + z \cdot x && \text{(Commutative)} \\ &= (y + z) \cdot x && \text{(Distributive)} \end{aligned}$$

Lemma 1.2 (Properties of Addition)

The properties of addition hold in a field.

1. If $x + y = x + z$, then $y = z$.
2. If $x + y = x$, then $y = 0$.
3. If $x + y = 0$, then $y = -x$.
4. $(-(-x)) = x$.

Proof.

For the first, we have

$$\begin{aligned} x + y = x + z &\implies -x + (x + y) = -x + (x + z) && \text{(addition is a function)} \\ &\implies (-x + x) + y = (-x + x) + z && \text{(+ is associative)} \\ &\implies 0 + y = 0 + z && \text{(definition of additive inverse)} \\ &\implies y = z && \text{(definition of identity)} \end{aligned}$$

For the second, we can set $z = 0$ and apply the first property. For the third, we have

$$\begin{aligned} x + y = 0 &\implies -x + (x + y) = -x + 0 && \text{(addition is a function)} \\ &\implies (-x + x) + y = -x + 0 && \text{(+ is associative)} \\ &\implies 0 + y = -x + 0 && \text{(definition of additive inverse)} \\ &\implies y = -x && \text{(definition of identity)} \end{aligned}$$

For the fourth, we simply follow that if y is an inverse of z , then z is an inverse of y . Therefore, $-x$ being an inverse of x implies that x is an inverse of $-x$. $-(-x)$ must also be an inverse of $-x$. Since inverses are unique^a, $x = -(-x)$.

^aThis is proved in algebra.

Lemma 1.3 (Properties of Multiplication)

The properties of multiplication hold in a field.

1. If $x \neq 0$ and $xy = xz$, then $y = z$.
2. If $x \neq 0$ and $xy = x$, then $y = 1$.
3. If $x \neq 0$ and $xy = 1$, then $y = x^{-1}$.
4. If $x \neq 0$, then $(x^{-1})^{-1} = x$.

Proof.

The proof is almost identical to the first. Since $x \neq 0$, we can always assume that x^{-1} exists. For the first, we have

$$\begin{aligned} xy &= xz \implies x^{-1}(xy) = x^{-1}(xz) && (\text{multiplication is a function}) \\ &\implies (x^{-1}x)y = (x^{-1}x)z && (\times \text{ is associative}) \\ &\implies 1y = 1z && (\text{definition of multiplicative inverse}) \\ &\implies y = z && (\text{definition of identity}) \end{aligned}$$

For the second, we can set $z = 1$ and apply the first property. For the third, we have

$$\begin{aligned} xy &= 1 \implies x^{-1}(xy) = x^{-1}1 && (\text{multiplication is a function}) \\ &\implies (x^{-1}x)y = x^{-1}1 && (\times \text{ is associative}) \\ &\implies 1y = x^{-1}1 && (\text{definition of multiplicative inverse}) \\ &\implies y = x^{-1} && (\text{definition of identity}) \end{aligned}$$

For the fourth, we simply see that x^{-1} is a multiplicative inverse of both x and $(x^{-1})^{-1}$ in the group $(\mathbb{F} \setminus \{0\}, \times)$, and since inverses are unique, they must be equal.

Lemma 1.4 (Properties of Distribution)

For any $x, y, z \in \mathbb{F}$, the field axioms satisfy

1. $0 \cdot x = 0$.
2. If $x \neq 0$ and $y \neq 0$, then $xy \neq 0$.
3. $-1 \cdot x = -x$.
4. $(-x)y = -(xy) = x(-y)$.
5. $(-x)(-y) = xy$.

Proof.

For the first, note that

$$0x = (0 + 0) \cdot x = 0x + 0x \tag{9}$$

and subtracting $0x$ from both sides gives $0 = 0x$. For the second, we can claim that $xy \neq 0$ equivalently claiming that it will have an identity. Since $x, y \neq 0$, their inverses exist, and we claim

that $(xy)^{-1} = y^{-1}x^{-1}$ is an inverse. We can see that by associativity,

$$(y^{-1}x^{-1})(xy) = y^{-1}(x^{-1}x)y = y^{-1}y = 1 \quad (10)$$

For the third, we see that

$$0 = 0 \cdot x = (1 + (-1)) \cdot x = 1 \cdot x + (-1) \cdot x = x + (-1) \cdot x \quad (11)$$

which implies that $-1 \cdot x$ is the additive inverse. The fourth follows immediately from the third by the associative property. For the fifth we can see that

$$\begin{aligned} (-x)(-y) &= (-1)x(-1)y && \text{(property 3)} \\ &= (-1)(-1)xy && (\times \text{ is commutative}) \\ &= -1 \cdot (-xy) && \text{(property 3)} \\ &= -(-xy) && \text{(property 3)} \\ &= xy && \text{(addition property 4)} \end{aligned}$$

Now that we've reviewed some fields, let's construct \mathbb{Q} from \mathbb{Z} and verify it's a field.

Definition 1.2 (Rationals)

Given the ordered ring of integers $(\mathbb{Z}, +_{\mathbb{Z}}, \times_{\mathbb{Z}}, \leq_{\mathbb{Z}})$ the **rational numbers** $(\mathbb{Q}, +_{\mathbb{Q}}, \times_{\mathbb{Q}})$ are defined as such.

1. \mathbb{Q} is the quotient space on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ with the equivalence relation \sim

$$(a, b) \sim (c, d) \iff a \times_{\mathbb{Z}} d = b \times_{\mathbb{Z}} c \quad (12)$$

We denote this class as (a, b) , where $b > 0$, since if $b < 0$, we know that $(-a, -b)$ are also in this order.

2. The additive and multiplicative identities are

$$0_{\mathbb{Q}} := (0_{\mathbb{Z}}, a), \quad 1_{\mathbb{Q}} := (a, a) \quad (13)$$

3. Addition on \mathbb{Q} is defined

$$(a, b) +_{\mathbb{Q}} (c, d) := ((a \times_{\mathbb{Z}} d) +_{\mathbb{Z}} (b \times_{\mathbb{Z}} c), b \times_{\mathbb{Z}} d) \quad (14)$$

4. The additive inverse is defined

$$-(a, b) := (-a, b) \quad (15)$$

5. Multiplication on \mathbb{Q} is defined

$$(a, b) \times_{\mathbb{Q}} (c, d) := (a \times_{\mathbb{Z}} c, b \times_{\mathbb{Z}} d) \quad (16)$$

6. The multiplicative inverse is defined

$$(a, b)^{-1} := (b, a) \quad (17)$$

Theorem 1.1 (Rationals are a Field)

\mathbb{Q} is a field.

Proof.

We do a few things.

1. Verify the additive identity.

$$(a, b) + (0, c) = (ac + 0b, bc) = (ac, bc) \sim (a, b) \quad (18)$$

2. Verify the multiplicative identity.

$$(a, b) \times (c, c) = (ac, bc) \sim (a, b) \quad (19)$$

3. Additive inverse is actually an inverse.

$$(a, b) + (-a, b) = (ab + (-ba), bb) = (0, bb) \sim (0, 1) \quad (20)$$

4. Multiplicative inverse is actually an inverse.

$$(a, b) \times (b, a) = (ab, ba) = (ab, ab) \sim (1, 1) \quad (21)$$

5. Addition is commutative.

$$(a, b) + (c, d) = (ad + bc, bd) = (cb + ad, bd) = (c, d) + (a, b) \quad (22)$$

6. Addition is associative.

$$(a, b) + ((c, d) + (e, f)) = (a, b) + (cf + de, df) \quad (23)$$

$$= (adf + bcf + bde, bdf) \quad (24)$$

$$= (ad + bc, bd) + (e, f) \quad (25)$$

$$= ((a, b) + (c, d)) + (e, f) \quad (26)$$

7. Multiplication is commutative.

$$(a, b) \times (c, d) = (ac, bd) = (ca, db) = (c, d) \times (a, b) \quad (27)$$

8. Multiplication is associative.

$$(a, b) \times ((c, d) \times (e, f)) = (a, b) \times (ce, df) \quad (28)$$

$$= (ace, bdf) \quad (29)$$

$$= (ac, bd) \times (e, f) \quad (30)$$

$$= ((a, b) \times (c, d)) \times (e, f) \quad (31)$$

9. Multiplication distributes over addition.

$$(a, b) \times ((c, d) + (e, f)) = (a, b) \times (c, d) + (a, b) \times (e, f) \quad (32)$$

$$= (ac, bd) + (ae, bf) \quad (33)$$

$$= (abcf + abde, b^2df) \quad (34)$$

$$= (acf + ade, bdf) \quad = (a, b) \times (cf + de, df) \quad (35)$$

We have successfully defined the rationals, but now these are almost completely separate elements. We know that all integers are rational numbers, and so to show that the rationals are an extension of \mathbb{Z} we want to identify a *canonical injection* $\iota : \mathbb{Z} \rightarrow \mathbb{Q}$. This can't just be any canonical injection; it must preserve the algebraic structure between the two sets and must therefore be a *ring homomorphism*.

Theorem 1.2 (Canonical Injection of \mathbb{Z} to \mathbb{Q} is a Ring Homomorphism)

Let us define the canonical injection $\iota : \mathbb{Z} \rightarrow \mathbb{Q}$ to be $\iota(a) = (a, 1)$. This is a ring homomorphism.

Proof.

We show a few things.

1. Preservation of addition.

$$\iota(a) +_{\mathbb{Q}} \iota(b) = (a, 1) +_{\mathbb{Q}} (b, 1) \quad (36)$$

$$= (1a +_{\mathbb{Z}} 1b, 1^2) \quad (37)$$

$$= (a +_{\mathbb{Z}} b, 1) \quad (38)$$

$$= \iota(a +_{\mathbb{Z}} b) \quad (39)$$

2. Preservation of multiplication.

$$\iota(a) \times_{\mathbb{Q}} \iota(b) = (a, 1) \times_{\mathbb{Q}} (b, 1) \quad (40)$$

$$= (a \times_{\mathbb{Z}} b, 1^2) \quad (41)$$

$$= (a \times_{\mathbb{Z}} b, 1) \quad (42)$$

$$= \iota(a \times_{\mathbb{Z}} b, 1) \quad (43)$$

3. Preservation of multiplicative identity.

$$\iota(1_{\mathbb{Z}}) = (1, 1) = 1_{\mathbb{Q}} \quad (44)$$

1.1.2 Ordered Field Properties

Great, so we have established that \mathbb{Q} is a field. The next property we want to formalize is order.

Definition 1.3 (Partial, Total/Linear Order)

A **partial order** on a set X is a relation \leq satisfying.

1. Reflexive: $x \leq x$
2. Antisymmetric: $x \leq y, y \leq x \implies x = y$
3. Transitivity: $x \leq y, y \leq z \implies x \leq z$

Note that when we say $x \leq y$, this means " x is related to y " (but does not necessarily mean that y is related to x), or " x is less than or equal to y ." A set X with a partial order is called a partially ordered set.

Additionally, given elements x, y of partially ordered set X , if either $x \leq y$ or $y \leq x$, then x and y are **comparable**. Otherwise, they are **incomparable**. A partial order in which every pair of elements is comparable is called a **total order**, or **linear order**. Note that from this \leq relation, we can similarly define

1. \leq : less than or equal to
2. \geq : greater than or equal to
3. $<$: strictly less than ($x < y$ iff $x \leq y, x \neq y$)
4. $>$: strictly greater than ($x > y$ iff $x \geq y, x \neq y$)

Example 1.1 (Partially Ordered Sets)

We list some examples of partially ordered sets.

1. The real numbers ordered by the standard "less-than-or-equal" relation \leq (totally ordered set)

- as well).
2. The set of subsets of a given set X ordered by inclusion. That is, the power set 2^X with the partial order \subseteq is partially ordered.
 3. The set of natural numbers equipped with the relation of divisibility.
 4. The set of subspaces of a vector space ordered by inclusion.
 5. For a partially ordered set P , the sequence space containing all sequences of elements from P , where sequence a precedes sequence b if every item in a precedes the corresponding item in b .

We now want to define the natural ordering of the rationals. There are countless ways to do it, but I just take the difference and claim that it is greater than 0.

Theorem 1.3 (Order on Rationals)

The order $\leq_{\mathbb{Q}}$ defined on the rationals as

$$(a, b) \leq_{\mathbb{Q}} (c, d) \iff ad \leq_{\mathbb{Z}} bc \quad (45)$$

is a total order. Remember that we have defined $b, d > 0$.

Proof.

We prove the three properties.

1. Reflexive.

$$(a, b) \leq_{\mathbb{Q}} (a, b) \iff ab \leq_{\mathbb{Z}} ab \quad (46)$$

2. Antisymmetric.

$$(a, b) \leq_{\mathbb{Q}} (c, d) \implies ad \leq_{\mathbb{Z}} bc \quad (c, d) \leq_{\mathbb{Q}} (a, b) \implies bc \leq_{\mathbb{Z}} ad \quad (47)$$

This implies that both $ad = bc$, which by definition means that they are in the same equivalence class.

3. Transitivity. Assume that $(a, b) \leq (c, d)$ and $(c, d) \leq (e, f)$. Then, we notice that $b, d, f > 0$ and therefore by the ordered ring property^a of \mathbb{Z} , we have

$$(a, b) \leq_{\mathbb{Q}} (c, d) \implies ad \leq_{\mathbb{Z}} bc \implies adf \leq_{\mathbb{Z}} bcf \quad (48)$$

$$(c, d) \leq_{\mathbb{Q}} (e, f) \implies cf \leq_{\mathbb{Z}} de \implies bcf \leq_{\mathbb{Z}} bde \quad (49)$$

Therefore from transitivity of the ordering on \mathbb{Z} we have $adf \leq bde$. By the ordered ring property^b we have $0 \leq bde - adf = d(be - af)$. But notice that $d > 0$ from our definition of rationals, and therefore it must be the case that $0 \leq be - af \implies af \leq_{\mathbb{Z}} be$, which by definition means $(a, b) \leq_{\mathbb{Q}} (e, f)$.

^aIf $a \leq b$ and $0 \leq c$, then $ac \leq bc$.

^bIf $a \leq b$, then $a + c \leq b + c$.

As soon as we define an order the concept of extrema and bounds are well defined. Let's define them too.

Definition 1.4 (Extrema, Bounds)

Given a set X ,

1. $x \in X$ is a **maximum** X if $y \leq x$ for all $y \in X$.
2. $x \in X$ is a **minimum** X if $x \leq y$ for all $y \in X$.

Given a totally ordered set X and some subset $S \subset X$.

1. $x \in X$ is an **upper bound** of S if $x \geq y$ for all $y \in S$.

2. $x \in X$ is a **lower bound** of S if $x \leq y$ for all $y \in S$.
3. $x \in X$ is a **supremum**, or **least upper bound**, of S if x is the minimum of the set of all upper bounds of S .
4. $x \in X$ is a **infimum**, or **greatest lower bound**, of S if x is the maximum of the set of all lower bounds of S .

Note that we have defined max/min separately from the concept of bounds. You can define the maximum of a set with just knowing the set, but the bounds require *both* some subset S with respect to an enclosing set X .^a Intuitively, the main difference between the supremum/infimum and maximum/minimum is that the supremum/infimum accounts for limit points of the subset S .

^aFor example, it makes sense to define the maximum of a set $S = [0, 1]$ by itself, but not an upper bound for it. If $X = \mathbb{Q}$, then the supremum is 1, but if X was the set of all irrationals, then this has no supremum.

Note that given a set, we can really put whatever order we want on it. However, consider the field with the following order.

$$\mathbb{F} = \{0, 1\}, 0 < 1 \quad (50)$$

This does not behave well with respect to its operations because for example if we have $0 < 1$, then adding the same element to both sides should preserve the ordering. But this is not the case since $0 + 1 = 1 > 1 + 1 = 0$. While it may be easy to define an order, we would like it to be an ordered field.

Definition 1.5 (Ordered Field)

An **ordered field** is a field that has an order satisfying

1. $y < z \implies x + y < x + z$ for all $x \in \mathbb{F}$.
2. $x > 0, y > 0 \implies xy > 0$.

Theorem 1.4 (Properties)

In an totally ordered field,

1. $x > 0 \implies -x < 0$.
2. $x \neq 0 \implies x^2 > 0$.
3. If $x > 0$, then $y < z \implies xy < xz$.

Proof.

The first property is a single-liner

$$0 < x \implies 0 + -x < x + -x \implies -x < 0 \quad (51)$$

For the second property, it must be the case that $x > 0$ or $x < 0$. If $x > 0$, then by definition $x^2 > 0$. If $x < 0$, then

$$x^2 = 1 \cdot x^2 = (-1)^2 \cdot x^2 = (-1 \cdot x)^2 = (-x)^2 \quad (52)$$

and since $-x > 0$ from the first property, we have $x^2 = (-x)^2 > 0$. For the third, we use the distributive property.

$$y < z \implies 0 < z - y \quad (53)$$

$$\implies 0 = x0 < x(z - y) = xz - xy \quad (54)$$

$$\implies xy < xz \quad (55)$$

As we have hinted, the rationals is an ordered field.

Theorem 1.5 (Rationals are an Ordered Field)

\mathbb{Q} is an ordered field.

Proof.

We show that our defined order satisfies the definition.

1. Assume that $y = (a, b) \leq (c, d) = z$. Let $x = (e, f)$. Then $x + y = (af + be, bf)$, $x + z = (cf + de, df)$. Therefore

$$(af + be)df = adf^2 + bedf \quad (56)$$

$$\leq bcf^2 + bedf \quad (57)$$

$$= (cf + de)bf \quad (58)$$

But $(af + be)df = (cf + de)bf$ is equivalent to saying $(af + be, bf) \leq_{\mathbb{Q}} (cf + de, df)$, i.e. $x + y \leq x + z$!

2. Let $x = (a, b), y = (c, d)$. Since $0 < x, 0 < y$, by construction this means that $0 < a, 0 < c$ (since $b, d > 0$ in the canonical rational form). By the ordered ring property of the integers, $0 < ac$. So

$$0 < ac \iff 0 \cdot bd < ac \cdot 1 \iff (0, 1) < (ac, bd) \iff 0_{\mathbb{Q}} < (a, c) \times_{\mathbb{Q}} (b, d) = xy \quad (59)$$

Not only is it an ordered field, but it also is consistent with the ordering on \mathbb{Z} ! It's nice how all these properties seem to fit together.

Theorem 1.6 (Preservation of Order)

The canonical injection ι is an *order homomorphism*. That is, for $a, b \in \mathbb{Z}$,

$$a \leq_{\mathbb{Z}} b \iff \iota(a) \leq_{\mathbb{Q}} \iota(b) \quad (60)$$

Proof.

$$a \leq_{\mathbb{Z}} b \iff a \cdot 1 \leq_{\mathbb{Z}} b \cdot 1 \quad (61)$$

$$\iff (a, 1) \leq_{\mathbb{Q}} (b, 1) \quad (62)$$

$$\iff \iota(a) \leq_{\mathbb{Q}} \iota(b) \quad (63)$$

Note that an order can be used to generate an order topology, which we will define below.

Definition 1.6 (Order Topology on \mathbb{Q})

The order topology on \mathbb{Q} is the topology generated by the set \mathcal{B} of all open intervals

$$(a, b) := \{x \in \mathbb{Q} \mid a < x < b\} \quad (64)$$

Theorem 1.7 (Finite Fields)

There are no finite ordered fields.

Proof.

Assume \mathbb{F} is such an ordered field. It must be the case that $0, 1 \in \mathbb{F}$, with $0 < 1$. Therefore, we also have $0 + 1 < 1 + 1 \implies 1 < 1 + 1$. Repeating this we get

$$0 < 1 < 1 + 1 < 1 + 1 + 1 < \dots \quad (65)$$

where these elements must be distinct (since only one of $>$, $<$, $=$ must be true for a totally ordered set). Since this can be done for a countably infinite number of times, \mathbb{F} cannot be finite.

1.1.3 Norm

Note that we can also define a norm on the rationals with just the order and algebraic properties.

Theorem 1.8 (Norm on \mathbb{Q})

The following is indeed a norm on \mathbb{Q} .

$$|x| := \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases} \quad (66)$$

It is well known that the metric induced by any norm is indeed a metric. Therefore we state the metric as a definition.

Definition 1.7 (Metric on \mathbb{Q})

The Euclidean metric on \mathbb{Q} is defined

$$d(x, y) := |x - y| = \begin{cases} x - y & \text{if } x \geq y \\ y - x & \text{if } x < y \end{cases} \quad (67)$$

Thus we get to what we want: the induced topology of open balls. Again, since we know from point-set topology that metric topologies are indeed topologies, we will state this as a definition rather than a theorem.

Definition 1.8 (Open-Ball Topology on \mathbb{Q})

The Euclidean topology on \mathbb{Q} is the topology generated by the set \mathcal{B} of all open balls

$$B(x, r) := \{y \in \mathbb{Q} \mid |x - y| < r\} \quad (68)$$

Note that this is the same topology as the order topology. This should however be proved.

Theorem 1.9 (Metric and Order Topologies on \mathbb{Q})

The metric and order topologies on \mathbb{Q} are the same topologies.

Proof.**1.2 The Reals**

By constructing the topology of \mathbb{Q} earlier, we can talk about convergence. The first question to ask (if you were the first person inventing the reals) is “how do I know that there exists some other numbers at all?”

The first clue is trying to find the side length of a square with area 2. As we see, this number is not rational.

Theorem 1.10 ($\sqrt{2}$ is Not Rational)

There exists no $x \in \mathbb{Q}$ s.t. $x^2 = 2$.

Proof.

We can “imagine” that a square with area 2 certainly exists, but the fact that its side length is undefined is certainly unsettling. I don’t know about you, but I would try to “invent” $\sqrt{2}$. We can maybe do this in multiple ways.

1. I write out the decimal expansion one by one, which gives our first exposure to sequences.

$$1, 1.4, 1.41, 1.414, \dots \quad (69)$$

It is clear that on \mathbb{Q} , this sequence does not converge. Our intuition tells that if the terms get closer and closer to each other, they must be getting closer and closer to *something*, though that something is not in \mathbb{Q} . This motivates the definition for *Cauchy completeness*.

2. I would write out maybe some nested intervals so that $\sqrt{2}$ *must* lie within each interval.

$$[1, 2] \supset [1.4, 1.5] \supset [1.41, 1.42] \supset \dots \quad (70)$$

This motivates the definition of *nested-interval completeness*.

3. I would define the set of all rationals such that $x^2 < 2$, and try to define $\sqrt{2}$ as the max or supremum of this set. We will quickly find that neither the max nor the supremum exists in \mathbb{Q} , and this motivates the definition for *Dedekind completeness*.

All three of these methods points at the same intuition that there should not be any "gaps" or "missing points" in the set that we will construct to be \mathbb{R} . This contrasts with the rational numbers, whose corresponding number line has a "gap" at each irrational value.

1.2.1 Dedekind Completeness

Definition 1.9 (Dedekind Cut)

A **Dedekind cut** is a partition of the rationals $\mathbb{Q} = A \sqcup A'$ satisfying the three properties.^a

1. $A \neq \emptyset$ and $A \neq \mathbb{Q}$.^b
2. $x < y$ for all $x \in A, y \in A'$.
3. The maximum element of A does not exist in \mathbb{Q} .

The minimum of A' may exist in \mathbb{Q} , and if it does, the cut is said to be **generated** by $\min A'$.

^aThis can really be defined for any totally ordered set.

^bBy relaxing this property, we can actually complete \mathbb{Q} to the extended real number line.

Note that in \mathbb{Q} , there will be two types of cuts:

1. ones that are generated by rational numbers, such as

$$A = \{x \in \mathbb{Q} \mid x < 2/3\}, A' = \{x \in \mathbb{Q} \mid x \geq 2/3\} \quad (71)$$

2. and the ones that are not

$$A = \{x \in \mathbb{Q} \mid x^2 < 2\}, A' = \{x \in \mathbb{Q} \mid x^2 \geq 2\} \quad (72)$$

We can intuitively see that the set of all Dedekind cuts (A, A') will “extend” the rationals into a bigger set. We can then define some operations and an order to construct this into an ordered field, and finally it will have the property that we call “completeness.”

Definition 1.10 (Dedekind Completeness)

A totally ordered algebraic field \mathbb{F} is **complete** if every Dedekind cut of \mathbb{F} is generated by an element of \mathbb{F} .

Theorem 1.11 ()

\mathbb{Q} is not Dedekind-complete.

Proof.

The counter-example is given above for the cut

$$A = \{x \in \mathbb{Q} \mid x^2 < 2\}, A' = \{x \in \mathbb{Q} \mid x^2 \geq 2\} \quad (73)$$

Now we have the tools to define the reals, giving us the beefy theorem.

Theorem 1.12 (Reals as the Dedekind-Completion of Rationals)

Let \mathbb{R} be the set of all Dedekind cuts (A, A') of \mathbb{Q} of \mathbb{Q} . For convenience we can uniquely represent (A, A') with just A since $A' = \mathbb{Q} \setminus A$. By doing this we can intuitively think of a real number as being represented by the set of all smaller rational numbers. Let A, B be two Dedekind cuts. Then, we define the following order and operations.

1. *Order.* $A \leq_{\mathbb{R}} B \iff A \subset B$.
2. *Addition.* $A +_{\mathbb{R}} B := \{a +_{\mathbb{Q}} b \mid a \in A, b \in B\}$.
3. *Additive Identity.* $0_{\mathbb{R}} := \{x \in \mathbb{Q} \mid x < 0\}$.
4. *Additive Inverse.* $-B := \{a - b \mid a < 0, b \in (\mathbb{Q} \setminus B)\}$.
5. *Multiplication.* If $A, B \geq 0$, then we define $A \times_{\mathbb{R}} B := \{a \times_{\mathbb{Q}} b \mid a \in A, b \in B, a, b \geq 0\} \cup 0_{\mathbb{R}}$. If A or B is negative, then we use the identity $A \times B = -(A \times_{\mathbb{R}} -B) = -(-A \times_{\mathbb{R}} B) = (-A \times_{\mathbb{R}} -B)$ to convert A, B to both positives and apply the previous definition.
6. *Multiplicative Identity.* $1_{\mathbb{R}} = \{x \in \mathbb{Q} \mid x < 1\}$.
7. *Multiplicative Inverse.* If $B > 0$, $B^{-1} := \{a \times_{\mathbb{Q}} b^{-1} \mid a \in 1_{\mathbb{R}}, b \in (\mathbb{Q} \setminus B)\}$. If B is negative, then we compute $B^{-1} = -((-B)^{-1})$ by first converting to a positive number and then applying the definition above.

We claim that $(\mathbb{R}, +_{\mathbb{R}}, \times_{\mathbb{R}}, \leq_{\mathbb{R}})$ is a totally ordered field, and the canonical injection $\iota : \mathbb{Q} \rightarrow \mathbb{R}$ defined

$$\iota(q) = \{x \in \mathbb{Q} \mid x < q\} \quad (74)$$

is an ordered field isomorphism. Finally, by construction \mathbb{R} is Dedekind-complete.

Definition 1.11 (Least Upper Bound Property)

A totally ordered algebraic field \mathbb{F} (must it be a field?) is complete if every nonempty set of F having an upper bound must have a least upper bound (supremum) in F .

Theorem 1.13 (LUB is Equivalent to GLB)

A set (X, \leq) has the least upper bound property iff it has the greatest lower bound property.^a

^aEvery set bounded below has a greatest lower bound.

Proof.

We will prove one direction since the other is the same logic. Let $S \subset X$ be a nonempty set that is bounded below by some $l \in X$. Let $L \subset X$ be the set of all lower bounds of S . Since l exists, it is nonempty. Furthermore, L is bounded above by any element of S . Due to LUB property L has a least upper bound, call it $z = \sup L$. We claim that $z = \inf S$.

1. z is a lower bound of S . Assume that it is not. Then there exists $s \in S$ s.t. $s < z$. But by construction s is an upper bound for L and so z is not the *least* upper bound, a contradiction.
2. z is a *greatest* lower bound. Assume that z is not. Then there exists a $z' \in X$ s.t. $z < z' \leq s$ for all $s \in S$. But since z', z are lower bounds, this means $z, z' \in L$ by definition and $z < z'$ contradicts the fact that z is an upper bound of L .

We are done.

Theorem 1.14 (Dedekind Completeness Equals Least-Upper-Bound Property)

Dedekind completeness is equivalent to the least upper bound property.

Proof.

Definition 1.12 (Archimedean Principle)

An ordered ring $(X, +, \cdot, \leq)$ that embeds the naturals \mathbb{N}^a is said to obey the **Archimedean principle** if given any $x, y \in X$ s.t. $x, y > 0$, there exists an $n \in \mathbb{N}$ s.t. $\iota(n) \cdot x > y$. Usually, we don't care about the canonical injection and write $nx > y$.

^aas in, there exists an ordered ring homomorphism $\iota : \mathbb{N} \rightarrow X$

By the canonical injections $\mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q} \rightarrow \mathbb{R}$, we can talk about whether this set has the Archimedean property. In fact Dedekind completeness does imply it.

Theorem 1.15 ()

\mathbb{R} satisfies the Archimedean principle.

Proof.

Assume that this property doesn't hold. Then for any fixed x , $nx < y$ for all $n \in \mathbb{N}$. Consider the set

$$A = \bigcup_{n \in \mathbb{N}} (-\infty, nx), \quad B = \mathbb{R} \setminus A \tag{75}$$

A by definition is nonempty, and B is nonempty since it contains y . Then, we can show that $a \in A, b \in B \implies a < b$ using proof by contradiction. Assume that there exists $a' \in A, b' \in B$ s.t. $a' > b'$. Since $a' \in A$, there exists a $n' \in \mathbb{N}$ s.t. $a' \in (-\infty, n'x) \iff a' < n'x$. But by transitivity of order, this means $b' < n'x \iff b' \in (-\infty, n'x) \implies b' \in A$.

Going back to the main proof, we see that A is upper bounded by y , and so by the least upper bound property it has a supremum $z = \sup A$.

1. If $z \in A$, then by the induction principle^a $z + x \in A$, contradicting that z is an upper bound.
2. If $z \notin A$, then by the induction principle^b $z - x \notin A \implies z - x \in B$. Since every element of B upper bounds A and since $x > 0$, this means that $z - x < z$ is a smaller upper bound of A ,

contradicting that z is a least upper bound.
Therefore, it must be the case that $nx > y$ for some $n \in \mathbb{N}$.

^aNote that \mathbb{N} is defined recursively as $1 \in \mathbb{N}$ and if $n \in \mathbb{N}$, then $n + 1 \in \mathbb{N}$.

^bThe contrapositive of the recursive definition of \mathbb{N} is: if $n \notin \mathbb{N}$, then $n - 1 \notin \mathbb{N}$.

1.2.2 Cauchy Completeness

Definition 1.13 (Cauchy Sequence)

A sequence a_n in a metric space (X, d) is a **Cauchy sequence** if for every $\epsilon > 0$, there exists an N s.t.

$$d(a_i, a_j) < \epsilon \quad (76)$$

for every $i, j > N$. We call this **Cauchy convergence**.

Note that it is not sufficient to say that a sequence is Cauchy by claiming that each term becomes arbitrarily close to the preceding term. That is,

$$\lim_{n \rightarrow \infty} |x_{n+1} - x_n| = 0 \quad (77)$$

For example, look at the sequence

$$a_n = \sqrt{n} \implies a_{n+1} - a_n = \frac{1}{\sqrt{n+1} + \sqrt{n}} < \frac{1}{2\sqrt{n}} \quad (78)$$

However, it is clear that a_n gets arbitrarily large, meaning that a finite interval can contain at most a finite number of terms in $\{a_n\}$.

It is trivial that convergence implies Cauchy convergence, but the other direction is not true. Therefore, we would like to work in a space where these two are equivalent, and this is called completeness.

Definition 1.14 (Cauchy Completeness)

A metric space (X, d) is complete if every Cauchy sequence in that space converges to an element in X .

Theorem 1.16 ()

\mathbb{Q} is not Cauchy-complete.

Proof.

Let a_n be the largest number x up to the n th decimal expansion such that x^2 does not exceed 2. The first few terms are

$$1.4, 1.41, 1.414, \dots \quad (79)$$

Therefore, we can construct the reals as equivalence classes over Cauchy sequences. Rather than using the order, we take advantage of the metric.

Theorem 1.17 (Reals as the Cauchy-Completion of the Rationals)

Let \mathbb{R} be the quotient space of all Cauchy sequences (x_n) of rational numbers with the equivalence relation $(x_n) = (y_n)$ iff their difference tends to 0.^a That is, for every rational $\epsilon > 0$, there exists an integer N s.t. for all naturals $n > N$, $|x_n - y_n| < \epsilon$.

1. Order. $(x_n) \leq_{\mathbb{R}} (y_n)$ iff $x = y$ or there exists $N \in \mathbb{N}$ s.t. $x_n \leq_{\mathbb{Q}} y_n$ for all $n > N$.

2. *Addition.* $(x_n) + (y_n) := (x_n + y_n)$.
3. *Additive Identity.* $0_{\mathbb{R}} := (0_{\mathbb{Q}})$.
4. *Additive Inverse.* $-(x_n) := (-x_n)$.
5. *Multiplication.* $(x_n) \times_{\mathbb{R}} (y_n) = (x_n \times_{\mathbb{Q}} y_n)$.
6. *Multiplicative Identity.* $1_{\mathbb{R}} := (1)$.
7. *Multiplicative Inverse.* $(x_n)^{-1} := (x_n^{-1})$.

We claim that $(\mathbb{R}, +_{\mathbb{R}}, \times_{\mathbb{R}}, \leq_{\mathbb{R}})$ is a totally ordered field, and the canonical injection $\iota : \mathbb{Q} \rightarrow \mathbb{R}$ defined

$$\iota(q) = (q) \quad (80)$$

is an ordered field isomorphism. Finally, by construction \mathbb{R} is Cauchy-complete.

^aThis equivalence class reflects that the same real number can be approximated in many different sequences. In fact, this shows *by definition* that $1, 1, \dots$ and $0.9, 0.99, 0.999, \dots$ are the same number!

1.2.3 Nested Intervals Completeness

The final way we prove is using nested-intervals completeness.

Definition 1.15 (Nested Interval Completeness, Cantor's Intersection Theorem)

Let F be a totally ordered algebraic field. Let $I_n = [a_n, b_n]$ ($a_n < b_n$) be a sequence of closed intervals, and suppose that these intervals are nested in the sense that

$$I_1 \supset I_2 \supset I_3 \supset \dots$$

where

$$\lim_{n \rightarrow +\infty} b_n - a_n = 0$$

F is complete if the intersection of all of these intervals I_n contains exactly one point. That is,

$$\bigcap_{n=1}^{\infty} I_n \in F$$

Note that defining nested intervals requires only an ordered field. One may look at this and try to ask if this is a specific instance of the following conjecture: The intersection of a nested sequence of nonempty closed sets in a topological space has exactly 1 point. This claim may not even make sense, actually. If we define nested in terms of proper subsets, then for a finite topological space a sequence cannot exist since we will run out of open sets and so this claim is vacuously true and false. If we allow $S_n = S_{n+1}$ then we can just select $X \supset X \supset \dots$, which is obviously not true. However, a slightly weaker claim is that every proper nested non-empty closed sets has a non-empty intersection is a consequence of compactness.

Theorem 1.18 ()

\mathbb{Q} is not nested-interval complete.

Proof.

Consider the intervals $[a_i, b_i]$ where a_i is the largest number x up to the n th decimal expansion such that x^2 does not exceed 2, and b_i is the smallest number x up to the n th decimal expansion such that x^2 is not smaller than 2. The first few terms are

$$[1.4, 1.5], [1.41, 1.42], [1.414, 1.415], \dots \quad (81)$$

Therefore, we can complete \mathbb{Q} . It turns out that this is equivalent to the construction using Dedekind cuts, and by definition this new set is nested interval complete. However, like Cauchy completeness, this actually does not imply the Archimedean property.

1.2.4 Properties of the Real Line

Now that we have completed it, we can define the real numbers.

Definition 1.16 (The Real Numbers)

The **set of real numbers**, denoted \mathbb{R} , is a totally ordered complete Archimedean field.

It seems that the real numbers is *any* set that satisfies the definition above. Does this mean that there are multiple real number lines?

Example 1.2 (Multiple Reals?)

For example, let us construct three distinct sets satisfying these axioms:

1. A line \mathbb{L} with $+$ associated with the translation of \mathbb{L} along itself and \cdot associated with the "stretching/compressing" of the line around the additive origin 0.
2. An uncountable list of numbers with possibly infinite decimal points, known as the decimal number system.

$$\dots, -2.583 \dots, \dots, 0, \dots, 1.2343 \dots, \dots, \sqrt{2}, \dots \quad (82)$$

3. A circle with a point removed, with addition and multiplication defined similarly as the line.

We will show that there is only one set, up to isomorphism, that satisfies all these properties.

Theorem 1.19 (Uniqueness)

\mathbb{R} is unique up to field isomorphism. That is, if two individuals construct two ordered complete Archimedean fields \mathbb{R}_A and \mathbb{R}_B , then

$$\mathbb{R}_A \simeq \mathbb{R}_B \quad (83)$$

Proof.

The proof is actually much longer than I expected, so I draw a general outline.^a We want to show how to construct an isomorphism $f : \mathbb{R}_A \rightarrow \mathbb{R}_B$.

1. Realize that there are unique embeddings of \mathbb{N} in \mathbb{R}_A and \mathbb{R}_B that preserve the inductive principle, the order, closure of addition, and closure of multiplication, the additive identity, and the multiplicative identity. Call these ordered doubly-monoid (since it's a monoid w.r.t. $+$ and \times) homomorphisms ι_A, ι_B .
2. Construct an isomorphism $f_1 : \iota_A(\mathbb{N}) \rightarrow \iota_B(\mathbb{N})$ that preserves the inductive principle, order, addition, and multiplication. This is easy to do by just constructing $f_1 = \iota_B \circ \iota_A^{-1}$.
3. Extend f_1 to the ordered ring isomorphism f_2 by explicitly defining what it means to map additive inverses, i.e. negative numbers.
4. Extend f_2 to the ordered field isomorphism f_3 by explicitly defining what it means to map multiplicative inverses, i.e. reciprocals.
5. Extend f_3 to the ordered field isomorphism on the entire domain \mathbb{R}_A and codomain \mathbb{R}_B . There is no additional operations that we need to support, but we should explicitly show that this is both injective and surjective, which completes our proof.

^aFollowed from here.

Corollary 1.1 (Dedekind and Cauchy Completeness are Equivalent for Reals)

Let $\mathbb{R}_D, \mathbb{R}_C$ be the Dedekind and Cauchy completion of \mathbb{Q} . Then $\mathbb{R}_D \simeq \mathbb{R}_C$.^a

^aNote that this is only true for totally ordered Archimedean fields! The two completeness properties are not equal in general!

The second new property is that the reals are uncountable.

Theorem 1.20 (Cantor's Diagonalization)

The real numbers are uncountable.

Proof.

We proceed by contradiction. Suppose the real numbers are countable. Then there exists a bijection $f : \mathbb{N} \rightarrow \mathbb{R}$. This means we can list all real numbers in $[0, 1]$ as an infinite sequence.^a

$$f(1) = 0.a_{11}a_{12}a_{13}\dots$$

$$f(2) = 0.a_{21}a_{22}a_{23}\dots$$

$$f(3) = 0.a_{31}a_{32}a_{33}\dots$$

⋮

where each a_{ij} is a digit between 0 and 9.

Now construct a new real number $r = 0.r_1r_2r_3\dots$ where:

$$r_n = \begin{cases} 1 & \text{if } a_{nn} \neq 1 \\ 2 & \text{if } a_{nn} = 1 \end{cases} \quad (84)$$

This number r is different from $f(n)$ for every $n \in \mathbb{N}$, since r differs from $f(n)$ in the n th decimal place. Therefore $r \in [0, 1]$ but $r \notin \text{range}(f)$, contradicting that f is surjective. Thus our assumption that the real numbers are countable must be false.

^aThis must be explicitly proven, but we can take the set of all Cauchy sequences of rationals in their decimal expansion and construct the reals this way.

Provide examples of ordered, Cauchy-complete fields that are not Archimedean.

1.2.5 Roots, Exponentials, and Logarithms

Now we will focus on some other operations that become well-defined in the reals. We know that x^n for $n \in \mathbb{N}$ denotes repeated multiplication and x^{-1} denotes the multiplicative inverse. We need to build up on this notation. As a general outline, we will show that x^{-n} is well defined, then $x^q, q \in \mathbb{Q}$ is well-defined, and finally $x^r, r \in \mathbb{R}$ is well-defined. For the naturals, we have defined x^n as the repeated multiplication of n . It is trivial that the canonical injection $\iota_0 : \mathbb{N} \rightarrow \mathbb{R}$ commutes with the exponential map of naturals. We prove that $\iota_1 : \mathbb{Z} \rightarrow \mathbb{R}$ also commutes.

Lemma 1.5 (Integer Exponents)

We have

1. For $x_1, \dots, x_n \in \mathbb{R}$, $(x_1 \dots x_n)^{-1} = x_n^{-1} \dots x_1^{-1}$.
2. For $x \in \mathbb{R}$, $x > 0$, $(x^n)^{-1} = (x^{-1})^n$. This value is denoted x^{-n} .
3. For $x \in \mathbb{R}$ and $w, z \in \mathbb{Z}$, $x^{w+z} = x^w x^z$.

4. For $w, z \in \mathbb{Z}$, $x^{wz} = (x^z)^w = (x^w)^z$.

Proof.

Listed.

1. The proof is trivial, but for $n = 2$ and $x_1 = x, x_2 = y$, we see that by associativity, $(x^{-1}x^{-1})(xy) = y^{-1}(x^{-1}x)y = y^{-1}y = 1$ and we know inverses are unique.
2. Set $x_i = x$ using (1).
3. If $w, z > 0$ this is trivial by the associative property. If either or both are negative, say $w < 0 < z$, then we set $w' = -w > 0$ and using (2) we know that

$$x^w x^z = (x^{-1})^{w'} x^z = x^{-w'+z} = x^{w+z} \quad (85)$$

by associativity in the second last equality.

Therefore, we have successfully defined x^z for all $z \in \mathbb{Z}$, and if z is negative, we're allowed to “swap” the -1 and $|z|$ in the exponents. Now we want to extend this into rational exponents, first by proving the existence and uniqueness of n th roots for any real. The proof is a little involved, but the general idea is that we want to use the LUB property to define the n th root as the supremum of a set.

Theorem 1.21 (Existence of Nth Roots)

For any real $x > 0$ and every $n \in \mathbb{N}$ there is one and only one positive real $y \in \mathbb{R}$ s.t. $y^n = x$. This is denoted $x^{1/n}$.

Proof.

Let E be the set consisting of all reals $t \in \mathbb{R}$ s.t. $t^n < x$. We show that

1. it is nonempty. Consider $t = x/(1+x)$. Then $0 \leq t < 1 \implies t^n \leq t < x$. Thus $t \in E$ and E is nonempty.
2. it is bounded. Consider any number $s = 1 + x$. Then $s^n \geq s > x$, so $s \notin E$, and $s = 1 + x$ is an upper bound of E .

Therefore, E is a nonempty set that is upper bounded, so it has a least upper bound, called $y = \sup E$. We claim that $y^n = x$, proving by contradiction. For both cases, we use the fact that the identity $b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \dots + a^{n-1})$ gives the inequality

$$b^n - a^n < (b-a)n b^{n-1} \text{ for } 0 < a < b \quad (86)$$

1. Assume $y^n < x$. Then we choose a fixed $0 < h < 1$ s.t.

$$h < \frac{x - y^n}{n(y+1)^{n-1}} \quad (87)$$

Then by putting $a = y, b = y + h$, we have

$$(y+h)^n - y^n < hn(y+h)^{n-1} < hn(y+1)^{n-1} < x - y^n \quad (88)$$

and thus $y^n < (y+h)^n < x$. This means that $y+h \in E$, and so y is not an upper bound.

2. Assume $y^n > x$. Then we set a fixed number

$$k = \frac{y^n - x}{ny^{n-1}} \quad (89)$$

Then $0 < k < y$. If we take any $t \in \mathbb{R}$ s.t. $t \geq y-k$, this implies that $t^n \geq (y-k)^n \implies -t^n \geq -(y-k)^n$, and so

$$y^n - t^n \leq y^n - (y-k)^n < kny^{n-1} = y^n - x \quad (90)$$

Thus $t^n > x$ and $t \notin E$. So it must be the case that $t < y - k$, and so $y - k$ is an upper bound of E , contradicting that y is least.

At this point, rooting has been introduced as sort of an independent map from exponentiation. We show that they have the nice property of commuting.

Lemma 1.6 (Rooting and Exponentiation Commute)

For $p \in \mathbb{Z}, q \in \mathbb{N}$ and $x \in \mathbb{R}$ with $x > 0$, we have

$$(x^p)^{1/q} = (x^{1/q})^p \quad (91)$$

Proof.

If $p > 0$, then let $r = (x^p)^{1/q}$. By definition $r^q = x^p$. Let $s = x^{1/q}$. By definition $s^q = x$. Therefore $r^q = (s^q)^p = s^{qp}$ from the lemma on integer exponents. But since roots are well-defined and unique

$$r = (r^q)^{1/q} = (s^{qp})^{1/q} = s^p \implies (x^p)^{1/q} = (x^{1/q})^p \quad (92)$$

If $p = 0$, this is trivially 0, and if $p < 0$ the by the same logic with $p = -p'$ for $p' > 0$ and $y = x^{-1} > 0$. we know

$$(x^p)^{1/q} = ((y^{-1})^{-p'})^{1/q} = (y^{-(-p')})^{1/q} = (y^{p'})^{1/q} \quad (93)$$

$$= (y^{1/q})^{p'} = ((x^{-1})^{1/q})^{p'} = (x^{1/q})^{-p'} = (x^{1/q})^p \quad (94)$$

Theorem 1.22 (Rational Exponential Function)

Given $m, p \in \mathbb{Z}$ and $n, q \in \mathbb{N}$, prove that

$$(b^m)^{1/n} = (b^p)^{1/q} \quad (95)$$

Hence it makes sens to define $b^r = (b^m)^{1/n}$, since every element of the equivalence class r of each rational number maps to the same value.

Proof.

Since $m/n = p/q \implies mq = np$,

$$b^{mq} = b^{np} \implies (b^m)^q = (b^p)^n \quad (96)$$

$$\implies b^m = ((b^m)^q)^{1/q} = ((b^p)^n)^{1/q} \quad (97)$$

$$\implies b^m = ((b^p)^{1/q})^n \quad (98)$$

$$\implies (b^m)^{1/n} = (b^p)^{1/q} \quad (99)$$

Therefore we can define for any $r \in \mathbb{Q}$

$$x^r = x^{m/n} = (x^m)^{1/n} = (x^{1/n})^m \quad (100)$$

where the final equality holds from the commutativity of rooting and exponentiation.

It turns out that this is a homomorphism.

Corollary 1.2 (Rational Exponential Function is a Homomorphism)

The rational exponential function is a homomorphism. That is, given $r, s \in \mathbb{Q}$ and $x \in \mathbb{R}$,

$$x^{r+s} = x^r \cdot x^s \quad (101)$$

Proof.

Let $r = m/n, s = p/q$. Then

$$\begin{aligned} x^{r+s} &= x^{m/n+p/q} = x^{\frac{mq+np}{nq}} && (\text{addition on } \mathbb{Q}) \\ &= (x^{mq+np})^{1/nq} && (\text{exp and roots commute}) \\ &= (x^{mq} + x^{np})^{1/nq} && (\text{int exp lemma}) \\ &= (x^{mq})^{1/nq} (x^{np})^{1/nq} && (\text{int exp lemma}) \\ &= x^{mq/nq} x^{np/nq} && (\text{exp and roots commute}) \\ &= x^{m/n} x^{p/q} && (\text{relation from } \mathbb{Q}) \end{aligned}$$

With rational exponents defined, we can use the least upper bound property to define a consistent extension of a real exponent.

Lemma 1.7 ()

If $r \in \mathbb{Q}$ with $r \geq 0$, then for $x \in \mathbb{R}, x > 1, 1 \leq b^r$.

Proof.

Let $r = m/n$. Then $x^r = x^{m/n} = (x^m)^{1/n}$. Since $1 < x$, and $m \geq 0$, we have

$$1 \leq x \leq x^2 \leq \dots \leq x^m \implies 1 \leq b^m \quad (102)$$

Now set $y = x^{m/n}$ and assume that $y < 1$. Then

$$x^m = y^n < y^{n-1} < \dots < y < 1 \quad (103)$$

and so $x^m < 1$, which is a contradiction. So it must be the case that $y > 1$.

Lemma 1.8 (Monotonicity of Rational Exponents)

If $x, y \in \mathbb{R}$, then for any rational $r \in \mathbb{Q}$ with $r < x + y$, there exists a $p, q \in \mathbb{Q}$ s.t. $p < x, q < y$ and $p + q = r$. The converse is true as well.

Proof.

$r < x + y \implies r - y < x$. By density of \mathbb{Q} in \mathbb{R} , we can choose $r - y < p < x$. Then $-r + y > -p > x \implies r - r + y > r - p > r - x \implies y > r - p > r - x$, and we set $q = r - p$. We are done. The converse is trivial since given $p, q \in \mathbb{Q}$ with $p < x, q < y$, then by the ordered field properties $p + q < x + y$.

Corollary 1.3 (Real Exponential Function)

Given $x \in \mathbb{R}$, we define

$$B(x) := \{x^q \in \mathbb{R} \mid q \in \mathbb{Q}, q \leq x\} \quad (104)$$

We claim that given $r \in \mathbb{R}$,

$$x^r := \sup B(r) \quad (105)$$

is well-defined and is a homomorphism extension of the rational exponential function. That is,

$$\sup B(x+y) = \sup B(x) \cdot \sup B(y) \quad (106)$$

Proof.

To show that $x^r := \sup B(r)$ where $B(r) = \{x^t \in \mathbb{R} \mid t \in \mathbb{Q}, t \leq r\}$,

1. We show it's an upper bound. Assume it wasn't. Then $x^r < x^t$ for some $t \in \mathbb{Q}$ satisfying $t \leq r$. But $t \leq r \implies 0 \leq r - t$, and by the previous lemma, $1 \leq x^{r-t}$. So $1 \leq x^{r-t} = x^r x^{-t} = x^r (x^t)^{-1} \implies x^t \leq x^r$, which is a contradiction.
2. We show that it is least. Assume that it is not. Then $\exists r' \in \mathbb{Q}$ s.t. $x^t \leq x^{r'}$ and $r' < r$. Now let $s \in \mathbb{Q}$ be an element between r' and r , which is guaranteed to exist due to density of rationals in reals. But $s < r$, so by definition $x^s \in B(r)$, but

$$0 < s - r' \implies 1 < b^{s-r'} \quad (107)$$

$$\implies b^{r'}(b^{r'})^{-1} < b^s(b^{-r'}) \quad (108)$$

$$\implies 1 < b^s(b^{r'})^{-1} \quad (109)$$

$$\implies b^{r'} < b^s \quad (110)$$

and so $b^{r'}$ is not an upper bound for $B(r)$. By contradiction, b^r is least.

Since this is defined, the analogous definition for real numbers is consistent with that of the rationals, and it is upper bounded by the Archimedean principle, so such a supremum must exist. Note that t is rational. For the second part, from the previous lemma and the homomorphism properties of the rational exponent,

$$B(x+y) = B'(x+y) := \{b^{p+q} \in \mathbb{R} \mid p, q \in \mathbb{Q}, p \leq x, q \leq y\} \quad (111)$$

$$= \{b^p b^q \in \mathbb{R} \mid p, q \in \mathbb{Q}, p \leq x, q \leq y\} \quad (112)$$

$$(113)$$

Therefore we can treat B and B' as the same set.

1. Prove upper bound $\sup B(x+y) \leq \sup B(x) \sup B(y)$. Given $\alpha \in B'(x+y)$, there exists $p_\alpha, q_\alpha \in \mathbb{Q}$ (with $p_\alpha < x, q_\alpha < y$) s.t. $b^{p_\alpha} b^{q_\alpha} = \alpha$. But

$$b^{p_\alpha} b^{q_\alpha} \leq \sup_{p_\alpha} \{b^{p_\alpha}\} \cdot \sup_{q_\alpha} \{b^{q_\alpha}\} = \sup B(x) \sup B(y) \quad (114)$$

2. To prove least, assume there exists $K \in \mathbb{R}$ s.t. $\sup B'(x+y) \leq K < \sup B(x) \sup B(y)$. Then, since the image of b^x is always positive, we assume $0 < K$. We bound its factors as so: $K < \sup B(x) \sup B(y) \implies K/\sup B(x) < \sup B(y)$. By density of the rationals, there exists a $\beta \in \mathbb{Q}$, s.t.

$$\frac{K}{\sup B(x)} < \beta < \sup B(y) \quad (115)$$

This means $K/\beta < \sup B(x)$ and $\beta < \sup B(y)$. But this means that there exists $\phi, \gamma \in B(x), B(y)$ s.t. $K/\beta < \phi, \beta < \gamma \implies K = (K/\beta) \cdot \beta < \phi\gamma \implies \phi\gamma \in B'(x+y)$ by definition. So K is not an upper bound.

Furthermore, this is an isomorphism, and the inverse is defined. Let's define this analytically.

Theorem 1.23 (Logarithm)

For $b > 1$ and $y > 0$, there is a unique real number x s.t. $b^x = y$. We claim

$$x = \sup\{w \in \mathbb{R} : b^w < y\} \quad (116)$$

x is called the **logarithm of y to the base b** .

Proof.

We use the inequality $b^n - 1 \leq n(b - 1)$ for all $n \in \mathbb{N}$.^a By substituting $b = b^{1/n}$ (valid since $b > 1 \iff b^{1/n} > 1$) so $b - 1 \geq n(b^{1/n} - 1)$. Now set some $t > 1$, and by Archimedean principle, we can choose some $n \in \mathbb{N}$ s.t. $n > \frac{b-1}{t-1}$. Then $n(t-1) > b-1$, and with the inequality derived we get

$$n(t-1) > b-1 \geq n(b^{1/n} - 1) \implies t > b^{1/n} \quad (117)$$

This allows us to prove 2 things.

1. If w satisfies $b^w < y$, then $b^{w+(1/n)} < y$ for sufficiently large n . Setting $t = yb^{-w}$ (which is greater than 1 since $b^w < y$) gives $y \cdot b^{-w} > b^{1/n} \implies b^w b^{1/n} < y \implies b^{w+(1/n)} < y$.
2. If w satisfies $b^w > y$, then $b^{w-(1/n)} > y$ for sufficiently large n . Setting $t = b^w y^{-1}$ (which is greater than 1 since $b^w > y$) gives $b^w y^{-1} > b^{1/n} \implies b^{w-(1/n)} > y$.

Now we can prove existence. Let A the set of all w s.t. $b^w < y$. We claim that $x = \sup A$.

1. Assume that $b^x < y$. We know that there exists $n \in \mathbb{N}$ s.t. $b^{x+(1/n)} < y \implies x + (1/n) \in A$, contradicting that x is an upper bound.
2. Assume that $b^x >$. We know that there exists $n \in \mathbb{N}$ s.t. $b^{x-(1/n)} > y \implies x - (1/n)$ is also an upper bound for A , contradicting that x is least. Therefore $b^x = y$.

We now prove uniqueness. Assume that there are two such x 's, call them x, x' . By total ordering and $x \neq x'$, WLOG let $x > x' \implies x - x' > 0 \implies b^{x-x'} > 1$. By density of rationals, since we can choose $r \in \mathbb{R}$ s.t. $0 < r < x - x'$, we have $1 < b^r < b^{x-x'}$ and so $B(r) \subset B(x - x')$. Since $1 < b^{x-x'} \implies 1 \cdot b^{x'} < b^{x-x'} \cdot b^x = b^x$, we have $b^{x'} < b^x$ and they cannot both by y . So $x = x'$.

^aWe prove by induction. For $n = 1$ $b^1 - 1 \leq 1(b - 1)$. Assume that this holds for some n . Then $b^{n+1} - 1 = b^{n+1} - b + b - 1 = b(b^n - 1) + (b - 1) \geq bn(b - 1) + (b - 1) = (bn + 1)(b - 1) \geq (n + 1)(b - 1)$, where the last step follows since $b \geq 1 \implies bn \geq n \implies bn + 1 \geq n + 1$.

1.3 The Extended Reals and Hyperreals

Great! We have officially constructed the reals, and we can finally feel satisfied about defining metrics, norms, and inner products as mappings to the codomain \mathbb{R} . Now let's make the concept of infinite numbers a bit more rigorous. In short, what we do is just add the numbers $\pm\infty$ to \mathbb{R} , which we call the extended reals, and try and extend the properties from \mathbb{R} to the extended reals. We will see that not all properties can be transferred.

Theorem 1.24 (Extended Real Number Line)

The **extended real number line** is the set $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. We define the following operations.

1. *Order.* $-\infty \leq x$ and $x \leq +\infty$ for all $x \in \bar{\mathbb{R}}$.
2. *Addition.* $+\infty - \infty = 0$. $x + \infty = +\infty$ and $x - \infty = -\infty$ for all $x \in \mathbb{R}$.

3. *Multiplication.*

$$x \times \infty = \begin{cases} +\infty & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\infty & \text{if } x < 0 \end{cases} \quad (118)$$

It turns out that this is still Dedekind-complete, which is nice. Unfortunately we lose a lot of structure.

1. this is not even a field since the multiplicative inverse of $\pm\infty$ is not defined.
2. the Archimedean principle does not hold
3. we cannot define a metric nor a norm.
4. we can define the order topology, however.

The loss of the field property is quite bad, and we might want to recover this. Therefore, we can add more elements that serve to be the multiplicative inverse of infinity. We call these inverses *infinitesimals* and the new set the *hyperreal numbers*.

Theorem 1.25 (Hyperreals)

The hyperreals

In fact, when Newton first invented calculus, the hyperreals were what he worked with, and you can surprisingly build a good chunk of calculus with this. Even though this is a dead topic at this point, a lot of modern notation is based off of this number system, so it's good to see how it works. For example, when we write the integral

$$\int f(x) dx \quad (119)$$

we are saying that we are taking the uncountable sum of the terms $f(x) dx$, the multiplication of the real number $f(x)$ and the infinitesimal number dx living in the hyperreals. Unfortunately, we cannot fully construct a rigorous theory of calculus with only infinitesimals. However, in practice (especially physics) people tend to manipulate and do algebra with infinitesimals, so having a good foundation on what you can and cannot do with them is practical. While the focus won't be on *smooth infinitesimal analysis (SIA)*, I will include some alternate constructions later purely with infinitesimals.

1.4 Complex Numbers

The next field that will be particularly important is the complex numbers. It is straightforward to construct \mathbb{C} , but let's motivate this for a minute.

Example 1.3 (Polynomial Roots)

The roots of the polynomial

$$f(x) = x^2 + 1 \quad (120)$$

does not exist in \mathbb{R} .

Therefore, we would like to construct a new space that contains all possible roots for all possible polynomials with real coefficients. We call this \mathbb{C} . Clearly, by constructing polynomials of the form $x^2 - r^2$ for some $r \in \mathbb{R}$, we know that $\mathbb{R} \subset \mathbb{C}$. Therefore, we want to create a further extension of \mathbb{R} , along with some canonical injection $\iota : \mathbb{R} \rightarrow \mathbb{C}$ that is also a field homomorphism. It turns out that once we construct this field, there is no possible way that we can make it an ordered field. However, the norm extends naturally into \mathbb{C} such that ι is isometric. Finally, we can define a new operator called *conjugation* that gives us additional structure.

This is not the only way to construct the complex plane however. Rather than defining all these from scratch, we could just define the addition operations with an isometric vector space isomorphism from \mathbb{R}^2 to \mathbb{C} actually, and then define multiplication. Another way is to start again with $\mathbb{Q} \times \mathbb{Q}$, define a norm on it, complete it, and finally define the addition and multiplication operations that satisfy the field property.

Theorem 1.26 (Construction of the Complex Numbers)

Let \mathbb{C} be defined as the space $\mathbb{R} \times \mathbb{R}$ with the following operations.

1. *Addition.* $x = (a, b), y = (c, d) \implies x +_{\mathbb{C}} y = (a + c, b + d).$
2. *Additive Identity.* $0_{\mathbb{C}} = (0, 0).$
3. *Additive Inverse.* $x = (a, b) \implies -x = (-a, -b).$
4. *Multiplication.* $x = (a, b), y = (c, d) \implies x \times_{\mathbb{C}} y = (ac - bd, ad + bc).$
5. *Multiplicative Identity.* $1_{\mathbb{C}} = (1, 0).$
6. *Multiplicative Inverse.*

$$x = (a, b) \implies x^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) \quad (121)$$

Our first claim is that $(\mathbb{C}, +_{\mathbb{C}}, \times_{\mathbb{C}})$ is a field. Furthermore, we define the additional structures

1. *Conjugate.* $x = (a, b) \implies \bar{x} = (a, -b).$
2. *Norm.* $|x|_{\mathbb{C}} = x \times_{\mathbb{C}} \bar{x} = a^2 + b^2.$
3. *Metric.* This is the norm-induced metric. $d_{\mathbb{C}}(x, y) = |x - y|_{\mathbb{C}}.$
4. *Topology.* This is the metric-induced topology generated by the open balls $B(x, r) := \{y \in \mathbb{C} \mid d(x, y) < r\}$, where $x \in \mathbb{C}, r \in \mathbb{R}$.

Our second claim is that the canonical injection $\iota : \mathbb{R} \rightarrow \mathbb{C}$ defined

$$\iota(r) = (r, 0) \quad (122)$$

is an isometric field isomorphism. Our third claim is that \mathbb{C} is Cauchy-complete with respect to this metric.

Note that we do not talk about order \mathbb{C} , and so the concepts of Dedekind completeness, least upper bound properties, or Archimedean principle is meaningless in the complex plane.

Definition 1.17 (Imaginary Number)

Let us denote $i = (0, 1)$ which we call the **imaginary number**, which has the property that $i^2 = 1$. With this notation, we can see through abuse of notation that

$$(a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1) = a + bi \quad (123)$$

Therefore, we generally write complex numbers as $z = a + bi$, and we define the real and imaginary components as $\text{Re}(z)$ and $\text{Im}(z)$, respectively.

Note that the identity $x^2 + 1 \equiv (x + i)(x - i)$ implies that the equation $x^2 = -1$ has exactly two solutions in \mathbb{C} , i and $-i$. Therefore, if a subfield of \mathbb{C} contains one of these solutions, it must contain the other (since i and $-i$ are additive and multiplicative inverses).

Furthermore, since i is defined to be $\sqrt{-1}$, we could replace i with $-i$ and our calculations would still be consistent throughout the rest of mathematics. In fact, i and $-i$ behave **exactly** identically and cannot be distinguished in an abstract sense. Visually, the complex plane "flipped" across the real number axis produces the same complex plane.

Theorem 1.27 (Uniqueness of \mathbb{C})

\mathbb{C} is unique up to an isomorphism that maps all real numbers to themselves. Every complex number can be uniquely written as $a + bi$, where $a, b \in \mathbb{R}$ and i is a fixed element such that $i^2 = -1$.

Proof.

Consider the subset of \mathbb{C}

$$K \equiv \{a + bi \mid a, b \in \mathbb{R}\} \quad (124)$$

By evaluating its operations, we can check for closure, identity, and invertibility of nonzero elements to conclude that K is a subfield of $\mathbb{C} \implies$ by prop. (iii), $K = \mathbb{C} \implies$ every element in \mathbb{C} can be written in form $a + bi$. To prove uniqueness, we assume that $p \in \mathbb{C}$ can be written in distinct forms $p = a + bi = a' + b'i$. Then

$$\begin{aligned} a + bi = a' + b'i &\implies (a - a')^2 = (b'i - bi)^2 = -(b' - b)^2 \\ &\implies a - a' = b' - b = 0 \end{aligned}$$

To prove uniqueness of \mathbb{C} up to isomorphism, we assume that \mathbb{C}' exists with i' such that i'^2 containing elements $a + bi'$. Let $f : \mathbb{C} \longrightarrow \mathbb{C}'$ defined

$$f(a + bi) = a + bi' \quad (125)$$

Then,

$$\begin{aligned} f((a + bi) + (c + di)) &= f((a + c) + (b + d)i) \\ &= (a + c) + (b + d)i' \\ &= (a + bi') + (c + di') \\ &= f(a + bi) + f(c + di) \\ f(\kappa(a + bi)) &= f(\kappa a + \kappa bi) \\ &= \kappa a + \kappa bi' \\ &= \kappa(a + bi') \\ &= \kappa f(a + bi) \end{aligned}$$

So, f is an isomorphism, and $\mathbb{C} \simeq \mathbb{C}'$. From analysis, we can construct and prove the existence of \mathbb{R} . We then define the map

$$\rho : \mathbb{R}^2 \longrightarrow \mathbb{C}, \rho(a, b) \equiv a + bi \quad (126)$$

with $\rho(1, 0)$ as the multiplicative identity and $\rho(0, 1) \equiv i$. Therefore, every element of \mathbb{C} can be uniquely represented as an element of \mathbb{R}^2 .

Unfortunately, we lose the ordering.

Theorem 1.28 (Order on Complex Plane)

There exists no order on \mathbb{C} that makes it a totally ordered field.

Proof.

We attempt to construct an order on i and 0 in \mathbb{C} .

1. If $i = 0$, then $i^4 = 0 \cdot i^3 \implies 1 = 0$, which contradicts that $0 < 1$.
2. If $i \neq 0$, then $i^2 > 0$ from the field axioms, and so $-1 > 0$. But this also means that $1 = i^4 > 0$.

This contradicts the ordered field property that $x > 0 \iff -x < 0$.

Therefore \mathbb{C} cannot be turned into an ordered field.

Theorem 1.29 (Conjugation is an Isomorphism)

Conjugation is an isometric field automorphism of \mathbb{C} .

$$c = a + bi \mapsto \bar{c} = a - bi \quad (127)$$

This is identically defined by replacing i with $-i$. Clearly, $\bar{\bar{c}} = c$.

Proof.**Proposition 1.1 (Properties of Conjugation)**

For any $c \in \mathbb{C}$, $c + \bar{c}$ and $c\bar{c}$ are real.

Proof.

Using the fact that the complex conjugate is an isomorphism,

$$\begin{aligned} c + \bar{c} &= \bar{c} + \bar{\bar{c}} = \bar{c} + c = c + \bar{c} \\ c\bar{c} &= \bar{c}\bar{c} = \bar{c}c = c\bar{c} \end{aligned}$$

Note that we proved this abstractly using only the properties given above, and did not decompose c to its **algebraic form** $a + bi$.

If $c = a + bi$, $a, b \in \mathbb{R}$, then

$$c + \bar{c} = 2a, \quad c\bar{c} = a^2 + b^2 \quad (128)$$

In case the reader is unaware, it is common to interpret complex numbers $c = a + bi$ as points or vectors (a, b) on the complex plane.

Definition 1.18 (Polar Form of Complex Numbers)

The **polar representation**, or **trigonometric representation**, of a complex number $c = a + bi$ is defined using the equations

$$a = r \cos \varphi, \quad b = r \sin \varphi \implies c = r(\cos \varphi + i \sin \varphi) \quad (129)$$

where $r = |c|$ and φ is the **argument** of c , which is the angle formed by the corresponding vector with the polar axis defined within the interval $[0, 2\pi]$.

$$\arg(c) \equiv \tan^{-1} \frac{b}{a} \quad (130)$$

This mapping can be defined

$$\rho : \mathbb{R} \times \frac{\mathbb{R}}{2\pi} \longrightarrow \mathbb{C}, \quad \rho(r, \varphi) = r(\cos \varphi + i \sin \varphi) \quad (131)$$

Theorem 1.30 ()

ρ is "similar" to a homomorphism in the following way. By defining the domain and codomain as

groups,

$$\rho : (\mathbb{R}, \times) \times \left(\frac{\mathbb{R}}{2\pi} \right) \longrightarrow (\mathbb{C}, \times) \quad (132)$$

we can see that

$$\rho(r_1, \varphi_1) \times \rho(r_2, \varphi_2) = \rho(r_1 \times r_2, \varphi_1 + \varphi_2) \quad (133)$$

or equivalently,

$$r_1(\cos \varphi_1 + i \sin \varphi_1) \cdot r_2(\cos \varphi_2 + i \sin \varphi_2) = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)) \quad (134)$$

Corollary 1.4 ()

The formula for the ratio of complex numbers is defined

$$\frac{r_1(\cos \varphi_1 + i \sin \varphi_1)}{r_2(\cos \varphi_2 + i \sin \varphi_2)} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)) \quad (135)$$

Corollary 1.5 ()

The positive integer power of a complex number can be written using **De Moivre's formula**.

$$(r(\cos \varphi + i \sin \varphi))^n = r^n(\cos n\varphi + i \sin n\varphi) \quad (136)$$

We can use this formula to extract a root of n th degree from a complex number $c = r(\cos \varphi + i \sin \varphi)$, which means to solve the equation $z^n = c$. Let $z = s(\cos \psi + i \sin \psi)$. Then by De Moivre's formula,

$$\begin{aligned} z^n &= s^n(\cos n\psi + i \sin n\psi) = r(\cos \varphi + i \sin \varphi) \\ \implies s &= \sqrt[n]{r}, \psi = \frac{\varphi + 2\pi k}{n} \\ \implies z &= \sqrt[n]{r} \left(\cos \frac{\varphi + 2\pi k}{n} + i \sin \frac{\varphi + 2\pi k}{n} \right) \text{ for } k = 0, 1, \dots, n-1 \end{aligned}$$

Geometrically, the n solutions lie at the vertices of a regular n -gon centered at the origin. When $c = 1$, the solutions are the n th roots of unity.

1.5 Dual Numbers

Another similar number system.

1.6 Euclidean Space

Congratulations! We have rigorously constructed both the reals and complex numbers, and this becomes the cornerstone to construct other fundamental sets. Now we consider spaces of the form \mathbb{R}^n or \mathbb{C}^n , which we call *Euclidean spaces*, and construct them. This is actually quite easy since we understand linear algebra.

Definition 1.19 (Convex Sets)

A set S is convex if for every point $x, y \in S$, the point

$$z = tx + (1-t)y \in S \quad (137)$$

where $0 \leq t \leq 1$.

1.7 Exercises

Exercise 1.1 (Math 531 Spring 2025, PS2.1)

Prove that the set of all matrices of the form:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad (138)$$

with $a, b \in \mathbb{R}$ forms a field with the usual sum and product operations of matrices. What does this field resemble? Give extensions to 3×3 and 4×4 matrices.

Solution 1.1

Exercise 1.2 (Math 531 Spring 2025, PS2.2)

Why can't the field of complex numbers (with its usual operations) be made into an ordered field?

Solution 1.2

Solution is shown as theorem.

Exercise 1.3 (Math 531 Spring 2025, PS2.3)

Prove there are no finite ordered fields.

Solution 1.3

Solution is shown as theorem.

Exercise 1.4 (Math 531 Spring 2025, PS2.4)

Prove that if x and n are natural numbers, then

$$x^n - 1 = (x - 1)(1 + x + x^2 + \dots + x^{n-1}). \quad (139)$$

Solution 1.4

We use the commutative addition and multiplication, plus distributive property in \mathbb{Z} .

$$(x - 1) \left(\sum_{i=0}^{n-1} x^i \right) = x \sum_{i=0}^{n-1} x^i - \sum_{i=0}^{n-1} x_i \quad (140)$$

$$= \sum_{i=1}^n x^i - \sum_{i=0}^{n-1} x^i \quad (141)$$

$$= x^n + \sum_{i=1}^{n-1} x^i - \sum_{i=1}^{n-1} x^i - 1 \quad (142)$$

$$= x^n - 1 \quad (143)$$

Exercise 1.5 (Math 531 Spring 2025, PS2.7)

Prove that there is no $q \in \mathbb{Q}$ for which

$$q^2 + q = 4. \quad (144)$$

Solution 1.5

Assume that such a $q \in \mathbb{Q}$ in canonical form exists. Then by the field properties, since $\frac{1}{4} \in \mathbb{Q}$,

$$q^2 + q + \frac{1}{4} = 4 + \frac{1}{4} = \frac{17}{4} \in \mathbb{Q} \quad (145)$$

But by distributive properties, $(q + \frac{1}{2})^2 \in \mathbb{Q}$. We claim that there exists no rational $x = a/b$ (a, b coprime) s.t. $x^2 = 17/4$. If there were, then clearly $a \neq 0$ and

$$\frac{a^2}{b^2} = \frac{17}{4} \implies 4a^2 = 17b^2 \quad (146)$$

$$\implies 2|b, \text{ and so } b = 2b' \text{ for some } b' \in \mathbb{N} \quad (147)$$

$$\implies a^2 = 17(b')^2 \quad (148)$$

$$\implies 17|a \text{ and so } a = 17a' \text{ for some } a' \in \mathbb{Z} \quad (149)$$

$$\implies 17(a')^2 = (b')^2 \quad (150)$$

which implies that $17|b'$, but this contradicts the assumption that a, b are coprime. Therefore $q + \frac{1}{2} \notin \mathbb{Q} \implies q \notin \mathbb{Q}$.

Exercise 1.6 (Math 531 Spring 2025, PS2.8)

Let X be an ordered set with the least upper bound property. Prove that X has the greatest lower bound property.

Solution 1.6

Shown in theorem above.

Exercise 1.7 (Math 531 Spring 2025, PS2.9)

Prove that if $x, y \in \mathbb{Q}$ we have that

$$||x| - |y|| \leq |x - y|. \quad (151)$$

Solution 1.7

By subadditivity of the norm we have

$$|x| \leq |x - y| + |y| \implies |x| - |y| \leq |x - y| \quad (152)$$

$$|y| \leq |y - x| + |x| \implies |y| - |x| \leq |y - x| \quad (153)$$

But $|y - x| = |-1(x - y)| = |-1| \cdot |x - y| = |x - y|$, and so

$$\max\{|x| - |y|, |y| - |x|\} \leq |x - y| \quad (154)$$

and the LHS is the definition of the norm $||x| - |y||$ in \mathbb{Q} .

Exercise 1.8 (Rudin 1.1)

If r is rational ($r \neq 0$) and x is irrational, prove that $r + x$ and rx are irrational.

Solution 1.8

If we assume that $rx = t$ and $r + x = s$ are rational, then this violates the field axioms of \mathbb{Q} since then $x = tr^{-1}$ and $x = s + (-r)$ are rational.

Exercise 1.9 (Rudin 1.2)

Prove that there is no rational number whose square is 12.

Solution 1.9

Assume that there exists a number p/q such that p and q are both not even. Then,

$$\left(\frac{p}{q}\right)^2 = 12 \implies p^2 = 12q^2 = 3(2q)^2 \quad (155)$$

So p must be even $p = 2p'$. Therefore, $p'^2 = 3q^2$, and q must be odd. This means that p' must be odd. We can rewrite the equation

$$p'^2 - q^2 = 2q^2 \implies (p' + q)(p' - q) = 2q^2 \quad (156)$$

where the left hand side is divisible by 4 but the right hand side is divisible by at most 2, leading to a contradiction.

Exercise 1.10 (Rudin 1.3)

Prove that the axioms of multiplication imply the following.

1. If $x \neq 0$ and $xy = xz$, then $y = z$.
2. If $x \neq 0$ and $xy = x$, then $y = 1$.
3. If $x \neq 0$ and $xy = 1$, then $y = x^{-1}$.
4. If $x \neq 0$, then $(x^{-1})^{-1} = x$.

Solution 1.10

Listed.

1. $xy = xz \implies \frac{1}{x} \cdot xy = \frac{1}{x}xz \implies y = z$
2. $xy = x \implies \frac{1}{x}xy = \frac{1}{x}x \implies y = 1$
3. $xy = 1 \implies \frac{1}{x}xy = \frac{1}{x}1 \implies y = \frac{1}{x}$
4. $(x^{-1})^{-1} \cdot x^{-1} = 1 \implies (x^{-1})^{-1} \cdot x^{-1} \cdot x = 1 \cdot x \implies (x^{-1})^{-1} = x$

Exercise 1.11 (Rudin 1.4)

Let E be a nonempty subset of an ordered set; suppose α is a lower bound of E and β is an upper bound of E . Prove that $\alpha \leq \beta$.

Solution 1.11

Since E is nonempty, we choose any $x \in E$. By definition, $\alpha \leq x$ and $x \leq \beta$, and by transitive property of orderings, we have $\alpha \leq \beta$.

Exercise 1.12 (Rudin 1.5)

Let A be a nonempty set of real numbers which is bounded below. Let $-A$ be the set of all numbers $-x$, where $x \in A$. Prove that

$$\inf A = -\sup(-A) \quad (157)$$

Solution 1.12

We would like to prove that $\inf A \leq -\sup(-A)$ and $\inf A \geq -\sup(-A)$. For the first part, we start off with the definition of the infimum.

$$\begin{aligned} \inf A \leq x \forall x \in A &\implies -\inf A \geq -x \forall x \in A \\ &\implies -\inf A \geq x \forall x \in -A \\ &\implies -\inf A \geq \sup(-A) \\ &\implies \inf A \leq -\sup(-A) \end{aligned}$$

For the second part, we start with the definition of the supremum.

$$\begin{aligned} \sup(-A) \geq x \forall x \in -A &\implies \sup(-A) \geq -x \forall x \in A \\ &\implies -\sup(-A) \geq x \forall x \in A \\ &\implies -\sup(-A) \leq \inf A \end{aligned}$$

Exercise 1.13 (Rudin 1.6)

Fix $b > 1$.

1. If m, n, p, q are integers, $n > 0$, $q > 0$, and $r = m/n = p/q$, prove that

$$(b^m)^{1/n} = (b^p)^{1/q}. \quad (158)$$

Hence it makes sense to define $b^r = (b^m)^{1/n}$.

2. Prove that $b^{r+s} = b^r b^s$ if r and s are rational.
3. If x is real, define $B(x)$ to be the set of all numbers b^t , where t is rational and $t \leq x$. Prove that

$$b^r = \sup B(r) \quad (159)$$

when r is rational. Hence it makes sense to define

$$b^x = \sup B(x) \quad (160)$$

for every real x .

4. Prove that $b^{x+y} = b^x b^y$ for all real x and y .

Solution 1.13

Proved in theorem above.

Exercise 1.14 (Rudin 1.7)

Fix $b > 1$, $y > 0$, and prove that there is a unique real x such that $b^x = y$, by completing the following outline. (This x is called the logarithm of y to the base b .)

1. For any positive integer n , $b^n - 1 \geq n(b - 1)$.
2. Hence $b - 1 \geq n(b^{1/n} - 1)$.
3. If $t > 1$ and $n > (b - 1)/(t - 1)$, then $b^{1/n} < t$.
4. If w is such that $b^w > y$, then $b^{-1/n} < y$ for sufficiently large n ; to see this, apply part (c) with $t = y \cdot b^{-w}$.
5. If $b^w > y$, then $b^{w-(1/n)} > y$ for sufficiently large n .
6. Let A be the set of all w such that $b^w < y$, and show that $x = \sup A$ satisfies $b^x = y$.
7. Prove that this x is unique.

Solution 1.14

Proved in theorem above.

Exercise 1.15 (Rudin 1.8)

Prove that no order can be defined in the complex field that turns it into an ordered field.

Solution 1.15

Note that if $x \geq 0$, then $-x \leq 0$ for all x of any ordered field. Since if $x \geq 0$ and $-x > 0$, then $x - x > 0$, which is absurd. Therefore, one of either i or $-i$ should be greater than 0. But $i^2 = (-i)^2 = -1$, so this means that $-1 > 0$, which implies that $0 < 1$. But either 1 or -1 must ≥ 0 .

Exercise 1.16 (Rudin 1.9)

Equip \mathbb{C} with the dictionary order. That is, given $z = a + bi$ and $w = c + di$, $z < w$ if $a < c$, or if $a = c$ and $b < d$. Does this ordered set have a least upper bound property?

Solution 1.16

No it does not. Consider the set $S = \{a + bi \in \mathbb{C} \mid a \leq 3\}$. S is bounded by 4, but it doesn't have a least upper bound. Given any $3 + bi$, this is not an upper bound since we can construct $3 + (b + \epsilon)i \in S$. Given any $a + bi$ where $a > 3$, we can always find a lower bound of form $a + (b - \epsilon)i$ that also bounds S .

Exercise 1.17 (Rudin 1.10)

Suppose $z = a + bi$, $w = u + iv$, and

$$a = \left(\frac{|w| + u}{2} \right)^{1/2} \text{ and } b = \left(\frac{|w| - u}{2} \right)^{1/2} \quad (161)$$

Prove that $z^2 = w$ if $v \geq 0$ and that $(\bar{z})^2 = w$ if $v \leq 0$. Conclude that every complex number (with one exception!) has two complex square roots.

Solution 1.17

We can calculate

$$z^2 = (a^2 - b^2) + 2abi = u + \sqrt{v^2}i = \begin{cases} u + vi & \text{if } v \geq 0 \\ u - vi & \text{if } v \leq 0 \end{cases} \quad (162)$$

Since if we assume $v \geq 0$, then we have $z^2 = w$. We also get

$$\bar{z}^2 = (a^2 - b^2) - 2abi = u - \sqrt{v^2}i = \begin{cases} u - vi & \text{if } v \geq 0 \\ u + vi & \text{if } v \leq 0 \end{cases} \quad (163)$$

and assuming $v \leq 0$, we have $\bar{z}^2 = w$. Therefore, every complex number w has both $\pm z$ as its square root if $v \geq 0$, $\pm \bar{z}$ if $v \leq 0$, and just one root if $z = 0$.

Exercise 1.18 (Rudin 1.11)

If z is a complex number, prove that there exists an $r \geq 0$ and a complex number w with $|w| = 1$ s.t. $z = rw$. Are w and r always uniquely determined by z ?

Solution 1.18

If $z = 0$, then $r = 0$ and there is no unique w . If $z = a + bi \neq 0$, then define

$$r = |z| = (a^2 + b^2)^{1/2}, \quad w = \frac{1}{r}z \quad (164)$$

which proves existence. As for uniqueness, assume that there are two forms

$$z = rw = r'w' \quad (165)$$

Then, $w = \frac{r'}{r}w' \implies |w| = \left|\frac{r'}{r}\right||w'| = 1$, which implies that $r'/r = 1$ and so $r = r'$. This means that $w = w'$.

Exercise 1.19 (Rudin 1.12)

If z_1, \dots, z_n are complex, prove that

$$|z_1 + z_2 + \dots + z_n| \leq |z_1| + \dots + |z_n| \quad (166)$$

Solution 1.19

By induction, it suffices to prove $|z_1 + z_2| \leq |z_1| + |z_2|$. We have

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1 + z_2}) \\ &= (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) \\ &= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2 \\ &= |z_1|^2 + |z_2|^2 + z_1\bar{z}_2 + z_2\bar{z}_1 \\ &= |z_1|^2 + |z_2|^2 + 2(ac + bd) \\ &\leq |z_1|^2 + |z_2|^2 + 2\sqrt{a^2 + b^2}\sqrt{c^2 + d^2} \quad (\text{Schwartz}) \\ &= |z_1|^2 + |z_2|^2 + 2|z_1||z_2| \\ &= (|z_1| + |z_2|)^2 \end{aligned}$$

since both sides are positive, we can take their square root to get the desired result.

Exercise 1.20 (Rudin 1.13)

If x, y are complex, prove that

$$||x| - |y|| \leq |x - y| \quad (167)$$

Solution 1.20

Since both sides are nonnegative, we can square both sides. Note that due to Cauchy Schwartz inequality, $2|x||y| \geq x\bar{y} + y\bar{x}$ since expanding them gives

$$2\sqrt{a^2 + b^2}\sqrt{c^2 + d^2} \geq 2(ac + bd) \quad (168)$$

Therefore, the following inequality is true:

$$|x|^2 + |y|^2 - 2|x||y| \leq x\bar{x} + y\bar{y} - x\bar{y} - y\bar{x} \quad (169)$$

which reduces to form $(|x| - |y|)^2 \leq |x - y|^2$.

Exercise 1.21 (Rudin 1.14)

If z is a complex number s.t. $|z| = 1$, that is such that $z\bar{z} = 1$, compute

$$|1 + z|^2 + |1 - z|^2 \quad (170)$$

Solution 1.21

Compute.

$$(1 + z)(1 + \bar{z}) + (1 - z)(1 - \bar{z}) = 1 + z + \bar{z} + z\bar{z} + 1 - z - \bar{z} + z\bar{z} = 4 \quad (171)$$

Exercise 1.22 (Rudin 1.15)

Under what conditions does equality hold in the Schwarz inequality?

Solution 1.22

If they are antiparallel, since

$$\langle x, y \rangle = ||x|| ||y|| \cos \theta \quad (172)$$

Exercise 1.23 (Rudin 1.16)

Suppose $k \geq 3$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, $|\mathbf{x} - \mathbf{y}| = d > 0$, and $r > 0$. Prove:

a) If $2r > d$, there are infinitely many $\mathbf{z} \in \mathbb{R}^k$ s.t.

$$|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$$

b) If $2r = d$, there is exactly one such \mathbf{z} .

c) If $2r < d$, there is no such \mathbf{z} .

Exercise 1.24 (Rudin 1.17)

Prove that

$$|\mathbf{x} + \mathbf{y}|^2 + |\mathbf{x} - \mathbf{y}|^2 = 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2 \quad (173)$$

Solution 1.23

This is trivial if we simply expand

$$|\mathbf{x} + \mathbf{y}|^2 + |\mathbf{x} - \mathbf{y}|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle \quad (174)$$

$$= \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \quad (175)$$

$$= 2\langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{y}, \mathbf{y} \rangle \quad (176)$$

$$= 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2 \quad (177)$$

Exercise 1.25 (Rudin 1.18)

If $k \geq 2$ and $\mathbf{x} \in \mathbb{R}^k$, prove that there exists $\mathbf{y} \in \mathbb{R}^k$ s.t. $\mathbf{y} \neq \mathbf{0}$, but $\mathbf{x} \cdot \mathbf{y} = 0$. Is this also true if $k = 1$?

Solution 1.24

Let $x \in \mathbb{R}^k$ and $\ell \in \mathbb{R}^{k*}$, the dual space. By Riesz representation theorem, we can define the canonical isomorphism $\ell \mapsto y$ between these two spaces as

$$\ell(x) = (x, y) \quad (178)$$

Since $y \neq 0$ by assumption, $\ell \neq 0$, and so its rank is at least 1. Since ℓ maps to \mathbb{R} , the rank has to be 1. By rank nullity theorem, we have

$$\dim N(\ell) = k - \text{rank}(\ell) = k - 1 \quad (179)$$

and so there exists nontrivial annihilators ℓ of x , which can be mapped to a nontrivial $y \in \mathbb{R}^k$.

Exercise 1.26 (Rudin 1.19)

Suppose $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$. Find $\mathbf{c} \in \mathbb{R}^k$ and $r > 0$ s.t.

$$|\mathbf{x} - \mathbf{a}| = 2|\mathbf{x} - \mathbf{b}| \quad (180)$$

if and only if $|\mathbf{x} - \mathbf{c}| = r$.

Solution 1.25

If we draw out the circle, it must contain two points on the line drawn by connecting A and B . Since it must be symmetric, its center and radius can then be easily calculated to be

$$r = \frac{2}{3} |b - a|, \quad c = \frac{1}{3} (4b - a) \quad (181)$$

Exercise 1.27 (Zorich 2.2.1)

Using the principle of induction, show that

1. the sum $x_1 + \dots + x_n$ of real numbers is defined independently of the insertion of parentheses to specify the order of addition.
2. the same is true of the product $x_1 \dots x_n$
3. $|x_1 + \dots + x_n| \leq |x_1| + \dots + |x_n|$
4. $|x_1 \dots x_n| \leq |x_1| \dots |x_n|$
5. For any $n, m \in \mathbb{N}$ such that $m < n$, $(n - m) \in \mathbb{N}$.
6. $(1 + x)^n \geq 1 + nx$ for $x > -1$ and $n \in \mathbb{N}$, equality holding for when $n = 1$ or $x = 0$.
7. $(a + b)^n = a^n +_n C_1 a^{n-1} b^1 + \dots + b^n$ (aka binomial theorem).

Solution 1.26

Listed.

1. Let n denote the number of elements in the sum. We prove by strong law of induction. The base case for when $n = 1, 2, 3$ is trivially true.

$$\begin{aligned} x_1 &= x_1 && \text{(identity)} \\ x_1 + x_2 &= x_1 + x_2 && \text{(identity)} \\ (x_1 + x_2) + x_3 &= x_1 + (x_2 + x_3) && \text{(associativity)} \end{aligned}$$

Then, the sum of $n = k$ parameters is defined by $k - 2$ pairs of parentheses defining the order of the sum. These parentheses define a sequence of $k - 1$ 2-fold additions. Now, assume that the claim is true for

$$S_n \equiv x_1 + \dots + x_n \text{ for } n = 1, 2, \dots, k \quad (182)$$

Then, for a specific sum S_{k+1} of $k + 1$ elements with $k - 1$ parentheses, we can reduce the sum to its final 2-fold addition

$$S_{k+1} \equiv \underbrace{(x_1 + \dots + x_i)}_{\varphi_1} + \underbrace{(x_{i+1} + \dots + x_{k+1})}_{\varphi_2} \quad (183)$$

Since $i, k - i + 1 < k$, by the strong law φ_1 and φ_2 are independent of the order of sum.

2. Exactly identical to (a).
3. By the triangle inequality $|x_1 + x_2| \leq |x_1| + |x_2|$. Now, assume for $n = k$ is true. Then, let $S_k = x_1 + \dots + x_k$, so

$$|x_1 + \dots + x_k + x_{k+1}| = |S_k + x_{k+1}| \leq |S_k| + |x_{k+1}| \leq \sum_{i=1}^{k+1} |x_i| \quad (184)$$

4. Same as (c).
5. Let us fix m to be any element of \mathbb{N} . Then, the base case is for $n = m + 1$ (which is in \mathbb{N} since it is inductive), so

$$n - m = (m + 1) - m = 1 \in \mathbb{N} \quad (185)$$

Now, given that for some integer $n \geq m + 1$, $n - m \in \mathbb{N}$ is true, we have

$$\begin{aligned} (n + 1) - m &= n + (1 - m) && \text{(associativity)} \\ &= n + (-m + 1) && \text{(commutativity)} \\ &= (n - m) + 1 && \text{(associativity)} \end{aligned}$$

where $(n - m) + 1 \in \mathbb{N}$ by inductive property of \mathbb{N} .

6. We prove by induction. For $n = 1$, it is trivial that $(1+x)^1 \geq 1 + 1 \cdot x$. Now assume that the claim is true for some $k \in \mathbb{N}$. Then,

$$\begin{aligned}(1+x)^{k+1} &= (1+x)^k(1+x) \geq (1+kx)(1+x) \\ &= 1 + (k+1)x + kx^2 \\ &\geq 1 + (k+1)x\end{aligned}$$

where equality holds if $x = 0 \implies 1^{k+1} = 1^k \cdot 1 = 1$ or $n = 1 \implies$ trivial case.

7. The base case for $n = 1$ is trivial since $(a+b)^1 = \binom{1}{0}a + \binom{1}{1}b$. We introduce Newton's identity.

$$\begin{aligned}\binom{k}{j-1} + \binom{k}{j} &= \frac{k!}{(j-1)!(k-j+1)!} + \frac{k!}{j!(k-j)!} \\ &= k! \left(\frac{j}{j!(k-j+1)!} + \frac{k-j+1}{j!(k-j+1)!} \right) \\ &= k! \cdot \frac{k+1}{j!(k-j+1)!} \\ &= \frac{(k+1)!}{j!(k-j+1)!} = \binom{k+1}{j}\end{aligned}$$

Now assuming that the binomial formula holds for some $n = k$, we have

$$(a+b)^{k+1} = (a+b)^k(a+b) \quad (186)$$

$$= \left(\sum_{j=0}^k \binom{k}{j} a^j b^{k-j} \right) (a+b) \quad (187)$$

$$= \sum_{j=0}^k \binom{k}{j} a^{j+1} b^{k-j} + \sum_{j=0}^k \binom{k}{j} a^j b^{k-j+1} \quad (188)$$

$$= \binom{k}{0} a^0 b^{k+1} + \binom{k}{k} a^{j+1} b^0 + \sum_{j=0}^{k-1} \binom{k}{j} a^{j+1} b^{k-j} + \sum_{j=1}^k \binom{k}{j} a^j b^{k-j+1} \quad (189)$$

$$= \binom{k+1}{0} a^0 b^{k+1} + \binom{k+1}{k+1} a^{j+1} b^0 + \sum_{j=1}^k \left[\binom{k}{j-1} + \binom{k}{j} \right] a^j b^{k-j+1} \quad (190)$$

$$= \sum_{j=0}^{k+1} \binom{k+1}{j} a^j b^{k-j+1} \quad (191)$$

Exercise 1.28 (Zorich 2.2.3)

Show that an inductive set is not bounded above.

Solution 1.27

Assume that a X is a nonempty inductive set that is bounded above. By definition, there exists a number $B \in \mathbb{R}$ such that $\max X < B$. Then, this means that there exists no numbers in $[B, B+1]$. Since X is inductive, this means that there cannot exist any elements of X in the interval $[B-1, B)$, and similarly for the interval $[B-2, B)$, and so on, meaning that if $x \in X$, then $x \notin [B-k, B-k+1]$ for all $k \in \mathbb{Z}$. By the Archimedean principle, this implies that $X = \emptyset$, contradicting our assumption.

Exercise 1.29 (Zorich 2.2.4)

Prove the following.

1. An inductive set is infinite (that is, equipotent with one of its subsets different from itself).
2. The set $E_n = \{x \in \mathbb{N} \mid x \leq n\}$ is finite.

Solution 1.28

Listed.

1. Assume that an inductive set X is finite $\implies X$ is bounded above (we can choose upper bound $B = \max X + 1$). But from 2.2.3, an inductive set cannot be bounded above, contradicting our assumption.
2. It is trivial that $E_1 = \{1\}$ is finite since $\text{card } E_1 = 1$. Now, if for some k , E_k is finite with cardinality e_k , then $\text{card } E_{k+1} = e_k + 1$, which implies finiteness.

Exercise 1.30 (Zorich 2.2.5)

Listed.

1. Let $m, n \in \mathbb{N}$ and $m > n$. Their greatest common divisor $\gcd(m, n) = d \in \mathbb{N}$ can be found in a finite number of steps using the following algorithm of Euclid involving successive divisions with remainder.

$$\begin{aligned} m &= q_1 n + r_1 \\ n &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\dots = \dots \\ r_{k-2} &= q_k r_{k-1} + r_k \\ r_{k-1} &= q_{k+1} r_k + 0 \end{aligned}$$

Then $d = r_k$

2. If $d = \gcd(m, n)$, one can choose numbers $p, q \in \mathbb{Z}$ such that $pm + qn = d$.

Solution 1.29

Listed.

- 1.
2. Letting $n = r_0$, notice that the equations above satisfy for $i = 0, 1, \dots$

$$r_i = q_{i+2} r_{i+1} + r_{i+2} \implies r_i - q_{i+2} r_{i+1} = r_{i+2} \quad (1)$$

Note that the second-to-last equation allows us to write r_k as a linear combination of r_{k-2} and r_{k-1} : $r_k = r_{k-2} - q_k r_{k-1}$. Now by applying (1), we can reduce the above to a linear combination of r_{k-3} and r_{k-2} .

$$\begin{aligned} r_k &= r_{k-2} - q_k r_{k-1} \\ &= r_{k-2} - q_k(r_{k-3} - q_{k-1} r_{k-3}) \\ &= (1 + q_{k-1} q_k) r_{k-2} - q_k r_{k-3} \end{aligned}$$

and repeatedly doing this allows us to reduce r_k to a linear combination $q_0 r_0 + q_1 r_1$. By the ring properties of \mathbb{Z} , the new linear coefficients are also in \mathbb{Z} . Reducing one last time using the

first equation in the Euclidean algorithm gives

$$\begin{aligned} r_k &= q_0 r_0 + q_1 r_1 \\ &= q_0 n + q_1(m - q_1 n) \\ &= q_1 m + (q_0 - q_1)n \\ &= pm + qn \end{aligned}$$

Exercise 1.31 (Zorich 2.2.9)

Show that if the natural number n is not of the form k^m , where $k, m \in \mathbb{N}$, then the equation $x^m = n$ has no rational roots.

Solution 1.30

Assume that there is a rational solution $x = p/q$, with $p, q \in \mathbb{N}$ of the equation. Then,

$$\left(\frac{p}{q}\right)^m = \frac{p^m}{q^m} = n \implies p^m = q^m n \quad (192)$$

By the fundamental theorem of arithmetic, the exponents of the prime factors of p^m must all be multiples of m , and so it must be so for the right hand side $\implies x$ must be of form $x = k^m$ for some k . This is a contradiction.

Exercise 1.32 (Zorich 2.2.12)

Knowing that $\frac{m}{n} \equiv m \cdot n^{-1}$ by definition, where $m \in \mathbb{Z}$ and $n \in \mathbb{N}$, derive the “rules” for addition, multiplication, and division of fractions, and also the condition for two fractions to be equal.

Solution 1.31

We can construct a \mathbb{Q} as a quotient space $\mathbb{Z} \times \mathbb{N} / \sim$, where \sim is an equivalent relation where

$$(q_1, p_1) \sim (q_2, p_2) \text{ iff } q_1 p_2 = p_1 q_2 \quad (193)$$

which is the familiar equivalence relation from “simplifying” a fraction. We define addition and multiplication as the following

$$\begin{aligned} (a, b) + (c, d) &= (ad + bc, bd) \\ (a, b) \cdot (c, d) &= (ac, bd) \end{aligned}$$

which turns out to be algebraically closed in \mathbb{Q} . The additive identity is the equivalence class $0 = \{(0, c) \mid c \in \mathbb{N}\}$, and the multiplicative identity is the equivalence class $1 = \{(c, c) \mid c \in \mathbb{N}\}$. It is easy to check that $+$ is commutative, the additive inverse is $-(a, b) = (-a, b)$, and the multiplicative inverse is $(a, b)^{-1} = (b, a)$. We can subtract and divide these elements of \mathbb{Q} , called “fractions,” as such:

$$\begin{aligned} (a, b) - (c, d) &= (a, b) + (-c, d) = (a, b) + (-c, d) = (ad - bc, bd) \\ (a, b) \div (c, d) &= (a, b) \cdot (c, d)^{-1} = (a, b) \cdot (d, c) = (ad, bc) \end{aligned}$$

Exercise 1.33 (Zorich 2.2.13)

Verify that the rational numbers \mathbb{Q} satisfy all the axioms for real numbers except for the axiom of completeness.

Solution 1.32

From continuing the steps of 2.2.14, we can prove \mathbb{Q} is an algebraic field (associativity, commutativity of addition and multiplication, along with distributive property). We can actually define the order relation $\leq_{\mathbb{Q}}$ in two ways:

1. $(a, b) \leq (c, d)$ iff $ad \leq bc$, where $\leq_{\mathbb{Z}}$ is the order relation on \mathbb{Z} (which can be defined much more simply).
2. Recognizing that $\mathbb{Q} \subset \mathbb{R}$, we define the canonical injection map $i : \mathbb{Q} \rightarrow \mathbb{R}$ and by abuse of language, endow the relation $\leq_{\mathbb{Q}}$ as the restriction of $\leq_{\mathbb{R}}$ onto \mathbb{Q} . That is, for $(a, b), (c, d) \in \mathbb{Q}$,

$$(a, b) \leq_{\mathbb{Q}} (c, d) \text{ iff } i(a, b) \leq_{\mathbb{R}} i(c, d) \quad (194)$$

The ordering for the 1st step can be checked for consistency.

1. $(a, b) \leq (a, b)$ since $ab \leq ab$ (true in \mathbb{Z})
2. $(a, b) \leq (c, d), (c, d) \leq (a, b)$ means that $ad \leq bc$ and $bc \leq ad \implies ad = bc$ (true in \mathbb{Z})
3. $(a, b) \leq (c, d) \leq (e, f)$ implies $ad \leq bc, cf \leq de$. Multiplying positive (important that $f > 0$!) to the first inequality gives $adf \leq bcf$, and multiplying positive b to the second gives $bcf \leq bde$, and by interpreting \leq as the ordering defined on \mathbb{Z} , we use transitive property of $\leq_{\mathbb{Z}}$ to get $adf \leq bde \implies af \leq be \iff (a, b) \leq (e, f)$.
4. For any $(a, b), (c, d) \in \mathbb{Q}$, $(a, b) \leq (c, d)$ or $(a, b) \geq (c, d)$, which is equivalent to $ad \leq bc$ or $ad \geq bc$, which is true in \mathbb{Z} .

It is easy to prove $(a, b) \leq (c, d) \implies (a, b) + (p, q) \leq (c, d) + (p, q)$, and $0_{\mathbb{Q}} \leq (a, b), (c, d) \implies 0_{\mathbb{Q}} \leq (a, b) \cdot (c, d)$. However, \mathbb{Q} is **not** complete. We prove this by showing that the subset $X = \{x \in \mathbb{Q} \mid x^2 \leq 2\} \subset \mathbb{Q}$ does not satisfy the least upper bound property. Assume that there is a least upper bound $c \in \mathbb{Q}$. $c \neq \sqrt{2}$ (you should know how to prove irrationality of $\sqrt{2}$!), we have either $c > \sqrt{2}$ or $c < \sqrt{2}$.

1. Let $c < \sqrt{2} \iff c - \sqrt{2} > 0$. By the Archimedean principle, there exists a $k \in \mathbb{N}$ such that $0 < \frac{1}{k} < c - \sqrt{2}$. Then, $\frac{1}{k} \in \mathbb{Q}$ and \mathbb{Q} is a field, so $c - \frac{1}{k} \in \mathbb{Q}$.

$$c - \frac{1}{k} < c - c + \sqrt{2} = \sqrt{2} \quad (195)$$

So c is not least and so it must be the case that $c < \sqrt{2}$.

2. Let $c < \sqrt{2} \iff \sqrt{2} - c > 0$. By the Archimedean principle, there exists a $k \in \mathbb{N}$ such that $0 < \frac{1}{k} < \sqrt{2} - c$. Then, $c + \frac{1}{k} \in \mathbb{Q}$ and

$$c + \frac{1}{k} < c + (\sqrt{2} - c) = c \quad (196)$$

So c is not an upper bound.

Note that given a well-defined $c = \sup X$ and in the case where $c < \sqrt{2}$, we have $2 - c^2 > 0$, so we can choose a well-defined δ satisfying (by Archimedean principle)

$$0 < \delta < \min \left\{ 1, \frac{2 - c^2}{2c + 1} \right\} \quad (197)$$

which gives us

$$\begin{aligned} (c + \delta)^2 &= c^2 + \delta(2c + \delta) \\ &< c^2 + \delta(2c + 1) && (\delta < 1) \\ &< c^2 + (2 - c^2) = 2 \end{aligned}$$

meaning that c is not an upper bound. Similarly for when $c > \sqrt{2}$.

Exercise 1.34 (Zorich 2.2.15)

Prove the equivalence of these two statements.

1. If X and Y are nonempty sets of \mathbb{R} having the property that $x \leq y$ for every $x \in X, y \in Y$, then there exists $c \in \mathbb{R}$ such that $x \leq c \leq y$ for all $x \in X$ and $y \in Y$.
2. Every set $X \subset \mathbb{R}$ that is bounded above has a least upper bound.

Solution 1.33

Let S_1 be the first statement and S_2 the second.

1. ($S_2 \implies S_1$). Let $X \subset \mathbb{R}$ be a set that is bounded above, and Y is a set such that $x \leq y$ for all $x \in X, y \in Y$. Then, by LUB principle, there exists $c = \sup X \in \mathbb{R}$. Now, we claim that $c \leq y$ for all $y \in Y$. Assume it doesn't: then there exists $y' \in Y$ such that $y' < c$. But since we assumed $x \leq y$ for all $x \in X, y \in Y$, we have $x \leq y'$ for all $x \in X$, which means that y' is an upper bound of X . But $y' < c$, contradicting the given fact that c was the least upper bound.
2. ($S_1 \implies S_2$). Given a nonempty set $X \subset \mathbb{R}$, we wish to show the existence of $\sup X$. We are guaranteed the existence of nonempty set $Y \subset \mathbb{R}$ such that $x \leq y$ for all $x \in X, y \in Y$, which implies that X must be bounded above. Then, by S_1 , there must exist a $c \in \mathbb{R}$ such that

$$x \leq c \leq y \text{ for all } x \in X, y \in Y \quad (198)$$

We claim that $c = \sup X$. It is an upper bound of X since $x \leq c$ for all $x \in X$. It is least since the set of all upper bounds of X is Y , and $c \leq y$ for all $y \in Y$.

Exercise 1.35 (Olmsted 1.15)

Prove **Dedekind's Theorem**: Let the real numbers be divided into two nonempty sets A and B such that (i) if $x \in A$ and if $y \in B$, then $x < y$ and (ii) if $x \in \mathbb{R}$ then either $x \in A$ or $x \in B$, then there exists a number c (which may belong to either A or B) such that any number less than c belongs to A and any number greater than c belongs to B .

Solution 1.34

This is really the same statement as Zorich 2.2.15.a, the original statement of completeness, but with the extra condition that the sets $A = X, B = Y$ must be disjoint.

Exercise 1.36 (Olmsted 1.7)

If x is an irrational number, under what conditions on the rational numbers a, b, c, d is $(ax+b)/(cx+d)$ rational?

Solution 1.35

Note that a trivial solution is $a = b = c = d = 1$ which gives 1. Since

$$\frac{ax + b}{cx + d} = \frac{acx + ad - ad + bc}{cx + d} = a + \frac{bc - ad}{cx + d} \quad (199)$$

for the above to be rational it is necessary that $1/(cx + d)$ is rational. But this cannot be the case, which leaves us with the condition that $bc = ad$.

Exercise 1.37 (Olmsted 1.8)

Prove that the system of integers satisfies the axiom of completeness.

Solution 1.36

Let $S \subset \mathbb{Z}$ be bounded from above. It must have a maximum element (justify?), call it c . Then we claim that $c \in \mathbb{Z}$ is the least upper bound. Being the maximum, it is an upper bound, and c is least since the next smallest element is $c - 1$, which is less than $c \in S$, and therefore cannot be an upper bound.

Exercise 1.38 (Zorich 2.2.16/Olmsted 1.16)

Prove the following.

1. If $A \subset B \subset \mathbb{R}$, then $\sup A \leq \sup B$ and $\inf A \geq \inf B$.
2. Let $\mathbb{R} \supset X \neq \emptyset$ and $\mathbb{R} \supset Y \neq \emptyset$. If $x \leq y$ for all $x \in X, y \in Y$, then X is bounded above, Y is bounded below, and $\sup X \leq \inf Y$.
3. If the sets X, Y in (b), are such that $X \cup Y = \mathbb{R}$, then $\sup X = \inf Y$.
4. If X and Y are the sets defined in (c), then either X has a maximal element or Y as a minimal element.
5. Show that Dedekind's theorem is equivalent to the axiom of completeness.

Solution 1.37

Listed.

1. Let

$$\begin{aligned} A' &= \{x \in \mathbb{R} \mid x \geq a \ \forall a \in A\} \\ B' &= \{x \in \mathbb{R} \mid x \geq b \ \forall b \in B\} \end{aligned}$$

where we can easily verify that $B' \subset A'$. By definition, we get $\sup B = \min B'$ and $\sup A = \min A'$. But since $B' \subset A'$, for any $b' \in B'$, there exists an $a' \in A'$ such that $a' \leq b'$, which implies that $\sup B = \min B' \leq \min A' = \sup A$.

2. X is bounded above by any element of Y . Y is bounded below by any element of X . By the completion axiom, there exists a $c \in \mathbb{R}$ such that

$$x \leq c \leq y \text{ for all } x \in X, y \in Y \tag{200}$$

Since c is an upper bound of X , $\sup X \leq c$ by definition, and since c is a lower bound of Y , $\inf Y \geq c$ by definition. Therefore, $\sup X \leq c \leq \inf Y$.

3. From completeness there exists a $c \in \mathbb{R}$ such that $x \leq c \leq y$ for all $x \in X, y \in Y$. Y is, by definition, the set of *all* upper bounds of X (i.e. *every* upper bound of X is in Y , unlike Y defined in 2.2.16.b). Since $c \leq y$ for all $y \in Y$, c is minimal and so $c = \sup X$. X is the set of all lower bounds of Y by definition, so $c \geq x$ for all $x \in X \implies c = \inf Y$. So, $\inf Y = c = \sup X$.
4. We know that there exists $c = \inf Y = \sup X$. Since $X \cup Y = \mathbb{R}$, c must be in at least X or Y . If $c \in X$, then $c = \sup X = \max X$, and if $c \in Y$, then $c = \inf Y = \min Y$.
5. This is the same statement as Zorich 2.2.15.a (an iff equivalence, not just one way implying).

Exercise 1.39 (Olmsted 1.13)

Let S be a nonempty set of numbers bounded above, and let x be the least upper bound of S . Prove that x has the two properties corresponding to an arbitrary positive number ϵ :

1. every element $s \in S$ satisfies the inequality $s < x + \epsilon$

2. at least one element $s \in S$ satisfies the inequality $s > x - \epsilon$

Solution 1.38

Listed.

1. x is an upper bound $\implies s \leq x$ for all $s \in S$, which implies that $s \leq x < x + \epsilon$.
2. By definition, $x - \epsilon$ cannot be an upper bound, so $x - \epsilon \geq s$ for all $s \in S$ is not true. Therefore, there must exist one $s \in S$ such that $s > x - \epsilon$.

Exercise 1.40 (Zorich 2.2.18)

Let $-A$ be the set of numbers of the form $-a$, where $a \in A \subset \mathbb{R}$. Show that $\sup(-A) = -\inf(A)$.

Solution 1.39

If A is unbounded below, then $-\inf A = \infty$ and $-A$ is unbounded above, implying that $\sup A = \infty$. Now assume that A is bounded below, then by completeness, it must have a greatest lower bound. Let us define the set $B = \{b \in \mathbb{R} \mid b \leq a \ \forall a \in A\}$. From 2.2.16.b, we have $b \leq \inf A \leq a$ for all $a \in A, b \in B$. Multiplying by -1 gives $-b \geq -\inf A \geq -a$ for all $a \in A, b \in B$, which is equivalent to saying

$$a \leq -\inf A \leq b \text{ for all } a \in -A, b \in -B \quad (201)$$

by definition of $-A, -B$. $-\inf A$ is clearly an upper bound of $-A$, and since

$$\begin{aligned} B &= \{b \in \mathbb{R} \mid b \leq a \ \forall a \in A\} \\ &= \{b \in \mathbb{R} \mid -b \geq -a \ \forall a \in A\} \\ &= \{b \in \mathbb{R} \mid -b \geq a \ \forall a \in -A\} \end{aligned}$$

implies that $-B = \{b \in \mathbb{R} \mid b \geq a \ \forall a \in -A\}$ is the set of all upper bounds of A . So, $-\inf A$ is the least upper bound of $-A$, i.e. $-\inf A = \sup(-A)$.

Exercise 1.41 (Zorich 2.2.21)

Show that the set $\mathbb{Q}(\sqrt{n})$ of numbers of the form $a + b\sqrt{n}$ where $a, b \in \mathbb{Q}$, n is a fixed natural number that is not the square of any integer, is an ordered set satisfying the principle of Archimedes but not the axiom of completeness.

Solution 1.40

The order on $\mathbb{Q}(\sqrt{n})$ can be embedded from the ordering on the reals by defining the canonical injection map $i : \mathbb{Q}(\sqrt{n}) \rightarrow \mathbb{R}$ and defining for any $x, y \in \mathbb{Q}(\sqrt{n})$,

$$x \leq_{\mathbb{Q}(\sqrt{n})} y \iff i(x) \leq_{\mathbb{R}} i(y) \quad (202)$$

Now, let $h > 0$ be any fixed real number, and $x = (a, b) = a + b\sqrt{n}$. By the Archimedean principle, we can find a $k \in \mathbb{Z}$ such that

$$(k-1)h \leq x \leq kh \text{ for some } x \in \mathbb{Q}(\sqrt{n}) \subset \mathbb{R} \quad (203)$$

We now show that $\mathbb{Q}(\sqrt{n})$ is not complete since it doesn't satisfy the LUB property. Since there are infinite prime numbers in \mathbb{N} , choose a prime number p that is not a factor of n . Then, we are guaranteed that pn is not a perfect square, and can define the set

$$X = \{x \in \mathbb{Q}(\sqrt{n}) \mid x < \sqrt{pn}\} \subset \mathbb{Q}(\sqrt{n}) \quad (204)$$

and assume that $c = c_1 + c_2\sqrt{n} = \sup X$ exists ($c_1, c_2 \in \mathbb{Q}$). Clearly, $c \neq \sqrt{pn} \notin \mathbb{Q}(\sqrt{n})$.

1. Assume $c < \sqrt{pn} \iff 0 < \sqrt{pn} - c \in \mathbb{R}$. By the Archimedean principle, there exists a $k \in \mathbb{N}$ such that $0 < \frac{1}{k} < \sqrt{pn} - c$. Then, we can verify that $c + \frac{1}{k} = (c_1 + \frac{1}{k}) + c_2\sqrt{n} \in \mathbb{Q}(\sqrt{n})$ and

$$c + \frac{1}{k} < c + \sqrt{pn} - c = \sqrt{pn} \implies c + \frac{1}{k} \in X \quad (205)$$

implies that c is not an upper bound. So we must turn to case 2.

2. Assume $c > \sqrt{pn} \iff c - \sqrt{pn} > 0$. By AP, there exists a $k \in \mathbb{N}$ such that $0 < \frac{1}{k} < c - \sqrt{pn}$. Then, we can verify that $c - \frac{1}{k} \in \mathbb{Q}(\sqrt{n})$ and

$$c - \frac{1}{k} > c - c + \sqrt{pn} = \sqrt{pn} \quad (206)$$

implies that $c - \frac{1}{k}$ is an upper bound of X , so c is not least.

Therefore, by contradiction, c does not exist.

Exercise 1.42 (Zorich 2.2.22)

Let $n \in \mathbb{N}$ and $n > 1$. In the set $E_n = \{0, 1, \dots, n-1\}$, we define the sum and product of two elements as the remainders when the usual sum and product in \mathbb{R} are divided by n . With these operations on it, the set E_n is denoted \mathbb{Z}_n .

1. Show that if n is not a prime number, then there are nonzero numbers $m, k \in \mathbb{Z}_n$ such that $m \cdot k = 0$, i.e. there exist nonzero zero divisors.
2. Show that if p is prime, then there are no zero divisors in \mathbb{Z}_p and \mathbb{Z}_p is a field.
3. Show that, no matter what the prime p , \mathbb{Z}_p cannot be ordered in a way consistent with the arithmetic operations on it.

Solution 1.41

Listed.

1. n is composite implies that there exist $1 < m, k < n$ such that $n = mk$. These factors m, k are precisely the zero divisors of \mathbb{Z}_n since $mk = n \equiv 0 \pmod{n}$.
2. With p prime, assume that there are nontrivial zero divisors $1 < m, k < p$ in \mathbb{Z}_p . Then, $mk \equiv 0 \pmod{p} \implies mk = lp$ for some $l \in \mathbb{N}$. But this implies that m or k must divide p , which is impossible since $1 < m, k < p$. Then prove field axioms.
3. For any field, we must have $0 \leq 1$, because if not, then

$$0 > 1 \implies 0 < 1^{-1} \cdot 1 = 1^{-1} \implies 0 \cdot 0 < 1^{-1} \cdot 1^{-1} = 1 \quad (207)$$

So, $0 \leq 1$ implies that $0 \leq 1 \leq 2 \leq \dots \leq p-1$. But

$$0 + 1 \leq (p-1) + 1 = 0 \quad (208)$$

is false, so any ordering is impossible.

Exercise 1.43 (Zorich 2.2.23)

Show that if \mathbb{R} and \mathbb{R}' are two models of the set of real numbers and $f : \mathbb{R} \rightarrow \mathbb{R}'$ (with $f \not\equiv 0'$) is a mapping such that $f(x+y) = f(x) + f(y)$ and $f(x \cdot y) = f(x) \cdot f(y)$ for any $x, y \in \mathbb{R}$. Prove that f is an order-preserving isomorphism.

Solution 1.42

Let $0, 0'$ be the additive identity of \mathbb{R}, \mathbb{R}' , respectively, and $1, 1'$ the multiplicative identity. We claim that $f(0) = 0'$ since

$$\begin{aligned} f(0) &= f(0 + 0) && \text{(definition of additive identity)} \\ &= f(0) + f(0) && \text{(homomorphism over +)} \end{aligned}$$

which implies that $f(0) + f(0) = f(0) = 0' + f(0)$. Since $f(0)$ lives in field \mathbb{R}' , its additive identity $-f(0)$ is well defined, and we get $f(0) = f(0) + f(0) + (-f(0)) = 0' + f(0) + (-f(0)) = 0'$. We also claim that $f(1) = 1'$ since

$$\begin{aligned} f(1) &= f(1 \cdot 1) && \text{(definition of multiplicative identity)} \\ &= f(1) \cdot f(1) && \text{(homomorphism over \cdot)} \end{aligned}$$

which implies that $f(1) \cdot f(1) = 1' \cdot f(1)$. Since $f(1)$ lives in field \mathbb{R}' , its multiplicative identity $f(1)^{-1}$ is well defined, and we get $f(1) = f(1) \cdot f(1) \cdot f(1)^{-1} = 1' \cdot f(1) \cdot f(1)^{-1} = 1'$. Now that we have proved mapping of identities, this implies the mapping of inverses.

$$\begin{aligned} 0' &= f(0) = f(x - x) = f(x) + f(-x) \implies f(-x) = -f(x) \\ 1' &= f(1) = f(x \cdot x^{-1}) = f(x) \cdot f(x^{-1}) \implies f(x^{-1}) = f(x)^{-1} \end{aligned}$$

With these conditions, we have proved that f is a homomorphism of fields. Now we prove that f is a bijection, but first, we claim that $f(x) = 0' \implies x = 0$. Assume that there exists a nonzero $x \in \mathbb{R}$ such that $f(x) = 0'$. Then, x^{-1} is well defined, and

$$\begin{aligned} f(x) \cdot f(x^{-1}) &= f(x) \cdot f(x)^{-1} = 0' \\ f(x) \cdot f(x^{-1}) &= f(x \cdot x^{-1}) = f(1) = 1' \end{aligned}$$

which implies that $0' = 1'$. So, $f(1) = 1' = 0'$, and so for all $k \in \mathbb{R}$, $f(k) = f(k \cdot 1) = f(k) \cdot f(1) = f(k) \cdot 0' = 0' \implies f \equiv 0'$, leading to a contradiction of the assumption that $f' \not\equiv 0'$.

1. (f injective). Assume f is not injective, i.e. there exists distinct $x_1, x_2 \in \mathbb{R}$ s.t. $f(x_1) = f(x_2)$. Then, using that fact $f(x) = 0 \implies x = 0$,

$$0 = f(x_1) - f(x_2) = f(x_1 - x_2) \implies x_1 - x_2 = 0 \implies x_1 = x_2 \quad (209)$$

2. (f surjective). Let y be any nonzero element in \mathbb{R}' (clearly if $y = 0'$ then its preimage is 0) and y^{-1} its multiplicative inverse. Assume there exists no $x \in \mathbb{R}$ satisfying $f(x) = y$, meaning that there exist no x satisfying

$$f(x) \cdot y^{-1} = 1' \quad (210)$$

But since f maps inverses to inverses, we can choose $x = (y^{-1})^{-1}$, which leads to

$$f(x) \cdot y^{-1} = (y^{-1})^{-1} \cdot y^{-1} = 1' \quad (211)$$

Finally, we prove that f is order preserving. Assume that $x \leq y \iff 0 \leq y - x$, we wish to prove that

$$f(x) \leq f(y) \iff 0 \leq f(y) - f(x) = f(y - x) \quad (212)$$

Therefore, since this preservation of ordering is really the statement $0 \leq y - x \implies 0 \leq f(y - x)$, it suffices to prove that $0 \leq x \implies 0 \leq f(x)$. Now, assume that we have a x such that $f(x) < 0'$. Adding it with the equation $f(1) = 1'$ gives us

$$f(x + 1) < 1' \quad (213)$$

It is easy to prove that $0 \leq x \iff 0 \leq x^{-1}$. Now assume that $0 > f(x)$. **INCOMPLETE**

Exercise 1.44 (Density of Rationals in \mathbb{R})

Prove that for any two distinct $a < b \in \mathbb{R}$, there exists an infinite number of rational numbers between a and b .

Solution 1.43

Since $a < b$, then $b - a > 0$ and by the Archimedean principle, there exists a $k \in \mathbb{N}$ such that

$$0 < \frac{1}{k} < b - a \implies 1 < kb - ka \quad (214)$$

which implies that the length of $[ka, kb]$ greater than 1. By the inductive property of \mathbb{Z} , there must be an integer $p \in [ka, kb]$. If there were not, then this would imply that $[ka+1, kb+1)$ and $[ka-1, kb-1)$ had no integers and repeating would mean that there were no integers in \mathbb{R} . Therefore,

$$ka \leq p < kb \implies a \leq \frac{p}{k} < b \quad (215)$$

for all $a, b \in \mathbb{R}$, with $p/k \in \mathbb{Q}$. If a is irrational we can replace the \leq to $<$, leaving $a \leq \frac{p}{k} < b$, and if a is rational, we can construct another rational $a + \frac{1}{k} \in (a, b)$.

Exercise 1.45 (Nested Interval Lemma)

With the fact that \mathbb{R} is complete, prove the following.

1. For a sequence of closed nested intervals $I_1 \supset I_2 \supset \dots$ of \mathbb{R} , there exists a point $c \in \mathbb{R}$ belonging to all these intervals.
2. Furthermore, if the hypothesis also satisfies the fact that for any $\epsilon > 0$, there exists a $k \in \mathbb{N}$ such that $|I_k| < \epsilon$ (i.e. the length of the intervals decreases to 0), then the point c common to all sets is unique.

Solution 1.44

Listed.

1. Let $I_n = [a_n, b_n]$, with $a_n < b_n$ finite for all $n \in \mathbb{N}$. For all $n \in \mathbb{N}$, we have $I_n = [a_n, b_n]$ and can take the two subsets $X_n = (-\infty, a_n]$ and $Y_n = [b_n, \infty)$, where $x \leq y$ for every $x \in X_n, y \in Y_n$. We also have the fact that $\mathbb{R} = X_n \cup I_n \cup Y_n$. Since \mathbb{R} is complete, there exists a c such that $x \leq c \leq y$ for all $x \in X, y \in Y$. But $x \leq c \iff c \notin X_n$ and $c \leq y \iff c \notin Y_n$, so for all $n \in \mathbb{N}$, c must be in I_n .
2. Since we have proved (a), it now suffices to prove uniqueness of c . Let there be two distinct points $c_1, c_2 \in \mathbb{R}$ belonging to these intervals. Without loss of generality, assume $c_1 - c_2 > 0$, and choose

$$\epsilon = \frac{c_1 - c_2}{3} \quad (216)$$

Then, there should exist a $k \in \mathbb{N}$ such that $|I_k| < \epsilon$. Since I_k must contain c_1 , it must be a subset of $[c_1 - \epsilon, c_1 + \epsilon]$ (should be able to see why) and similarly for c_2 .

$$\begin{aligned} I_k &\subset \left[c_1 - \frac{c_1 - c_2}{3}, c_1 + \frac{c_1 - c_2}{3} \right] = \left[\frac{2c_1 + c_2}{3}, \frac{4c_1 - c_2}{3} \right] = L \\ I_k &\subset \left[c_2 - \frac{c_1 - c_2}{3}, c_2 + \frac{c_1 - c_2}{3} \right] = \left[\frac{-c_1 + 4c_2}{3}, \frac{c_1 + 2c_2}{3} \right] = M \end{aligned}$$

But since $c_1 > c_2 \implies \frac{c_1 + 2c_2}{3} < \frac{2c_1 + c_2}{3}$, L and M are disjoint $\implies I_k$, as a subset of both, leaves us with $I_k = \emptyset$, contradicting that it is a closed interval.

Exercise 1.46 ()

Compactness of Closed Interval in \mathbb{R} Prove that any system of open intervals covering (i.e. an open cover of) a closed interval contains a finite subsystem that covers the closed interval. Another way to state this is by saying that every closed interval of \mathbb{R} is compact.

Solution 1.45

A closed interval with a finite open covering is trivially compact since any subcovering is also finite. We only need to deal with when a closed interval $I = [a, b]$ has an infinite open covering $\{U_\alpha\}_{\alpha \in A}$, which means that the set of indices A is infinite. Assume that there exists no finite covering of I . Then, we divide I into two halves

$$I_1 = \left[a, \frac{a+b}{2}\right], \quad I_2 = \left[\frac{a+b}{2}, b\right] \quad (217)$$

and define a subcovering for each of them. That is, we can define $A_1 \subset A$ and $A_2 \subset A$ such that $\{U_\alpha\}_{\alpha \in A_1} \subset \{U_\alpha\}_{\alpha \in A}$ is a covering of I_1 and $\{U_\alpha\}_{\alpha \in A_2} \subset \{U_\alpha\}_{\alpha \in A}$ is a covering of I_2 . At least one of A_1 or A_2 must be infinite, since if they were both finite, then we can define a finite covering $\{U_\alpha\}_{\alpha \in A_1 \cup A_2}$ of I . Choose the interval with the infinite covering and repeat this procedure, which will result in a nested interval that decreases in length by a half.

$$I \supset I_1 \supset I_2 \supset \dots \quad (218)$$

By the nested interval lemma, there exists a unique point c common to all these intervals. But since $c \in [a, b]$, the open cover $\{U\}$ should contain an open interval $(c - \delta_1, c + \delta_2)$ containing c . We wish to prove that this interval is a superset of some I_k in the sequence above, contradicting the fact that I_k has an infinite cover. Since the length of each I_i decreases arbitrarily (i.e. we can choose any $\epsilon > 0$ and find a I_k with length less than ϵ), we choose $\epsilon = \frac{1}{2} \min\{\delta_1, \delta_2\}$, and we should be able to find some I_k that is a subinterval of $[c - \epsilon, c + \epsilon]$, which itself is a subinterval of $(c - \delta_1, c + \delta_2)$.

$$I_k \subset \left[c - \frac{1}{2} \min\{\delta_1, \delta_2\}, c + \frac{1}{2} \min\{\delta_1, \delta_2\}\right] \subset (c - \delta_1, c + \delta_2) \quad (219)$$

Therefore, $(c - \delta_1, c + \delta_2)$ is a finite cover of I_k , contradicting the fact that all I_k 's have infinite covers.

Exercise 1.47 ()

Bolzano-Weierstrass Theorem Prove that every bounded infinite set of real numbers has at least one limit point. (A limit point p of set X is a point such that every open neighborhood of p contains an infinite number of elements of X).

Solution 1.46

Let the set of points be denoted X , and let a be the lower bound and b be the upper bound. Then, $X \subset [a, b] = I$. Now divide $[a, b]$ into halves $[a, \frac{a+b}{2}] \cup [\frac{a+b}{2}, b]$. At least one of the halves must have an infinite number of points; choose the interval with infinite points as I_1 and doing this repeatedly gives the nested sequence

$$I \supset I_1 \supset I_2 \supset \dots \quad (220)$$

By the nested interval lemma, there exists at least one point $c \in \mathbb{R}$ that is in all these intervals. Furthermore, since $|I_i| = \frac{1}{2^i}(b - a)$ decreases to 0, we can choose a $\epsilon > 0$ and find an interval I_k with $|I_k| < \epsilon$. We claim that c is a limit point of X . Given an ϵ , we wish to prove that there are an infinite number of points within the ϵ -neighborhood $(c - \epsilon, c + \epsilon)$ of c . Since we can find some I_k with $|I_k| < \epsilon$,

we can see that

$$I_k \subset (c - \epsilon, c + \epsilon) \quad (221)$$

and therefore the ϵ -neighborhood of c contains I_k , which contains an infinite number of points in X .

Solution 1.47

We can construct another proof that is dependent on the compactness lemma. This construction will be useful for problem 2.3.4. Let X be a given subset of \mathbb{R} , and it follows from the definition of boundedness that X is contained in some closed interval $I \subset \mathbb{R}$. We show that at least one point of I is a limit point of X . Assume that it is not. Then each point $x \in I$ would have a neighborhood $U(x)$ containing at most a finite number of points from X . The totality of such neighborhoods $\{U(x)\}$ constructed for the points $x \in I$ forms an open covering of X . Since I is closed, it is compact and therefore we can find a finite subcovering $\{U_i(x)\}_i$ of open intervals that cover I and therefore cover X . This open cover $\{U_i(x)\}_i$ of X is a finite union of sets that each contain at most a finite number of points from X , so the covering of X contains a finite number of points from X , a contradiction that X contains infinite points.

Exercise 1.48 (Zorich 2.3.1)

Show that

1. if I is any system of nested closed intervals, then

$$\sup\{a \in \mathbb{R} \mid [a, b] \in I\} = \alpha \leq \beta = \inf\{b \in \mathbb{R} \mid [a, b] \in I\}$$

and

$$[\alpha, \beta] = \bigcap_{[a,b] \in I} [a, b]$$

2. if I is a system of nested open intervals (a, b) , the intersection

$$\bigcap_{(a,b) \in I} (a, b)$$

may happen to be empty.

Solution 1.48

Listed.

1. (May be tempted to say that $a_1 \leq a_2 \leq \dots$, but this assumes that the indexing set I is countable). We claim that for any two intervals $[a_n, b_n]$ and $[a_m, b_m]$ in I ,

$$a_n \leq b_m$$

Assume that $a_n > b_m$. Then $b_n \geq a_n > b_m \geq a_m$ implies that $[a_n, b_n]$ and $[a_m, b_m]$ are disjoint, contradicting the fact that they are nested. Now given that X is the set of a_n 's and Y is the set of b_n 's, we have $x \leq y$ for all $x \in X, y \in Y$. So by 2.2.16.b, we have $\sup X \geq \inf Y$.

To prove the second statement, we show that trying to “expand” the interval $[\alpha, \beta]$ will lead to a contradiction. Since α is the LUB, given any $\epsilon > 0$, there exists a $(a_l, b_l) \in I$ such that $\alpha - \epsilon < a_l < \alpha$, which implies that $[\alpha, \beta] \subset [a_l, \beta] \subset [\alpha - \epsilon, \beta]$. Assuming that this extended interval is the intersection, we should be able to choose any point in $[\alpha - \epsilon, \beta]$ and find that it is in every element of I . We choose a point in $[\alpha - \epsilon, a_l]$, which is not in the interval (a_l, b_l) . We do the same for $\beta \mapsto \beta + \epsilon$. We also check that “shrinking” the interval $[\alpha, \beta] \mapsto [\alpha + \epsilon, \beta]$ is no good, since we can find an element in $[\alpha, \alpha + \epsilon]$ that is in every interval in I .

2. Take the system of nested open intervals

$$(0, 1) \supset (0, \frac{1}{2}) \supset (0, \frac{1}{3}) \dots (0, \frac{1}{n}) \supset \dots$$

Take their infinite intersection, denote it S , and assume that some $\epsilon \in (0, 1)$ is in S . Since ϵ is a real number, by the Archimedean principle there exists a $k \in \mathbb{N}$ such that $\frac{1}{k} < \epsilon$. Therefore, $\epsilon \notin (0, \frac{1}{k}) \implies \epsilon \notin S$.

Exercise 1.49 (Zorich 2.3.2)

Show that

1. from a system of closed intervals covering a closed interval it is not always possible to choose a finite subsystem covering the interval.
2. from a system of open intervals covering a open interval it is not always possible to choose a finite subsystem covering the interval.
3. from a system of closed intervals covering a open interval it is not always possible to choose a finite subsystem covering the interval.

Solution 1.49

We show with the interval $(0, 1)$ or $[0, 1]$. Using linear transformations it is easy to generalize this to any other interval (a, b) or $[a, b]$.

1. Consider the infinite covering

$$[0, 1] = [0, \frac{1}{2}] \cup [\frac{1}{2}, \frac{3}{4}] \cup [\frac{3}{4}, \frac{7}{8}] \cup \dots$$

2. Consider the infinite covering

$$(0, 1) = (0, \frac{1}{2}) \cup (\frac{1}{2}, \frac{3}{4}) \cup (\frac{3}{4}, \frac{7}{8}) \cup \dots$$

3. Consider the infinite covering

$$(0, 1) = [0, \frac{1}{2}] \cup [\frac{1}{2}, \frac{3}{4}] \cup [\frac{3}{4}, \frac{7}{8}] \cup \dots$$

Exercise 1.50 (Zorich 2.3.3)

Show that if we only take the set \mathbb{Q} of rational numbers instead of the complete set \mathbb{R} of real numbers, with the definitions of closed, open, and neighborhood of a point $r \in \mathbb{Q}$ to mean respectively the corresponding subsets of \mathbb{Q} , then none of the three lemmas is true.

Solution 1.50

We prove only for the nested interval lemma. We choose the series of nested intervals

$$\left(\sqrt{2} - \frac{1}{n}, \sqrt{2} + \frac{1}{n}\right)$$

with $n \in \mathbb{N}$. Assume that there is a $r \in \mathbb{Q}$ such that

$$r \in \left(\sqrt{2} - \frac{1}{n}, \sqrt{2} + \frac{1}{n}\right) \text{ for all } n \in \mathbb{N}$$

which is equivalent to saying that $|r - \sqrt{2}| < \frac{1}{n}$ for all $n \in \mathbb{N}$. Clearly, $r \neq \sqrt{2}$, and by the Archimedean principle, there exists a $k \in \mathbb{N}$ such that

$$0 < \frac{1}{k} < |r - \sqrt{2}|$$

which contradicts the above.

Exercise 1.51 (Zorich 2.3.4)

Show that the three lemmas above are equivalent to the axiom of completeness.

Solution 1.51

Note that from the proofs, completeness implies nested interval lemma, which implies compactness of closed intervals, which implies the Bolzano-Weierstrass theorem. So, it is sufficient to prove that Bolzano-Weierstrass theorem implies completeness to determine equivalence. There are not a lot of direct proofs, so we prove that Weierstrass implies nested interval, which implies completeness.

1. (Weierstrass \implies Nested) Assume that we have \mathbb{R} with the Bolzano-Weierstrass theorem. Take the series of nested closed intervals

$$I = [a, b] \supset I_1 = [a_1, b_1] \supset I_2 = [a_2, b_2] \supset \dots$$

We see that $a \leq a_i \leq b$, so the infinite sequence of monotonically nondecreasing values a_i is bounded. Therefore, it must have a limit point, which we will denote as c . We claim that $a_i \leq c$ for all a_i . Since if it were not, then $c < a_i$ for some i , and choosing $\epsilon = 0.5(a_i - c)$, the ϵ -neighborhood of c will not contain a_j for $j \geq i$ since

$$c < a_i \implies 0.5c < 0.5a_i \implies c + \epsilon = 0.5c + 0.5a_i < a_i < a_{i+1} < \dots$$

. With similar reasoning, we can conclude that $b_i \geq c$ for all b_i . This implies that $a_i \leq c \leq b_i$ for all i which is equivalent to saying that $c \in [a_i, b_i] = I_i$ for all $i \in \mathbb{N}$.

2. (Nested \implies LUB Principle) Let $X \subset \mathbb{R}$ be a set that is bounded above, with b_1 any upper bound. Since X is nonempty, there exists $a_1 \in X$ that is not an upper bound (otherwise, X would be a singleton set and it trivially has a least upper bound). Consider the well-defined interval $[a_0, b_0]$. Take the mean $m_0 = 0.5(a_0 + b_0)$, and if m_0 is an upper bound, set it to b_1 (with $a_1 = a_0$) and a_1 if else (with $b_1 = b_0$). Then, we have a sequence of nested intervals

$$[a_0, b_0] \supset [a_1, b_1] \supset [a_2, b_2] \supset \dots$$

of decreasing lengths $|I_k| = \frac{1}{2^k}(b - a)$. All of them must contain a unique common point $c \in \mathbb{R}$ by the nested intervals lemma, which implies that

$$a_0 \leq a_1 \leq a_2 \leq \dots \leq c \leq \dots \leq b_2 \leq b_1 \leq b_0$$

I claim two things:

- (a) c is an upper bound for X . Suppose it were not, then there exists some $x \in X$ such that $c < x$, and let the distance between them be $\epsilon = x - c > 0$. By AP, we can choose $k \in \mathbb{N}$ such that $\frac{1}{k} < \epsilon$. All the b_n are upper bounds of X , so we have $x \leq b_n$. Subtracting c on both sides gives

$$0 < x - c = \epsilon \leq b_n - c \leq |I_n| = \frac{1}{2^n}(b_0 - a_0)$$

where the last inequality follows from $c \in I_n = [a_n, b_n]$, so the maximum distance it can be from the endpoint b_n is $|I_n|$. The inequality above holds for all $n \in \mathbb{N}$, so increasing n

arbitrarily should decrease $\frac{1}{2^n}(b_0 - a_0)$ past ϵ . To formalize this, we use the inequality

$$\frac{1}{2^n} < \frac{1}{n} \text{ for all } n \in \mathbb{N}$$

and so we have

$$\epsilon \leq b_n - c < \frac{1}{n}(b_0 - a_0)$$

We choose the natural number $n = \lceil \frac{2(b_0 - a_0)}{\epsilon} \rceil$, which does not satisfy the inequality above since

$$\epsilon < \frac{1}{n}(b_0 - a_0) = \frac{1}{\lceil 2(b_0 - a_0)/\epsilon \rceil}(b_0 - a_0) \leq \frac{\epsilon}{2(b_0 - a_0)}(b_0 - a_0) = \frac{\epsilon}{2}$$

This leads to a contradiction.

- (b) We now prove that c is least. Assume that c is not least \implies there exists an upper bound B such that $B < c$ and $x \leq B$ for all $x \in X$. **INCOMPLETE**

Exercise 1.52 (Zorich 2.4.1)

Show that the set of real numbers has the same cardinality as the points of the interval $(-1, 1)$.

Solution 1.52

We define the bijective map $\rho : (-1, 1) \rightarrow \mathbb{R}$

$$\rho(x) = \begin{cases} 0 & \text{if } x = 0 \\ \frac{1}{x} & \text{if } x \neq 0 \end{cases}$$

Exercise 1.53 (Zorich 2.4.2)

Give an explicit one-to-one correspondence between

1. the points of two open intervals
2. the points of two closed intervals
3. the point of a closed interval and an open interval
4. the points of the closed interval $[0, 1]$ and \mathbb{R}

Solution 1.53

Listed.

1. $\rho : (a, b) \rightarrow (c, d)$ defined

$$\rho(x) = \frac{d-c}{b-a}(x-a) + c$$

2. the extension of ρ defined on (a) to $[a, b]$
3. From (a) and (b), it suffices to prove a bijection from $(0, 1)$ to $[0, 1]$. We extract a countably infinite sequence from $(0, 1)$, say

$$x_1 = \frac{1}{3}, x_2 = \frac{1}{4}, \dots, x_i = \frac{1}{i+2}$$

Then, we define bijection $\rho : (0, 1) \rightarrow [0, 1]$ as

$$\rho(x) = \begin{cases} x & \text{if } x \notin \{x_i\} \\ 0 & \text{if } x = x_1 = \frac{1}{2} \\ 1 & \text{if } x = x_2 = \frac{1}{3} \\ x_{i-2} & \text{if } x = x_i \text{ for } i > 2 \end{cases}$$

Colloquially, we extract a copy of \mathbb{N} from $(0, 1)$ and use the bijection $\mathbb{N} \simeq \mathbb{N} \cup \{0, -1\}$ to “shift” the terms.

4. We compose the bijections $\rho_1 : [0, 1] \rightarrow (0, 1)$ and $\rho_2 : (0, 1) \rightarrow \mathbb{R}$.

Exercise 1.54 (Zorich 2.4.3)

Show that

1. every infinite set contains a countable subset
2. the set of even integers has the same cardinality as the set of all natural numbers
3. the union of an infinite set and an at most countable set has the same cardinality as the original infinite set.
4. the set of irrational numbers has the cardinality of the continuum
5. the set of transcendental numbers has the cardinality of the continuum

Solution 1.54

Listed.

1. Let A be an infinite set. By axiom of choice, choose $a_0 \in A$. Then, $A \setminus \{a_0\} \neq \emptyset$ since A is infinite. By induction, assume you have chosen $a_0, a_1, \dots, a_k \in A$. Then, since A is infinite, $A \setminus \{a_0, a_1, \dots, a_k\} \neq \emptyset$, so we can choose $a_{k+1} \in A \setminus \{a_0, \dots, a_k\}$. Thus, we have constructed a countable subset $\{a_k\}_{k \in \mathbb{N}}$ of A .
2. Given the quotient ring $2\mathbb{Z}$, define the bijection $\rho : 2\mathbb{Z} \rightarrow \mathbb{N}$ as

$$\rho(x) = \begin{cases} x + 2 & \text{if } x \geq 0 \\ -x - 1 & \text{if } x < 0 \end{cases}$$

3. From (a), we can extract a countable set from original set A , call it X . Since the product of countable sets is countable ($\mathbb{N} \cup \mathbb{N}$ is countable), we can define a bijection $\tilde{\rho} : X \rightarrow X \cup B$. Therefore, we can define a bijection $\rho : A \rightarrow A \cup B$ as

$$\rho(x) = \begin{cases} x & \text{if } x \in A \setminus X \\ \tilde{\rho}(x) & \text{if } x \in X \end{cases}$$

4. \mathbb{Q} is countable and \mathbb{R} is uncountable. So, $\mathbb{R} \setminus \mathbb{Q}$ must be uncountable since if it were countable, then the union of the rationals and irrationals, which is \mathbb{R} , would be countable.
5. It suffices to prove that the set of algebraic numbers (numbers that are possible roots of a polynomial with integer coefficients with leading coefficient nonzero) is countable, since we can apply (d) right after. The set of all k th degree polynomials with integer coefficients is isomorphic to \mathbb{Z}^k through the map

$$a_k x^k + a_{k-1} x^{k-1} + \dots + a_2 x^2 + a_1 x^1 + a_0 \mapsto (a^{k-1}, a^{k-2}, \dots, a_1, a_0)$$

and the union of these countable sets (minus the 0 map)

$$P = \left(\bigcup_{k=1}^{\infty} \mathbb{Z}^k \right) \setminus \{0\} = (\mathbb{Z} \setminus \{0\}) \cup \mathbb{Z}^2 \cup \dots$$

is countable. For any element in \mathbb{Z}^k , there are at most k real roots, and so we can define the set of roots of an element $z \in \mathbb{Z}^k \subset P$ as a j -tuple of algebraic numbers, which can have at most $j = k$ roots.

$$r(z) = \underbrace{(r_{1z}, r_{2z}, \dots, r_{jz})}_{j \leq k}$$

Therefore, the union of all these j -tuples for all $z \in P$

$$\bigcup_{z \in P} r(z) = \bigcup_{k=1}^{\infty} \bigcup_{z \in \mathbb{Z}^k} r(z)$$

is a countable union of a countable union of finite sets, making it countable.

Exercise 1.55 (Zorich 2.4.4)

Show that

1. the set of increasing sequences of natural numbers has the same cardinality as the set of fractions of the form $0.\alpha_1\alpha_2\dots$
2. the set of all subsets of countable set has cardinality of the continuum

Solution 1.55

Listed.

1. Given a sequence of increasing naturals $S = (n_1, n_2, \dots)$, we can define a binary expansion $0.\alpha_1\alpha_2\dots$ where $\alpha_i = 1$ if and only if $i \in \mathbb{N}$ is in S and $\alpha_i = 0$ if not. This is clearly a bijection.
2. The set of all segments of increasing natural is equipotent with $2^{\mathbb{N}}$, since the elements of each sequence define a subset of \mathbb{N} . Cantor's diagonalization argument proves that the set of infinite binary expansions is uncountable, and by (a), this proves that $2^{\mathbb{N}}$ is uncountable.

This is very interesting since $\mathbb{N} \simeq \mathbb{R}$, but $2^{\mathbb{N}} \simeq \mathbb{R}$, and the set of all infinite q -ary expansions is equipotent to \mathbb{R} too.

Exercise 1.56 (Zorich 2.4.5)

Show that

1. the set $\mathcal{P}(X)$ of subsets of a set X has the same cardinality as the set of all functions $f : X \rightarrow \{0, 1\}$.
2. for a finite set X of n elements, $\text{card } \mathcal{P}(X) = 2^n$
3. one can write $\text{card } \mathcal{P}(X) = 2^{\text{card } X}$, which implies $\text{card } \mathcal{P}(\mathbb{N}) = 2^{\text{card } \mathbb{N}} = \text{card } \mathbb{R}$
4. for any set X , $\text{card } X < 2^{\text{card } X}$

Solution 1.56

Listed.

1. An element $Y \in \mathcal{P}(X)$ is a subset of X by definition. Letting

$$f_Y(x) = \begin{cases} 0 & \text{if } x \notin Y \\ 1 & \text{if } x \in Y \end{cases}$$

we can construct the bijective map $Y \mapsto f_Y$.

2. We can prove this using the identity (which can be proved using induction)

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

3. Let $F(X; \{0, 1\})$ be the set of all binary valued functions from X to $\{0, 1\}$. From (a), $\text{card } \mathcal{P}(X) \simeq F(X; \{0, 1\})$. Each binary-valued function f is determined by the assignment $f(x)$ for each $x \in X$. Since $f(x)$ has two possible values, the assignment of $f(x)$ for all $x \in X$ has $\{0, 1\}^{\text{card } X}$ possible choices. This gives another bijection $F(X; \{0, 1\}) \simeq \{0, 1\}^{\text{card } X}$, so

$$\mathcal{P}(X) \simeq \{0, 1\}^{\text{card } X} \implies \text{card } \mathcal{P}(X) = \text{card}(\{0, 1\}^{\text{card } X}) = 2^{\text{card } X}$$

4. If X is finite, then letting $n = \text{card } X$, we can simply prove $n < 2^n$ by induction (which we will not do here). If X is countable, then $\mathcal{P}(X)$ is uncountable (from 2.4.4) and so using (c),

$$\text{card } X = \text{card } \mathbb{N} < \text{card } \mathbb{R} = \text{card } \mathcal{P}(X) = 2^{\text{card } X}$$

For uncountable sets (and for the two cases mentioned above), we can use Cantor's theorem, which states that $\text{card } X < \text{card } \mathcal{P}(X)$, and so using (c), we have $\text{card } X < \text{card } \mathcal{P}(X) = 2^{\text{card } X}$.

Exercise 1.57 (Zorich 2.4.6)

Let X_1, \dots, X_m be a finite system of finite sets. Show that

$$\begin{aligned} \text{card} \left(\bigcup_{i=1}^m X_i \right) &= \sum_{i_1} \text{card } X_{i_1} - \sum_{i_1 < i_2} \text{card}(X_{i_1} \cap X_{i_2}) + \dots \\ &\quad \sum_{i_1 < i_2 < i_3} \text{card}(X_{i_1} \cap X_{i_2} \cap X_{i_3}) - \dots + (-1)^{m-1} \text{card}(X_1 \cap \dots \cap X_m) \\ &= \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) \end{aligned}$$

Solution 1.57

Ignoring Russell's paradox (defining the universe set of all sets), we can use the commutative, associative, and distributive properties of \cup, \cap on the algebra of sets. We prove using induction on m . For $m = 1$, we trivially have $\text{card } X_1 = \text{card } X_1$, and for $m = 2$, we claim

$$\text{card}(X_1 \cup X_2) = \text{card}(X_1) + \text{card}(X_2) - \text{card}(X_1 \cap X_2)$$

X_1 and $X_2 \setminus X_1$ are clearly exclusive sets by definition, with $X_1 \cup X_2 = X_1 \cup (X_2 \setminus X_1)$, so

$$\text{card}(X_1 \cup X_2) = \text{card}(X_1 \cup (X_2 \setminus X_1)) = \text{card}(X_1) + \text{card}(X_2 \setminus X_1) \tag{2}$$

By definition, the set $X_2 \setminus X_1$ and $X_1 \cap X_2$ are disjoint and satisfies $X_2 = (X_2 \setminus X_1) \cup (X_1 \cap X_2)$ (also by definition), so

$$\text{card}(X_2) = \text{card}(X_2 \setminus X_1) + \text{card}(X_1 \cap X_2) \tag{3}$$

and substituting (3) into (2) gives the claim for $m = 2$. Assuming that the claim is satisfied for some m , we have

$$\begin{aligned}
 \text{card} \left(\bigcup_{i=1}^{m+1} X_i \right) &= \text{card} \left(\left[\bigcup_{i=1}^m X_i \right] \cup X_{m+1} \right) \\
 &= \text{card} \left(\bigcup_{i=1}^k X_i \right) + \text{card}(X_{m+1}) - \text{card} \left(\left[\bigcup_{i=1}^m X_i \right] \cap X_{m+1} \right) \quad (\text{claim for } m = 2) \\
 &= \text{card} \left(\bigcup_{i=1}^k X_i \right) + \text{card}(X_{m+1}) - \text{card} \left(\bigcup_{i=1}^m (X_i \cap X_{m+1}) \right) \quad (\text{distributive prop.}) \\
 &= \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) + \text{card}(X_{m+1}) \\
 &\quad - \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k (X_{i_j} \cap X_{m+1}) \right)
 \end{aligned}$$

With a bit of thought, we can see that the k th term of the second summation contributes to adding another term to the $k+1$ th summation term of the first. Therefore, we must try to shift the summation over by 1 index. Let us simplify this by taking the summations and extracting the first and last term, respectively. We have

$$\begin{aligned}
 \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) &= \sum_{1 \leq i_1 \dots i_k \leq m} \text{card}(X_{i_1}) \\
 &\quad + \sum_{k=2}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right)
 \end{aligned}$$

and

$$\begin{aligned}
 \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k (X_{i_j} \cap X_{m+1}) \right) \\
 &= \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\left[\bigcap_{j=1}^k X_{i_j} \right] \cap X_{m+1} \right) \\
 &= \sum_{k=1}^{m-1} \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k (X_{i_j} \cap X_{m+1}) \right) + (-1)^{m-1} \text{card} \left(\bigcap_{j=1}^{m+1} X_j \right) \\
 &= \sum_{k=2}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-2} \text{card} \left(\bigcap_{j=1}^{k-1} (X_{i_j} \cap X_{m+1}) \right) + (-1)^{m-1} \text{card} \left(\bigcap_{j=1}^{m+1} X_j \right)
 \end{aligned}$$

So subtracting the summations gives

$$\begin{aligned}
 & \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) - \sum_{k=1}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k (X_{i_j} \cap X_{m+1}) \right) + |X_{m+1}| \\
 &= \sum_{1 \leq i_1 \dots i_k \leq m} \text{card}(x_i) + \sum_{k=2}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) + \text{card}(X_{m+1}) \\
 &\quad + \sum_{k=2}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^{k-1} (X_{i_j} \cap X_{m+1}) \right) + (-1)^m \text{card} \left(\bigcap_{j=1}^{m+1} X_j \right) \\
 &= \sum_{1 \leq i_1 \dots i_k \leq m+1} \text{card}(X_i) + \sum_{k=2}^m \sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \left[\text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) \right. \\
 &\quad \left. + \text{card} \left(\left[\bigcap_{j=1}^{k-1} X_{i_j} \right] \cap X_{m+1} \right) \right] + (-1)^m \text{card} \left(\bigcap_{j=1}^{m+1} X_j \right)
 \end{aligned}$$

and since the set of sequences of k terms bounded by $m+1$ (of form $1 \leq i_1 \dots i_k \leq m+1$) is the set of sequences of k terms bounded by m (of form $1 \leq i_1 \dots i_k \leq m$) unioned with the set of sequences of k terms with max element $m+1$ (of form $1 \leq i_1 \dots i_k = m+1$), we have

$$\sum_{1 \leq i_1 \dots i_k \leq m} (-1)^{k-1} \left[\text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) + \text{card} \left(\left[\bigcap_{j=1}^{k-1} X_{i_j} \right] \cap X_{m+1} \right) \right] = \sum_{1 \leq i_1 \dots i_k \leq m+1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right)$$

and therefore, substituting the above and observing that the independent terms are the first and last terms of the summation gives

$$\begin{aligned}
 \text{card} \left(\bigcup_{i=1}^{m+1} X_i \right) &= \sum_{1 \leq i_1 \dots i_k \leq m+1} \text{card}(X_i) + \sum_{k=2}^m \sum_{1 \leq i_1 \dots i_k \leq m+1} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right) \\
 &\quad + \dots + (-1)^m \text{card} \left(\bigcap_{j=1}^{m+1} X_j \right) \\
 &= \sum_{k=1}^{m+1} \sum_{1 \leq i_1 \dots i_k \leq m+1} (-1)^{k-1} \text{card} \left(\bigcap_{j=1}^k X_{i_j} \right)
 \end{aligned}$$

Exercise 1.58 (Zorich 2.4.7)

On the closed interval $[0, 1] \subset \mathbb{R}$, describe the sets of numbers $x \in [0, 1]$ whose ternary representation $x = 0.\alpha_1\alpha_2\dots$, $\alpha_i \in \{0, 1, 2\}$ has the property.

1. $\alpha_1 \neq 1$
2. $\alpha_1 \neq 1$ and $\alpha_2 \neq 1$
3. For all $i \in \mathbb{N}$, $\alpha_i \neq 1$ (the Cantor set)

Solution 1.58

Listed.

1. $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$
2. $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$
3. Made by recursively removing the middle third of every partitioned intervals.

Exercise 1.59 (Zorich 2.4.8)

Show that

1. the set of numbers $x \in [0, 1]$ whose ternary representation does not contain 1 has the same cardinality as the set of all numbers whose binary representation has the form $0.\beta_1\beta_2\dots$
2. the Cantor set has the same cardinality as the closed interval $[0, 1]$

Solution 1.59

Listed.

1. We can define a bijection $0.\alpha_1\alpha_2\dots \mapsto 0.\beta_1\beta_2\dots$ as $\alpha_i = 0 \iff \beta_i = 0$ and $\alpha_i = 1 \iff \beta_i = 2$.
2. The map above defines a bijection between the Cantor set and the set of all infinite binary expansions in $[0, 1]$, which is uncountable by Cantor's diagonalization theorem.

2 Euclidean Topology

With the construction of the real line and the real space, the extra properties of completeness, norm, and order (for the real line) allows us to restate these topological properties in terms of these “higher-order” properties. It also proves much more results than for general topological spaces. Therefore, the next few sections will focus on reiterating the topological properties of \mathbb{R} and \mathbb{R}^n (this can be done slightly more generally for metric spaces, but we talk about this in point-set topology). In this section, we will restate the notion of open sets, limit points, compactness, connectedness, and separability. Then we can continue in the next section sequences and their limits, and after that we describe continuity. Once this is done, we can focus constructing the derivative and integral, which are unique to Banach spaces.

2.1 Open Sets

It is well-known that the set of open-balls of a metric space (X, d) is indeed a topology, which we prove in point-set topology. Once we prove this, we have access to a whole suite of theorems on topological spaces that we can just apply to \mathbb{R}^n . We will restate many of these topological theorems for completeness but will not prove them. However, if any of these theorems use any other structure, such as order/metrics/norms/-completeness, we will have to prove them.

Definition 2.1 (Topology)

Let X be a set and \mathcal{T} be a family of subsets of X . Then \mathcal{T} is a **topology** on X ^a if it satisfies the following properties.

1. *Contains Empty and Whole Set:*

$$\emptyset, X \in \mathcal{T} \quad (222)$$

2. *Closure Under Union.* If $\{U_\alpha\}_{\alpha \in A}$ is a class of sets in \mathcal{T} , then

$$\bigcup_{\alpha \in A} U_\alpha \in \mathcal{T} \quad (223)$$

3. *Closure Under Finite Intersection:* If U_1, \dots, U_n is a finite class of sets in \mathcal{T} , then

$$\bigcap_{i=1}^n U_i \in \mathcal{T} \quad (224)$$

A **topological space** is denoted (X, \mathcal{T}) .

^aI will use script letters to denote topologies and capital letters to denote sets.

Theorem 2.1 (Euclidean Topology)

Let $\tau_{\mathbb{R}}$ (which we denote as \mathcal{T}) be the set of subsets S of $(\mathbb{R}^n, \|\cdot\|)$ satisfying the property that if $x \in S$, then there exists an open ϵ -ball $B(x, \epsilon)$ s.t. $B \subset S$. \mathcal{T} is a topology of \mathbb{R}^n .

Proof.

We prove the following three properties.

1. \emptyset, \mathbb{R}^n are open.
2. For any collection $\{G_\alpha\}_\alpha$ of open sets, $\bigcup_\alpha G_\alpha$ is open.
3. For any finite collection G_1, \dots, G_n of open sets, $\bigcap_{i=1}^n G_i$ is open.

Listed.

1. Let $x \in \bigcup_\alpha G_\alpha$. Then, $x \in G_k$ for some k and since G_k is open, there exists a $B_\epsilon(x) \subset G_k \subset \bigcup_\alpha G_\alpha$, proving that $\bigcup_\alpha G_\alpha$ is open.

2. Let $x \in \cap_{i=1}^n G_i$. Then, $x \in G_i$ for every i , and so for each G_i , there exists an $\epsilon_i > 0$ s.t. $B_{\epsilon_i}(x) \subset G_i$. Since the set $\{e_i\}$ is finite, we can take

$$\epsilon = \min_i \{\epsilon_i\}$$

and see that $B_\epsilon(x) \subset G_i$ for all i , which implies that $B_\epsilon(x) \subset \cap_{i=1}^n G_i$. Since we have proved the existence of ϵ , $\cap_{i=1}^n G_i$ is open.

Definition 2.2 (Open Set)

An **open set** is an element of \mathcal{T} .

1. An **open neighborhood**, or sometimes just the **neighborhood**, of $x \in \mathbb{R}^n$ is an open set U_x containing x .
2. A **punctured neighborhood** is $U_x^\circ = U_x \setminus \{x\}$.

Theorem 2.2 (Equivalence to Open Ball Topology)

\mathcal{T} is equal to the topology \mathcal{T}' generated by the basis \mathcal{B} of open balls

$$B(x, r) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\} \quad (225)$$

Proof.

Let \mathcal{T} be the Euclidean topology and \mathcal{T}' be the open ball topology.

1. We show $\mathcal{T} \subset \mathcal{T}'$.
2. We show $\mathcal{T}' \subset \mathcal{T}$.

By defining the topology, we have automatically defined a bunch of topological objects and properties. For clarification, we will restate them.

Corollary 2.1 ()

An open ball is an open set.

Proof.

Given $x \in B_r(p)$, we can imagine that x will always have some space between it and the boundary. We want to show that there exists some $\epsilon > 0$ s.t. $B_\epsilon(x) \subset B_r(p)$. That is, given any point $y \in B_\epsilon(x)$, we can show that $y \in B_r(p)$. Since $\|x - p\| < r$, there exists some space $0 < r - \|x - p\|$. There always exists a real number $0 < \epsilon < r - \|x - p\|$, so given $y \in B_\epsilon(x)$, we can bound

$$\|y - p\| = \|y - x + x - p\| \leq \|y - x\| + \|x - p\| \leq \epsilon + \|x - p\| < r \quad (226)$$

Example 2.1 ()

Here are some examples of sets which are open and not open.

1. $U = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \neq 1\}$ is open since for every point $x \in U$, we just need to find a radius $\epsilon > 0$ that is smaller than its distance to the unit circle.
2. $(a, b) \times (c, d) \subset \mathbb{R}^2$ is open since given a point x , we can take the minimum of its distance between the two sides of the rectangle and construct an open ball.
3. $S = \{(x, y) \in \mathbb{R}^2 : xy \neq 0\}$ is open since given a point $x \in S$, we can take the minimum of the distance between it and the x and y axes.

4. The set of all complex z such that $|z| \leq 1$ is not open since we cannot construct open balls at the boundary points that are fully contained in the set.
5. The set $S = \{1/n\}_{n \in \mathbb{N}}$ is not open since given any point $x = 1/n$, we can construct an open ball with radius $\epsilon < 1/(n+1)$, which contains irrationals that are not in S .

Definition 2.3 (Interior Point)

A point $p \in S$ is an **interior point** if there exists a neighborhood N of p such that $N \subset S$.

An interior point means that we can always contain the point in S with some "breathing room." By definition an open set is a set where all of its points are interior points. A set is then said to be open if every point has this breathing room. This can be useful when defining differentiation at a point within an open set, since we can always find a neighborhood to take limits on.

Now that we have defined the Euclidean topology, we will prove that the features of topological objects can be reduced to features in \mathbb{R}^n .

Theorem 2.3 (Convexity)

An open ball is convex in a normed vector space.

Proof.

The normed part is important here, as the properties of the metric is not sufficient. Given $B_r(p)$, $x, y \in B_r(p)$ implies that $\|x - p\| < r$ and $\|y - p\| < r$. Therefore,

$$\|tx + (1-t)y - p\| = \|tx - tp + (1-t)y - (1-t)p\| \quad (227)$$

$$\leq t\|x - p\| + (1-t)\|y - p\| \quad (228)$$

$$= tr + (1-t)r = r \quad (229)$$

What happens if we weaken it to a metric?

2.2 Limit Points and Closure

Definition 2.4 (Limit Point)

A point $p \in \mathbb{R}^n$ is a **limit point** of $S \subset \mathbb{R}^n$ if every punctured neighborhood of p has a nontrivial intersection with S .^a The set of all limit points of S is denoted S' .

^aThe definition just means that if we take a point and draw smaller and smaller circles around it, the circle itself should still overlap with S , no matter how small it gets.

Theorem 2.4 ()

Let A_1, \dots, A_n be a finite collection of sets. Then

$$\bigcup_{i=1}^n A'_i = \left(\bigcup_{i=1}^n A_i \right)'$$

Proof.

Let the LHS be W and the RHS be V . If $x \in W$, $x \in A'_i$ for some i , and so for all $\epsilon > 0$, there exists a $B_\epsilon^\circ(x)$ s.t.

$$B_\epsilon^\circ(x) \cap A_i \neq \emptyset \implies B_\epsilon^\circ(x) \cap \left(\bigcup_{i=1}^n A_i \right) \neq \emptyset$$

which means that $x \in V$. Now assume that $x \in V$. Then for all $\epsilon > 0$, there exists a $B_\epsilon^\circ(x)$ s.t.

$$B_\epsilon^\circ(x) \cap \left(\bigcup_{i=1}^n A_i \right) \neq \emptyset$$

which implies that $B_\epsilon^\circ(x) \cap A_i \neq \emptyset$ for some i , which means that $x \in A'_i \subset W$.

A closed set can be defined in many equivalent ways for arbitrary topological spaces. The more general proof is done in topology, but we still prove it in the context of analysis.

Definition 2.5 (Closed Set)

A **closed set** $S \in \mathbb{R}^n$ is a set that contains all of its limit points.

Theorem 2.5 (Alternative Definition of Closed Set)

A set S is closed iff S^c is open.

Proof.

We prove both ways:

1. (\rightarrow) Given that S is closed, then let $x \in S^c$. x is not a limit point of S since if it were, then it would be in S , and so there exists a punctured open neighborhood $B_\epsilon^\circ(x)$ of x s.t. $S \cap B_\epsilon^\circ(x) = \emptyset$. Since $x \notin S$, we also have $S \cap B_\epsilon^\circ(x) = \emptyset$, which implies that $B_\epsilon^\circ(x) \subset S^c$. Since for every $x \in S^c$, there exists a $B_\epsilon^\circ(x) \subset S^c$, S^c is open.
2. (\leftarrow) For simplicity, it suffices to prove if S open, then S^c is closed. Given that S is open, we have for every $x \in S$, there exists $B_\epsilon^\circ(x) \subset S$, which implies that $B_\epsilon^\circ(x) \cap S^c = \emptyset$. Since there exists an $B_\epsilon^\circ(x)$ that does not contain points in S^c , x cannot be a limit point of S^c , and so there exists no limit points of S^c in S . Therefore, all limit points of S^c are in S^c , proving that S^c is closed.

Theorem 2.6 ()

We have the following topological properties:

1. For any collection $\{F_\alpha\}_\alpha$ of closed sets, $\bigcap_\alpha F_\alpha$ is closed.
2. For any finite collection F_1, \dots, F_n of open sets, $\bigcup_{i=1}^n F_i$ is closed.

Proof.

Listed.

1. Let x be a limit point of $\bigcap_\alpha F_\alpha$, and we want to show that $x \in \bigcap_\alpha F_\alpha$. By definition of limit points, for every $\epsilon > 0$, we have

$$B_\epsilon^\circ(x) \cap \left(\bigcap_\alpha F_\alpha \right)$$

which means that $B_\epsilon^\circ(x) \cap F_\alpha \neq \emptyset$ for all α . This means that x is a limit point for every F_α ,

and since they are all closed, $x \in F_\alpha$ for all α , which implies that $x \in \cap_\alpha F_\alpha$.

We can intuitively see a few properties about this. First, a finite set S of points does not have any limit points, since if we draw small enough circles around a $p \in S$, then at some point the circle will not contain any more points (remember that we're talking about deleted neighborhoods). Following this, we can deduce that a limit point must always have an infinite number of points close to it, as in no matter how small the circle gets, there are always an infinite number of points contained within that circle. This also means that if p is a limit point, then we can construct a sequence of points in S that converges to p , since every open ball with smaller and smaller radii will still have points in S .

Theorem 2.7 ()

If p is a limit point of S , then every neighborhood of p contains infinitely many points of S . The converse is also true trivially.

Proof.

Assume p is a limit point and that there exists a finite number of points within a deleted neighborhood $B_r^\circ(p)$. Then, we can enumerate them p_1, p_2, \dots, p_n by their distances to p , with

$$d(p_1, p) \leq d(p_2, p) \leq \dots \leq d(p_n, p) \quad (230)$$

Since $p_1 \neq p$, we have $d(p_1, p) > 0$ and so, we can choose an $0 < \epsilon < d(p_1, p)$ s.t. $B_\epsilon^\circ(p)$ does not contain any of the p_i 's. This neighborhood does not contain any elements of S and so p is not a limit point.

Corollary 2.2 ()

A finite set has no limit points.

Proof.

If S is a finite set, then every neighborhood of every point p in \mathbb{R}^n will have at most finite points, which, by the previous theorem, is not a limit point.

We show a very useful result that will make things much more convenient when proving the following theorems and exercises. This is quite intuitive, since it shows that the limit points of a finite union of sets is the same as the finite union of the limit points of each set. This is clearly not true for infinite unions:

1. Look at the countable set $\mathbb{Q} \subset \mathbb{R}$. Each $\{q\}' = \emptyset$, but $\mathbb{Q}' = \mathbb{R}$.
2. Look at the uncountable set \mathbb{R} . Each $\{x \in \mathbb{R}\}' = \emptyset$, but $\mathbb{R}' = \mathbb{R}$.

Now, we give two more definitions for convenience of deriving open and closed sets from any arbitrary set.

Definition 2.6 (Closure)

Given a set S , let the set of all limit points of S be denoted S' . The **closure** of S is the set $\bar{S} = S \cup S'$. It is the smallest closed set that contains S .

Definition 2.7 (Interior)

Given a set S , the **interior** of S is denoted S° , the set of all interior points of S . It is the largest open set that is within S .

Theorem 2.8 ()

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

Proof.

Assume that y is not a limit point of E . Then, there exists some $\epsilon > 0$ s.t. $(y - \epsilon, y + \epsilon)$ does not intersect with E . This means that $y - \epsilon$ is an upper bound of E , and so y is not the supremum.

Theorem 2.9 ()

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

1. \overline{E} is closed.
2. $E = \overline{E}$ if and only if E is closed.
3. $\overline{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$. That is, if $E \subset F$ closed, then “increasing” the size of E to its closure will not make it greater than F .

Proof.

Listed.

1. Let x be a limit point of \overline{E} . Then, for every $\epsilon > 0$, we have $B_\epsilon(x) \cap \overline{E} \neq \emptyset$, which means that either $B_\epsilon(x) \cap E \neq \emptyset$ (in which case $x \in E' \implies x \in \overline{E}$ and we are done) or $B_\epsilon(x) \cap E' \neq \emptyset$. We wish to prove that in the latter case, x being a limit point of E' still implies that x is a limit point of E . Since $B_\epsilon(x) \cap E' \neq \emptyset$, there must exist a $y \in B_\epsilon(x) \cap E'$. Since $y \in E'$, we can construct an open ball $B_\delta(y)$ containing elements of E , and since $B_\epsilon(x)$ is open, we can contain $B_\delta(y)$ entirely within $B_\epsilon(x)$. Therefore,

$$B_\delta(y) \cap E \neq \emptyset \implies B_\epsilon(x) \cap E \neq \emptyset$$

therefore, $x \in E' \implies x \in \overline{E}$.

2. If E is closed, then $E' \subset E \implies \overline{E} = E \cup E' = E$. If $E = \overline{E} = E \cup E'$, then $E' \subset E \implies E$ is closed.
3. Since $E \subset F$, it suffices to prove that $E' \subset F$. Consider a limit point x of E . Then every punctured open neighborhood of x satisfies $B_\epsilon^\circ(x) \cap E \neq \emptyset$. But since $E \subset F$, we have

$$B_\epsilon^\circ(x) \cap F \neq \emptyset$$

and so x is also a limit point of F . But since F is closed, $x \in F$. Therefore, $\overline{E} = E \cup E' \subset F$.

The first two statements (1) and (2) imply the following.

Corollary 2.3 ()

The closure of the closure of E is equal to the closure of E .

Proof.

We know that $\overline{\overline{E}} \supset \overline{E}$, so we must prove that $\overline{\overline{E}} \subset \overline{E}$, which is equivalent to proving that $\overline{E}' \subset \overline{E}$. Let $x \in \overline{E}'$, i.e. is a limit point of \overline{E} . Then, for every $\epsilon > 0$, we have $B_\epsilon(x) \cap \overline{E} \neq \emptyset$. Pick a point y from this intersection, and since $B_\epsilon(x)$ is open, we can construct an open ball $B_\delta(y)$ fully contained

in $B_\epsilon(x)$. Since $y \in \overline{E}$, y is a limit point of E , which implies

$$B_\delta(y) \cap E \neq \emptyset \implies B_\epsilon(x) \cap E \neq \emptyset \quad (231)$$

and therefore x is a limit point of E , $x \in \overline{E}$.

2.3 Compactness

Definition 2.8 (Open Cover)

An **open cover** of a set E in a metric space X is a collection $\{G_\alpha\}$ of open subsets of X such that $E \subset \cup_\alpha G_\alpha$.

Definition 2.9 (Compact Set)

A subset S of a metric space X is said to be **compact** if every open cover of S contains a finite subcover.

While openness behaves differently depending on its embedding space, compactness stays constant. Therefore, we don't have to worry about talking about which space a compact set is embedded in.

Theorem 2.10 (Compactness is Preserved Under Subspace Topology)

Suppose $K \subseteq Y \subseteq X$. Then K is compact relative to X if and only if K is compact relative to Y .

Proof.

We can prove bidirectionally.

1. Suppose that K is compact in X . Then given any open cover $\{U_\alpha\}_\alpha$ of K , there exists a finite subcover $\{U_i\}_i$. Now let there exist an open cover $\{V_\alpha\}$ in Y , but every $V_\alpha = U_\alpha \cap Y$ for some U_α open in X . Therefore, we can take the finite subcover $\{V_i = V_i \cap Y\}_i$.
2. Suppose that K is compact in Y . Then given any open cover $\{V_\alpha\}$ of K , there exists a finite subcover $\{V_i\}_i$. Now let there exist an open cover $\{U_\alpha\}$ in X . Then we set $\{V_\alpha = U_\alpha \cap Y\}_\alpha$, which has a finite subcover $\{V_i = U_i \cap Y\}$, and therefore we can take $\{U_i\}$ as our finite subcover in X .

Theorem 2.11 ()

A finite union of compact sets is compact.

Proof.

It suffices to prove for two sets A, B by induction. Take an arbitrary cover \mathcal{L} of $A \cup B$. Then \mathcal{L} is a cover of A , so it has a finite subcover $\mathcal{F} \subset \mathcal{L}$. It is also a cover of B , so it has a finite subcover $\mathcal{G} \subset \mathcal{L}$. Therefore, $\mathcal{F} \cup \mathcal{G} \subset \mathcal{L}$ is a cover of $A \cup B$, and since it is the union of finite covers, it is finite.

As we will see in the following theorems, compact sets behave well with closed sets. In fact, compactness is in a form a stronger notion than closedness.

Theorem 2.12 ()

Compact subsets of metric spaces are closed.

Proof.

We would like to show that if A is compact in X , then A^c is open. What we would like to do is if we have some $x \in A^c$, then we must prove that there exists some open set $B_\epsilon(x)$ that is disjoint with A . For every point $a \in A$, we can construct an open balls $V_a = B_{d(x,a)/2}(a)$ and $U_a = B_{d(x,a)/2}(x)$. We know that if $y \in B_{d(x,a)/2}(a)$, then assuming $y \in B_{d(x,a)/2}(x)$ will give

$$d(x, a) \leq d(x, y) + d(y, a) < \frac{d(x, a)}{2} + \frac{d(x, a)}{2} = d(x, a) \quad (232)$$

which is absurd. Since $\{V_a\}_{a \in A}$ forms an open covering of A , then by compactness we can take a finite subcover V_{a_1}, \dots, V_{a_n} , along with the respective neighborhoods of x U_{a_1}, \dots, U_{a_n} . Since we have established

$$V_{a_i} \cap U_{a_i} = \emptyset \implies \bigcap_{i=1}^n V_{a_i} \cap \left(\bigcup_{i=1}^n U_{a_i} \right) = \emptyset \quad (233)$$

and since $\bigcap_{i=1}^n V_{a_i}$ is open (as it is the intersection of open sets) and disjoint from an open cover of A and hence from A , we have proved that A^c is open, and so A is closed.

Theorem 2.13 ()

Closed subsets of compact sets are compact.

Proof.

Let $C \subset K \subset X$ with K compact and C closed. Then let $\{U_\alpha\}$ be an open cover of C . Then C^c is open in X , and so $\{U_\alpha\} \cup \{C^c\}$ is an open cover of K , so it has a finite subcover S .

1. If $C^c \notin S$, then we have a finite subcover of C .
2. If $C^c \in S$, then we can take the element out to get a finite subcover of C .

Therefore we have constructed a way to make a finite subcover. C is compact.

Corollary 2.4 ()

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

The general notion of compactness⁶ for topological spaces is not needed for analysis. Rather, we make use of the following theorem which allows us to focus on the compactness of subsets in Euclidean spaces \mathbb{R}^n .

Theorem 2.14 (Heine-Borel)

Let $E \subset \mathbb{R}^k$. The following are equivalent.

1. E is closed and bounded
2. E is compact.

⁶According to Terry Tao, a compact set is "small," in the sense that it is easy to deal with. While this may sound counterintuitive at first, since $[0, 1]$ is considered compact while $(0, 1)$, a subset of $[0, 1]$, is considered noncompact. More generally, a set that is compact may be large in area and complicated, but the fact that it is compact means we can interact with it in a finite way using open sets, the building blocks of topology. That finite collection of open sets makes it possible to account for all the points in a set in a finite way. This is easily noticed, since functions defined over compact sets have more controlled behavior than those defined over noncompact sets. Similarly, classifying noncompact spaces are more difficult and less satisfying.

3. Every infinite subset of E has a limit point in E .

Example 2.2 ()

An open set in \mathbb{R}^2 is not compact. Take the open rectangle $R = (0, 1)^2 \subset \mathbb{R}^2$. There exists an infinite cover of R

$$R = \bigcup_{n=0}^{\infty} (0, 1) \times \left(0, \frac{2^{n+1} - 1}{2^{n+1}}\right)$$

that does not have a finite subcover.

Clearly, the limit point of an open set is its boundary points. Note that a sequence of points can also have a limit point.

Theorem 2.15 (Bolzano-Weierstrass Theorem)

Every bounded sequence in \mathbb{R}^n has a limit point.

Proof.

The fact that the infinite sequence is bounded means that there exists some closed subset $I \in \mathbb{R}^n$ that contains all point of the sequence. By definition I is compact, so by the Heine-Borel theorem, every cover of I has a finite subcover.

Now, assume that there exists an infinite sequence in I that is not convergent, i.e. has no limit point. Then, each point $x_i \in I$ would have a neighborhood $U(x_i)$ containing at most a finite number of points in the sequence. We can define I such that the union of the neighborhoods is a cover of I . That is,

$$I \subset \bigcup_{i=1}^{\infty} U(x_i)$$

However, since every $U(x_i)$ contains at most a finite number of points, we must have an infinite open neighborhoods to cover $I \implies$ we cannot have a finite subcover. This contradicts the fact that I is compact.

In fact, compactness actually implies completeness.

Theorem 2.16 ()

Compact metric spaces are complete.

So far, we've been pretty abstract about compact sets. In general, it's pretty easy to prove that a set is not compact. We just need to find one example of an open cover that does not have a finite subcover. To prove that set *is* compact, we must show that for *every* open cover, we can get a finite subcover. This sounds quite daunting, but here is a special theorem that can start us off, and the theorems above allow us to construct more compact sets. We will need the third interpretation of completeness of the reals: nested intervals completeness.

Theorem 2.17 (Nested Intervals Theorem)

If $\{I_n = [a_n, b_n]\}$ in \mathbb{R} is a sequence of nested closed intervals, then

$$\bigcap_{i=1}^{\infty} I_n \neq \emptyset \tag{234}$$

Proof.

Note that $\{a_n\}$ is bounded above by b_1 . Therefore by LUB property it must have a supremum, call it $x = \sup_n \{a_n\}$. Then, we see that $a_n \leq x \leq b_n$ for all n , and so x is in the intersection.

Corollary 2.5 ()

Every closed interval is compact.

Proof.

Let $I = [a, b]$. Then if $x, y \in I$, $|x - y| \leq b - a = \delta$. Now by contradiction, suppose that there exists an open cover $\{U_\alpha\}$ of I which contains no finite subcover of I . Then letting $c = (a + b)/2$, at least one of the two intervals $[a, c], [c, b]$ cannot have a finite subcovering (otherwise their finite union can be covered). WLOG let it be $[a, c]$. We keep subdividing and get the sequence of nested intervals.

$$I \supset I_1 \supset I_2 \supset \dots \quad (235)$$

We know that I_n is not covered by any finite subcollection of $\{U_\alpha\}$ and if $x, y \in I_n$, then $|x - y| < 2^{-n}\delta$. From the nested intervals theorem, there exists a point z lying in every I_n . There must then be an open neighborhood U_z in the open cover, and by definition of openness there exists a $\epsilon > 0$ s.t. $z \in B_\epsilon(z) \subset U_z$. By the Archimedean property, we can set n so large that $2^{-n}\delta < \epsilon$ and this means that $B_\epsilon(z) \supset I_n$, which contradicts the fact that I_n is not covered by a finite subcollection. Therefore I is compact.

2.4 Connectedness

Definition 2.10 (Separate, Connected Sets)

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

Example 2.3 ()

It is clear that separate sets imply disjointness. However, this is not true for the other way around.

1. $(0, 1)$ and $[1, 2)$ are disjoint but not separate.
2. The rationals and irrationals are disjoint, but not separate.

Theorem 2.18 ()

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

Proof.

2.5 Separability

2.6 Perfect Sets

Definition 2.11 (Perfect Sets)

A set P is perfect if it is closed and all of its points are limit points of P . In other words, the limit points of P and P itself coincide.

$$P' = P \quad (236)$$

Theorem 2.19 ()

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

2.7 Exercises

Exercise 2.1 (Math 531 Spring 2025, PS5.1)

We know what it means for a metric space (X, d) to be compact. We say that it is sequentially compact if every sequence in (X, d) has a convergent subsequence. Prove that a metric space is compact if and only if it is sequentially compact.

Solution 2.1

We prove bidirectionally.

1. (\rightarrow). Assume that X is compact and let (x_n) be a sequence in X . For any $\epsilon > 0$, let

$$\mathcal{C}_\epsilon := \{B_\epsilon(x) \mid x \in X\} \quad (237)$$

be an open cover of open balls. Then there exists a finite subcover $\mathcal{F}_\epsilon \subset \mathcal{C}_\epsilon$. There is a countable sequence (x_n) , with each point in at least one open set in \mathcal{C} . By the pigeonhole principle, at least one open set must be hit infinitely many times, call this $B(x^*, \epsilon)$. Now consider for $\epsilon = \frac{1}{n}$ for $n \in \mathbb{N}$.

2. (\leftarrow).

Exercise 2.2 (Math 531 Spring 2025, PS5.2)

Give an example of a sequence of real numbers x_n for which

$$|x_{n+1} - x_n| \rightarrow 0 \quad (238)$$

as $n \rightarrow \infty$, but x_n is not convergent.

Solution 2.2

Consider the sequence

$$x_n = \sum_{i=1}^n \frac{1}{i} \quad (239)$$

It is the case that $x_{n+1} - x_n = 1/(n+1)$ which tends to 0, but this is a harmonic series which is not convergent.

Exercise 2.3 (Math 531 Spring 2025, PS5.3)

Let $\{x_n\}_{n=0}^{\infty}$ be a sequence of real numbers. Assume that

$$|x_{n+1} - x_n| \leq c|x_n - x_{n-1}| \quad (240)$$

for all $n \geq 1$, for some fixed $c < 1$. Prove that x_n is convergent. Hint: you may want to use the formula that you proved in a previous homework for $1 + c + c^2 + \dots + c^N$.

Solution 2.3**Exercise 2.4 (Math 531 Spring 2025, PS5.4)**

Let x_n be a sequence of rational numbers defined recursively by:

$$x_0 = 1 \quad (241)$$

$$x_{n+1} = \frac{1}{x_n + 2} \quad (242)$$

when $n \geq 0$. The first few terms of this sequence are $1, \frac{1}{3}, \frac{3}{7}, \dots$. Prove that the sequence is convergent and find its limit.

Solution 2.4**Exercise 2.5 (Math 531 Spring 2025, PS5.5)**

As we know, every bounded sequence of real numbers has a convergence subsequence.

1. Let's say we have two sequences a_n and b_n that are bounded. Find a single sequence of indices $\{n_k\}$ so that both a_{n_k} and b_{n_k} are convergent. This is called a common convergent subsequence for a_n and b_n .
2. Show that for any finite number of bounded sequences of real numbers, we can find a common convergent subsequence.
3. Now suppose have a sequence of bounded sequences. Find a common convergent subsequence. What does this remind you of?

Solution 2.5**Exercise 2.6 (Math 531 Spring 2025, PS5.6)**

Given a sequence of real numbers x_n , we can define the sequence of its means by:

$$x_1, \frac{x_1 + x_2}{2}, \frac{x_1 + x_2 + x_3}{3}, \dots \quad (243)$$

Call the sequence of means y_n . Prove that if $x_n \rightarrow x$, then $y_n \rightarrow x$. Discuss the examples $x_n = (-1)^n$, $x_n = n$, $x_n = (-1)^n n$, and $x_n = (-1)^n \sqrt{n}$. The fourth example shows that it is possible to average out chaotic behavior (so long as it isn't focused in one direction). The third example shows that this is impossible if the system becomes too chaotic.

Solution 2.6**Exercise 2.7 (Math 531 Spring 2025, PS4.1)**

Let (X, d) be a metric space. Assume that K is compact and F is closed in (X, d) . Assume $K \cap F = \emptyset$. Prove that

$$\inf_{x \in F, y \in K} d(x, y) > 0. \quad (244)$$

Show by an example that this number could be zero if K is only assumed to be closed (rather than compact).

Solution 2.7

$K \cap F = \emptyset \implies K \subset F^c$ with F^c open. This means that for every $x \in K \subset F^c$, there exists a $r_x > 0$ s.t. $B(x, r_x) \subset F^c \iff B(x, r_x) \cap F = \emptyset$.

Now we take the covering $\{B(x, \frac{r_x}{2}) \mid x \in K\}$, and since K is compact there must be a finite subcovering, which we denote $C = \{B(x_i, \frac{r_i}{2}) \mid i = 1, \dots, n\}$. Denote $r^* = \min\{r_i\}$, which is positive since we take the minimum of a finite number of positive elements.

Now for any $x \in K$ and $y \in F$, x must be in some $B(x_i, \frac{r_i}{2}) \iff d(x, x_i) < \frac{r_i}{2} \iff -d(x, x_i) > -\frac{r_i}{2}$. With the same i , since $B(x_i, r_i)$ is disjoint from F , we have $d(x_i, y) \geq r_i$. Therefore, by the triangle inequality,

$$d(x, y) \geq d(x_i, y) - d(x_i, x) = r_i - \frac{r_i}{2} = \frac{r_i}{2} \geq \frac{r^*}{2} \quad (245)$$

and thus we have found a nontrivial lower bound.

Exercise 2.8 (Math 531 Spring 2025, PS4.2)

Consider \mathbb{R} with the usual metric. Find an open cover of \mathbb{Q} that does not cover \mathbb{R} .

Solution 2.8

$$\mathbb{Q} = (-\infty, \sqrt{2}) \cup (\sqrt{2}, +\infty).$$

Exercise 2.9 (Math 531 Spring 2025, PS4.3)

\mathbb{R} is not compact with the usual metric since it is not bounded. Let us, however, define the following metric on \mathbb{R} :

$$d_*(x, y) = \frac{|x - y|}{(1 + |x|)(1 + |y|)}. \quad (246)$$

Verify that (\mathbb{R}, d_*) is a metric space. Prove that all subsets of (\mathbb{R}, d_*) are bounded. Show that \mathbb{R} still isn't compact with this metric. What is the problem?

Solution 2.9

We first verify metric.

- Since $|x - y| \geq 0, 1 + |x| \geq 1, 1 + |y| \geq 1, d_*(x, y) \geq 0$. We also see that

$$d_*(x, y) = 0 \iff |x - y| = 0 \iff x = y \quad (247)$$

2. It is symmetric since

$$d_*(x, y) = \frac{|x - y|}{(1 + |x|)(1 + |y|)} = \frac{|y - x|}{(1 + |y|)(1 + |x|)} = d_*(y, x) \quad (248)$$

3. It satisfies triangle inequality since

$$d_*(x, y) + d_*(y, z) = \frac{|x - y|}{(1 + |x|)(1 + |y|)} + \frac{|y - z|}{(1 + |y|)(1 + |z|)} \quad (249)$$

$$= \frac{|x - y|(1 + |z|) + |y - z|(1 + |x|)}{(1 + |x|)(1 + |y|)(1 + |z|)} \quad (250)$$

$$= \frac{|x - y| + |x - y| \cdot |z| + |y - z| + |y - z| \cdot |x|}{(1 + |x|)(1 + |y|)(1 + |z|)} \quad (251)$$

$$\geq \frac{|x - z| + (|x| - |y|) \cdot |z| + (|y| - |z|) \cdot |x|}{(1 + |x|)(1 + |y|)(1 + |z|)} \quad (252)$$

$$= \frac{|x - z| + |x||y| - |y||z|}{(1 + |x|)(1 + |y|)(1 + |z|)} \quad (253)$$

$$= \frac{(1 + |y|)|x - z|}{(1 + |x|)(1 + |y|)(1 + |z|)} \quad (254)$$

$$= \frac{|x - z|}{(1 + |x|)(1 + |z|)} \quad (255)$$

$$= d_*(x, z) \quad (256)$$

where the inequality comes from $|x - y| + |y - z| \geq |x - z|$, $|x - y| \geq ||x| - |y|| \geq |x| - |y|$, and $|y - z| \geq ||y| - |z|| \geq |y| - |z|$.

This is bounded since for any $x, y \in \mathbb{R}$, we can show that the numerator is bounded by the denominator. Since both are positive from (1), it suffices to prove $|x - y|^2 \leq (1 + |x|)^2(1 + |y|)^2$.

$$(1 + |x|)^2(1 + |y|)^2 = (1 + 2|x| + |x|^2)(1 + 2|y| + |y|^2) \quad (257)$$

$$= 1 + 2|x| + 2|y| + 4|x||y| + |x|^2 + |y|^2 + 2|x||y|^2 + 2|y||x|^2 + |x|^2|y|^2 \quad (258)$$

$$\geq |x|^2 + |y|^2 + 2|x||y| \quad (259)$$

$$\geq |x|^2 + |y|^2 - 2xy \quad (260)$$

$$= |x - y|^2 \quad (261)$$

where the first inequality holds since all the terms in the expansion are nonnegative and the second holds since $|x||y| \geq xy$. Therefore, $d_*(x, y) \leq 1$. \mathbb{R} is still not compact since we can construct the set of open balls $B_r(0)$ around 0 w.r.t. d_* . Consider the cover

$$\mathcal{C} = \{B_{1-\frac{1}{n}}(0)\}_{n \in \mathbb{N}, n \geq 2} \quad (262)$$

Now assume that there is a finite subcover. Then there must be a maximum index $N \in \mathbb{N}$ in this cover. I claim that this does not cover \mathbb{R} . Consider the element $y = N - 1 \in \mathbb{R}$. The distance is

$$d_*(x, y) = \frac{|0 - (N - 1)|}{(1 + |0|)(1 + |N - 1|)} = \frac{N - 1}{N} = 1 - \frac{1}{N} \quad (263)$$

and so $y \notin \mathcal{C}$. Hence \mathcal{C} is not a cover of \mathbb{R} .

Exercise 2.10 (Math 531 Spring 2025, PS4.4)

Consider the set $X = \mathbb{R} \cup \{Gandalf\}$. Define a metric d_* on X by:

$$d_*(x, y) = \frac{|x - y|}{(1 + |x|)(1 + |y|)} \quad (264)$$

for $x, y \in \mathbb{R}$, while

$$d_*(Gandalf, x) = \frac{1}{1 + |x|}, \quad (265)$$

for all $x \in \mathbb{R}$. Verify that d_* is a metric on X (you don't have to do much for this, since you already did part of it in the previous problem). Prove that (X, d_*) is a compact metric space.

Solution 2.10**Exercise 2.11 (Math 531 Spring 2025, PS4.5)**

Let X be any set and endow it with the metric $d(x, y) = 1$ if $x \neq y$ and $d(x, x) = 0$. Check that this is a metric. Find all compact sets in (X, d) .

Solution 2.11

It trivially satisfies nonnegativity since it's either 0 or 1, and $d(x, x) = 0$. It is symmetric as well. As for triangle inequality this is trivial. All compact sets are finite sets.

Exercise 2.12 (Math 531 Spring 2025, PS3.1)

Determine for each of the following sets, whether or not it is countable. Justify your answers

1. The set of all functions $f : \{0, 1\} \rightarrow \mathbb{N}$.
2. The set B_n of all functions $f : \{1, \dots, n\} \rightarrow \mathbb{N}$
3. The set $C = \bigcup_{n \in \mathbb{N}} B_n$
4. The set of all functions $f : \mathbb{N} \rightarrow \{0, 1\}$.
5. The set of all functions $f : \mathbb{N} \rightarrow \{0, 1\}$ that are “eventually zero” (We say that f is eventually zero if there exists some $N \geq 1$ so that $f(n) = 0$ for all $n \geq N$.)
6. G the set of all functions $f : \mathbb{N} \rightarrow \mathbb{N}$ that are eventually constant.

Solution 2.12

Listed.

1. Countable since bijective to $\mathbb{N} \times \mathbb{N}$. We define the bijection as

$$(a_0, a_1) \in \mathbb{N} \times \mathbb{N} \mapsto f(i) = a_i \quad (266)$$

2. Countable since bijective to \mathbb{N}^n . We define the bijection as

$$(a_1, \dots, a_n) \in \mathbb{N}^n \mapsto f(i) = a_i \quad (267)$$

3. Countable since we proved that B_n is countable, and a countable union of countable sets are countable from Rudin Theorem 2.12.

4. Uncountable since we can create a bijection from the set of all sequences (a_i) of 0 or 1, which from Rudin Theorem 2.14 is uncountable.

$$(a_i)_{i \in \mathbb{N}} \mapsto f(i) = a_i \quad (268)$$

5. Countable. Call this set B , and call the set of functions f that have their final 1 at index k to be A_k . Then,

$$B = \bigcup_{k=1}^{\infty} A_k \quad (269)$$

where $A_0 = A_1 = \{1\}$, and $|A_k| = 2^{k-1}$ for $k \geq 2$. Since B is the countable union of at most countable sets, B must be countable.

6. Countable. Call this set B . Let A_k be the set of functions that are eventually constant to value k . Let A_{ki} be the set of functions that are always k starting from index i (where i is the smallest element). Since everything is determined to be k at i and beyond, A_{ki} can be divided up into the first $i - 1$ elements of any natural number, followed by a sequence of k 's. Therefore $|A_{ki}| \approx \aleph^{i-1}$, where \approx means equipotent, and so A_{ki} is countable. Therefore since countable unions of countable sets are countable,

$$A_k = \bigcup_{i=1}^{\infty} A_{ki} \text{ is countable} \implies B = \sum_{k=1}^{\infty} A_k \text{ is countable} \quad (270)$$

Exercise 2.13 (Math 531 Spring 2025, PS3.2)

Tell if the following subsets $A \subset \mathbb{R}$ (with the usual metric $d(x, y) = |x - y|$) are open or closed. Also, find (i) the limit points of A , (ii) the interior of A , (iii) \bar{A} .

1. $A = \mathbb{Q}$
2. $A = (0, 1]$
3. $A = \{1, \frac{1}{2}, \frac{1}{4}, \dots\}$
4. $A = \{0, 1, \frac{1}{2}, \frac{1}{4}, \dots\}$
5. $A = \mathbb{Z}$

Solution 2.13

Listed. We denote A' as the limit points of A and the interior as A° .

1. Not open nor closed. $A' = \mathbb{R}$, $A^\circ = \emptyset$. $\bar{A} = \mathbb{R}$.
2. Not open nor closed. $A' = [0, 1]$, $A^\circ = (0, 1)$. $\bar{A} = [0, 1]$.
3. Not open nor closed. $A' = \{0\}$, $A^\circ = \emptyset$. $\bar{A} = \{0\} \cup A$.
4. Closed. $A' = \{0\}$, $A^\circ = \emptyset$. $\bar{A} = A$.
5. Closed. $A' = \emptyset$. $A^\circ = \emptyset$. $\bar{A} = A$.

Exercise 2.14 (Math 531 Spring 2025, PS3.3)

Prove the following statements about subsets A, B of a general metric space (X, d) .

- $\overline{A \cup B} = \bar{A} \cup \bar{B}$.
- Show by example that $\overline{A \cap B} \neq \bar{A} \cap \bar{B}$.

Solution 2.14

For the first part, we show bidirectionally.

1. $\overline{A \cup B} \subset \bar{A} \cup \bar{B}$. Let $x \in \overline{A \cup B}$. If $x \in A \cup B$, then it must be the case that either $x \in A \subset (A \cup A') = \bar{A}$ or $x \in B \subset (B \cup B') = \bar{B}$, which means $x \in \bar{A} \cup \bar{B}$. Now assume not. Then $x \in (A \cup B)'$. Therefore, for any $r > 0$, we know that $B(x, r) \cap (A \cup B) \neq \emptyset$. Now let us take a sequence $(r_n = \frac{1}{n})_{n \in \mathbb{N}}$, and for each r_n we have some element $x_n \in (A \cup B)$. Given that we have a countably infinite sequence of x_n , each which may be in A or B , by the pigeonhole principle either A or B must be hit infinitely many times. If $x_n \in A$ infinitely many times, then $x \in \bar{A}$, and analogous for B .
2. $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$. WLOG let $x \in \bar{A}$. If $x \in A$, then $x \in (A \cup B) \subset \overline{A \cup B}$. If $x \notin A$, then $x \in A'$.

Therefore for every $r > 0$, $B(x, r) \cap A \neq \emptyset$. But this means

$$\emptyset \neq (B(x, r) \cap A) \cup (B(x, r) \cap B) = B(x, r) \cap (A \cup B) \implies x \in (A \cup B)' \subset \overline{A \cup B} \quad (271)$$

For a counterexample, consider the sequences

$$A = (x_n) = \frac{1}{n} \quad B = (y_n) = -\frac{1}{n} \quad (272)$$

for $n \in \mathbb{N}$. $\overline{A} = A \cup \{0\}$, $\overline{B} = B \cup \{0\}$, and so $\overline{A} \cap \overline{B} = \{0\}$. However, $A \cap B = \emptyset \implies \overline{A} \cap \overline{B} = \emptyset$.

Exercise 2.15 (Math 531 Spring 2025, PS3.4)

Consider the set of rationals in canonical form (such that numerator and denominator are relatively prime) with potential distance:

$$d_1\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) = |q_1 - q_2|. \quad (273)$$

Is this a metric? Prove that the following defines a metric

$$d_2\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) = |p_1 - p_2| + |q_1 - q_2|. \quad (274)$$

Solution 2.15

This is not a metric since

$$d_1\left(\frac{2}{1}, \frac{3}{1}\right) = 1 - 1 = 0 \quad (275)$$

when $2/1 \neq 3/1$. For d_2 , we show that it satisfies the three properties.

1. *Nonnegativity.* Since it is the sum of 2 absolute values which are norms and therefore nonnegative, it must be nonnegative by ordered field properties. We see that

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} \iff p_1 = p_2 \text{ and } q_1 = q_2 \quad (276)$$

$$\iff |p_1 - p_2| = |q_1 - q_2| = 0 \quad (277)$$

$$\iff |p_1 - p_2| + |q_1 - q_2| = 0 \quad (278)$$

2. For symmetry, note that

$$d_2\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) = |p_1 - p_2| + |q_1 - q_2| = |p_2 - p_1| + |q_2 - q_1| = d_2\left(\frac{p_2}{q_2}, \frac{p_1}{q_1}\right) \quad (279)$$

3. For triangle inequality, we see that for any $p_1/q_1, p_2/q_2, p_3/q_3$,

$$d_2\left(\frac{p_1}{q_1}, \frac{p_3}{q_3}\right) = |p_1 - p_3| + |q_1 - q_3| \quad (280)$$

$$= |(p_1 - p_2) + (p_2 - p_3)| + |(q_1 - q_2) + (q_2 - q_3)| \quad (281)$$

$$\leq |p_1 - p_2| + |p_2 - p_3| + |q_1 - q_2| + |q_2 - q_3| \quad (\text{subadditivity of norm})$$

$$= d_2\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) + d_2\left(\frac{p_2}{q_2}, \frac{p_3}{q_3}\right) \quad (282)$$

Exercise 2.16 (Math 531 Spring 2025, PS3.5)

Let $M = \{x_1, \dots, x_3\}$ be a set with three points. Describe the set of all metrics on M . What if M has four points?

Solution 2.16

If M has 3 points call them x_1, x_2, x_3 , then the metric is completely defined by the three values

$$d(x_1, x_2) = d(x_2, x_1) \quad (283)$$

$$d(x_2, x_3) = d(x_3, x_2) \quad (284)$$

$$d(x_3, x_1) = d(x_1, x_3) \quad (285)$$

where $d(x, x) = 0$. We must make sure that the triangle inequality satisfies for these 3 numbers. Therefore we can think of this as the set of all triangles in \mathbb{R}^2 (that are equivalent under translation and rotation, but not permutation of points).

Similarly for 4 points, we can visualize the metrics as the set of all tetrahedra in \mathbb{R}^3 (since each face is a triangle, and therefore for any three points the triangle inequality is guaranteed to be satisfied), equivalent under translation and rotation, but not permutation of the 4 points.

Exercise 2.17 (Math 531 Spring 2025, PS3.6)

Let P be a polynomial of degree $n \geq 1$. Prove that if $P(0) = 0$, then $P(x) = xQ(x)$, for some polynomial Q of degree $n - 1$. Deduce that if $P(a) = 0$, then we can write $P(x) = (x - a)Q(x)$ for some Q of degree $n - 1$.

Solution 2.17

A n th degree polynomial will have the form

$$p(x) = \sum_{i=0}^n c_i x^i \quad (286)$$

Since $p(0) = c_0 = 0 \implies c_0 = 0$. This means that

$$p(x) = \sum_{i=1}^n c_i x^i = x \sum_{i=0}^{n-1} c_{i+1} x^i \text{ where } Q(x) = \sum_{i=0}^{n-1} c_{i+1} x^i \quad (287)$$

If $p(a) = 0$, we can construct $f(x) = p(x + a)$, where f is a polynomial since the expansion does not increase its degree. Since $f(x) = p(a) = 0$, by above f can be factorized $f(x) = xg(x)$ for some $(n-1)$ th degree polynomial g , and by substitution this means that $p(x) = f(x-a) = (x-a)g(x-a)$.

Exercise 2.18 (Math 531 Spring 2025, PS3.7)

Consider all polynomials $P : \mathbb{R} \rightarrow \mathbb{R}$ of degree less than or equal to n . Call this set \mathcal{P}_n . Let's define potential distances on \mathcal{P}_n .

$$d_1(p, q) = |p(0) - q(0)|. \quad (288)$$

Show this defines a distance on \mathcal{P}_0 but not on \mathcal{P}_n for $n \geq 1$. Now consider

$$d_N(p, q) = \sum_{j=0}^N |p(j) - q(j)| \quad (289)$$

Show that this defines a distance on \mathcal{P}_n , for every $n \leq N$. What does the solution say about polynomials of degree N ?

Solution 2.18

If $n = 0$, \mathcal{P}_n is a set of constant functions P , where each constant function P is determined completely by its value at any point, e.g. 0. We check the properties.

1. $d_1(p, q) \geq 0$ since we take the norm at the end. We can see that

$$d_1(p, q) = 0 \iff |p(0) - q(0)| \quad (290)$$

$$\iff p(0) = q(0) \quad (291)$$

$$\iff p = q \quad (292)$$

2. It is clearly symmetric.

$$d_1(p, q) = |p(0) - q(0)| = |q(0) - p(0)| = d_1(q, p) \quad (293)$$

3. It satisfies the triangle inequality by subadditivity of the norm.

$$d_1(p, r) = |p(0) - r(0)| \quad (294)$$

$$= |(p(0) - q(0)) + (q(0) - r(0))| \quad (295)$$

$$\leq |p(0) - q(0)| + |q(0) - r(0)| \quad (296)$$

$$= d_1(p, q) + d_1(q, r) \quad (297)$$

It doesn't satisfy for \mathcal{P}_n because consider $p(x) = x$ and $q(x) = x^2$. They are not the same function but $d_1(p, q) = |p(0) - q(0)| = 0$. For d_N defined on \mathcal{P}_n for $n \leq N$, we verify the properties.

1. This is the sum of norms, so it must be nonnegative. Now we see that if $p = q$, then $p(x) = q(x) \implies |p(x) - q(x)| = 0 \implies d_N(p, q) = 0$. For the other way around, suppose $d_N(p, q) = 0$. Then from problem 3.8, we are solving the linear equation $0 = Vb - Vc$, where b, c are the vectors representing the coefficients of p, q , and V is the Vandermonde matrix with $a_i = i$. By linearity, this is equivalent to solving $0 = V(b - c)$, and since we showed that V is invertible (since a_i 's are distinct), V has a trivial kernel and therefore $b - c = 0 \iff b = c \implies p = q$.
2. Symmetricity is trivial.

$$d_N(p, q) = \sum_{j=0}^N |p(j) - q(j)| = \sum_{j=0}^N |q(j) - p(j)| = d_N(q, p) \quad (298)$$

3. For triangle inequality,

$$d_N(p, r) = \sum_{j=0}^N |p(j) - r(j)| \quad (299)$$

$$= \sum_{j=0}^N |(p(j) - q(j)) + (q(j) - r(j))| \quad (300)$$

$$\leq \sum_{j=0}^N |p(j) - q(j)| + |q(j) - r(j)| \quad (301)$$

$$= \sum_{j=0}^N |p(j) - q(j)| + \sum_{j=0}^N |q(j) - r(j)| \quad (302)$$

$$= d_N(p, q) + d_N(q, r) \quad (303)$$

This shows that we need to "sample" more points from higher-degree polynomials to get the metric as they are higher-dimensional.

Exercise 2.19 (Math 531 Spring 2025, PS3.8)

Given distinct numbers a_0, \dots, a_N and numbers b_0, \dots, b_N , prove that there exists a polynomial P of degree N with the property that

$$P(a_i) = b_i, \quad (304)$$

for $0 \leq i \leq N$. The most direct way to solve this problem, in my view, is to write the system equations you are trying to solve as a linear system for the coefficients of P . This will give you some matrix M that depends on the numbers a_0, \dots, a_N . The key is to show that the determinant of this matrix is non-zero. It turns out that the determinant of this matrix is equal to

$$\prod_{0 \leq i < j \leq N} (a_i - a_j), \quad (305)$$

up to a potential $-$ sign depending on how you defined M . Prove this and deduce the result.

Solution 2.19

We can write the system of equations using the Vandermonde matrix $V \in \mathbb{R}^{(N+1) \times (N+1)}$ and c is the vector of coefficients of P .

$$b = Vc \iff \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_N \end{bmatrix} = \begin{bmatrix} 1 & a_0 & a_0^2 & \dots & a_0^N \\ 1 & a_1 & a_1^2 & \dots & a_1^N \\ \vdots & \vdots & \ddots & & \vdots \\ 1 & a_N & a_N^2 & \dots & a_N^N \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_N \end{bmatrix} \quad (306)$$

To calculate the determinant of V , we prove using induction. Clearly for $N = 1$ we have

$$\det \begin{pmatrix} 1 & a_0 \\ 1 & a_1 \end{pmatrix} = a_1 - a_0 \quad (307)$$

Now assume that this formula holds for some $N - 1 \in \mathbb{N}$. Then for N , we can take V and subtract a_0 times the i th column from the $(i + 1)$ st column. This gives us

$$V = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_1 - a_0 & a_1^2 - a_0 a_1 & \dots & a_1^N - a_0 a_1^{N-1} \\ 1 & a_2 - a_0 & a_2^2 - a_0 a_2 & \dots & a_2^N - a_0 a_2^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_N - a_0 & a_N^2 - a_0 a_N & \dots & a_N^N - a_0 a_N^{N-1} \end{bmatrix} \quad (308)$$

When calculating the determinant, we can perform the cofactor expansion by the first row, and then for each i th row factor out $(a_i - a_0)$ to get

$$\det V = \prod_{j=1}^N (a_j - a_0) \det \begin{bmatrix} 1 & a_1 & \dots & a_1^{N-1} \\ 1 & a_2 & \dots & a_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_N & \dots & a_N^{N-1} \end{bmatrix} \quad (309)$$

which is the $(N - 1) \times (N - 1)$ Vandermonde matrix. Therefore, we can apply our inductive hypothesis to get

$$\det V = \prod_{j=1}^N (a_j - a_0) \prod_{1 \leq i < j \leq N} (a_j - a_i) = \prod_{0 \leq i \leq j \leq N} (a_j - a_i) \quad (310)$$

Note that this has a 0 determinant iff $a_i = a_j$ for some $i \neq j$. Therefore since a_i 's are distinct, it must be nonzero. Therefore, this matrix is nonsingular, i.e. invertible, and we can solve the matrix equation to get

$$c = V^{-1}b \quad (311)$$

which from linear algebra is guaranteed to exist and is unique.

Exercise 2.20 (Rudin 2.1)

Prove that the empty set is a subset of every set.

Solution 2.20

It must suffice that if $x \in \emptyset$, then $x \in A$ for any arbitrary set A . This is vacuously true, since the initial condition is never met.

Exercise 2.21 ()

Show that the empty function $f : \emptyset \rightarrow X$, where X is an arbitrary set, is always injective. If $X = \emptyset$, then f is bijective.

Solution 2.21

Given distinct $x, y \in \emptyset$, $f(x) \neq f(y)$ is vacuously true, but if $X \neq \emptyset$, then there exists a $w \in X$ with no preimage. If $X = \emptyset$, then the statement for all $w \in X$, there exists an $x \in \emptyset$ s.t. $f(x) = w$ is vacuously true.

Exercise 2.22 (Rudin 2.2)

A complex number z is said to be algebraic if there are integers a_0, a_1, \dots, a_n , not all zero, such that

$$a_0 z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n = 0. \quad (312)$$

Prove that the set of all algebraic complex numbers is countable. Hint: For every positive integer N there are only finitely many equations with

$$n + |a_0| + |a_1| + \dots + |a_n| = N \quad (313)$$

Solution 2.22

Consider all polynomials s.t. $n + \sum_{i=0}^n |a_i| = N$. There is only a finite number of them, and each polynomial has at most n distinct complex roots. So this set is finite, an unioning over all $N \in \mathbb{N}$ gives an at most countable set of roots.

Exercise 2.23 (Rudin 2.3)

Prove there exists real numbers which are not algebraic.

Solution 2.23

From the previous exercise, if there were no real numbers which are not algebraic, then every real number is algebraic. This contradicts the fact that the set of all complex numbers is countable.

Exercise 2.24 (Rudin 2.4)

Is the set of all irrational real numbers countable?

Solution 2.24

No. Assume that it is countable. We have \mathbb{Q} countable. Then, by assumption, we must have $\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}^c$ be the union of countable sets, which must be countable, contradicting the fact that it is uncountable.

Exercise 2.25 (Rudin 2.5)

Construct a bounded set of real numbers which exactly 3 limit points.

Solution 2.25

We can construct the union of 3 sequences that converge onto the limit points 0, 1, 2.

$$\left\{\frac{1}{n}\right\}_{n \in \mathbb{N}} \cup \left\{\frac{1}{n} + 1\right\}_{n \in \mathbb{N}} \cup \left\{\frac{1}{n} + 2\right\}_{n \in \mathbb{N}} \quad (314)$$

Exercise 2.26 ()

Prove that the union of the limit points of sets is equal to the limit points of the union of the sets.

$$\bigcup_{k=1}^m A'_k = \left(\bigcup_{k=1}^m A_k \right)' \quad (315)$$

Exercise 2.27 (Rudin 2.6)

Let E' be the set of all limit points of a set E . Prove that E' is closed. Prove that E and \overline{E} have the same limit points. (Recall that $\overline{E} = E \cup E'$). Do E and E' always have the same limit points?

Solution 2.26

Listed.

1. Let x be a limit point of E' . Then, for every $\epsilon > 0$, $U = B_\epsilon(x) \cap E' \neq \emptyset$. Take a $y \in U$. Since $y \in B_\epsilon(x)$, which is open, we can construct an open ball $B_\delta(y) \subset B_\epsilon(x)$. Since $y \in E'$, $B_\delta(y)$ must contain elements of E , which means that $B_\epsilon(x)$ must also contain elements of E , and so x is a limit point of $E \implies x \in E'$ and E' is closed.
2. To prove that $E' \subset \overline{E}'$, we know that if $x \in E'$, then for every $\epsilon > 0$, there exists a $B_\epsilon^\circ(x)$ that has a nontrivial intersection with E which means that it has a nontrivial intersection with $E \cup E'$. To prove that $\overline{E}' \subset E'$, we know that if $y \in \overline{E}'$, then for every $\delta > 0$ there exists a $B_\delta(x)$ that has a nontrivial intersection with \overline{E} . If $B_\delta(x)$ intersects E then we are done. If $B_\delta(x)$ intersects E' , then we can find a $y \in E' \cap B_\delta(x)$. Since $B_\delta(x)$ is open, we can construct $B_\epsilon(y) \subset B_\delta(x)$ and since $y \in E'$, we know that $B_\epsilon(y)$ contains an element of E , which means that $B_\delta(x)$ contains an element of E . Therefore, $E' = \overline{E}'$.
3. No. Consider the set $E = \{1/n\}_{n \in \mathbb{N}}$. $E' = \{0\}$, but $E'' = \emptyset$.

Exercise 2.28 (Rudin 2.7)

Let A_1, A_2, \dots be subsets of a metric space.

1. If $B_n = \bigcup_{i=1}^n A_i$, prove that $\bar{B}_n = \bigcup_{i=1}^n \bar{A}_i$ for $n = 1, 2, 3, \dots$
2. If $B = \bigcup_{i=1}^{\infty} A_i$, prove that $\bar{B} \supset \bigcup_{i=1}^{\infty} \bar{A}_i$.

Solution 2.27

Listed.

1. We will prove that $\overline{B_n} \subseteq \bigcup_{i=1}^n \overline{A_i}$ and $\bigcup_{i=1}^n \overline{A_i} \subseteq \overline{B_n}$. If $x \in B_n$, then $x \in \bigcup_{i=1}^n A_i$. Therefore, assume that $x \in B'_n$. Then for every $\epsilon > 0$, there exists a $B_\epsilon^\circ(x)$ s.t.

$$B_\epsilon^\circ(x) \cap B_n \neq \emptyset \implies B_\epsilon^\circ(x) \cap \left(\bigcup_{i=1}^n A_i \right) \neq \emptyset$$

This means that there exists some $i = i(\epsilon)$, a function of ϵ , s.t. $B_\epsilon^\circ(x) \cap A_i \neq \emptyset$. However, this i may change if we unfix ϵ . We have so far proved that just for one $\epsilon > 0$ there exists an i . Now if we take a sequence of $\epsilon = 1, \frac{1}{2}, \frac{1}{3}, \dots$, we have a sequence of $i(\epsilon)$'s living in $\{1, \dots, n\}$. By the pigeonhole principle, there must be at least one i that is hit infinitely many times, and so we can choose this i , that works for all $\epsilon > 0 \implies x \in A'_i \subseteq \bigcup_{i=1}^n \overline{A_i}$. If $x \in \bigcup_{i=1}^n \overline{A_i}$, then there exists an $\overline{A_i}$ s.t. $x \in \overline{A_i}$. If $x \in A_i$, then we are done. If $x \in A'_i$, then for every $\epsilon > 0$, there exists a $B_\epsilon^\circ(x)$ s.t.

$$B_\epsilon^\circ(x) \cap A_i \neq \emptyset \implies B_\epsilon^\circ(x) \cap \left(\bigcup_{i=1}^n A_i \right) \neq \emptyset$$

and so $x \in B'_n \subset \overline{B_n}$.

2. $x \in \bigcup_{i=1}^{\infty} \overline{A_i} \implies x \in \overline{A_i}$ for some i . If $x \in A_i$, then $x \in B$ and we are done. If $x \in A'_i$, then for every $\epsilon > 0$ there exists $B_\epsilon^\circ(x)$ s.t.

$$B_\epsilon^\circ(x) \cap A_i \neq \emptyset \implies B_\epsilon^\circ(x) \cap \left(\bigcup_{i=1}^{\infty} A_i \right) \neq \emptyset$$

and so $B_\epsilon^\circ(x) \cap B \neq \emptyset \implies x \in B' \subset \overline{B}$.

Exercise 2.29 (Rudin 2.8)

Is every point of every open set $E \subset \mathbb{R}^2$ a limit point of E ? Answer the same question for closed sets in \mathbb{R}^2 .

Solution 2.28

Yes for open. Given any $x \in U$ open, there always exists an $\epsilon > 0$ s.t.

$$B_\epsilon^\circ(x) \subset B_\epsilon(x) \subset U \tag{316}$$

and so $B_\epsilon^\circ(x)$ has a nontrivial intersection with U . If U is closed, then no. Note that for closed U , we have that every limit point is in U , but not every point in U is a limit point. Consider the isolated point $U = \{x\}$. x is not a limit point of U .

Exercise 2.30 (Rudin 2.9)

Let E° denote the set of all interior points of E in X . Prove the following:

- E° is always open.
- E is open if and only if $E^\circ = E$.
- If $G \subseteq E$ and G is open, then $G \subset E^\circ$.
- Prove that the complement of E° is the closure of the complement of E .
- Do E and \bar{E} always have the same interiors?
- Do E and E° always have the same closures?

Solution 2.29

Listed.

- We assume that E° is not open (this does not mean that E° is necessarily closed!). That is, there exists an $x \in E^\circ$ s.t. we can't construct an open ball $B_\epsilon(x) \subseteq E^\circ$. Since $x \in E^\circ \subset E$, by definition of an interior point we can construct a $B_\epsilon(x) \subset E$. But from our assumption $B_\epsilon(x) \not\subseteq E^\circ$. We choose a $y \in B_\epsilon(x) \setminus E^\circ$. Since $B_\epsilon(x)$ is open, there exists a $\delta > 0$ s.t.

$$B_\delta(y) \subset B_\epsilon(x) \subset E$$

But the fact that we can construct an open ball around y means that $y \in E^\circ$, leading to a contradiction.

- If E is open, then by definition $E \subset E^\circ$. Now $E^\circ \subset E$ holds for all sets since E° must be composed of points from E . If $E = E^\circ$, then for every $x \in E$, $x \in E^\circ$, so by definition there exists an $\epsilon > 0$ s.t. $B_\epsilon(x) \subset E$, which means that E is open.
- Let $x \in G$ open. Then there exists an $\epsilon > 0$ s.t. $B_\epsilon(x) \subset G$, and so $B_\epsilon(x) \subset E$. Since we can always construct an open ball around x contained within E , $x \in E^\circ$ and $G \subset E^\circ$.
- ((E°)^c ⊂ $\overline{E^c}$) If $x \in (E^\circ)^c$, then there exists no $\epsilon > 0$ s.t. $B_\epsilon(x) \subset E$. Then, for any $\epsilon > 0$, $B_\epsilon(x) \not\subseteq E \implies B_\epsilon(x) \cap E^c \neq \emptyset \implies x \in E^c \subset \overline{E^c}$. ($\overline{E^c} \subset (E^\circ)^c$) If $x \in \overline{E^c}$, then $x \in E^c$ or $x \in E^{c'}$. If $x \in E^c$, note $E^\circ \subset E \implies (E^\circ)^c \supset E^c \implies x \in (E^\circ)^c$. If $x \in E^{c'}$, then for all $\epsilon > 0$ $B_\epsilon(x) \cap E^c \neq \emptyset \implies B_\epsilon(x) \not\subseteq E \implies x \in E^\circ$.
- No. Consider the rationals $\mathbb{Q} \subset \mathbb{R}$. $\mathbb{Q}^\circ = \emptyset$ but $\overline{\mathbb{Q}^\circ} = \mathbb{R}^\circ = \mathbb{R}$. It is true and straightforward to prove that $E^\circ \subset \overline{E^\circ}$. Let $x \in E^\circ$. Then there exists an $\epsilon > 0$ s.t. $B_\epsilon(x) \subset E \implies B_\epsilon(x) \subset \overline{E} \implies x \in \overline{E^\circ}$.
- No. Consider $\mathbb{Q} \subset \mathbb{R}$. Then $\overline{\mathbb{Q}} = \mathbb{R}$ and $\overline{\mathbb{Q}^\circ} = \emptyset = \emptyset$.

Exercise 2.31 (Rudin 2.10)

Let X be an infinite set. For $p \in X$ and $q \in X$, define

$$d(p, q) = \begin{cases} 1 & \text{if } p \neq q \\ 0 & \text{if } p = q \end{cases} \quad (317)$$

Prove that this is a metric. Which subsets of the resulting metric space are open? Which are closed? Which are compact?

Solution 2.30

This is a metric since clearly it satisfies symmetry and the fact that $d(p, p) = 0$. The triangle inequality

$$d(p, r) \leq d(p, q) + d(q, r) \quad (318)$$

is trivially satisfied if $p = r$, and if $p \neq r$, then either $p \neq q$ or $q \neq r$, and so the RHS ≥ 1 . An open ϵ -ball around $x \in X$ is either X , when $\epsilon > 1$, or $\{x\}$ when $\epsilon \leq 1$. Therefore

Exercise 2.32 (Rudin 2.11)

For $x \in \mathbb{R}$ and $y \in \mathbb{R}$, define

$$\begin{aligned}d_1(x, y) &= (x - y)^2 \\d_2(x, y) &= \sqrt{|x - y|} \\d_3(x, y) &= |x^2 - y^2| \\d_4(x, y) &= |x - 2y| \\d_5(x, y) &= \frac{|x - y|}{1 + |x - y|}\end{aligned}$$

Determine, for each of these, whether it is a metric or not.

Solution 2.31

Listed. Positive semidefiniteness and symmetry are easy to check.

1. The triangle inequality gives

$$\begin{aligned}d_1(x, z) \leq d_1(x, y) + d_1(y, z) &\iff (x - z)^2 \leq (x - y)^2 + (y - z)^2 \\&\iff 0 \leq (x - y)(y - z)\end{aligned}$$

which is not satisfied if $x < y < z$, so this is not a valid metric.

2. The triangle inequality gives $\sqrt{|x - z|} \leq \sqrt{|x - y|} + \sqrt{|y - z|}$, and since both sides are positive this inequality is equivalent to squaring both sides to get

$$|x - z| \leq |x - y| + |y - z| + 2\sqrt{|x - y||y - z|}$$

which is true since $|x - z| \leq |x - y| + |y - z|$ of the Euclidean distance satisfies the triangle inequality and $0 \leq \sqrt{|x - y||y - z|}$.

3. This does not satisfy triangle inequality, as taking 0, 1, 2 gives

$$d_3(0, 2) = 4 > 1 + 1 = d_3(0, 1) + d_3(1, 2)$$

4. This does not satisfy symmetry.

5. For simplicity, let us set $A = |x - y|$, $B = |y - z|$, $C = |x - z|$. Then, we get

$$\frac{C}{1 + C} \leq \frac{A}{1 + A} + \frac{B}{1 + B} \iff C \leq A + B + 2AB + ABC$$

where $C \leq A + B$ is true by triangle inequality of Euclidean distance, $0 \leq AB$, and $0 \leq ABC$. Intuitively, we want a metric that doesn't "blow up" the distance between x and y . More precisely, we want a valid metric $d(x, y)$ to be $O(|x - y|)$. Having something like a quadratic growth rate $(x - y)^2$ will blow the distance $d(x, z)$ up too much overpowering the individual $d(x, y) + d(y, z)$.

Exercise 2.33 (Rudin 2.12)

Let $K \subset \mathbb{R}$ consist of 0 and the numbers $1/n$ for $n = 1, 2, 3, \dots$. Prove that K is compact directly from the definition (without using the Heine-Borel theorem).

Solution 2.32

Every open cover of K must have an open set G s.t. $0 \in G$. Since G is open, there exists an open neighborhood $B_\epsilon(0) \subset G$ around 0. By the Archimedean principle, there exists an $N \in \mathbb{N}$ s.t.

$$\epsilon N > 1 \implies \epsilon > \frac{1}{N} \quad (319)$$

and so, $B_\epsilon(0)$ contains all points $\{1/n\}$ for $n > N$. For the rest of the points $1, 1/2, \dots, 1/N$, we can simply construct a finite cover over each of them, hence getting a finite cover.

Exercise 2.34 (Rudin 2.13)

Construct a compact set of real numbers whose limit points form a countable set.

Solution 2.33

Consider the set

$$E = \left\{ \left(\frac{1}{10}\right)^n + \left(\frac{1}{10}\right)^{n+k} : n \in \{0\} \cup \mathbb{N}, k \in \mathbb{N} \right\} \cup \{0\} \quad (320)$$

This is clearly bounded by 0 and 1.1. Let us represent the elements of this set by (n, k) . We can show that

$$(n_1, k_1) > (n_2, k_2) \quad (321)$$

if $n_1 < n_2$ or $n_1 = n_2$ and $k_1 < k_2$. Therefore, to prove closedness, we must prove that every limit point is a point in E . We can do this by proving that a point not in E cannot be a limit point. Choose any $x \notin E$. Then, due to the ordering, we can see that there exists a (n, k) s.t.

$$A = \left(\frac{1}{10}\right)^n + \left(\frac{1}{10}\right)^{n+k} < k < \left(\frac{1}{10}\right)^n + \left(\frac{1}{10}\right)^{n+k+1} = B \quad (322)$$

and so we can take $\epsilon = \min\{k - A, B - k\}$ and show that $B_\epsilon(x)$ does not contain A nor B , and so has an empty intersection with E . Therefore, it cannot be a limit point of E and is closed. Since E is bounded and closed in \mathbb{R} , it is compact. Its limit points contain $1, 0.1, 0.01, \dots, 0$ (simply by fixing n and letting $k \rightarrow \infty$, and so E' is infinite. We have just shown that since E is closed, $E' \subset E$. But E is countable, so E' is countable.

Exercise 2.35 (Rudin 2.14)

Given an example of an open cover of the segment $(0, 1)$ which has no finite subcover.

Solution 2.34

Consider

$$(0, 1/2) \cup \left(\bigcup_{i=1}^{\infty} \left[1 - \frac{1}{2^i}, 1 - \frac{1}{2^{i+1}} \right) \right) \quad (323)$$

Exercise 2.36 (Rudin 2.15)

Exercise 2.37 (Rudin 2.16)

Regard \mathbb{Q} , the set of all rational numbers, as a metric space, with $d(p, q) = |p - q|$. Let E be the set of all $p \in \mathbb{Q}$ s.t. $2 < p^2 < 3$. Show that E is closed and bounded in \mathbb{Q} but is not compact. Is E open in \mathbb{Q} ?

Solution 2.35

E is clearly bounded by 0 and 2 since $0^2 < 2 < p^2 < 3 < 2^2$. It is closed and we can show this by showing that E^c is open. Let $x \in E^c$. Then, $x^2 < 2$ or $x^2 > 3$.

1. $x^2 < 2 \iff -\sqrt{2} < x < \sqrt{2}$. Now let $\epsilon = \min\{\sqrt{2} - x, x + \sqrt{2}\} > 0$. Then by the Archimedean property there exists a $n \in \mathbb{N}$ s.t. $0 < \frac{1}{n} < \epsilon$. Therefore, the image of $B_{1/n}(x) \subset \mathbb{Q}$ will map onto $(0, 2)$.
2. $x^2 > 3 \iff x > \sqrt{3}$ or $x < -\sqrt{3}$. If $x > \sqrt{3}$, then by AP there exists a $n \in \mathbb{N}$ s.t. $x - \frac{1}{n} > \sqrt{3} \implies (x - \frac{1}{n})^2 > 3$. If $x < -\sqrt{3}$, then by AP there exist $n \in \mathbb{N}$ s.t. $x + \frac{1}{n} < -\sqrt{3} \implies (x + \frac{1}{n})^2 > 3$. Either way, the image of $B_{1/n}(x)$ will map within E^c .

It is not compact because E is not closed in \mathbb{R} . The limit points of E in \mathbb{R} is $[\sqrt{2}, \sqrt{3}] \cup [-\sqrt{3}, -\sqrt{2}]$, which contains irrationals and is clearly not a subset of E . Since it is not closed in \mathbb{R} , it is not compact in \mathbb{R} , and it is not compact in $\mathbb{Q} \subset \mathbb{R}$. It is open because

$$E = ((\sqrt{2}, \sqrt{3}) \cup (-\sqrt{3}, \sqrt{2})) \cup \mathbb{Q} \subset \mathbb{R} \quad (324)$$

which is the union of open $(\sqrt{2}, \sqrt{3}) \cup (-\sqrt{3}, \sqrt{2})$ and subset $\mathbb{Q} \subset \mathbb{R}$, and so it is open.

Exercise 2.38 (Rudin 2.17)

Let E be the set of all $x \in [0, 1]$ whose decimal expansion consists of only the digits 4 and 7. Is E countable? Is E dense in $[0, 1]$? Is E compact? Is E perfect?

Exercise 2.39 (Rudin 2.18)

Is there a nonempty perfect set in \mathbb{R} which contains no rational number?

Exercise 2.40 (Rudin 2.19)

Listed.

1. If A and B are disjoint closed sets in some metric space X , prove that they are separated.
2. Prove the same for disjoint open sets.
3. Fix $p \in X$, $\delta > 0$, define A to be the set of all $q \in X$ for which $d(p, q) < \delta$. Define B similarly, with $>$ in place of $<$. Prove that A and B are separated.
4. Prove that every connected metric space with at least two points is uncountable.

Solution 2.36

Listed.

1. This is trivial with the fact that the closure of the closure of A is the closure of A .
2. Let A, B be open. We wish to show that if $x \in A'$, then $x \notin B$. Assume $x \in B$. Then there exists $\epsilon > 0$ s.t. $B_\epsilon(x) \subset B$. But $B \cap A = \emptyset \implies B_\epsilon(x) \cap A = \emptyset$ and so $x \notin A'$, which is a contradiction.
3. Clearly, $A \cap B = \emptyset$. Not let $x \in A \implies$ there exists $\epsilon > 0$ s.t. $B_\epsilon(x) \subset A \implies B_\epsilon(x) \cap B = \emptyset \implies x \in B'$. We can prove similarly to show that $x \in B \implies x \notin A'$.

4. Assume X is countable (solution is very similar for finite). Then, we can enumerate a $X = \{x_i\}_{i=1}^{\infty}$. We wish to show that X can be decomposed into the union of an open ball and the interior of its complement as shown in (3). We fix $p \in X$. Then, we take the set $D = \{d(p, x)\}_{x \neq p} \subset \mathbb{R}$. Since D is a countable subset of \mathbb{R} , there must exist some $\alpha > 0$ s.t. $\alpha \notin D$. This α partitions the distances into two sets, and we can define

$$X = \{q \in X \mid d(p, q) < \alpha\} \cup \{q \in X \mid d(p, q) > \alpha\}$$

and by (3), these two sets are separated, which means that X is not connected, leading to a contradiction.

Exercise 2.41 (Rudin 2.20)

Are closures and interiors of connected sets always connected? Look at subsets of \mathbb{R}^2 .

Solution 2.37

The interiors are not always connected. Consider the two closed balls $\overline{B_1((1, 0))}$ and $\overline{B_1((-1, 0))}$ as subsets of \mathbb{R}^2 . They are connected but their interiors, which are the two open balls, are not connected. As for closures, they are always connected. Let W be connected. Then for any partition $A \cup B = W$, $\overline{A} \cap B \neq \emptyset$ WLOG. Consider $\overline{W} = W \cup W'$ and take any partition $\overline{W} = C \cup D$. Then, label $A = C \cap W$, $A^* = C \cap W'$, $B = D \cap W$, $B^* = D \cap W'$. This implies that $C = A \cup A^*$, $D = B \cup B^*$, and $A \cup B = W$ (which is connected). Then, we can show that

$$\begin{aligned}\overline{C} \cap D &= (\overline{A \cup A^*} \cap D) = (\overline{A} \cup \overline{A^*}) \cap D = (\overline{A} \cap D) \cup (\overline{A^*} \cap D) \\ &= (\overline{A} \cap B) \cup (\overline{A} \cap B^*) \cup (\overline{A^*} \cap D)\end{aligned}$$

which cannot be empty since by connectedness of W , $\overline{A} \cap B \neq \emptyset$. Therefore, \overline{W} is connected.

Exercise 2.42 (Rudin 2.21)

Let A and B be separated subsets of some \mathbb{R}^k . Suppose $a \in A, b \in B$ and define

$$p(t) = (1 - t)a + tb \tag{325}$$

for $t \in \mathbb{R}$. Put $A_0 = p^{-1}(A)$, $B_0 = p^{-1}(B)$.

1. Prove that A_0 and B_0 are separated subsets of \mathbb{R} .
2. Prove that there exists a $t_0 \in (0, 1)$ s.t. $p(t_0) \notin A \cup B$.
3. Prove that every convex subset of \mathbb{R}^k is connected.

Exercise 2.43 (Rudin 2.22)

A metric space is called *separable* if it contains a countable dense subset. Show that \mathbb{R}^k is separable.

Solution 2.38

Consider the set $\mathbb{Q}^k \subset \mathbb{R}^k$. It is a finite Cartesian product (and hence, a countable union) of countable \mathbb{Q} , and so it is countable. \mathbb{Q}^k is dense in \mathbb{R}^k since given any $x \in \mathbb{R}^k$, we claim x is a limit point of \mathbb{Q}^k . Given any $\epsilon > 0$, we can construct $B_{\epsilon}^o(x)$. For each coordinate x_i , by density of rationals in \mathbb{R} we can choose a $q_i \in \mathbb{Q}$ s.t. $0 < d(x_i, q_i) < \epsilon/k$. Then, using triangle inequality, we can take the distances between each coordinate changed from x_i to q_i . Let q^k be the vector x with the components

x_1, \dots, x_k changed to q_1, \dots, q_k , respectively.

$$d(x, q) = d(x, q^1) + d(q^1, q^2) + \dots + d(q^{k-1}, q_k) < \frac{\epsilon}{k} + \dots + \frac{\epsilon}{k} = \epsilon \quad (326)$$

and so $q \in B_\epsilon^\circ(x)$. Hence the intersection of \mathbb{Q}^k and $B_\epsilon^\circ(x)$ for any $\epsilon > 0$ is nontrivial, so x is a limit point of \mathbb{Q}^k .

Exercise 2.44 (Rudin 2.23)

A collection $\{V_\alpha\}$ of open subsets of X is said to be a *base* for X if the following is true: For every $x \in X$ and every open set $G \subset X$ such that $x \in G$, we have $x \in V_\alpha \subset G$ for some α . In other words, every open set in X is the union of a subcollection of $\{V_\alpha\}$. Prove that every separable metric space has a countable base.

Solution 2.39

Since X is separable it contains a countable dense subset, call it S . Then for every $x \in S$, we can look at the set of all open balls with center x and rational radii, call it \mathcal{B} . Then \mathcal{B} is countable. Now consider an open set U . By definition, for every $x \in U$, there exists an $\epsilon > 0$ s.t. $B_\epsilon(x) \subset U$. By AP, we can find a $n \in \mathbb{N}$ s.t. $0 < \frac{1}{n} < \epsilon$, and therefore we can find an open ball $B \in \mathcal{B}$ s.t. $B(x) \subset U$. We claim that

$$W := \bigcup_{x \in U} B(x) = U \quad (327)$$

If $x \in U$, then by construction it is contained in $B(x) \subset \bigcup_{x \in U} B(x)$, and so $U \subset W$. If $x \in W$, then it is in $B(x)$, which is fully contained in U and so $W \subset U$. Therefore every open set can be constructed by a countable union of open balls in countable \mathcal{B} .

Exercise 2.45 (Rudin 2.24)

Let X be a metric space in which every infinite subset has a limit point. Prove that X is separable.

Solution 2.40

We fix $\delta > 0$. Choose $x_1 \in X$. Then choose $x_2 \in X$ s.t. $d(x_1, x_2) \geq \delta$, and keep doing this until we choose $x_{j+1} \in X$ s.t. $d(x_{j+1}, x_i) \geq \delta$ for all $i \in 1, \dots, j$.

1. We claim that this must stop after a finite number of steps. Assume it doesn't. Then by assumption $V = \{x_i\}_{i=1}^\infty$ should have a limit point in X , denote it x . Choose $\frac{\delta}{2} > 0$. Then, $B_{\delta/2}^\circ(x) \cap V \neq \emptyset$. This intersection can only have one point since if it had two x', x'' , then since both are in $B_{\delta/2}(x)$, then

$$d(x', x'') \leq d(x', x) + d(x, x'') \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

and since they are both in V , then $d(x', x'') \geq \delta$, which is a contradiction. Since there is a finite number of points in $B_{\delta/2}(x)$ of V , x cannot be a limit point. So this must terminate at some finite $J < \infty$.

2. Denote $W = \{x_i\}_{i=1}^J$. Then, $\mathcal{B}_\delta = \{B_\delta(x) \mid x \in W\}$ must cover X , since if it didn't, there would exist a $y \in X$ s.t. $d(y, x) \geq \delta$ for all $x \in W$, and we can add another element in W .
3. Consider $\delta = 1, 1/2, 1/3, \dots$ and construct the same cover

$$\mathcal{B}_k = \{B_{1/k}(x_{ki}) \mid i = 1, \dots, J_k\}$$

which is finite. Therefore, $\mathcal{B} = \bigcup_{k=1}^\infty \mathcal{B}_k$ must be countable.

4. We claim that countable $\{x_{ki}\}_{k,i}$ is dense. Consider any $x \in X$. For every $\epsilon > 0$, we can find an arbitrarily large $n \in \mathbb{N}$ s.t. $0 < \frac{1}{n} < \epsilon$. Since \mathcal{B}_n is an open cover, there must exist some x_{ni} s.t. $x \in B_{1/n}(x_{ni})$, which by symmetry implies that $x_{ni} \in B_{1/n}(x) \subset B_\epsilon(x)$. Therefore, there always exists an x_{ni} in every $B_\epsilon(x)$, and so $B_\epsilon(x) \cap \{x_{ki}\} \neq \emptyset \implies x$ is a limit point of $\{x_{ki}\}$ and so it is dense.

Exercise 2.46 (Rudin 2.25)

Prove that every compact metric space K has a countable base, and that K is therefore separable.

Solution 2.41

For every $n \in \mathbb{N}$, let us consider an open covering $\mathcal{F}_n := \{B_{1/n}(x_n) \mid x_n \in K\}$. Since K is compact, it has a finite subcovering

$$\mathcal{G}_n := \{B_{1/n}(x_{ni}) \mid i = 1, \dots, k(n)\} \quad (328)$$

Now consider the union $\mathcal{G} = \bigcup_{i=1}^n \mathcal{G}_n$, which is countable. We claim that \mathcal{G} is a base. Consider any open set U . Then for every $x \in U$, we want to show that x is contained in a $B_{1/n}(x_{ni}) \subset U$. Since U is open, there exists a $\epsilon > 0$ s.t. $B_\epsilon(x) \subset U$. Now by AP, there exists a $n \in \mathbb{N}$ s.t. $0 < \frac{1}{n} < \frac{\epsilon}{2}$. Therefore $B_{1/n}(x) \subset B_\epsilon(x)$. Since \mathcal{G} is an open covering, there must exist some $B_{1/n}(x_{ni})$ that contains x . Now we wish to show that $B_{1/n}(x_{ni})$ is fully contained in U . Let $y \in B_{1/n}(x_{ni})$. Then, by triangle inequality,

$$d(y, x) = d(y, x_{ni}) + d(x_{ni}, x) < \frac{1}{n} + \frac{1}{n} < \epsilon \quad (329)$$

and therefore $x \in B_{1/n}(x_{ni}) \subset B_\epsilon(x)$. Therefore, for every $x \in U$, we can construct an open ball of \mathcal{G} containing x and contained in U , proving that this is a base.

We claim that the set of all $\mathcal{P} = \{x_{ni}\}_{n,i}$ forms a countable dense subset. This is clearly countable since \mathcal{G} is countable. We must prove that the closure of $\mathcal{P} = K$. Let $x \in K$. Given any $\epsilon > 0$, we wish to show that $B_\epsilon(x) \cap \mathcal{P} \neq \emptyset$. Since $B_\epsilon(x)$ is open, it can be covered by a subcollection of \mathcal{G} , and so their centers must be in $B_\epsilon(x)$, proving that $B_\epsilon(x) \cap \mathcal{P} \neq \emptyset$. Therefore, x is a limit point of \mathcal{P} .

Exercise 2.47 (Rudin 2.26)

Exercise 2.48 (Rudin 2.27)

Exercise 2.49 (Rudin 2.28)

Exercise 2.50 (Rudin 2.29)

Prove that every open set in \mathbb{R} is the union of an at most countable collection of disjoint segments.

Solution 2.42

Let $U \subset \mathbb{R}$ be open. Then for all $x \in U$ there exists $\epsilon > 0$ s.t. $(x - \epsilon, x + \epsilon) \subset U$. Now since \mathbb{R} is separable (by exercise Rudin 2.22), it has a countable dense subset \mathbb{Q} . Consider all segments of rational radius and rational centers

$$\mathcal{B} = \{(q - p, q + p) \subset \mathbb{R} \mid q, p \in \mathbb{Q}\} \quad (330)$$

This is clearly countable. We claim that every open U can be expressed as the union of a subset of \mathcal{B} . Now by AP, there exists $n \in \mathbb{N}$ s.t. $0 < \frac{1}{n} < \frac{\epsilon}{2}$, so for all $x \in U$, there exists $n \in \mathbb{N}$ s.t. $(x - \frac{1}{n}, x + \frac{1}{n}) \subset U$. Now since \mathbb{Q} is dense in \mathbb{R} , $x \in \mathbb{Q}' \implies (x - \frac{1}{n}, x + \frac{1}{n}) \cap \mathbb{Q} \neq \emptyset$. Say r is in this intersection. Then, by symmetry of metric, $x \in (r - \frac{1}{n}, r + \frac{1}{n})$. Therefore, for all $x \in U$, we have found an open ball in \mathcal{B} that contains x . Now, we must show that this actually is fully contained in U . This is easy, since if $y \in B_{1/n}(r)$, then

$$d(y, x) \leq d(y, r) + d(r, x) \leq \frac{1}{n} + \frac{1}{n} < \epsilon \quad (331)$$

and so $B_{1/n}(r)$ is completely contained in the ϵ -ball around x , which is a subset of U . So for all $x \in U$, we found an open set $U_x \in \mathcal{B}$ covering x and fully contained in U , which means that $\cup_{x \in U} U_x = U$. Now for some intervals $B_1, B_2 \in \mathcal{B}$, if $B_1 \cap B_2 \neq \emptyset$, take their union, which is another segment, and keep doing this until $B_i \cap B_j = \emptyset$ for all i, j . The cardinality of this new pruned set will be less than or equal to \mathcal{B} , which is countable, and so this must be at most countable.

3 Sequences in Euclidean Space

We have already defined sequences, but we'll add to our vocabulary to describe and classify sequences and series. To make this chapter a bit more self-contained, we redefine sequences and series.

Definition 3.1 (Sequence)

A sequence on a space X is a function $f : \mathbb{N} \rightarrow X$. We usually denote the sequence as (x_n) , where $x_n = f(n)$.^a

^aNote that (x_n) as a sequence is different from $\{x_n\}$ as a set.

Definition 3.2 (Constant Sequence)

Let X be any set.

1. $\{x_i\}$ is a **constant sequence** if $x_i = x$ for all i
2. $\{x_i\}$ is an **ultimately constant sequence** if $x_i = x$ for all $i > N$ for some $N \in \mathbb{N}$. If $x = 0$ (assuming that an identity exists), then $\{a_i\}$ is **finary**.

3.1 Convergence in Metric Topologies

Now that we have defined sequences, we want to talk about their convergence, and if they do converge, *what* they converge to.

Definition 3.3 (Limit of Sequence in Topological Space)

A number $x \in \mathbb{R}$ is called the **limit of the sequence** $\{x_n\}$, written

$$\lim_{n \rightarrow \infty} x_n = x \quad (332)$$

if for every neighborhood U_x there exists an index N such that $x_n \in U_x$ for all $n > N$. If x is the limit of (x_n) , then we say that (x_n) **converges** to x . If the limit of (x_n) is not well defined or finite, then we say that (x_n) is **divergent**.

Note that x being the limit of a sequence (x_i) is stronger than the claim that x is a limit point of $\{x_i\}$. If we consider the sequence $0, 1, 0, 1, \dots$, we can see that both 0 and 1 are limit points, but the limit does not exist. We would like to define some notion of limit points in the language of sequences. We can precisely do this by treating a sequence as a set and talking about subsequential limits.

Definition 3.4 (Subsequences)

A **subsequence** of $\{a_n\}$ is a sequence $\{a_{\gamma_k}\}$, where $\{\gamma_k\}$ is a strictly increasing infinite subset of \mathbb{N} .

Definition 3.5 (Partial Limits)

The **partial limit** of a sequence $\{x_n\}$ is the limit of any of its subsequence.

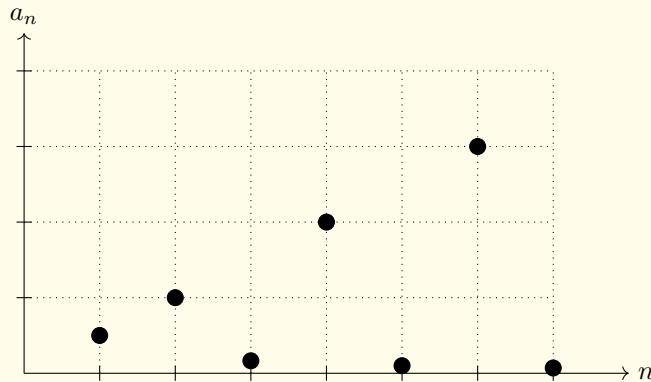


Figure 1: Two partial limits of the sequence $a_n = 1/n$ for n odd and $n/2$ for n even, is $+\infty$ and 0.

Lemma 3.1 (Partial Limits Equivalent to Limit Point)

Given a sequence (x_n) , x is a limit point of $\{x_n\}$ iff there exists a subsequence of (x_n) that converges to x .

Immediately from the properties of the metric topology we know the following.

Definition 3.6 (Limit of a Sequence in Metric Space)

x is the limit of (x_n) if for every $\epsilon > 0$, there exists an index $N \in \mathbb{N}$ such that

$$d(x_n, x) < \epsilon \forall n > N \quad (333)$$

The next few theorems help us develop some intuition behind convergence of sequences. This is where the concept of limit points from topology becomes connected to sequences. In here, a limit point of a sequence is a viable candidate for which a sequence converges to.

Theorem 3.1 (Sufficient Conditions for Convergence in Metric Space)

Let $\{x_n\}$ be a sequence in a metric space X .

1. $\{x_n\}$ converges to $x \in X$ if and only if every neighborhood of x contains x_n for all but finitely many n .
2. If $\{x_n\}$ converges to x , then x is unique.
3. If $\{x_n\}$ converges, then $\{x_n\}$ is bounded.
4. If $E \subset X$ and x is a limit point of E , then there exists a sequence $\{x_n\}$ in E that converges to x .

Proof.

Listed.

1. (\Rightarrow) Let $p_n \rightarrow p \in X$. Then, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ s.t. $d(p, p_n) < \epsilon$ for all $n > N$. Given neighborhood $B_\epsilon(p)$, $x_n \in B_\epsilon(p)$ for all $n > N \Rightarrow$ at most N elements are not in $B_\epsilon(p)$. (\Leftarrow) Now for any $\epsilon > 0$, let every $B_\epsilon(p)$ contain all but finitely many p_n . Enumerate them $\{x_{n_k}\}_{k=1}^K$, and let

$$\alpha = \max\{n_k\}$$

- This means that there exists an $\alpha \in \mathbb{N}$ s.t. $p_n \in B_\epsilon(p)$ for all $n > \alpha$, which implies that $p_n \rightarrow p$.
2. Assume $\{p_n\}$ converges to $p, p' \in X$, with $p \neq p'$. Then for all $\epsilon > 0$ there exists $N_1, N_2 \in \mathbb{N}$

s.t. $d(p, p_n) < \epsilon$ for all $n > N_1$ and $d(p', p_n) < \epsilon$ for all $n > N_2$. This means that there exists a $N = \max\{N_1, N_2\}$ satisfying the above. Since $p \neq p'$, set $\epsilon = d(p, p')/2$. Then,

$$d(p, p_n) < \frac{d(p, p')}{2} \text{ and } d(p', p_n) < \frac{d(p, p')}{2}$$

which implies that by adding both sides and invoking triangle inequality, we have

$$d(p, p') \leq d(p, p_n) + d(p', p_n) < d(p, p')$$

which is absurd.

3. Choose any $\epsilon > 0$. Then from (a), $B_\epsilon(x)$ contains all but finitely many $V = \{x_{n_k}\}_{k=1}^K$. Take

$$M = \max\{\epsilon, d(x_{n_1}, x), \dots, d(x_{n_K}, x)\}$$

and so $d(x, x_n) < M$ for all $n \in \mathbb{N}$.^a

4. We can explicitly construct one. Let $p \in E'$. Then choose $\epsilon = 1, \frac{1}{2}, \frac{1}{3}, \dots$ and for every $\epsilon > 0$, $B_\epsilon(p) \cap E \neq \emptyset$. Choose a p_n within this intersection for every $\epsilon = \frac{1}{n}$. Then, we have $\{p_n\}$ contained in E . We want to show that this converges to p . Take any $\epsilon > 0$, then there exists $N \in \mathbb{N}$ s.t. $0 < \frac{1}{N} < \epsilon$, and for every $n > N$, $\frac{1}{n} < \frac{1}{N} < \epsilon$. Therefore, for every $n > N$,

$$p_n \in B_{1/n}(p) \subset B_\epsilon(p)$$

which means that $p_n \in B_\epsilon(p)$ for all $n > N$, implying that $\lim_{n \rightarrow \infty} p_n = p$.

^aThis is also a direct result of every metric topology being Hausdorff.

Lemma 3.2 ()

$\{x_n\}$ converges to x if and only if every subsequence of $\{x_n\}$ converges to x .

Proof.

Let $p_n \rightarrow p$. Then, take any subsequence $\{p_{n_k}\}$ of $\{p_n\}$. For any $\epsilon > 0$, there exists a $N \in \mathbb{N}$ s.t. $d(p, p_n) < \epsilon$ for all $n > N$. Since N is finite and the n_k 's are unbounded, there must exist a $K \in \mathbb{N}$ s.t. $n_k > N$ if $k > K$. Therefore, given any $\epsilon > 0$, we have proved the existence of a $K \in \mathbb{N}$ s.t. $k > K \implies n_k > N$, which implies by convergence of p_n , that

$$d(p_{n_k}, p) < \epsilon \tag{334}$$

which by definition means that p_{n_k} converges to p . Now, for the other direction, given $\{p_n\}$ with every subsequence converging to p , we can take the subsequence $\{p_n\}$ itself ($n_k = k$), which converges to p .

Theorem 3.2 ()

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

Proof.

Corollary 3.1 (Bolzano-Weierstrass Theorem)

Every bounded sequence in \mathbb{R}^n contains a convergent subsequence.

Proof.

It suffices to prove that there exists a monotonic sequence within a bounded sequence $\{x_n\}$.

Theorem 3.3 (Nested Compact Sets)

Listed.

1. If \overline{E} is the closure of a set E in a metric space X , then

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

2. If K_n is a sequence of compact sets in X s.t. $K_n \supset K_{n+1}$ for $n \in \mathbb{N}$ and if

$$\lim_{n \rightarrow \infty} \text{diam } K_n = 0$$

then $\cap_{n=1}^{\infty} K_n$ consists of exactly one point.

So far, we have talked about the limits of a sequence, which may or may not exist. But we do know from 3.2.2 that there *always* exists a subsequence that is either convergent or tends to $\pm\infty$. In this section, we focus on these subsequential limits. Here is our first result.

Theorem 3.4 ()

The subsequential limits of a sequence (x_n) in a metric space form a closed subset of X .

Proof.

Let E be the set of subsequential limits and $y \in E'$ be a limit point of E . We must show that $y \in E$.^a We will construct a subsequence (x_{n_k}) that converges to y . Given any $\epsilon > 0$, we can see that the $B_{\epsilon/2}(y) \cap E \neq \emptyset$, so choose an element z_ϵ . Furthermore, z_ϵ means that it is a limit point of $\{x_n\}$, and so $B_{\epsilon/2}(z_\epsilon) \cap \{x_n\} \neq \emptyset$, call this $x(\epsilon)$. Therefore, by the triangle inequality,

$$|y - x(\epsilon)| \leq |y - z_\epsilon| + |z_\epsilon - x(\epsilon)| = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad (335)$$

and so we can take a point from the sequence (x_n) for every ϵ . Now we do this for $\epsilon = 1/n$, and choose n_k that is greater than its previous by restricting the sequence to that past n_{k-1} . Doing this gives a subsequence which converges to y . Therefore $y \in E$.

^aIntuitively, we can see that y is infinitesimally close to E , which consists of points infinitesimally close to $\{x_n\}$, and so y should be infinitesimally close to $\{x_n\}$.

3.2 Convergence in Reals

The properties of a general metric space give us some general conditions to determine whether a sequence converges in \mathbb{R}^n . To add to our toolbox in determining convergence, we will focus on sufficient conditions for convergence in \mathbb{R} , which has the additional properties of being an ordered field. These additional structures unlock a whole new suite of theorems in convergence. Why do we want to focus on just real-valued, i.e. numerical, sequences? First is that since \mathbb{R}^n is constructed as the product topology of \mathbb{R} , we can prove a lot about continuity of functions in \mathbb{R}^n by proving limits in \mathbb{R} , and letting the construction of the product

topology do the rest. Second, the codomain of many natural structures such as inner products, norms, and measures lie in the reals, and we often need to prove convergence of these values.

3.2.1 Completeness

So far, in order to show that a sequence is convergent, we must identify a real number first and then show using the ϵ - δ definition that it converges. This might be overkill in a case where we just want to prove that a sequence converges, but we don't care what it converges to. Unsurprisingly, we use the fact that the sequence lives in the reals. We can determine convergence by using Cauchy-completeness, which gives us the "theorem" (though it is really a fact by construction).

Theorem 3.5 (Cauchy-Convergence Criterion)

A cauchy sequence in \mathbb{R} converges.

The second result is an immediate consequence of Dedekind completeness, which is equivalent to Cauchy completeness in the reals.

Definition 3.7 (Monotonic Sequences)

Let X be an ordered set. $\{x_n\}$ is

1. **increasing** if $x_{n+1} > x_n$ for all n .
2. **decreasing** if $x_{n+1} < x_n$ for all n .
3. **nondecreasing** if $x_{n+1} \geq x_n$ for all n .
4. **nonincreasing** if $x_{n+1} \leq x_n$ for all n .

Sequences of these types are called **monotonic**.

Lemma 3.3 (Convergence Criterion for Monotonic Sequences)

In order for a nondecreasing (nonincreasing) sequence to be convergent, it is necessary and sufficient that it is bounded above (or below).

Proof.

It satisfies to prove the first case, as the second case can be done similarly without much difficulty. Let $x_n \leq x_{n+1}$. Then the set $\{x_n\}$ is bounded above in \mathbb{R} , which has the least upper bound property, and so there exists a least upper bound x . We claim that the sequence converges to x . For every $\epsilon > 0$, since it is least, there exists at least one $x_N \in (x - \epsilon, x)$. By monotonicity, this means that $x_n \in (x - \epsilon, x)$ for all $n \geq N$, and so the sequence converges to x .

3.2.2 Properties of Order

We are able to see how both Cauchy and Dedekind completeness of the reals define convergence in \mathbb{R} . Now let's squeeze a bit more out of the total ordering to gain some properties of convergence and divergence.

Theorem 3.6 (Preservation of Ordering Between Sequences and Limits)

Given convergent sequences $\{x_n\}$ and $\{y_n\}$, if

$$\lim_{n \rightarrow \infty} x_n < \lim_{n \rightarrow \infty} y_n \tag{336}$$

then there exists an index $N \in \mathbb{N}$ such that $x_n < y_n$ for all $n > N$.

Proof.

Given $x_n \rightarrow x, y_n \rightarrow y$ and $x < y (\in \mathbb{R})$, then for every $\epsilon > 0$, there exists $N_1, N_2 \in \mathbb{N}$ s.t. $d(x, x_n) < \epsilon$ for all $n > N_1$ and $d(y, y_n) < \epsilon$ for all $n > N_2$. Setting $N = \max\{N_1, N_2\}$, we can say the same for all $n > N$. We choose $\epsilon = \frac{y-x}{2} > 0$. Then, there exists $N \in \mathbb{N}$ s.t. $x_n \in (x - \epsilon, x + \epsilon)$ and $y_n \in (y - \epsilon, y + \epsilon)$ for all $n > N$. Therefore, if $a \in (x - \epsilon, x + \epsilon)$ and $b \in (y - \epsilon, y + \epsilon)$, then

$$a < \sup B_\epsilon(x) = x + \epsilon = y - \epsilon = \inf B_\epsilon(y) < b$$

which implies that $x_n < y_n$ for all $n > N$.

Theorem 3.7 (Squeeze Theorem for Sequences)

Given sequences $\{x_n\}, \{y_n\}, \{z_n\}$ such that

$$x_n \leq y_n \leq z_n$$

for all $n > N$, if $\{x_n\}$ and $\{z_n\}$ both converge to the same limit, then the sequence $\{y_n\}$ also converges to that limit. That is,

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} z_n = A \implies \lim_{n \rightarrow \infty} y_n = A$$

Proof.

We first prove that if there exists a $N \in \mathbb{N}$ s.t. $a_n \leq b_n$ for all $n > N$, then $\lim_{n \rightarrow \infty} a_n = a \leq b = \lim_{n \rightarrow \infty} b_n$. Assume this weren't true, that $a > b$. Then for $\epsilon = \frac{a-b}{2} > 0$, there must exist $M \in \mathbb{N}$ s.t. $a_n \in (a - \epsilon, a + \epsilon)$ and $b_n \in (b - \epsilon, b + \epsilon)$ for all $n > M$. But

$$b_n < \sup(b - \epsilon, b + \epsilon) = b + \epsilon = a - \epsilon = \inf(a - \epsilon, a + \epsilon) < a_n \quad (337)$$

which contradicts $a_n \leq b_n$. Therefore, $a \leq b$. Therefore, we can use this to get

$$A = \lim_{n \rightarrow \infty} x_n \leq \lim_{n \rightarrow \infty} y_n \leq \lim_{n \rightarrow \infty} z_n = A \implies \lim_{n \rightarrow \infty} y_n = A \quad (338)$$

Note that while a convergent sequence can be visualized quite easily by the Cauchy convergence criterion, there are many ways in which a sequence can be divergent.

1. Increasing/decreasing indefinitely
2. Oscillating between two constant values
3. Oscillating between a value tending to $+\infty$ and a value tending to $-\infty$
4. Many other classes of divergence

Definition 3.8 (Sequence Tending to Infinity)

The sequence $\{x_n\}$ **tends to positive infinity** if for each number c there exists $N \in \mathbb{N}$ such that $x_n > c$ for all $n > N$. It is denoted

$$x_n \rightarrow +\infty \text{ or } \lim_{n \rightarrow \infty} x_n = +\infty \quad (339)$$

We define sequences that **tend to negative infinity** similarly. And $\{x_n\}$ **tends to infinity** if for each c there exists $N \in \mathbb{N}$ such that $|x_n| > c$ for all $n > N$, which is written

$$x_n \rightarrow \infty \quad (340)$$

Note that

$$x_n \rightarrow +\infty \text{ or } x_n \rightarrow -\infty \implies x_n \rightarrow \infty \quad (341)$$

but the converse is not necessarily true. The simple example is the sequence $x_n = (-1)^n n$. Also, it is important to know that a sequence may be unbounded and yet not tend to $+\infty$, $-\infty$, or ∞ .

Example 3.1 (Unbounded Sequence that Doesn't tend to ∞)

The sequence $x_n = n^{(-1)^n}$ is divergent yet does not tend to positive infinity, negative infinity, nor infinity.

However, if a sequence is unbounded we can find a convergent subsequence that does tend to infinity.

Theorem 3.8 ()

If (x_n) is not bounded above then it has a subsequence $x_{n_k} \rightarrow +\infty$.

Proof.

We can construct such a subsequence.

Therefore, we can construct a subsequential limit to $\pm\infty$ if (x_n) is not bounded. If it is bounded, then by the Bolzano-Weierstrass theorem it contains a convergent subsequence. Therefore, we have the following.

Corollary 3.2 ()

From each sequence of real numbers there exists either a convergent subsequence or a subsequence tending to infinity.

Example 3.2 ()

We claim that

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1 \quad (342)$$

3.3 Arithmetic

Theorem 3.9 (Arithmetic on Limits)

Given that $\{x_n\}, \{y_n\}$ are numerical sequences with $y_n \neq 0$ for all n , and let

$$\lim_{n \rightarrow \infty} x_n = A, \quad \lim_{n \rightarrow \infty} y_n = B \neq 0$$

then,

$$\begin{aligned} \lim_{n \rightarrow \infty} (x_n + y_n) &= A + B \\ \lim_{n \rightarrow \infty} (cx_n) &= cA \\ \lim_{n \rightarrow \infty} (x_n \cdot y_n) &= A \cdot B \\ \lim_{n \rightarrow \infty} \frac{x_n}{y_n} &= \frac{A}{B} \end{aligned}$$

It immediately follows that the set of all convergent sequences in $\mathbb{R}^{\mathbb{N}}$ is a subspace of $\mathbb{R}^{\mathbb{N}}$.

Proof.

Assume that

$$\lim_{n \rightarrow \infty} x_n = A \text{ and } \lim_{n \rightarrow \infty} y_n = B \neq 0$$

This means that for every $\epsilon > 0$, there exists $N_1, N_2 \in \mathbb{N}$ such that

$$\begin{aligned} |x_n - A| &< \epsilon \text{ for all } n > N_1 \\ |y_n - B| &< \epsilon \text{ for all } n > N_2 \end{aligned}$$

Therefore, for a given ϵ , we wish to prove that there exists a N such that for all $n > N$,

1. $|x_n + y_n - (A + B)| < \epsilon$
2. $|cx_n - cA| < \epsilon$
3. $|(x_n y_n) - (AB)| < \epsilon$
4. $\left| \frac{x_n}{y_n} - \frac{A}{B} \right| < \epsilon$

1. By the triangle inequality, we can see that

$$|(x_n + y_n) - (A + B)| = |x_n - A| + |y_n - B|$$

Since we can choose the error between x_n and A for $n > N_1$, and y_n and B for $n > N_2$ as small as we want, we set it to $\epsilon/2$. Then, we have

$$|(x_n + y_n) - (A + B)| = |x_n - A| + |y_n - B| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for all $n > N = \max\{N_1, N_2\}$. Therefore, for a given ϵ , there exists an N such that

$$|(x_n + y_n) - (A + B)| < \epsilon \text{ for all } n > N$$

2. This proof is easy. For a given ϵ , we choose the error to be $\frac{\epsilon}{c}$.

$$|x_n - A| < \frac{\epsilon}{c} \text{ for all } n > N_1$$

Then, there exists natural number N_1 such that

$$|cx_n - cA| < c|x_n - A| = c \frac{\epsilon}{c} = \epsilon \text{ for all } n > N_1$$

3. We first observe that since the limit of $\{y_n\}$ exists, it must be bounded by a value, say B . That is,

$$|y_n| < Y \text{ for all } n \in \mathbb{N}$$

Then, we see that

$$\begin{aligned} |x_n y_n - AB| &= |(x_n y_n - Ay_n) + (Ay_n - AB)| \\ &< |x_n y_n - Ay_n| + |Ay_n - AB| \\ &= |y_n||x_n - A| + |A||y_n - B| \end{aligned}$$

Suppose $\epsilon > 0$ is given. Then, we can set the error bounds freely; there exists $N_1, N_2 \in \mathbb{N}$ such that

$$\begin{aligned} |x_n - A| &< \frac{\epsilon}{2Y} \text{ for all } n > N_1 \\ |y_n - B| &< \frac{\epsilon}{2|A|} \text{ for all } n > N_2 \end{aligned}$$

Then, we can see that

$$|x_n y_n - AB| \leq |y_n||x_n - A| + |A||y_n - B| < Y \cdot \frac{\epsilon}{2Y} + |A| \frac{\epsilon}{2|A|} = \epsilon$$

for all $n > N = \max\{N_1, N_2\}$.

4. We use the estimate

$$\left| \frac{A}{B} - \frac{x_n}{y_n} \right| = \frac{|x_n||y_n - B| + |y_n||x_n - A|}{y_n^2} \cdot \frac{1}{1 - \delta(y_n)}, \quad \delta(y_n) = \frac{|y_n - B|}{|y_n|}$$

For a given $\epsilon > 0$, we find natural numbers N_1, N_2 such that

$$\begin{aligned} |x_n - A| &< \min \left\{ 1, \frac{\epsilon|B|}{8} \right\} \text{ for all } n > N_1 \\ |y_n - B| &< \min \left\{ \frac{|B|}{4}, \frac{\epsilon B^2}{16(|A|+1)} \right\} \text{ for all } n > N_2 \end{aligned}$$

From this we can deduce that

$$|x_n| = |x_n - A + A| < |x_n - A| + |A| < |A| + 1$$

and

$$\begin{aligned} |B| &= |y_n + B - y_n| < |y_n| + |B - y_n| \\ \implies |y_n| &> |B| - |y_n - B| > |B| - \frac{|B|}{4} > \frac{|B|}{2} \\ \implies \frac{1}{|y_n|} &< \frac{2}{|B|} \\ \implies 0 < \delta(y_n) &= \frac{|y_n - B|}{|y_n|} < \frac{|B|/4}{|B|/2} = \frac{1}{2} \\ \implies 1 - \delta(y_n) &> \frac{1}{2} \\ \implies 0 < \frac{1}{1 - \delta(y_n)} &< 2 \end{aligned}$$

So, we can substitute

$$\begin{aligned} |x_n| \cdot \frac{1}{y_n^2} \cdot |y_n - B| &< (|A| + 1) \cdot \frac{4}{B^2} \cdot \frac{\epsilon \cdot B^2}{16(|A| + 1)} = \frac{\epsilon}{4} \\ \left| \frac{1}{y_n} \right| \cdot |x_n - A| &< \frac{2}{|B|} \cdot \frac{\epsilon|B|}{8} = \frac{\epsilon}{4} \end{aligned}$$

into the final equation to get

$$\left| \frac{A}{B} - \frac{x_n}{y_n} \right| < \epsilon \text{ for all } n > N = \max\{N_1, N_2\}$$

Example 3.3 ()

We claim that

$$\lim_{n \rightarrow \infty} \frac{n}{q^n} = 0 \text{ if } q > 1$$

Since $x_n = \frac{n}{q^n} \implies x_{n+1} = \frac{n+1}{nq}x_n$ for $n \in \mathbb{N}$. Since

$$\lim_{n \rightarrow \infty} \frac{n+1}{nq} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right) \frac{1}{q} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right) \cdot \lim_{n \rightarrow \infty} \frac{1}{q} = 1 \cdot \frac{1}{q} = \frac{1}{q} < 1$$

there exists an index N such that $\frac{n+1}{nq} < 1$ for $n > N$. Thus, we have

$$x_n > x_{n+1} = x_n \cdot \frac{n+1}{nq} \text{ for } n > N$$

which means that the sequence will be monotonically decreasing from index N on. The terms of the sequence

$$x_{N+1} > x_{N+2} > x_{N+3} > \dots$$

are positive (bounded below) and are monotonically decreasing, so it must have a limit.

Finding the actual limit is easy. Let $x = \lim_{n \rightarrow \infty} x_n$. It follows from the relation $x_{n+1} = \frac{n+1}{nq}x_n$ that

$$x = \lim_{n \rightarrow \infty} (x_{n+1}) = \lim_{n \rightarrow \infty} \left(\frac{n+1}{nq} x_n \right) = \lim_{n \rightarrow \infty} \frac{n+1}{nq} \cdot \lim_{n \rightarrow \infty} x_n = \frac{1}{q}x$$

which implies that $(1 - \frac{1}{q}) = 0 \implies x = 0$.

Definition 3.9 (Cauchy Product)

The Cauchy Product is the direct convolution of two sequences.

Definition 3.10 (Recursive Sequence)

Sometimes, a sequence may be defined **recursively**, where the n th term contains a combination of the $n - 1$ terms before it.

3.4 Limsup and Liminf

The superior and inferior limits represent some sort of "bound" on the sequence in the long run. That is, on the long run, the terms of the sequence (x_n) cannot be greater than its superior limit and cannot be less than its inferior limit. With this interpretation, the following definition should be clear.

Definition 3.11 (Inferior, Superior Limits)

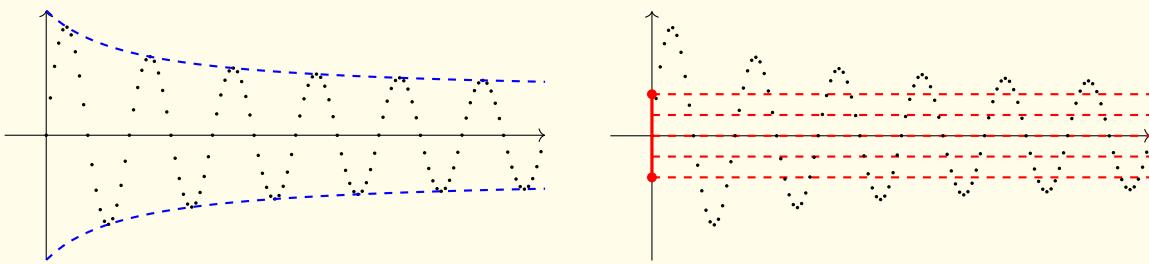
The **superior/inferior limit** of a sequence (x_n) is defined in the equivalent ways.

1. Given that E is the set of all partial limits, the limsup/liminf is the supremum/infimum of E .

$$\limsup_{n \rightarrow \infty} x_n := \sup\{E\} \quad \liminf_{n \rightarrow \infty} x_n := \inf\{E\} \quad (343)$$

2. The limsup/liminf is the limit of the sequence of supremums/infimums of the elements up to k .

$$\limsup_{n \rightarrow \infty} x_n := \lim_{n \rightarrow \infty} \sup_{k \geq n} x_k \quad \limsup_{n \rightarrow \infty} x_n := \lim_{n \rightarrow \infty} \inf_{k \geq n} x_k \quad (344)$$



(a) In order to find the limsup, we first look at the whole sequence in \mathbb{N} and find the supremum. We now "decrease" our domain from \mathbb{N} to $\{2, \dots\}$, then $\{3, \dots\}$, then $\{4, \dots\}$ and so on, continuing to label the supremum of the sequence. The limit of this sequence of suprema is the limsup.

(b) The 5 red lines marked in the middle (along with infinitely many others) are viable partial limits because one can choose a subsequence such that all of its points after a certain n lie in some ϵ -neighborhood of the limit. Therefore, we claim that the limsup/inf is the supremum of this set E .

Figure 2: Two ways to visualize the superior and inferior limits of the divergent sequence $x_n = (\frac{1}{x+1} + 0.5) \sin(2\pi x)$. The left is the limit of the supremum, and the right is the supremum of the closed set of subsequential limits.

Example 3.4 (Computing Limsup and Liminf)

We give some basic examples.

- Let $x_n = (-1)^n$. Then $E = \{-1, +1\}$ and

$$\limsup_{n \rightarrow \infty} x_n = 1, \quad \liminf_{n \rightarrow \infty} x_n = -1 \quad (345)$$

- Let $x_n = (-1)^n/[1 + (1/n)]$. Then

$$\limsup_{n \rightarrow \infty} x_n = 1, \quad \liminf_{n \rightarrow \infty} x_n = -1 \quad (346)$$

Let's give two warnings. First, limsup and liminf do *not* behave like limits under addition and multiplication. That is,

$$\limsup x_n + \limsup y_n \neq \limsup x_n + y_n \quad (347)$$

Example 3.5 (Counterexamples of Arithmetic Consistency of Limit superior)

Consider $(x_n) = (-1)^n$ and $y_n = (-1)^{n+1}$. Then

$$\limsup x_n = \limsup y_n = 1, \quad \liminf x_n = \liminf y_n = -1 \quad (348)$$

But $(x_n + y_n) = 0$, so

$$\limsup x_n + y_n = \liminf x_n + y_n = 0 \quad (349)$$

Second, note that even though we are talking about subsequential limits, the limsup and liminf are *not* subsequential limits! It is the supremum of subsequential limits E , which may or may not be in E .

Example 3.6 (Limsup that is not attained by any subsequential limit)

This should be a sequence not in \mathbb{R} .

However, in \mathbb{R} , it turns out that the limsup and liminf are both contained in E , so we are fine.

Lemma 3.4 ()

If (x_n) is a sequence in \mathbb{R} , then

1. the limsup is indeed a subsequential limit, i.e. $\limsup x_n \in E$.
2. If $x > \limsup x_n$, then $\exists N \in \mathbb{N}$ s.t. $n \geq N \implies x_n < x$.

Proof.

For the first claim, there are two cases to consider. If (x_n) is unbounded from above, then $\exists (x_{n_k})$ such that $x_{n_k} \rightarrow +\infty \implies \infty = \limsup x_n \in E$. If (x_n) is bounded from above, then the subsequential limits of (x_n) are either in (x_n) or they are limit points of x_n . This implies that the set E consists of points either in $\{x_n\}$ or are limit points of the set $\{x_n\} \implies \sup E$ is in E since it's a limit point. For the second claim, if there are infinitely many terms of the sequence larger than x , then we could find a subsequence (x_{n_k}) with $x_{n_k} > x$ for all k . Therefore (x_n) has a subsequential limit which must be $\geq x$. Every subsequential limit of (x_{n_k}) is also a subsequential limit of (x_n) . This contradicts $\limsup x_n = \sup E$.

Theorem 3.10 (Requirements of Partial Limits for Limit to Exist)

Here are two results in which we can use partial limits to determine if a sequence has a limit or not.

1. A sequence has a limit or tends to $\pm\infty$ if and only if its inferior and superior limits are the same.

$$\limsup x_n = \liminf x_n = x \implies \lim_{n \rightarrow +\infty} x_n = x \quad (350)$$

2. A sequence converges if and only if every subsequence of it converges.

Proof.

For (1), we pick $x + \epsilon > x$. Then every term past some N_1 must be less than $x + \epsilon$. By the same logic, we have N_2 for $x - \epsilon < x$. So take $N = \max\{N_1, N_2\}$, which is contained in the ϵ -ball around x .

Theorem 3.11 (Ordering on Subsequential Limits)

If $s_n \leq t_n$ for $n \geq N$, where N is fixed, then

$$\begin{aligned} \liminf_{n \rightarrow \infty} s_n &\leq \liminf_{n \rightarrow \infty} t_n \\ \limsup_{n \rightarrow \infty} s_n &\leq \liminf_{n \rightarrow \infty} t_n \end{aligned}$$

Example 3.7 ()

We claim

$$\lim_{n \rightarrow \infty} n^{1/n} = 1 \quad (351)$$

We can consider $x_n = n^{1/n} - 1$ and want to show that $x_n \rightarrow 0$. We have $x_n \geq 0$. If $n > 1$, then $n = (x_n + 1)^n \geq x_n^2 \cdot \frac{n(n-1)}{2}$ from the binomial theorem. This means that

$$x_n^2 \leq \frac{2}{n-1} \implies 0 \leq x_n \leq \sqrt{\frac{2}{n-1}} \rightarrow 0 \quad (352)$$

And so by the squeeze theorem, $x_n \rightarrow 0$.

Example 3.8 ()

If $x > 1, \alpha \in \mathbb{R}$, then

$$\lim_{n \rightarrow +\infty} \frac{n^\alpha}{x^n} = 0 \quad (353)$$

3.5 Convergence Tests for Real Series

Definition 3.12 (Series over \mathbb{R})

Given a sequence of real numbers (x_n) , the **series (of partial sums)** is the sequence

$$(s_n) = \sum_{k=1}^n x_k \quad (354)$$

The **sum of the series** is the limit of (s_n) . Usually we define (s_n) implicitly and use the summation notation.

$$\sum_{n=1}^{\infty} x_n := \lim_{n \rightarrow \infty} s_n \quad (355)$$

1. If the sequence (s_n) converges to s , the series is **convergent**, written

$$\sum x_n < +\infty \quad (356)$$

2. If the sequence does not converge, it is **divergent**.
3. If the series of partial sums of $(|x_n|)$ converges, then it is said to be **absolutely convergent**.^a

$$\sum_{n=1}^{\infty} |x_n| \quad (357)$$

^aClearly, every absolutely convergent series because $|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|$.

We must reiterate a few warnings here. Note that the series $\sum x_n$ is simply notation and should *not* be treated as an “infinite sum.” Such a thing does not exist for algebraic structures which have finary operations. More specifically, given a series, we cannot in general split nor combine series, and we cannot reindex nor rearrange (an infinite number of) terms. However, we can manipulate each term for a fixed index.

Example 3.9 (Disasters of Reindexing and Rearranging)

Let us take the series $\sum 0$. We clearly know that the corresponding sequence of partial sums $0, 0, \dots$

is convergent to 0. But if we do this series of steps.

$$\begin{aligned}
 \sum_{n=1}^{\infty} 0 &= \sum_{n=1}^{\infty} n - n && \text{(Can manipulate terms)} \\
 &= \sum_{n=1}^{\infty} n - \sum_{n=1}^{\infty} n && \text{(Cannot split series)} \\
 &= 1 + \sum_{n=2}^{\infty} n - \sum_{n=1}^{\infty} n && \text{(Can take 1st term out)} \\
 &= 1 + \sum_{n=1}^{\infty} (n+1) - \sum_{n=1}^{\infty} n && \text{(Cannot reindexing)} \\
 &= 1 + \sum_{n=1}^{\infty} (n+1) - n && \text{(Cannot combine series)} \\
 &= 1 + \sum_{n=1}^{\infty} 1 && \text{(Can manipulate terms)} \\
 &= 1 + \infty = +\infty && \tag{358}
 \end{aligned}$$

The wrong steps show that the series is divergent.

We have seen the consequences of these mistakes that beginners make and are often on popular media. However, note that we can always do splitting, combining, reindexing, and rearranging for *finite sums*, which are algebraically defined. Later on, we will show that some of these operations are allowed for series that we know are convergent.

Since the convergence of a series is equivalent to convergence of its sequence of partial sums, applying the Cauchy convergence criterion to the sequence $\{s_n\}$ leads to the following theorem.

Theorem 3.12 (Cauchy Convergence Criterion for Series)

The series $a_1 + \dots + a_n + \dots$ converges if and only if for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $m \geq n > N$,

$$|a_n + \dots + a_m| < \epsilon \tag{359}$$

Corollary 3.3 (nth Term Test)

A necessary (but not sufficient) condition for convergence of the series $a_1 + \dots + a_n + \dots$ is that the terms tend to 0 as $n \rightarrow \infty$. That is, it is necessary that

$$\lim_{n \rightarrow \infty} a_n = 0 \tag{360}$$

Proof.

It suffices to set $m = n$ in the Cauchy convergence criterion. This would mean that for every $\epsilon > 0$ there exists a $N \in \mathbb{N}$ such that

$$|a_n| = |a_n - 0| < \epsilon \text{ for all } n > N \tag{361}$$

which, by definition, means that $\{a_n\}$ converges to 0.

Nothing so far is really surprising here. The Cauchy convergence criterion really just follows from the definition

of Cauchy completeness, and the n th term test is pretty trivial. The way that we will build up convergence tests is by proving some special cases of convergence and then using the direct comparison test to then classify further series.

Example 3.10 (Telescoping Series)

A **telescoping series** is a series in which the partial sums can cancel out. An example is the series of partial sums of the sequence $(x_n) = \frac{1}{n(n+1)}$. In here, the series term is

$$s_n = \sum_{k=1}^n \frac{1}{k(k+1)} \quad (362)$$

$$= \sum_{k=1}^n \frac{1}{k} - \frac{1}{k+1} \quad (363)$$

$$= \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k+1} \quad (364)$$

$$= \sum_{k=1}^n \frac{1}{k} - \sum_{k=2}^{n+1} \frac{1}{k} \quad (365)$$

$$= \frac{1}{1} + \left(\sum_{k=2}^n \frac{1}{k} \right) - \left(\sum_{k=2}^n \frac{1}{k} \right) - \frac{1}{n+1} \quad (366)$$

$$= 1 + \left(\sum_{k=2}^n \frac{1}{k} - \frac{1}{k} \right) - \frac{1}{n+1} \quad (367)$$

$$= 1 - \left(\sum_{k=2}^n 0 \right) - \frac{1}{n+1} \quad (368)$$

$$= 1 - \frac{1}{n+1} \quad (369)$$

Note that all of the examples that we have done here are for finite sums, so they are all legal.

Example 3.11 (Geometric Series)

The series $\sum_{n=0}^{\infty} q^n$ is called a **geometric series**.

$$1 + q + q^2 + \dots + q^n + \dots \quad (370)$$

is called the **geometric series**. We can see that

1. $|q| \geq 1 \iff \sum q^n$ is divergent. $|q| \geq 1 \implies |q|^n \geq 1$, and so the terms q^n does not converge to 0, and the n th term test is not met.
2. $|q| < 1 \iff \sum q^n$ is convergent. We can use the identity

$$s_n = 1 + q + \dots + q^{n-1} = \frac{1 - q^n}{1 - q} \implies \lim_{n \rightarrow \infty} \frac{1 - q^n}{1 - q} = \frac{1}{1 - q} \quad (371)$$

since $\lim_{n \rightarrow \infty} q^n = 0$ if $|q| < 1$.

The Cauchy convergence criterion can be used to prove the direct comparison test.

Theorem 3.13 (Direct Comparison Test)

For some fixed N , if

1. If $|x_n| \leq y_n$ for all $n \geq N$ and $\sum y_n$ converges, then $\sum x_n$ converges.
2. If $x_n \geq y_n \geq 0$ for all $n \geq N$ and $\sum y_n$ diverges, then $\sum x_n$ diverges.

Example 3.12 (Comparison with Telescoping Series)

We can prove the special case a geometric series with the direct comparison test. We claim that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is finite. We can see that

$$\frac{1}{n^2} \leq \frac{2}{n(n+1)} \quad (372)$$

where the series of the terms in the RHS is telescoping and therefore converges. So by the direct comparison test, $\sum \frac{1}{n^2}$ converges.

Now we prove another corollary of the Cauchy convergence criterion.

Theorem 3.14 (Cauchy Condensation Test)

If $a_1 \geq a_2 \geq \dots \geq 0$, the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the series

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \dots \quad (373)$$

converges.

Proof.

Letting $A_k = a_1 + a_2 + \dots + a_k$ and $S_n = a_1 + 2a_2 + \dots + 2^n a_{2^n}$, it is clear that by adding up the inequalities

$$\begin{aligned} a_2 &\leq a_2 \leq a_1 \\ 2a_4 &\leq a_3 + a_4 \leq 2a_2 \\ 4a_8 &\leq a_5 + a_6 + a_7 + a_8 \leq 4a_4 \\ &\dots \\ 2^n a_{2^{n+1}} &\leq a_{2^n+1} + \dots + a_{2^{n+1}} \leq 2^n a_{2^n}, \end{aligned}$$

we get

$$\frac{1}{2}(S_{n+1} - a_1) \leq A_{2^{n+1}} - a_1 \leq S_n \quad (374)$$

Since the sequences $\{A_k\}$ and $\{S_k\}$ are nondecreasing, and hence from the inequalities we can conclude that they are either both bounded above (which means that they are both convergent since it is a bounded, nondecreasing series) or both unbounded above (which means that they are both divergent since they are nondecreasing and unbounded).

Corollary 3.4 (p-series Test)

The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \quad (375)$$

converges for $p > 1$ and diverges for $p \leq 1$.^a

^aThis sort of reminds you of u -substitution. For example, look at $\int_1^\infty f(t) dt = \int_0^\infty e^u f(e^u) du$, where the convergence of LHS \iff convergence of RHS.

Proof.

Suppose $p \geq 0$. By the previous theorem, the series converges or diverges simultaneously with the series

$$\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=0}^{\infty} (2^{1-p})^k \quad (376)$$

which is really just a geometric series. A necessary and sufficient condition for the convergence of this series is that $2^{1-p} < 1$, that is, $p > 1$.

Now suppose $p \leq 0$. The series is then clearly divergent since all of the terms are larger than 1.

Example 3.13 (Harmonic Series)

The **harmonic series**

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots \quad (377)$$

seems at first glance to be converging since the terms converge to 0. However, it does not pass the Cauchy condensation test since

$$\sum_{n=1}^{\infty} 2^n x_n = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \sum_{n=1}^{\infty} 1 = +\infty \quad (378)$$

As you can see, this increases logarithmically, so in early calculators it was hard to numerically detect divergence (you would have to double the number of series terms to get a linear increase).

3.6 Euler's Number and Irrationality

Definition 3.13 (Euler's Number)

We define **Euler's number** as

$$e := \sum_{n=0}^{\infty} \frac{1}{n!} \quad (379)$$

The first thing we should do is show that it converges, this is a one-liner.

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \sum_{n=2}^{\infty} \frac{1}{n!} \leq 2 + \sum_{n=2}^{\infty} \frac{1}{n(n-1)} \quad (380)$$

Theorem 3.15 (Euler's Number as a Limit)

We have

$$\lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^n = e \quad (381)$$

Proof.

Let us define the sequence

$$t_n = \sum_{k=0}^n \frac{1}{k!}, \quad s_n = \left(1 + \frac{1}{n}\right)^n \quad (382)$$

We know that $t_n \rightarrow e$, and we want to show that $s_n \rightarrow e$. We do this with the squeeze theorem.

1. We can see that

$$s_n = \left(1 + \frac{1}{n}\right)^n \quad (383)$$

$$= \sum_{k=0}^n \binom{n}{k} 1^{n-k} \left(\frac{1}{n}\right)^k \quad (384)$$

$$= 1 + 1 + \frac{n(n-1)}{2!} \frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!} \frac{1}{n^3} + \dots \quad (385)$$

$$= 1 + 1 + \frac{1}{2!}(1) \left(1 - \frac{1}{n}\right) + \frac{1}{3!}(1) \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) + \dots + \frac{1}{n!}(1) \prod_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \quad (386)$$

$$\leq \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} = t_n \quad (387)$$

and so $s_n < t_n \implies \limsup s_n \leq \limsup t_n = e$.

2. Let $m \leq n$ be fixed. Then,

$$s_n \geq 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \dots + \frac{1}{m!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \quad (388)$$

since we are just taking the first m positive terms of the element. Therefore, letting $n \rightarrow +\infty$ and keeping m fixed, we get

$$\liminf_{n \rightarrow +\infty} s_n \geq 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{n!} \text{ for all } m \in \mathbb{N} \quad (389)$$

which implies $\liminf s_n \geq t_m$ for all $m \in \mathbb{N}$, and now letting $m \rightarrow +\infty$, we have $\liminf s_n \geq \liminf t_m = e$.

Now we prove the irrationality of e . It is usually extremely difficult to prove that an arbitrary number is irrational, e.g. π^e or π^{e^e} .

Theorem 3.16 (e is Irrational)

e is irrational.

Proof.

Letting $t_n = \sum_{k=0}^n \frac{1}{k!}$, we have

$$e - t_n = \sum_{k=n+1}^{\infty} \frac{1}{k!} \quad (390)$$

$$= \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+3)(n+2)} + \dots \right) \quad (391)$$

$$< \frac{1}{(n+1)!} \underbrace{\left(1 + \frac{1}{n+2} + \frac{1}{(n+2)^2} + \dots \right)}_{\text{geometric}} \quad (392)$$

$$= \frac{1}{(n+1)!} \left(\frac{1}{1 - (1/(n+2)!)^2} \right) \quad (393)$$

$$= \frac{1}{n!n} \cdot \underbrace{\frac{(n+2)n}{(n+1)^2}}_{<1} \quad (394)$$

$$= \frac{1}{n!n} \quad (395)$$

Note that we can combine and split sums since we know that e is convergent. Now suppose that $e = p/q$. Then,

$$0 < q!(e - t_q) < \frac{1}{q} \quad (396)$$

But $q!e$ is an integer and $q!t_q$ is also an integer. So we have $q! \cdot \frac{p}{q}$, an integer, between 0 and 1, which is a contradiction.

3.7 Ratio and Root Tests

Now we introduce the root and ratio tests, which are derived by the comparison test with a geometric series. The ratio test is used more day-to-day, but not as decisive as the root test. Both tests have a similar flavor.

Theorem 3.17 (Ratio Test)

Suppose the limit $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \alpha$ exists for the series $\sum_{n=1}^{\infty} a_n$. Then,

1. $\alpha < 1 \implies \sum a_n$ converges absolutely.
2. $\alpha > 1 \implies \sum a_n$ diverges.
3. $\alpha = 1 \implies \sum a_n$ is inconclusive.

Alternatively, if

1. $\limsup |a_{n+1}/a_n| = \alpha < 1$, then $\sum a_n$ converges
2. If $\exists N$ s.t. $|a_{n+1}/a_n| \geq 1$ for all $n \geq N$, then $\sum a_n$ diverges.

Proof.

Since $\limsup \left| \frac{a_{n+1}}{a_n} \right| = \alpha < 1$, fix any $\alpha < \beta < 1$. Then $\exists N$ s.t. if $n > N$, $|a_{n+1}/a_n| < \beta$. So $|a_{N+1}| < \beta|a_N| \implies |a_{N+2}| < \beta^2|a_N|$. So letting $C = |a_N|$, for all $m \geq N$,

$$|a_m| \leq \frac{C}{\beta^N} \beta^m \implies |a_m| \leq \tilde{C} \beta^m \text{ for all } m \geq N \quad (397)$$

So $\sum a_n$ converges by comparison test since $\sum \beta^m < \infty$ when $\beta < 1$.

Theorem 3.18 (Root Test)

Let $\sum_{n=1}^{\infty} a_n$ be a given series and

$$\alpha = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \quad (398)$$

Then,

1. $\alpha < 1 \implies \sum a_n$ converges.
2. $\alpha > 1 \implies \sum a_n$ diverges.
3. $\alpha = 1 \implies \sum a_n$ is inconclusive.

Proof.

Listed.

1. If $\limsup \sqrt[n]{|a_n|} = \alpha < 1$, take any $\beta < \alpha < 1$. Then $\exists N \in \mathbb{N}$ s.t. if $n \geq N$, then $|a_n|^{1/n} < \beta \iff |a_n| < \beta^n$. Since $\beta < 1$, $\sum \beta^n < \infty$, and by comparison test, $\sum a_n$ converges.
2. Suppose $\alpha > 1$. Then $\limsup |a_n|^{1/n} = \alpha > 1$. So there exists a subsequence (a_{n_k}) s.t. $(|a_{n_k}|^{1/n_k}) \rightarrow \alpha > 1$. This means $\exists N$ s.t. for $n \geq N$, $|a_{n_k}|^{1/n_k} > 1 \implies |a_{n_k}| > 1$. But this fails the n th term test.
3. We do not claim anything and so there's nothing to prove.

Example 3.14 (Root Test Inconclusive Results)

Consider $\sum \frac{1}{n} = +\infty$, but from the root test

$$\sqrt[n]{\frac{1}{n}} \rightarrow 1, \text{ so } \alpha = 1 \quad (399)$$

Consider $\sum \frac{1}{n^2} < +\infty$, but from the root test

$$\sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{n^{1/n}} \right)^2 \rightarrow 1, \text{ so } \alpha = 1 \quad (400)$$

Example 3.15 ()

The sequence $\sum \frac{c^n}{n!}$ always converges for $c \in \mathbb{R}$.

Theorem 3.19 (Weierstrass M-test for Absolute Convergence)

Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be series. Suppose there exists an index $N \in \mathbb{N}$ such that $|a_n| \leq b_n$ for all $n > N$. Then,

$$\sum_{n=1}^{\infty} b_n \text{ converges} \implies \sum_{n=1}^{\infty} a_n \text{ converges absolutely} \quad (401)$$

We finally conclude by giving a theorem about the convergence of some special sequences.

Theorem 3.20 (Special Sequences)

Some special sequences:

1. If $p > 0$, then $\lim_{n \rightarrow \infty} \frac{1}{n^p} = 0$.
2. If $p > 0$, then $\lim_{n \rightarrow \infty} \sqrt[p]{n} = 1$.
3. $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

4. If $p > 0$ and α is real, then $\lim_{n \rightarrow \infty} \frac{n^\alpha}{(1+p)^n} = 0$.
5. If $|x| < 1$, then $\lim_{n \rightarrow \infty} x^n = 0$.

3.8 Exercises

Exercise 3.1 (Math 531 Spring 2025, PS4.6)

Consider the set of all bounded sequences of real numbers. That is, we consider sequences $\{x_n\}$ for which

$$\sup_{n \in \mathbb{N}} |x_n| \quad (402)$$

exists. For example, the sequence $\{1, 2, 3, \dots\}$ does not belong to the set, but the sequence $\{1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots\}$ does. Call this set X . Endow it with a metric:

$$d(\{x_n\}, \{y_n\}) = \sup_{n \in \mathbb{N}} |x_n - y_n|. \quad (403)$$

Explain why this is a metric. Make sure to explain why the supremum on the right hand side exists.

Exercise 3.2 (Math 531 Spring 2025, PS4.7)

Consider the metric space (X, d) from the previous problem. Is $\overline{B_1(\{0\})}$ a compact set? Here, $\{0\}$ is just the sequence of zeros: $\{0, 0, 0, 0, \dots\}$.

Exercise 3.3 (Rudin 3.1)

Prove that convergence of $\{x_n\}$ implies convergence of $\{|x_n|\}$. Is the converse true?

Solution 3.1

If $\{x_n\}$ converges to x , then for all $\epsilon > 0$, there exists a $N \in \mathbb{N}$ s.t. $|x_n - x| < \epsilon$ if $n > N$. We use the inequality $||x_n| - |x|| \leq |x_n - x|$ to show that then for every $\epsilon > 0$ there exists a $N \in \mathbb{N}$ s.t.

$$||x_n| - |x|| \leq |x_n - x| \leq \epsilon$$

and so $\{|x_n|\}$ converges to $|x|$.

Exercise 3.4 (Rudin 3.2)

Calculate

$$\lim_{n \rightarrow \infty} \sqrt{n^2 + n} - n$$

Solution 3.2

We can compute

$$\lim_{n \rightarrow \infty} \sqrt{n^2 + n} - n = \lim_{n \rightarrow \infty} (\sqrt{n^2 + n} - n) \cdot \frac{\sqrt{n^2 + n} + n}{\sqrt{n^2 + n} + n} = \lim_{n \rightarrow \infty} \frac{n}{\sqrt{n^2 + n} + n}$$

where

$$A_n = \frac{n}{\sqrt{n^2 + 2n + 1} + n} \leq \frac{n}{2n + 1} \leq \frac{n}{\sqrt{n^2 + n} + n} \leq \frac{n}{\sqrt{n^2 + n}} = \frac{n}{\sqrt{n^2}} = \frac{n}{2n} = \frac{1}{2} = C_n$$

C_n is ultimately constant. It suffices to prove that A_n limits to $\frac{1}{2}$ by showing that

$$\frac{n}{2n+1} = \frac{n/n}{(2n+1)/n} = \frac{1}{2 + \frac{1}{n}}$$

where $\{\frac{1}{n}\}$ is infinitesimal.

Exercise 3.5 (Rudin 3.3)

If $s_1 = \sqrt{2}$ and

$$s_{n+1} = \sqrt{2 + \sqrt{s_n}}$$

for $n = 1, 2, \dots$, prove that $\{s_n\}$ converges and that $s_n < 2$ for $n = 1, 2, \dots$

Solution 3.3

We can show that $s_n < 2$ by induction. $s_1 = \sqrt{2} < 2$, so the base case is proved. Now, given that $s_n < 2$, $\sqrt{s_n} < 2 \implies 2 + \sqrt{s_n} < 2 + \sqrt{2} < 4 \implies s_{n+1} = \sqrt{2 + \sqrt{s_n}} < 2$ and we are done.

Exercise 3.6 (Rudin 3.4)

Find the upper and lower limits of the sequence $\{s_n\}$ defined by

$$s_1 = 0; \quad s_{2m} = \frac{s_{2m-1}}{2}; \quad s_{2m+1} = \frac{1}{2} + s_{2m}.$$

Solution 3.4

Exercise 3.7 (Rudin 3.5)

For any two real sequences $\{a_n\}$, $\{b_n\}$, prove that

$$\limsup_{n \rightarrow \infty} (a_n + b_n) \leq \limsup_{n \rightarrow \infty} a_n + \limsup_{n \rightarrow \infty} b_n,$$

provided the sum on the right is not of the form $\infty - \infty$.

Solution 3.5

Exercise 3.8 (Rudin 3.6)

Investigate the behavior (convergence or divergence) of $\sum a_n$ if

- (a) $a_n = \sqrt{n+1} - \sqrt{n}$;
- (b) $a_n = (\sqrt{n+1} - \sqrt{n})/n$;
- (c) $a_n = (\sqrt[n]{n} - 1)^n$;
- (d) $a_n = \frac{1}{1+z^n}$, for complex values of z .

Solution 3.6**Exercise 3.9 (Rudin 3.7)**

Prove that the convergence of $\sum a_n$ implies the convergence of

$$\sum \frac{\sqrt{a_n}}{n},$$

if $a_n \geq 0$.

Solution 3.7**Exercise 3.10 (Rudin 3.8)**

If $\sum a_n$ converges, and if $\{b_n\}$ is monotonic and bounded, prove that $\sum a_n b_n$ converges.

Solution 3.8**Exercise 3.11 (Rudin 3.9)**

Find the radius of convergence of each of the following power series:

- (a) $\sum n^3 z^n$,
- (b) $\sum \frac{2^n}{n!} z^n$,
- (c) $\sum \frac{2^n}{n^2} z^n$,
- (d) $\sum \frac{n^3}{3^n} z^n$.

Solution 3.9**Exercise 3.12 (Rudin 3.10)**

Suppose that the coefficients of the power series $\sum a_n z^n$ are integers, infinitely many of which are distinct from zero. Prove that the radius of convergence is at most 1.

Solution 3.10**Exercise 3.13 (Rudin 3.11)**

Suppose $a_n > 0$, $s_n = a_1 + \dots + a_n$, and $\sum a_n$ diverges.

- (a) Prove that $\sum \frac{a_n}{1+a_n}$ diverges.
- (b) Prove that

$$\frac{a_{N+1}}{s_{N+1}} + \dots + \frac{a_{N+k}}{s_{N+k}} \geq 1 - \frac{s_N}{s_{N+k}}$$

and deduce that $\sum \frac{a_n}{s_n}$ diverges.

(c) Prove that

$$\frac{a_n}{s_n^2} \leq \frac{1}{s_{n-1}} - \frac{1}{s_n}$$

and deduce that $\sum \frac{a_n}{s_n^2}$ converges.

(d) What can be said about

$$\sum \frac{a_n}{1+na_n} \text{ and } \sum \frac{a_n}{1+n^2a_n}?$$

Solution 3.11

Exercise 3.14 (Rudin 3.12)

Suppose $a_n > 0$ and $\sum a_n$ converges. Put

$$r_n = \sum_{m=n}^{\infty} a_m.$$

(a) Prove that

$$\frac{a_m}{r_m} + \cdots + \frac{a_n}{r_n} > 1 - \frac{r_n}{r_m}$$

if $m < n$, and deduce that $\sum \frac{a_n}{r_n}$ diverges.

(b) Prove that

$$\frac{a_n}{\sqrt{r_n}} < 2(\sqrt{r_n} - \sqrt{r_{n+1}})$$

and deduce that $\sum \frac{a_n}{\sqrt{r_n}}$ converges.

Solution 3.12

Exercise 3.15 (Rudin 3.13)

Prove that the Cauchy product of two absolutely convergent series converges absolutely.

Solution 3.13

Exercise 3.16 (Rudin 3.14)

If $\{s_n\}$ is a complex sequence, define its arithmetic means σ_n by

$$\sigma_n = \frac{s_0 + s_1 + \cdots + s_n}{n+1} \quad (n = 0, 1, 2, \dots).$$

(a) If $\lim s_n = s$, prove that $\lim \sigma_n = s$.

(b) Construct a sequence $\{s_n\}$ which does not converge, although $\lim \sigma_n = 0$.

- (c) Can it happen that $s_n > 0$ for all n and that $\limsup s_n = \infty$, although $\lim \sigma_n = 0$?
 (d) Put $a_n = s_n - s_{n-1}$, for $n \geq 1$. Show that

$$s_n - \sigma_n = \frac{1}{n+1} \sum_{k=1}^n k a_k.$$

Assume that $\lim(na_n) = 0$ and that $\{\sigma_n\}$ converges. Prove that $\{s_n\}$ converges. [This gives a converse of (a), but under the additional assumption that $na_n \rightarrow 0$.]

- (e) Derive the last conclusion from a weaker hypothesis: Assume $M < \infty$, $|na_n| \leq M$ for all n , and $\lim \sigma_n = \sigma$. Prove that $\lim s_n = \sigma$, by completing the following outline:
 If $m < n$, then

$$s_n - \sigma_n = \frac{m+1}{n-m}(\sigma_n - \sigma_m) + \frac{1}{n-m} \sum_{i=m+1}^n (s_n - s_i).$$

For these i ,

$$|s_n - s_i| \leq \frac{(n-i)M}{i+1} \leq \frac{(n-m-1)M}{m+2}.$$

Fix $\varepsilon > 0$ and associate with each n the integer m that satisfies

$$m \leq \frac{n-\varepsilon}{1+\varepsilon} < m+1.$$

Then $(m+1)/(n-m) \leq 1/\varepsilon$ and $|s_n - s_i| < M\varepsilon$. Hence

$$\limsup_{n \rightarrow \infty} |s_n - \sigma| \leq M\varepsilon.$$

Since ε was arbitrary, $\lim s_n = \sigma$.

Solution 3.14

Exercise 3.17 (Rudin 3.15)

Definition 3.21 can be extended to the case in which the a_n lie in some fixed \mathbb{R}^k . Absolute convergence is defined as convergence of $\sum |\mathbf{a}_n|$. Show that Theorems 3.22, 3.23, 3.25(a), 3.33, 3.34, 3.42, 3.45, 3.47, and 3.55 are true in this more general setting. (Only slight modifications are required in any of the proofs.)

Solution 3.15

Exercise 3.18 (Rudin 3.16)

Fix a positive number α . Choose $x_1 > \sqrt{\alpha}$, and define x_2, x_3, x_4, \dots , by the recursion formula

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{\alpha}{x_n} \right).$$

- (a) Prove that $\{x_n\}$ decreases monotonically and that $\lim x_n = \sqrt{\alpha}$.

(b) Put $\varepsilon_n = x_n - \sqrt{\alpha}$, and show that

$$\varepsilon_{n+1} = \frac{\varepsilon_n^2}{2x_n} < \frac{\varepsilon_n^2}{2\sqrt{\alpha}}$$

so that, setting $\beta = 2\sqrt{\alpha}$,

$$\varepsilon_{n+1} < \beta \left(\frac{\varepsilon_1}{\beta} \right)^{2^n} \quad (n = 1, 2, 3, \dots).$$

(c) This is a good algorithm for computing square roots, since the recursion formula is simple and the convergence is extremely rapid. For example, if $\alpha = 3$ and $x_1 = 2$, show that $\varepsilon_1/\beta < \frac{1}{10}$ and that therefore

$$\varepsilon_5 < 4 \cdot 10^{-16}, \quad \varepsilon_6 < 4 \cdot 10^{-32}.$$

Solution 3.16

Exercise 3.19 (Rudin 3.17)

Fix $\alpha > 1$. Take $x_1 > \sqrt{\alpha}$, and define

$$x_{n+1} = \frac{\alpha + x_n}{1 + x_n} = x_n + \frac{\alpha - x_n^2}{1 + x_n}.$$

- (a) Prove that $x_1 > x_3 > x_5 > \dots$.
- (b) Prove that $x_2 < x_4 < x_6 < \dots$.
- (c) Prove that $\lim x_n = \sqrt{\alpha}$.
- (d) Compare the rapidity of convergence of this process with the one described in Exercise 16.

Solution 3.17

Exercise 3.20 (Rudin 3.18)

Replace the recursion formula of Exercise 16 by

$$x_{n+1} = \frac{p-1}{p}x_n + \frac{\alpha}{p}x_n^{-p+1}$$

where p is a fixed positive integer, and describe the behavior of the resulting sequences $\{x_n\}$.

Solution 3.18

Exercise 3.21 (Rudin 3.19)

Associate to each sequence $a = \{\alpha_n\}$, in which α_n is 0 or 2, the real number

$$x(a) = \sum_{n=1}^{\infty} \frac{\alpha_n}{3^n}.$$

Prove that the set of all $x(a)$ is precisely the Cantor set described in Sec. 2.44.

Solution 3.19**Exercise 3.22 (Rudin 3.20)**

Suppose $\{p_n\}$ is a Cauchy sequence in a metric space X , and some subsequence $\{p_{n_i}\}$ converges to a point $p \in X$. Prove that the full sequence $\{p_n\}$ converges to p .

Solution 3.20**Exercise 3.23 (Rudin 3.21)**

Prove the following analogue of Theorem 3.10(b): If $\{E_n\}$ is a sequence of closed nonempty and bounded sets in a complete metric space X , if $E_n \supset E_{n+1}$, and if

$$\lim_{n \rightarrow \infty} \text{diam } E_n = 0,$$

then $\bigcap_{i=1}^{\infty} E_n$ consists of exactly one point.

Solution 3.21**Exercise 3.24 (Rudin 3.22)**

Suppose X is a nonempty complete metric space, and $\{G_n\}$ is a sequence of dense open subsets of X . Prove Baire's theorem, namely, that $\bigcap_{i=1}^{\infty} G_n$ is not empty. (In fact, it is dense in X .) Hint: Find a shrinking sequence of neighborhoods E_n such that $E_n \subset G_n$, and apply Exercise 21.

Solution 3.22**Exercise 3.25 (Rudin 3.23)**

Suppose $\{p_n\}$ and $\{q_n\}$ are Cauchy sequences in a metric space X . Show that the sequence $\{d(p_n, q_n)\}$ converges. Hint: For any m, n ,

$$d(p_n, q_n) \leq d(p_n, p_m) + d(p_m, q_m) + d(q_m, q_n);$$

it follows that

$$|d(p_n, q_n) - d(p_m, q_m)|$$

is small if m and n are large.

Solution 3.23

Exercise 3.26 (Rudin 3.24)

Let X be a metric space.

- (a) Call two Cauchy sequences $\{p_n\}$, $\{q_n\}$ in X equivalent if

$$\lim_{n \rightarrow \infty} d(p_n, q_n) = 0.$$

Prove that this is an equivalence relation.

- (b) Let X^* be the set of all equivalence classes so obtained. If $P \in X^*$, $Q \in X^*$, $\{p_n\} \in P$, $\{q_n\} \in Q$, define

$$\Delta(P, Q) = \lim_{n \rightarrow \infty} d(p_n, q_n);$$

by Exercise 23, this limit exists. Show that the number $\Delta(P, Q)$ is unchanged if $\{p_n\}$ and $\{q_n\}$ are replaced by equivalent sequences, and hence that Δ is a distance function in X^* .

- (c) Prove that the resulting metric space X^* is complete.
 (d) For each $p \in X$, there is a Cauchy sequence all of whose terms are p ; let P_p be the element of X^* which contains this sequence. Prove that

$$\Delta(P_p, P_q) = d(p, q)$$

for all $p, q \in X$. In other words, the mapping φ defined by $\varphi(p) = P_p$ is an isometry (i.e., a distance-preserving mapping) of X into X^* .

- (e) Prove that $\varphi(X)$ is dense in X^* , and that $\varphi(X) = X^*$ if X is complete. By (d), we may identify X and $\varphi(X)$ and thus regard X as embedded in the complete metric space X^* . We call X^* the completion of X .

Solution 3.24

Exercise 3.27 (Rudin 3.25)

Let X be the metric space whose points are the rational numbers, with the metric $d(x, y) = |x - y|$. What is the completion of this space? (Compare Exercise 24.)

Solution 3.25

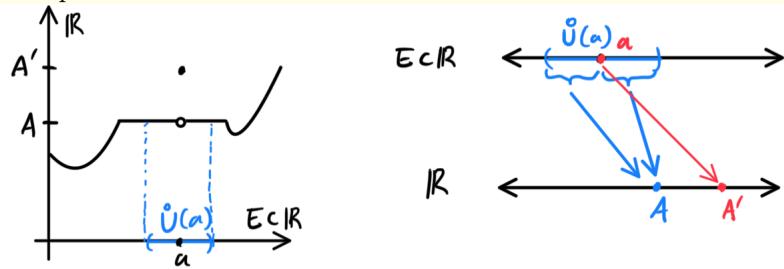
4 Continuous Functions

Even though we can generalize the concept of limits to functions mapping between arbitrary topological spaces $f : X \rightarrow Y$, we define this for functions $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$.

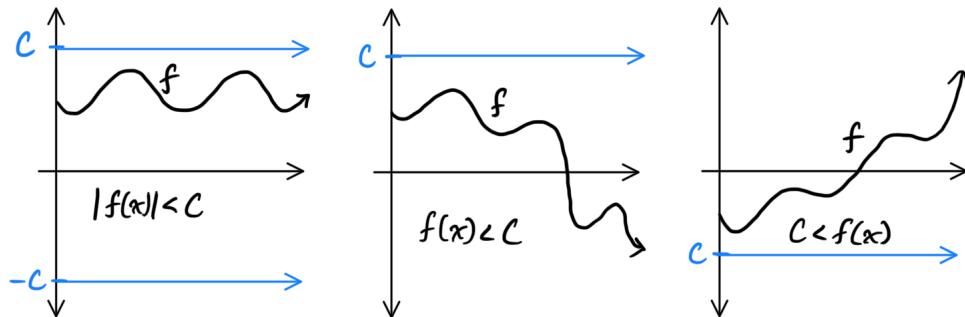
Definition 4.1 (Functions)

Given a real-valued function $f : E \rightarrow \mathbb{R}$ defined on domain $E \subset \mathbb{R}$,

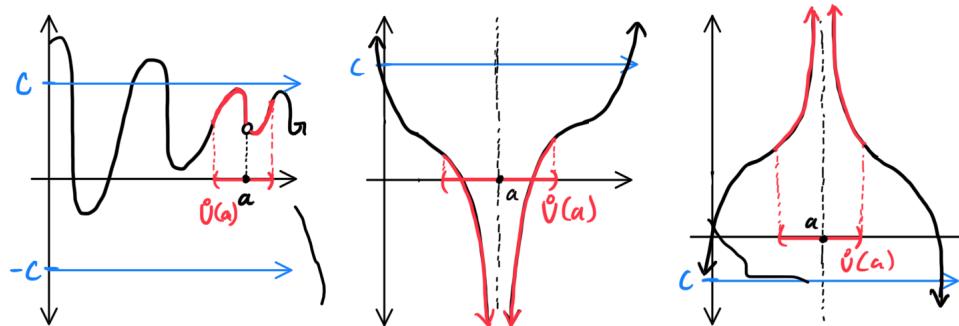
1. f is a **constant function** if $f(x) = A$ for all $x \in E$
2. f is called **ultimately constant** as $x \rightarrow a$ if it is constant in some deleted neighborhood $\dot{U}(a)$, where a is a limit point of E .



3. f is **bounded**, **bounded above**, or **bounded below** respectively if there is a number $C \in \mathbb{R}$ such that $|f(x)| < C$, $f(x) < C$, or $C < f(x)$ for all $x \in E$.



4. f is **ultimately bounded**, **ultimately bounded above**, or **ultimately bounded below** as $x \rightarrow a$ if it is bounded, bounded above, or bounded below in some deleted neighborhood $\dot{U}_E(a)$.



Example 4.1 ()

The function

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

for $x \neq 0$ is not bounded on the domain of definition, but it is ultimately bounded as $x \rightarrow 0$.

Definition 4.2 ($\epsilon - \delta$ Definition of a Limit)

The function $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$ tends to $A \in \mathbb{R}$ as x tends to a , or that

$$\lim_{x \rightarrow a} f(x) = A$$

if for every $\epsilon > 0$ there exists $\delta > 0$ such that

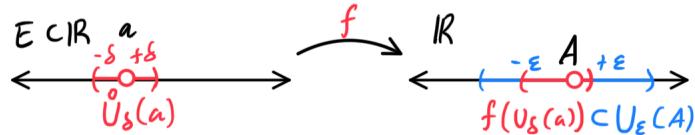
$$0 < |x - a| < \delta \implies |f(x) - A| < \epsilon$$

Note that we set the $0 < |x - a|$ to ensure that $x \neq a$.

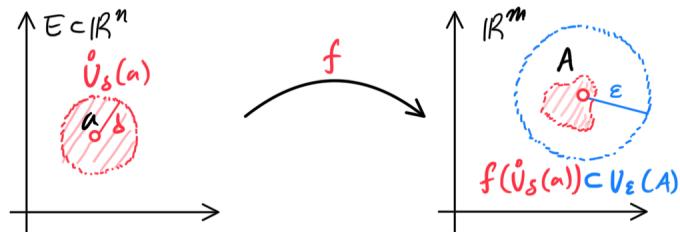
Therefore, in other words, for any arbitrarily small $\epsilon > 0$, if we can find a $\delta > 0$ such that the image of the **deleted δ -neighborhood of a** (defined to be $\mathring{U}_\delta(a) \equiv U_\delta(a) \setminus a$) is completely within the ϵ -neighborhood $U_\epsilon(A)$, then

$$\lim_{x \rightarrow a} f(x) = A$$

Visually,



In higher dimensional spaces, we have



Example 4.2 (Limit of the Signum Function)

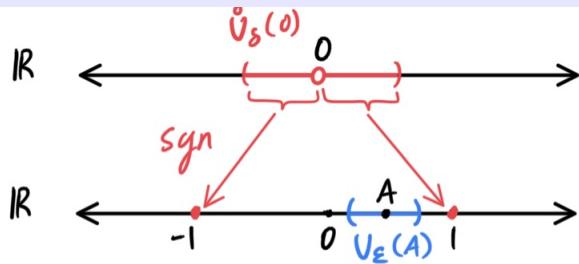
The function $\text{sgn}: \mathbb{R} \rightarrow \mathbb{R}$ defined

$$\text{sgn } x = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

has no limit as $x \rightarrow 0$.

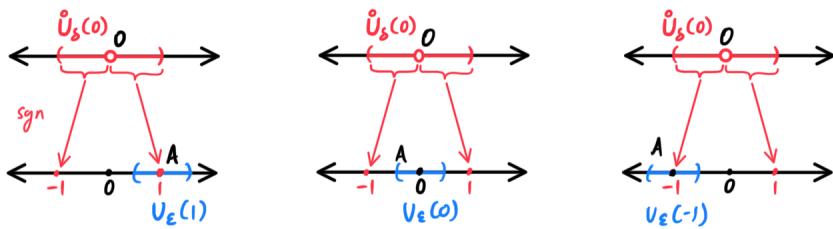
First, it is ludicrous that the limit would be any number that is not $\{-1, 0, 1\}$. If we assume that $A \notin \{-1, 0, 1\}$, then we can choose any arbitrarily small ϵ -neighborhood of A that does not include the three numbers. Clearly, there doesn't exist any $\delta > 0$ such that the deleted δ -neighborhood of 0 maps to a set completely contained in the ϵ -neighborhood of A . That is,

$$\text{sgn}(\mathring{U}_\delta(0)) = \{-1, 1\} \not\subset U_\epsilon(A)$$



It doesn't even intersect the ϵ -neighborhood at all.

1. If $A = 1$, we can construct a ϵ -neighborhood V_A for $\epsilon = \frac{1}{2}$. Clearly, there exists no open neighborhood U_0 of 0 that is entirely mapped to V , since U_0 contains both negative numbers and 0 and hence must be mapped to 0, -1.
2. Similarly, given the ($\epsilon = \frac{1}{2}$)-neighborhood of $A = -1$, there exists no open neighborhood U_0 of 0 that is entirely mapped to it, since U_0 contains both positive numbers and 0 and hence must be mapped to 0, 1.
3. Finally, given the ($\epsilon = \frac{1}{2}$)-neighborhood of $A = 0$, there exists no open neighborhood U_0 of 0 that is entirely mapped to it, since U_0 contains both positive and negative numbers and hence must be mapped to ± 1 .



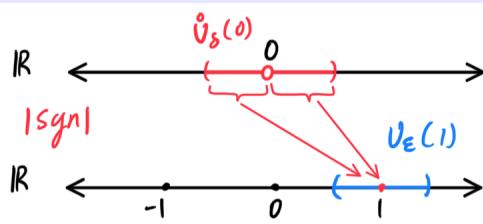
Therefore, the limit does not exist.

Example 4.3 (Limit of Absolute Value of Signum Function)

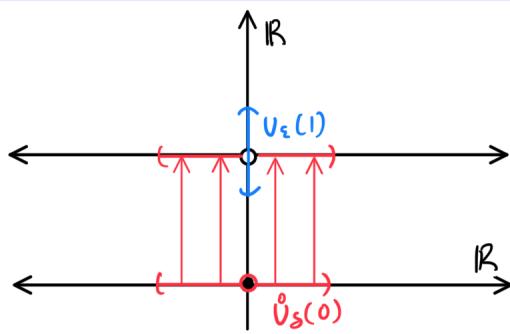
We will show that

$$\lim_{x \rightarrow 0} |\operatorname{sgn} x| = 1$$

We construct a ϵ -neighborhood $U_\epsilon(1)$ around 1. Given this neighborhood, we can imagine choosing the deleted δ -neighborhood $\dot{U}_\delta(0)$ around 0. Since every element in $\dot{U}_\delta(0)$ maps to 1, it is clearly in U_ϵ .



In fact, for arbitrarily small $\epsilon > 0$, we can choose **any** $\delta > 0$ since everything in $\mathbb{R} \setminus 0$ maps to 1. We can visualize this in \mathbb{R}^2 as



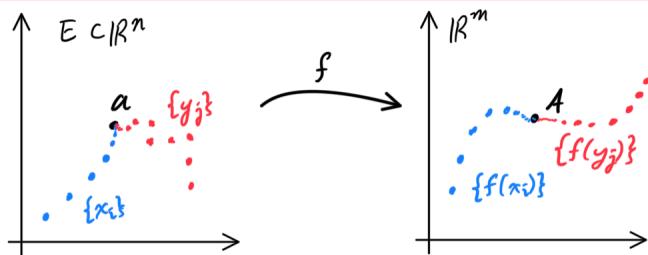
The following lemma nicely interweaves the concepts of limits of sequences and limits of functions. It can be nice for visualization to define the limit of a function using Cauchy sequences rather than the usual $\epsilon - \delta$ definition.

Lemma 4.1 (Cauchy Sequence Criterion of a Limit)

The relation

$$\lim_{x \rightarrow a} f(x) = A$$

holds if and only if for every sequence $\{x_i\}$ of points $x_n \in E \setminus a$ converging to a , the sequence $\{f(x_n)\}$ converges to A .



Note that we choose the points x_n to be in the "deleted" neighborhood $E \setminus a$ (neighborhood E with point a removed) to force us to choose a sequence that is not

$$a, a, a, a, a, a, a, \dots$$

That is, it forces us to choose different points for the sequence.

Theorem 4.1 (Properties of Limits)

Given two numerical valued functions $f, g : E \subset \mathbb{R} \rightarrow \mathbb{R}$ with a common domain where $g(x) \neq 0$ for all $x \in E$, let

$$\lim_{x \rightarrow a} f(x) = A, \quad \lim_{x \rightarrow a} g(x) = B$$

then,

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

$$\lim_{x \rightarrow a} (cf)(x) = cA$$

$$\lim_{x \rightarrow a} (f \cdot g)(x) = A \cdot B$$

$$\lim_{x \rightarrow a} \left(\frac{f}{g} \right)(x) = \frac{A}{B}$$

Proof.

The the Cauchy sequence criterion for a limit, this theorem is an immediate consequence on the corresponding theorem on limits of sequences.

We end this with a theorem connecting the relationship between a limit of a function as $x \rightarrow a$ and its ultimate behavior as $x \rightarrow a$.

Theorem 4.2 ()

Let $f : E \rightarrow \mathbb{R}$ be a function. Then,

1. f is ultimately the constant A as $x \rightarrow a$ implies that $\lim_{x \rightarrow a} f(x) = A$.
2. $\lim_{x \rightarrow a} f(x)$ implies that f is ultimately bounded as $x \rightarrow a$.

4.1 Infinitesimal Functions

Definition 4.3 (Infinitesimal Function)

A function $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$ is said to be **infinitesimal** as $x \rightarrow a$ if

$$\lim_{x \rightarrow a} f(x) = 0$$

Lemma 4.2 (Sums, Products of Infinitesimals)

It is clear that if α, β are infinitesimal as $x \rightarrow a$, then

1. $\alpha + \beta$ is infinitesimal as $x \rightarrow a$
2. $\alpha \cdot \beta$ is infinitesimal as $x \rightarrow a$

Furthermore, if α is infinitesimal and β is ultimately bounded as $x \rightarrow a$, then the product $\alpha \cdot \beta$ is infinitesimal as $x \rightarrow a$.

Proof.

We prove all three statements.

1. Assume that α and β are infinitesimal as $x \rightarrow a$. Then, let us fix a small $\epsilon > 0$. This means that for every $\frac{\epsilon}{2}$ there exists an open deleted neighborhood $\mathring{U}'(a)$ such that its image $\alpha(\mathring{U}'(a)) \subset U'_{\epsilon/2}(0) \subset \mathbb{R}$. Additionally, for every $\frac{\epsilon}{2}$ there exists an open deleted neighborhood $\mathring{U}''(a)$ such that its image $\beta(\mathring{U}''(a)) \subset U'_{\epsilon/2}(0) \subset \mathbb{R}$. Thus, for the deleted neighborhood

$$\mathring{U}(a) \subset \mathring{U}'(a) \cup \mathring{U}''(a)$$

we can see that for all $x \in \mathring{U}(a)$,

$$|(\alpha + \beta)(x)| = |\alpha(x) + \beta(x)| \leq |\alpha(x)| + |\beta(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

and hence $(\alpha + \beta)(\mathring{U}(a)) \subset U_\epsilon(0)$.

2. This case is a special case of assertion 3. That is, every function that has a limit is ultimately bounded.
3. Since $\beta(x)$ is ultimately bounded, this means that there exists a constant M and an open deleted neighborhood $\mathring{U}'(a) \subset E$ such that for all $x \in \mathring{U}'(a)$, its image is bounded: $|\beta(x)| < M$. Let us fix a small $\epsilon > 0$. Then, by definition of the limit, for every $\frac{\epsilon}{M}$ there exists an open deleted neighborhood $\mathring{U}''(a)$ such that its image $\beta(\mathring{U}''(a)) \subset U_{\epsilon/M}(0) \subset \mathbb{R}$. Therefore, for the deleted neighborhood

$$\mathring{U}(a) \subset \mathring{U}'(a) \cup \mathring{U}''(a)$$

we can see that for all $x \in \mathring{U}(a)$,

$$|(\alpha \cdot \beta)(x)| = |\alpha(x)\beta(x)| = |\alpha(x)||\beta(x)| < \frac{\epsilon}{M} \cdot M = \epsilon$$

Therefore, $(\alpha \cdot \beta)(\mathring{U}(a)) \subset U_\epsilon(0)$.

Note that in proving these properties of the limits, we have used the following fact about open deleted neighborhoods around a .

1. $\mathring{U}(a)$ is not the empty set.
2. Given open deleted neighborhoods $\mathring{U}'(a)$ and $\mathring{U}''(a)$, there exists an open deleted neighborhood in the intersection of these neighborhoods.

$$\mathring{U}(a) \subset \mathring{U}'(a) \cup \mathring{U}''(a)$$

These facts can be used to generalize the concept of limits as limits over a certain **filter base**.

Theorem 4.3 (Representation of a Convergent Function as a Shift of its Infinitesimal)

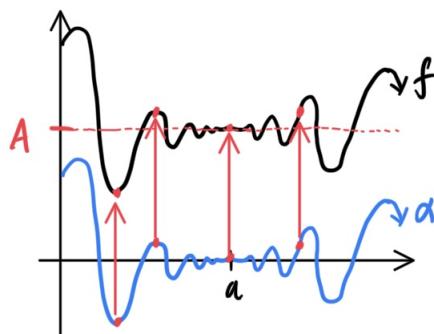
Given a function $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$, its limit exists and

$$\lim_{x \rightarrow a} f(x) = A$$

if and only if f can be represented as

$$f(x) = A + \alpha(x)$$

where α is infinitesimal as $x \rightarrow a$. We can visualize this theorem by thinking of a function f that results from a "shift" of an infinitesimal.



Finally, we reiterate some limit theorems already stated for sequences, but now corresponding to functions. Interpreting the function limit as the Cauchy sequence definition of limits renders the proofs of these theorems trivial.

Theorem 4.4 (Behavior of Functions with Different Limits)

If the functions $f, g : E \rightarrow \mathbb{R}$ are such that

$$\lim_{x \rightarrow a} f(x) = A < B = \lim_{x \rightarrow a} g(x)$$

then there exists a deleted neighborhood $U_\delta(a)$ in E at each point of which $f(x) < g(x)$.

Theorem 4.5 (Squeeze Theorem for Limits of Functions)

Given the functions $f, g, h : E \subset \mathbb{R} \rightarrow \mathbb{R}$ such that

$$f(x) \leq g(x) \leq h(x) \text{ for all } x \in E$$

then,

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x) = C \implies \lim_{x \rightarrow a} g(x) = C$$

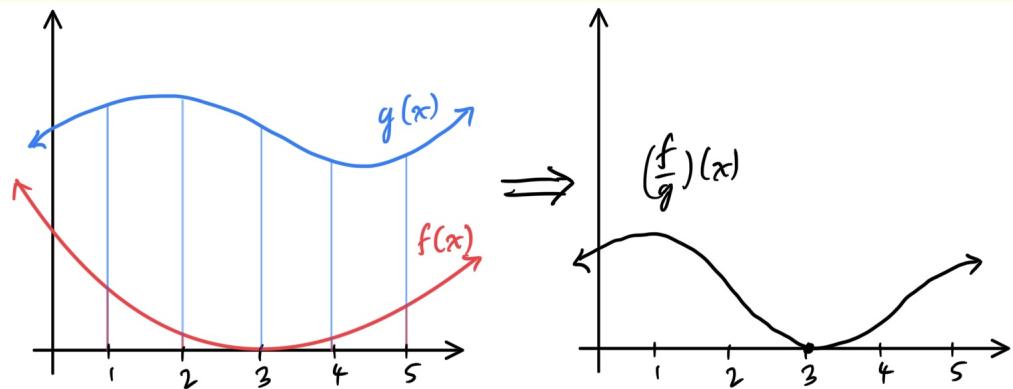
4.2 Asymptotic Behavior of Functions

Definition 4.4 (Little-O Notation)

The function $f : E \rightarrow \mathbb{R}$ is said to be **infinitesimal compared with the function $g : E \rightarrow \mathbb{R}$** as $x \rightarrow a$, written (by abuse of notation) $f = o(g)$ as $x \rightarrow a$, if

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$$

or in other words, if f/g is an infinitesimal function as $x \rightarrow a$. Therefore, $f = o(1)$ as $x \rightarrow a$ means that f is infinitesimal as $x \rightarrow a$.



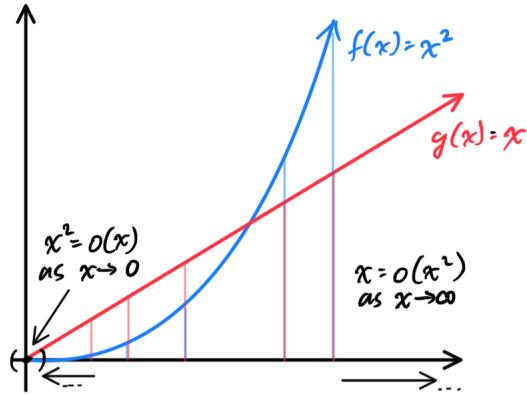
Note that writing $f = o(g)$ is again, an abuse of notation. $f = o(g)$ is really a shorthand way of writing that f is in the class of functions that is infinitesimal compared with the function g .

Intuitively, $f = o(g)$ means that the ratio between $f(x)$ and $g(x)$ will tend to infinity as $x \rightarrow a$ (this does not mean that f will be infinitely greater than g , however!). For example, looking at the two functions $f(x) = x^2$ and $g(x) = x$, we have

1. $x^2 = o(x)$ as $x \rightarrow 0$ (since $\frac{x^2}{x} = x$ is infinitesimal as $x \rightarrow 0$)

2. $x = o(x^2)$ as $x \rightarrow \infty$ (since $\frac{x}{x^2} = \frac{1}{x}$ is infinitesimal as $x \rightarrow \infty$)

We can visualize $g/f(x)$ tending to infinity within a neighborhood of 0 and $f/g(x)$ tending to infinity within a neighborhood of ∞ .



Definition 4.5 (Orders of Infinitesimals, Infinities)

If $f = o(g)$ and g is infinitesimal as $x \rightarrow a$, then f is an **infinitesimal of higher order than g** as $x \rightarrow a$. Furthermore, if f and g are infinite functions as $x \rightarrow a$ and $f = o(g)$ as $x \rightarrow a$, then g is a **higher order infinity than f** as $x \rightarrow a$.

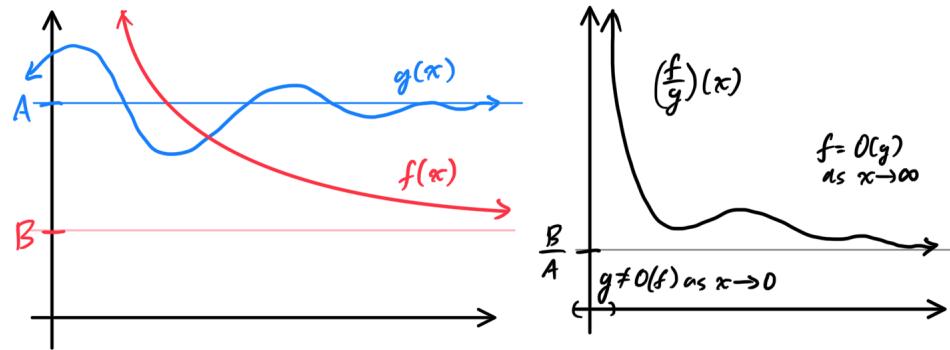
Definition 4.6 (Big-O Notation)

By abuse of notation, $f = O(g)$ as $x \rightarrow a$ means that

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \infty$$

or in other words, f/g is ultimately bounded as $x \rightarrow a$. In particular, $f = O(1)$ as $x \rightarrow a$ means that f is bounded within a certain neighborhood $U(a)$ of a .

In the visual below, we can see that $f = O(g)$ as $x \rightarrow +\infty$ since the limit converges to constant $\frac{B}{A}$ which is bounded. In fact, at any other positive real number x , $(f/g)(x)$ is finite and is therefore bounded. However, at every neighborhood of $x = 0$, $(f/g)(x)$ is unbounded, meaning that $g \neq O(f)$ as $x \rightarrow 0$.



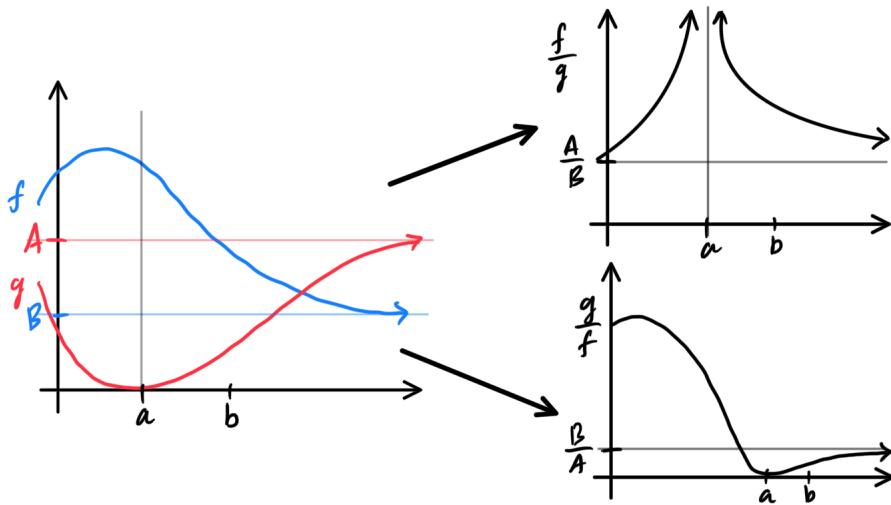
Definition 4.7 (Functions of Same Order)

The functions f and g are of the same order as $x \rightarrow a$, written

$$f \asymp g \text{ as } x \rightarrow a$$

if $f = O(g)$ and $g = O(f)$ as $x \rightarrow a$. Intuitively, this means that the ratio between f and g within some deleted neighborhood of a is finite.

In the visual below, we can see that as long as $k \neq a$, $f = O(g)$ as $x \rightarrow k$ and as $x \rightarrow \infty$. In other words, the function f/g becomes ultimately bounded at every other point other than a , and f/g is unbounded within every neighborhood of a . When looking at g/f , we can see that this function is bounded for all $x \in \mathbb{R}$ and therefore $g = O(f)$ as $x \rightarrow k$ for all k .



Therefore, we can see that as long as $k \neq a$, $f \asymp g$ as $x \rightarrow k$.

Note that the condition that f and g be of the same order as $x \rightarrow a$ is (by definition of ultimately bounded functions) equivalent to the condition that there exist $c_1, c_2 > 0$ and an open neighborhood $U(a)$ such that the relations

$$c_1|g(x)| \leq |f(x)| \leq c_2|g(x)|$$

is true for $x \in U(a)$.

Definition 4.8 (Asymptotic Equivalence of Functions)

For functions f and g , if

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$$

we say that f behaves asymptotically like g as $x \rightarrow a$, or that f is equivalent to g as $x \rightarrow a$, written

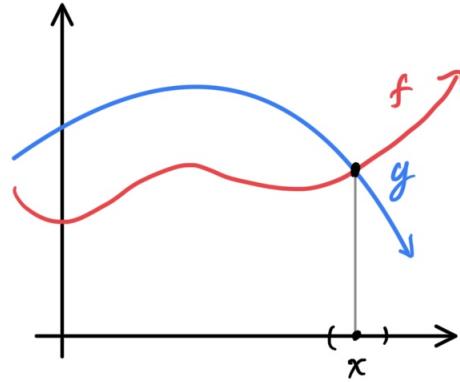
$$f \sim g \text{ as } x \rightarrow a$$

Moreover, \sim is an equivalence relation, which means that

1. $f \sim f$ as $x \rightarrow a$
2. $f \sim g$ as $x \rightarrow a \implies g \sim f$ as $x \rightarrow a$
3. $f \sim g$ and $g \sim h$ as $x \rightarrow a \implies f \sim h$ as $x \rightarrow a$

We list a few examples in order to develop some sort of visual intuition for when two functions are asymptotically equivalent.

1. If $f(a) = g(a) \neq 0$, then $f \sim g$ trivially since the ratio of f and g converges to 1 within a neighborhood of a .

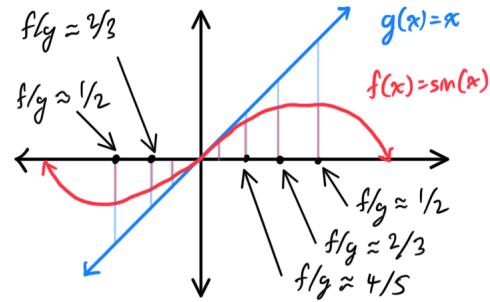


2. When $f(a) = g(a) = 0$, it may be f may be equivalent to g or one function may be infinitesimally smaller than the other.

- (a) When $f(x) = \sin x$ and $g(x) = x$, then $f \sim g$ since we see that

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

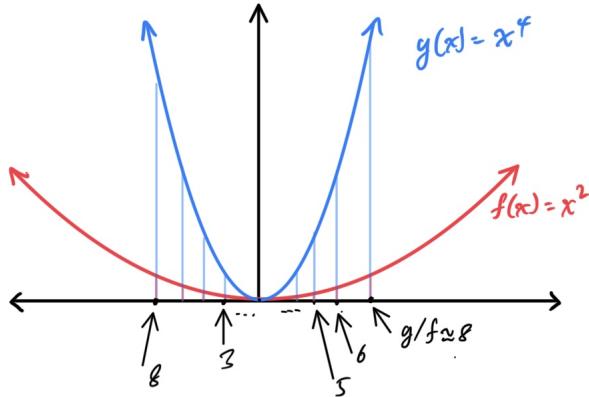
and so $\sin x \sim x$ as $x \rightarrow 0$



- (b) When $f(x) = x^2$ and $g(x) = x^4$, then

$$\lim_{x \rightarrow 0} \frac{x^4}{x^2} = 0$$

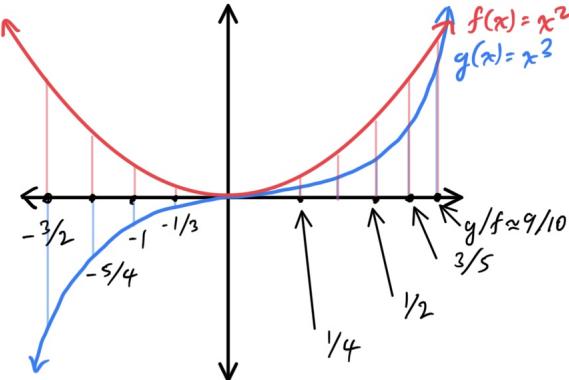
and so $x^4 \not\sim x^2$. In fact, $x^4 = o(x^2)$. Therefore, since x^4 decreases to 0 infinitely faster than x^2 , they are not equivalent.



(c) When $f(x) = x^2$ and $g(x) = x^3$, then

$$\lim_{x \rightarrow 0} \frac{x^3}{x^2} = 0$$

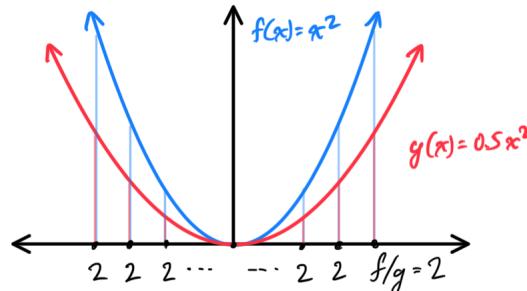
and so $x^3 \not\sim x^2$. In fact, $x^3 = o(x^2)$. Therefore, since x^4 decreases to 0 infinitely faster than x^2 , they are not equivalent.



(d) When $f(x) = x^2$ and $g(x) = 0.5x^2$, then

$$\lim_{x \rightarrow 0} \frac{0.5x^2}{x^2} = \frac{1}{2}$$

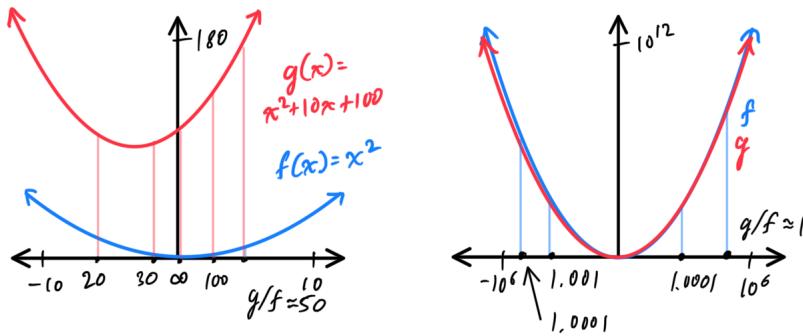
and so $0.5x^2 \not\sim x^2$. Therefore, since $0.5x^2$ is always as twice as small as x^2 , they are not equivalent.



3. When analyzing the behavior of functions as $x \rightarrow \infty$, we can picture the two graphs of f and g on the plane and "zoom out" to see if the ratio of the values converge to 1. This would mean that as $x \rightarrow \infty$, we should see the graphs overlapping more and more. For example, taking $f(x) = x^2$ and $g(x) = x^2 + 10x + 100$, we can see that the discrepancy is high around a neighborhood of $x = 0$. But as $x \rightarrow +\infty$, we get

$$\lim_{x \rightarrow +\infty} \frac{x^2 + 10x + 100}{x^2} = 1$$

and so the graphs look like they are overlapping.



Notice that even though the absolute difference $|(x^2 + 10x + 100) - x^2| = |10x + 100|$ tends to infinity, this difference increases infinitesimally compared to f and g .

From this, we can see that if $f \sim g$ as $x \rightarrow a$, then their difference

$$f - g = o(g) = o(f)$$

That is, $(f - g)(x)$ is infinitesimal compared to g or f (doesn't matter which one we compare it to). This leads to our next section, where we formalize this concept with absolute and relative errors.

4.2.1 Approximations of Functions

It is useful to note that since the relation $\lim_{x \rightarrow a} \gamma(x) = 1$ is equivalent to

$$\gamma(x) = 1 + \alpha(x), \text{ where } \lim_{x \rightarrow a} \alpha(x) = 0$$

the relation $f \sim g$ as $x \rightarrow a$ is equivalent to saying that

$$\frac{f(x)}{g(x)} = \gamma(x), \text{ where } \lim_{x \rightarrow a} \gamma(x) = 1$$

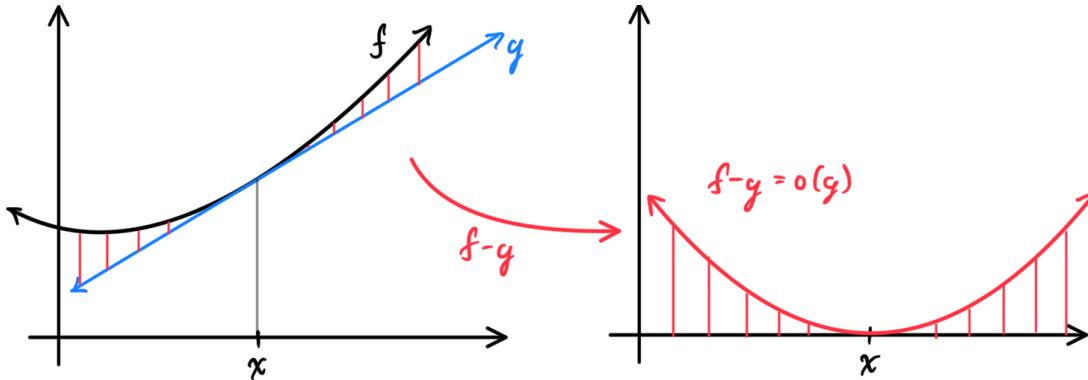
which implies

$$f(x) = g(x) + \alpha(x)g(x) = g(x) + o(g(x)) \text{ as } x \rightarrow a$$

or, symmetrically,

$$g(x) = f(x) + \alpha(x)f(x) = f(x) + o(f(x)) \text{ as } x \rightarrow a$$

This means that f can be exactly represented by another function g , plus another (error) function $o(g(x))$ that is infinitesimal compared to g .



Note that it is not a sufficient condition that the error function be infinitesimal! The error function $f - g$ must be infinitesimal compared to g ! This tells us that not only does the error function

decrease infinitesimally, but also is infinitesimal compared to the approximation function we already have, which is in general a much stronger claim. This representation of certain types functions will provide the foundation for differential calculus when we talk about "good" approximations for a function.

Definition 4.9 (Relative Error)

Since $f \sim g$ as $x \rightarrow a$ means that

$$f(x) = g(x) + o(g(x))$$

we can define the **relative error** of g as an approximation of f to be

$$|\alpha(x)| = \left| \frac{f(x) - g(x)}{g(x)} \right|$$

Clearly, since $f \sim g$, the relative error must be infinitesimal as $x \rightarrow a$.

We use the following lemma to check whether two functions are asymptotically equivalent.

Lemma 4.3 ()

$f \sim g$ as $x \rightarrow a$ if and only if the relative error of g is infinitesimal as $x \rightarrow a$.

Example 4.4 ()

We claim that

$$x^2 + x = \left(1 + \frac{1}{x}\right)x^2 \sim x^2 \text{ as } x \rightarrow \infty$$

We see that the absolute error of this approximation

$$|(x^2 + x) - x^2| = |x|$$

tends to infinity, but the relative error

$$\frac{|x|}{x^2} = \frac{1}{|x|}$$

tends to 0 as $x \rightarrow \infty$.

Theorem 4.6 (Prime Number Theorem)

Let $\pi(x)$ be the number of prime numbers strictly less than x . Then $\pi \sim \frac{x}{\ln x}$ as $x \rightarrow +\infty$, or more precisely,

$$\pi(x) = \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right) \text{ as } x \rightarrow +\infty$$

Example 4.5 ()

It is a fact that $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$, so we have $\sin x \sim x$ as $x \rightarrow 0$. So,

$$\sin x = x + o(x) \text{ as } x \rightarrow 0$$

The following theorem proves useful when computing limits.

Theorem 4.7 ()

If $f \sim \tilde{f}$ as $x \rightarrow a$, then

$$\lim_{x \rightarrow a} f(x)g(x) = \lim_{x \rightarrow a} \tilde{f}(x)g(x)$$

provided one of these limits exist.

Theorem 4.8 (Properties of $o(g)$ and $O(g)$ Functions)

For $x \rightarrow a$,

1. $o(f) + o(f) = o(f)$
2. $o(f)$ is also $O(f)$
3. $o(f) + O(f) = O(f)$
4. $O(f) + O(f) = O(f)$
5. If $g(x) \neq 0$, then

$$\frac{o(f(x))}{g(x)} = o\left(\frac{f(x)}{g(x)}\right), \text{ and } \frac{O(f(x))}{g(x)} = O\left(\frac{f(x)}{g(x)}\right)$$

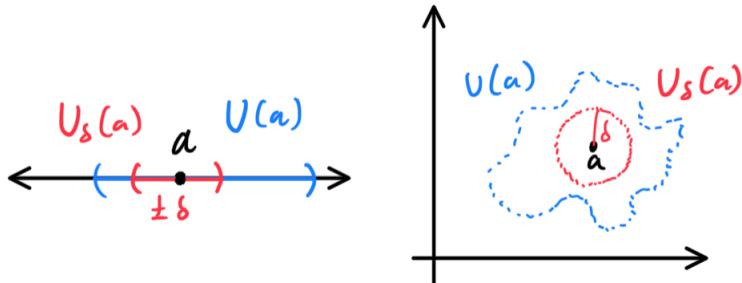
5 Continuous Functions

Definition 5.1 (Continuity of a Function)

A function f is **continuous at point a** if for any neighborhood $V(f(a))$ of $f(a)$, there is a neighborhood $U(a)$ of a whose image under the mapping f is contained in $V(f(a))$.

Generalizing this, we say that a function is **(globally) continuous** if the preimage of every neighborhood in its codomain is an open set in its domain.

The equivalent of these statements for functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ follows from the fact that any neighborhood of a point contains a symmetric neighborhood of the point.

**Lemma 5.1 (Existence of Limits of Continuous Functions)**

$f : E \rightarrow \mathbb{R}$ is continuous at $a \in E$, where a is a limit point of E if and only if

$$\lim_{x \rightarrow a} f(x) = f(a)$$

Proof.

The limit equalling $f(a)$ means that, by definition, for any arbitrarily small deleted neighborhood of $f(a)$, denoted $U_{f(a)} \setminus f(a)$, its preimage will be an open neighborhood of a , which itself will contain an open set.

This also means that we can use the Cauchy limit definition to define continuity of a function at a point. That is, for any sequence $\{a_n\}$ of points in codomain E which converges to point a , the function f is continuous at a if the corresponding sequence $\{f(a_n)\}$ converges to $f(a)$.

Theorem 5.1 ()

This means that the continuous functions commute with the operation of passing to the limit at a point.

$$\lim_{x \rightarrow a} f(x) = f\left(\lim_{x \rightarrow a} x\right)$$

Lemma 5.2 (Properties of Continuous Functions)

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ with $c \in \mathbb{R}$.

1. f continuous at $x_0 \implies cf$ continuous at x_0 .
2. f, g continuous at $x_0 \implies f + g$ continuous at x_0 .
3. Let $m = 1$. f, g continuous at $x_0 \implies fg$ continuous at x_0 .
4. f continuous at x_0 and $f(x) \neq 0 \forall x \in \mathbb{R}^n \implies 1/f$ continuous at x_0 .
5. If $f(x) = (f_1(x), f_2(x), \dots, f_n(x))$ coordinate-wise, then

$$f \text{ continuous at } x_0 \iff f_1, f_2, \dots, f_m \text{ continuous at } x_0$$

6. f continuous at x_0 and g continuous at $y_0 = f(x_0) \implies g \circ f$ continuous at x_0 .

Proof.

This is an immediate result of the equivalence of a function being continuous at point a and its limit at point a existing.

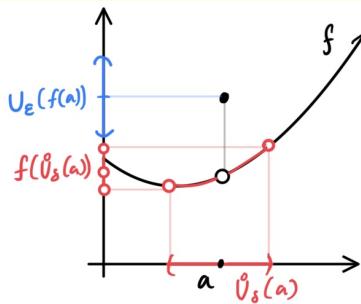
5.1 Points of Discontinuity

Definition 5.2 (Discontinuity)

If the function $f : E \rightarrow \mathbb{R}$ is not continuous at a point of E , then this point is called a **point of discontinuity**, or simply a **discontinuity** of f .

That is, a is a point of discontinuity of f if for some neighborhood $V(f(a))$ of $f(a)$, there exists no neighborhood of a whose image under the mapping f is contained in $V(f(a))$. There are three types of discontinuities:

1. A **removable discontinuity** is characterized by the fact that the limit $\lim_{x \rightarrow a} f(x) = A$ exists, but $A \neq f(a)$.



This means that we can modify f and define a new function $\tilde{f} : E \rightarrow \mathbb{R}$ as

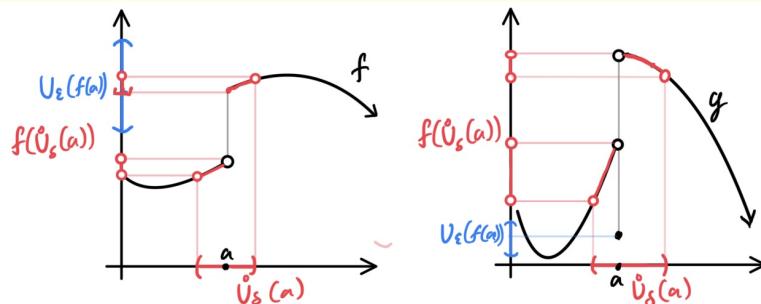
$$\tilde{f}(x) = \begin{cases} f(x), & x \in E \setminus a \\ A, & x = a \end{cases}$$

which would be continuous on E .

2. A **discontinuity of first kind**, also known as a jump/step discontinuity, is characterized by both the left and right-hand limits

$$\lim_{x \rightarrow a-0} f(x) \text{ and } \lim_{x \rightarrow a+0} f(x)$$

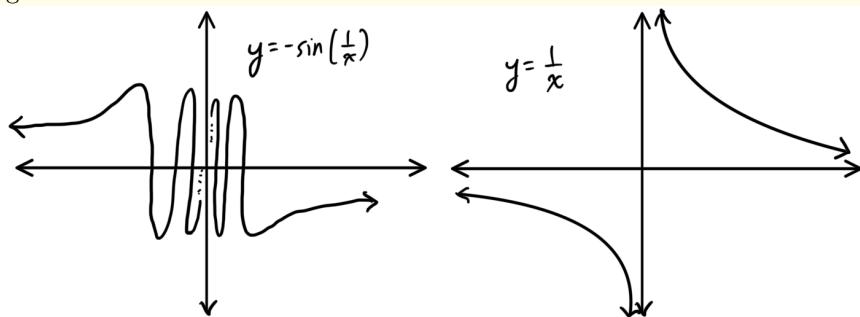
existing, but at least one of them is not equal to the value $f(a)$ that the function assumes at a .



3. A **discontinuity of second kind**, also known as an essential discontinuity, is characterized by at least one of the two limits

$$\lim_{x \rightarrow a-0} f(x) \text{ and } \lim_{x \rightarrow a+0} f(x)$$

not existing.



Note that strictly speaking, a removable discontinuity is really a discontinuity of first kind, but in this context we distinguish them.

Example 5.1 (Dirichlet Function)

The Dirichlet function, defined

$$\mathcal{D}(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is discontinuous at every point, and obviously all of its discontinuities are of second kind, since in every interval there are both rational and irrational numbers and therefore there exists no limit at any point $a \in \mathbb{R}$.

More specifically, given any point $a \in \mathbb{R}$, assume that a is rational. We can set $\epsilon = 0.1$ -neighborhood around the value 1, but no matter how small we let δ , the interval $(a - \delta, a + \delta)$ will contain both rationals and irrationals, meaning that it will map to $\{0, 1\}$ always, which is not fully contained in $(0.9, 1.1)$.

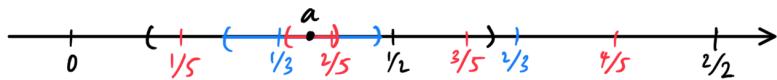
Here is a slightly more interesting example.

Example 5.2 (Riemann Function)

Let the Riemann function \mathcal{R} be defined

$$\mathcal{R}(x) = \begin{cases} \frac{1}{n}, & \text{if } x = \frac{m}{n} \in \mathbb{Q}, \text{ where } \gcd(m, n) = 1 \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

We first note that for any point $a \in \mathbb{R}$, any bounded neighborhood $U(a)$ of it, and any number $N \in \mathbb{N}$, the neighborhood $U(a)$ contains only a finite number of rational numbers $\frac{m}{n}$, where $n < N$. By shrinking the neighborhood, we can assume that the denominators of all rational numbers in the neighborhood are larger than N . We can visualize why this is by seeing that rational numbers with larger denominators have smaller "gaps" between them.



Thus, at any point $x \in U(a) \setminus a$, we have

$$|\mathcal{R}(x)| < \frac{1}{N}$$

and therefore

$$\lim_{x \rightarrow a} \mathcal{R}(x) = 0$$

at any point $a \in \mathbb{R} \setminus \mathbb{Q}$. Hence, the Riemann function is continuous at any irrational number.

5.2 Properties of Continuous Functions

Theorem 5.2 (Local Properties of Continuous Functions)

Let $f : E \rightarrow \mathbb{R}$ be a function that is continuous at the point $a \in E$. Then,

1. f is bounded in some neighborhood $U(a)$.
2. If $f(a) \neq 0$, then in some neighborhood $U(a)$ all the values of the function have the same sign as $f(a)$.

3. If the function $g : U(a) \subset E \rightarrow \mathbb{R}$ is defined in some neighborhood of a and is continuous at a , then the following functions

$$(f + g)(x)$$

$$(f \cdot g)(x)$$

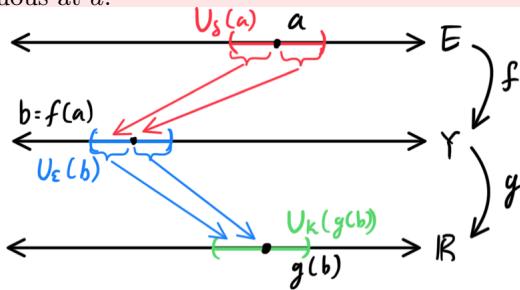
$$\left(\frac{f}{g}\right)(x) \text{ where } g(a) \neq 0$$

are also defined in $U(a)$ and continuous at a .

4. If the function $g : Y \rightarrow \mathbb{R}$ is continuous at a point $b \in Y$ and f is such that $f : E \rightarrow Y$, $f(a) = b$, and f is continuous at a , then the composite function

$$g \circ f : E \rightarrow \mathbb{R}$$

is defined on E and continuous at a .



This is easy to see because given the open neighborhood of $g(b)$, we know for a fact that $U_\delta(a)$ maps completely into $U_\epsilon(b)$, and that $U_\epsilon(b)$ maps completely into $U_\kappa(g(b))$ and so the composition of these mappings must mean that $U_\delta(a)$ maps completely into $U_\kappa(g(b))$.

Example 5.3 ()

An algebraic polynomial

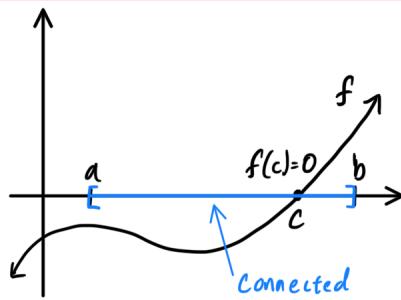
$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n$$

is a continuous function on \mathbb{R} . Since $f(x) = x$ and $f(x) = c$ are continuous functions, by induction on x , we can multiply them together to find that $f(x) = x^n$ is continuous, which implies that ax^n is continuous, which implies that the sums of these functions are also continuous.

Unlike local properties, the global property of a function is a property involving the entire domain of definition of the function.

Theorem 5.3 (Intermediate Value Theorem)

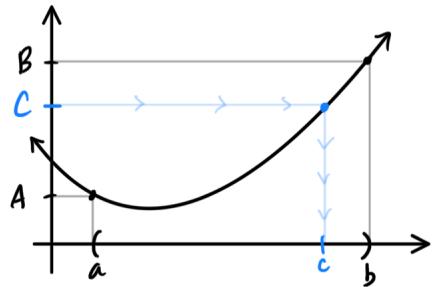
If a function that is continuous on a closed interval assumes values with different signs at the endpoints of the interval, then there is a point in the interval where it assumes the value 0.

**Proof.**

This following proof provides a very simple algorithm for finding the zero of the equation $f(x) = 0$ on an interval whose endpoints has values with opposite signs. Note that the colloquial description of the intermediate value theorem, that is is impossible to pass continuously from positive to negative values without assuming the value 0 along the way), assumes more than they state. That is, this theorem is actually dependent on the domain of definition: that is is a closed interval, or more generally, that it is **connected**.

Corollary 5.1 ()

If a function f is continuous on an open interval and assumes values $f(a) = A, f(b) = B$, then for any number $C \in (A, B)$, there is a point c between a and b such that $f(c) = C$.

**Theorem 5.4 (Weierstrass Maximum-Value Theorem)**

A function that is continuous on a closed interval is bounded in that interval, with a maximum and minimum.

5.2.1 Uniform Continuity

Roughly speaking, a function f is uniformly continuous if it is possible to guarantee that $f(x)$ and $f(y)$ be as close to each other as we please by requiring only that x and y be sufficiently close to each other.

Definition 5.3 (Uniform Continuity)

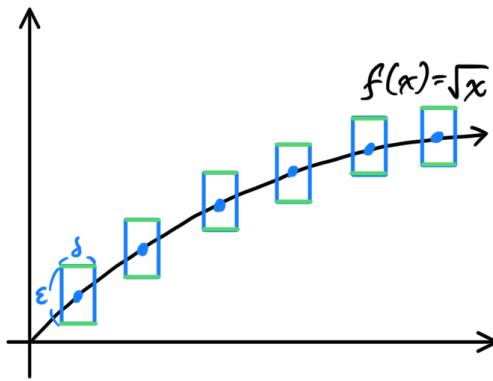
A function $f : E \rightarrow \mathbb{R}$ is **uniformly continuous** on a set $E \subset \mathbb{R}$ if for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|f(x_1) - f(x_2)| < \epsilon$$

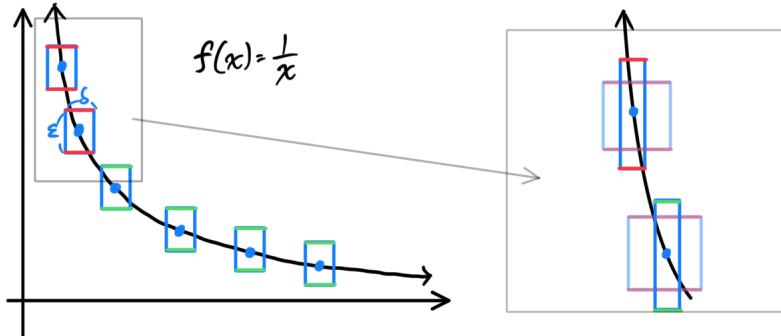
for all points $x_1, x_2 \in E$ such that $|x_1 - x_2| < \delta$.

Intuitively, uniform continuity says that given any two points x, y in the domain where their distance is arbitrarily small (δ apart), we can guarantee that the distance between $f(x), f(y)$ is at maximum some arbitrarily small ϵ .

The following visual shows the radical function $f(x) = \sqrt{x}$ defined on \mathbb{R}^+ . We can see that it satisfies uniform continuity because the graph does not escape the top and/or bottom of the $\epsilon \times \delta$ window, no matter where the box is located on the graph. More strictly speaking, no matter what we set the ϵ (how long the box is), uniform continuity says that we can choose a sufficient δ (width of the box) such that the graph does not escape the top/bottom of the window no matter where the window is.

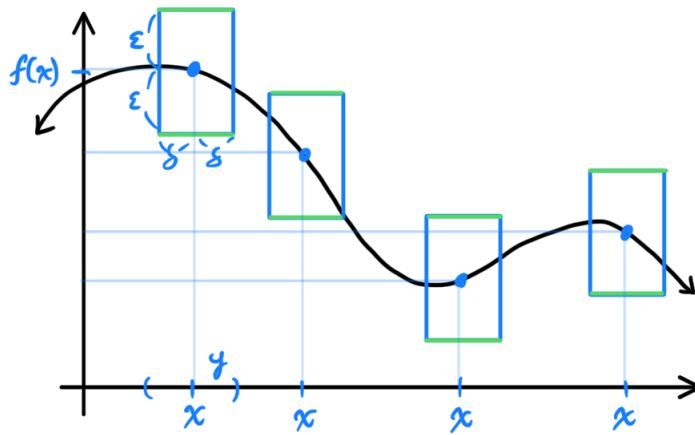


We can clearly see that the function $f(x) = 1/x$ is not uniformly continuous, since the graph escapes the $\epsilon \times \delta$ window at some point (marked in red). More strictly speaking, given any length ϵ of the window, we cannot create a thin-enough δ box that will contain the graph, since as $x \rightarrow 1$, the function becomes unbounded.



That is, arbitrarily thin boxes don't help when the slope is arbitrarily steep.

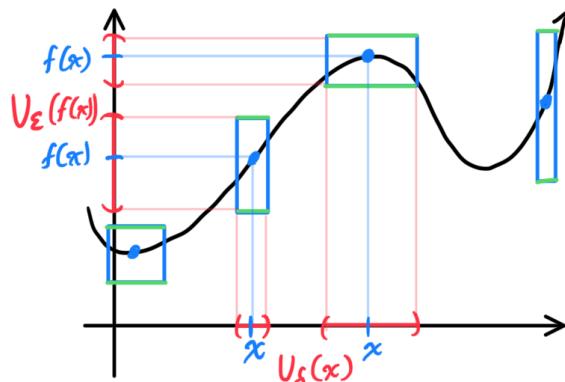
To compare uniform continuity with regular continuity, we can adapt this alternate (yet equivalent interpretation): Let there exist function $f : E \rightarrow \mathbb{R}$. Given any $\epsilon > 0$, we can choose a $\delta > 0$ such that given any point $x \in E$ and $f(x)$, as long as a second point y is δ away from x , then $f(y)$ is ϵ away from $f(x)$. This visualization would lead to there being a $2\epsilon \times 2\delta$ window around point x .



Uniform continuity means that the box above does not change dimensions no matter where the point is (hence, the name uniform). Therefore, given a certain $\epsilon > 0$, the way we choose δ is only dependent on ϵ , and so it must be a function of ϵ :

$$\delta = \delta(\epsilon)$$

However, in continuity, there just has to exist **some** δ -neighborhood of x such that its image is contained in the ϵ -neighborhood of $f(x)$. There are no restrictions on the dimensions of this box; it just has to exist.



Lemma 5.3 ()

If f is uniformly continuous on the set E , it is continuous at each point of that set. However, the converse is not generally true.

Theorem 5.5 (Cantor's Theorem on Uniform Continuity)

A function that is continuous on a closed interval is uniformly continuous on that interval.

Example 5.4 ()

Let $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 3x + 7$. Then f is uniformly continuous. Choose $\epsilon > 0$. Let $\delta = \epsilon/3$. Choose $x, y \in \mathbb{R}$ and assume $|x - y| < \delta$. Then,

$$|f(x) - f(y)| = |3x + 7 - 3y - 7| = 3|x - y| < 3\delta = \epsilon$$

Example 5.5 ()

Let $f : (0, 4) \subset \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$. Then f is uniformly continuous on $(0, 4)$. Choose $\epsilon > 0$. Let $\delta = \epsilon/8$. Choose $x, y \in (0, 4)$ and assume $|x - y| < \delta$. Then,

$$|f(x) - f(y)| = |x^2 - y^2| = (x + y)|x - y| < (4 + 4)|x - y| = 8\delta = \epsilon$$

In both examples, the function satisfied an inequality of form

$$|f(x_1) - f(x_2)| \leq M|x_1 - x_2|$$

this is called the Lipshitz inequality.

5.2.2 Lipshitz Continuity

Lipshitz continuity is a strong form of uniform continuity for functions. Intuitively, a Lipshitz continuous function is limited in how fast it can change (by the Lipshitz constant).

Definition 5.4 (Lipshitz Continuous Function)

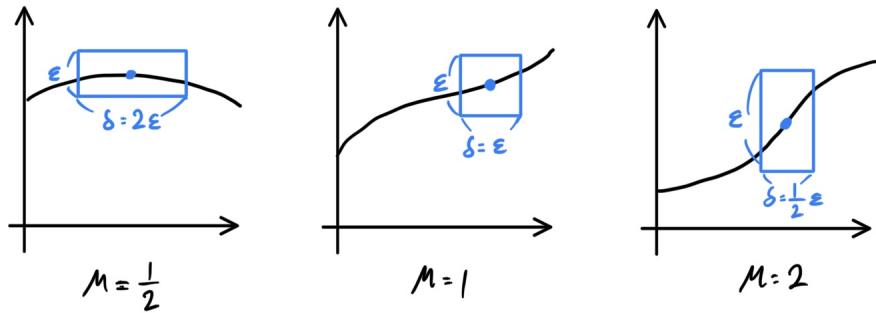
Given $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$, f is **Lipshitz continuous** if there exists a positive real constant M such that for all real $x, y \in E$,

$$|f(x) - f(y)| \leq M|x - y|$$

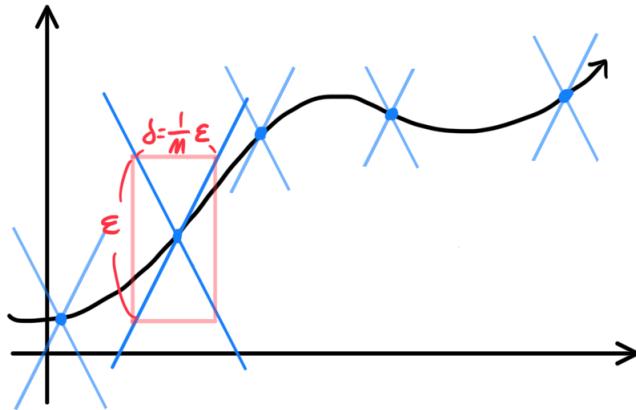
The corresponding M is called the **Lipshitz constant**, and the smallest constant M satisfying this inequality is called the **best Lipshitz constant**.

Note that Lipshitz continuity pops up as a very natural extension of uniform continuity. The inequality above just means that given an ϵ , we can choose a δ such that a linear multiple of δ is always greater than ϵ . This means that Lipshitz continuity is just uniform continuity such that the δ function is linear:

$$\delta = \delta(\epsilon) = \frac{1}{M}\epsilon$$



Another way to interpret uniform continuity is by seeing that the derivative of f is bounded by the slope M .



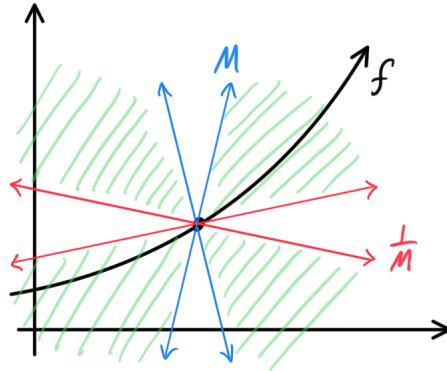
This slope bound implies that for every pair of points on the graph of this function, the absolute value of the slope of the line connecting them is not greater than M' . The smallest M' is the best Lipschitz constant.

Definition 5.5 (Bi-Lipschitz Continuity)

A function $f : E \subset \mathbb{R}$ is **Bi-Lipschitz continuous** if there exists constant $M \geq 1$ such that for all real $x, y \in E$,

$$\frac{1}{M}|x - y| \leq |f(x) - f(y)| \leq M|x - y|$$

A visual of this map is shown, where the function f must always land in the shaded green area.



It immediately follows that for $x \neq y$, $|f(x) - f(y)|$ cannot equal 0, which means that a bilipshitz map is injective. A bilipshitz map is really just Lipschitz map with its inverse also being Lipschitz.

Proposition 5.1 ()

A bilipshitz map f is a homeomorphism onto its image.

5.2.3 Inverse Function Theorem

We begin by introducing this intuitive lemma.

Lemma 5.4 ()

A continuous mapping $f : E \rightarrow \mathbb{R}$ of a closed interval $E = [a, b]$ into \mathbb{R} is injective if and only if the function f is strictly monotonic on $[a, b]$.

Furthermore, every strictly monotonic function $f : X \subset \mathbb{R} \rightarrow \mathbb{R}$ (for arbitrary X) has an inverse

$$f^{-1} : f(X) \subset \mathbb{R} \rightarrow \mathbb{R}$$

with the same kind of monotonicity on $f(X)$ that f has on X .

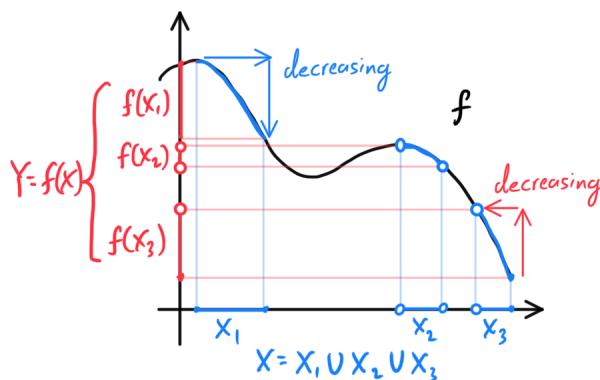
Lemma 5.5 (Criterion for Continuity of a Monotonic Function)

A monotonic function $f : E \rightarrow \mathbb{R}$ defined on a closed interval $E = [a, b]$ is continuous if and only if its set of values $f(E)$ is the closed interval with endpoints $f(a)$ and $f(b)$.

Note that both conditions imply that there are no points of discontinuities in the graph of f .

Theorem 5.6 (Inverse Function Theorem)

A function $f : X \rightarrow \mathbb{R}$ that is strictly monotonic on a set $X \subset \mathbb{R}$ has an inverse $f^{-1} : Y \rightarrow \mathbb{R}$ defined on the set $Y = f(X)$ of values of f . The function $f^{-1} : Y \rightarrow \mathbb{R}$ is monotonic and has the same type of monotonicity on Y that f has on X .



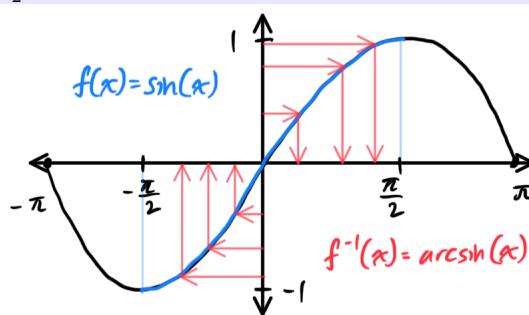
If in addition, X is a closed interval $[a, b]$ and f is continuous on X , then the set $Y = f(X)$ is the closed interval with endpoints $f(a)$ and $f(b)$ and the function $f^{-1} : Y \rightarrow \mathbb{R}$ is continuous on it.

Example 5.6 ()

The function $f(x) = \sin x$ is increasing and continuous on the closed interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Hence, the restriction to the closed interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$ has an inverse $x = f^{-1}(y)$, which is denoted by

$$x = \arcsin y$$

This function is defined on the closed interval $[-\sin(-\frac{\pi}{2}), \sin(\frac{\pi}{2})] = [-1, 1]$ and increases continuously from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.



6 Differentiation

6.1 Functions Differentiable at a Point

Definition 6.1 (Differentiable Function)

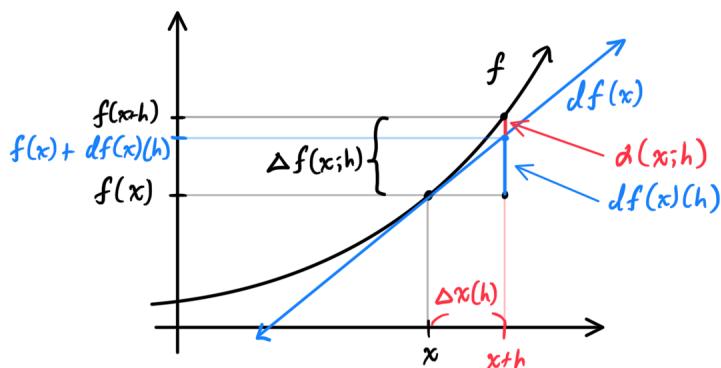
A function $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$ is **differentiable** at a given point x (that is a limit point of E) if there exists a linear function $h \mapsto df(x)h$ (called the **differential of f**) and an infinitesimal $\alpha(x; h) = o(h)$ as $h \rightarrow 0$, such that

$$f(x + h) - f(x) = df(x)(h) + \alpha(x; h)$$

Note that x is fixed; what we are really interested here is the h value. Furthermore,

1. $\Delta x(h) \equiv (x + h) - x = h$ is called the **increment of the argument**
2. $\Delta f(x; h) \equiv f(x + h) - f(x)$ is called the **increment of the function**

They are often denoted (inappropriately) by the symbols Δx and $\Delta f(x)$ representing functions of h . The differential and the infinitesimal can be visualized below.

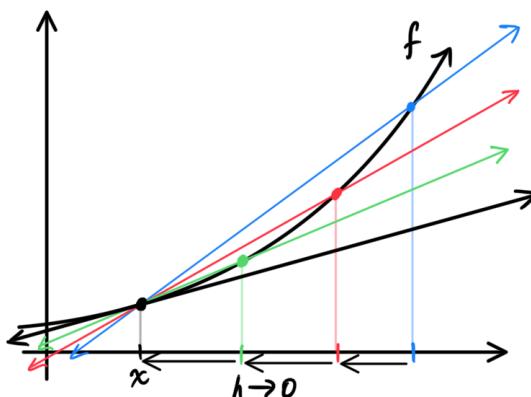


Definition 6.2 (Derivative)

Given function $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$, the number

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

is called the **derivative** of the function f at x . It can be visualized as the sequence of the slopes of the secant lines converging onto the slope of the black tangent line as shown.



This equality can also be written in the equivalent form:

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \alpha(h)$$

where α is infinitesimal as $h \rightarrow 0$. This is also equivalent to:

$$f(x+h) - f(x) = f'(x)h + o(h)$$

where the error term $o(h) \rightarrow 0$ as $h \rightarrow 0$.

Note that we have defined the differentiability of a function at a point and the existence of its derivative at a point completely separately. But it turns out that the existence of this arbitrary number $f'(x)$ we call the "derivative," defined

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

actually has an equivalent form of

$$f(x+h) - f(x) = f'(x)h + o(h)$$

But since $f'(x)$ is in \mathbb{R} , the function $h \mapsto f'(x)h$ is linear and $o(h)$ is infinitesimal, so it is in the form

$$f(x+h) - f(x) = df(x)(h) + \alpha(x; h)$$

which, by definition, means that it is differentiable! Therefore, we have determined the equivalence between the differentiability of a function at a point and the existence of its derivative at the same point. Furthermore, this function $h \mapsto f'(x)h$ is precisely the differential of f , meaning that

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

Furthermore,

$$\Delta f(x; h) - df(x)(h) = \alpha(x; h)$$

and $\alpha(x; h) = o(h)$ as $h \rightarrow 0$, or in other words, the difference between the increment of the function and the value of the function $df(x)$ in h is an infinitesimal of higher order than the first in h . For this reason, we say that the differential is the **principal linear part of the increment of the function**.

In particular, if $f(x) \equiv x$, then we have $f'(x) \equiv 1$ and

$$dx(h) = 1 \cdot h = h$$

Substituting this equality into $df(x)(h) = f'(x)h$, we get

$$df(x)(h) = f'(x) dx(h)$$

or without the input parameter h ,

$$df(x) = f'(x) dx$$

Note that this is an equality between two functions of h . From this, we obtain the familiar **Leibniz notation** of the derivative:

$$\frac{df(x)(h)}{dx(h)} = f'(x) \iff \frac{df(x)}{dx} = f'(x)$$

That is, the function $\frac{df(x)}{dx}$, which is the ratio of the functions $df(x)$ and dx , is constant and equals $f'(x)$.

6.2 Tangent Line: Geometric Meaning of the Derivative, Differential

Let us try to construct successive approximations to an arbitrary function $f : E \rightarrow \mathbb{R}$ at a given limit point x_0 . That is, we find a function g such that

$$f = g + o(g)$$

Depending on what g is, we can construct better approximations of f .

6.2.1 Constant Approximation

The 0th order approximation is when g is a constant. That is, $g \equiv c_0$ for some $c_0 \in \mathbb{R}$. This means

$$f(x) = c_0 + o(c_0) = c_0 + o(1) \text{ as } x \rightarrow x_0$$

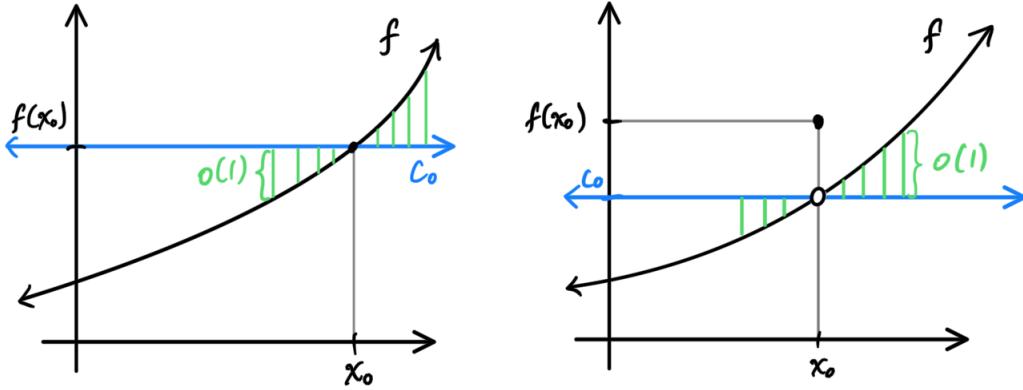
More precisely, we want this difference $f(x) - c_0$ to be $o(1)$ as $x \rightarrow x_0$, which means that it is simply infinitesimal. Visualizing this, we can see that given a constant approximation (labeled in blue) to a function at x_0 , its error term (labeled in green) is in fact, infinitesimal. All this boils down to the fact that

$$\lim_{x \rightarrow x_0} f(x) = c_0$$

If the function is continuous at x_0 , then

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

and naturally $c_0 = f(x_0)$. Both the continuous (left) and noncontinuous case (right) is shown, but in most cases, we will assume continuity.



6.2.2 Linear Approximation

The 1st order approximation is a linear function that approximates f as

$$f(x) = c_0 + c_1(x - x_0) + o(x - x_0) \text{ as } x \rightarrow x_0$$

Following the previous logic, assuming f continuous means that $c_0 = f(x_0)$. Furthermore, as $x \rightarrow x_0$

$$\begin{aligned} f(x) = c_0 + c_1(x - x_0) + o(x - x_0) &\implies c_1 = \frac{f(x) - c_0 - o(x - x_0)}{x - x_0} \\ &\implies c_1 = \frac{f(x) - c_0}{x - x_0} - \frac{o(x - x_0)}{x - x_0} \\ &\implies c_1 = \frac{f(x) - c_0}{x - x_0} - o(1) \\ &\implies c_1 = \lim_{x \rightarrow x_0} \frac{f(x) - c_0}{x - x_0} = f'(x_0) \end{aligned}$$

But this just means that $f'(x_0) = c_1$. Note that before, we have proved the equivalence of the existence of a derivative at x_0 with differentiability at x_0 (which itself means that there exists a linear approximation $df(x)(h)$ that is a function of h). Here, we have created a linear approximation with respect to $x = x_0 + h$, rather than h (shifted the function).

Therefore, the function

$$\alpha(x) = f(x_0) + f'(x_0)(x - x_0)$$

provides the best linear approximation to the function f in a neighborhood of x_0 in the sense that for any other function $\beta(x)$ of the form

$$\beta(x) = c_0 + c_1(x - x_0)$$

we have $f(x) - \beta(x) \neq o(x - x_0)$ as $x \rightarrow x_0$. The graph of the function α is the straight line

$$y - f(x_0) = f'(x_0)(x - x_0)$$

This leads to the definition of our familiar tangent line.

Definition 6.3 (Tangent Line)

If a function $f : E \rightarrow \mathbb{R}$ is differentiable at a point $x_0 \in E$, the line defined by

$$y - f(x_0) = f'(x_0)(x - x_0)$$

is called the **tangent** to the graph of f at the point $(x_0, f(x_0))$.

6.2.3 Higher Order Approximations

We can continue this pattern to get a quadratic approximation of f in the form

$$f(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + o((x - x_0)^2) \text{ as } x \rightarrow x_0$$

As we have done in the previous subsection, we can derive (assuming continuity of f) $c_0 = f(x_0)$, $c_1 = f'(x_0)$. To derive what c_2 should be, we see that the equation above implies

$$c_2 = \frac{f(x) - c_0 - c_1(x - x_0) - o((x - x_0)^2)}{(x - x_0)^2} = \frac{f(x) - c_0 - c_1(x - x_0)}{(x - x_0)^2} - o(1)$$

which means

$$c_2 = \lim_{x \rightarrow x_0} \frac{f(x) - c_0 - c_1(x - x_0)}{(x - x_0)^2}$$

Extending this, if we are seeking a polynomial $P_n(x_0; x) = c_0 + c_1(x - x_0) + \dots + c_n(x - x_0)^n$ such that

$$f(x) = c_0 + c_1(x - x_0) + \dots + c_n(x - x_0)^n + o((x - x_0)^n) \text{ as } x \rightarrow x_0$$

we would find

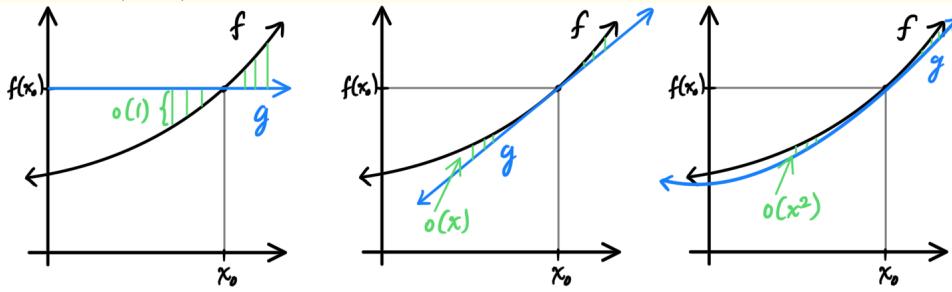
$$\begin{aligned} c_0 &= \lim_{x \rightarrow x_0} f(x) \\ c_1 &= \lim_{x \rightarrow x_0} \frac{f(x) - c_0}{x - x_0} \\ c_2 &= \lim_{x \rightarrow x_0} \frac{f(x) - c_0 - c_1(x - x_0)}{(x - x_0)^2} \\ &\dots = \dots \\ c_n &= \lim_{x \rightarrow x_0} \frac{f(x) - (c_0 + \dots + c_{n-1}(x - x_0)^{n-1})}{(x - x_0)^n} \end{aligned}$$

We formalize the order of these approximations by analyzing their error bound.

Definition 6.4 (nth Order Contact)

If $f, g : E \rightarrow \mathbb{R}$ are continuous at point x_0 and $(f - g)(x) = o((x - x_0)^n)$ as $x \rightarrow x_0$, then we say that f and g have **n th order contact at x_0** , or more precisely, **contact of order at least n** .

The following visual shows approximations g of an arbitrary function f that have 0th (left), 1st (middle), and 2nd (right) order contact at x_0 .

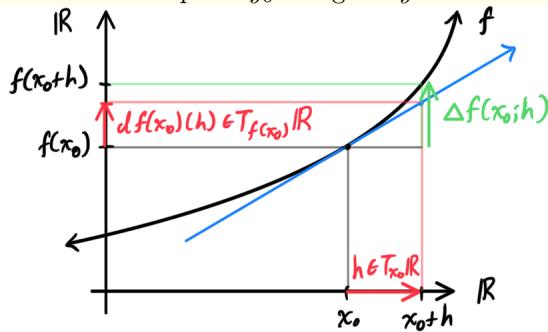


6.2.4 The Real Tangent Space

Tangent spaces.

Definition 6.5 (Tangent Space)

Given function $f : E \rightarrow \mathbb{R}$ and a point $x_0 \in E$, the increment of the argument $h = x - x_0$ can be regarded as a vector attached to the point x_0 and defining the transition from x_0 to $x_0 + h$. h is called a **tangent vector**, and the set of all such vectors as $T_{x_0}\mathbb{R}$. Similarly, we denote $T_{y_0}\mathbb{R}$ the set of all displacement vectors from the point y_0 along the y -axis.



Then, we can see that the differential is a mapping

$$df(x_0) : T_{x_0}\mathbb{R} \longrightarrow T_{f(x_0)}\mathbb{R}$$

Note that there are two functions to pay attention to here:

1. The true increment of f , defined $h \mapsto f(x_0 + h) - f(x_0) = \Delta f(x_0; h)$ (labeled in green).
2. The differential $h \mapsto f'(x_0)h = df(x_0)(h)$, which gives the increment of the tangent to the graph for increment h in the argument (labeled in red).

Example 6.1 ()

Let $f(x) = \sin x$. Then we will show that $f'(x) = \cos x$.

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin(x)}{h} &= \lim_{h \rightarrow 0} \frac{2 \sin\left(\frac{h}{2}\right) \cos\left(x + \frac{h}{2}\right)}{h} \\ &= \lim_{h \rightarrow 0} \cos\left(x + \frac{h}{2}\right) \cdot \lim_{h \rightarrow 0} \frac{\sin\left(\frac{h}{2}\right)}{\left(\frac{h}{2}\right)} = \cos(x) \end{aligned}$$

Here, we have used the theorem on the limit of a product, the continuity of the function $\cos(x)$, the equivalence $\sin t \sim t$ as $t \rightarrow 0$, and the theorem on the limit of a composite function.

Example 6.2 ()

We will show that $\cos'(x) = -\sin(x)$.

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{\cos(x+h) - \cos(x)}{h} &= \lim_{h \rightarrow 0} \frac{-2 \sin\left(\frac{h}{2}\right) \sin\left(x + \frac{h}{2}\right)}{h} \\ &= -\lim_{h \rightarrow 0} \sin\left(x + \frac{h}{2}\right) \cdot \lim_{h \rightarrow 0} \frac{\sin\left(\frac{h}{2}\right)}{\left(\frac{h}{2}\right)} = -\sin(x)\end{aligned}$$

6.3 Rules of Differentiation over \mathbb{R}

6.3.1 Basic Properties; Derivatives of Composite, Inverse Functions

Theorem 6.1 (Arithmetic Properties of Differentiation over \mathbb{R})

If functions $f, g : E \rightarrow \mathbb{R}$ are differentiable at a point $x \in E$, then

1. their sum is differentiable at x , and

$$d(f+g)(x) = df(x) + dg(x) \iff (f+g)'(x) = (f'+g')(x)$$

2. their product is differentiable at x , and

$$d(f \cdot g)(x) = g(x)df(x) + f(x)dg(x) \iff (f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

3. their quotient is differentiable at x if $g(x) \neq 0$, and

$$d\left(\frac{f}{g}\right)(x) = \frac{g(x)df(x) - f(x)dg(x)}{g^2(x)} \iff \left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$$

It is clear that $c \cdot df(x) = d(cf)(x)$, it is clear that the derivative is a linear operator from the space of all functions differentiable at x_0 the space of all functions.

Proof.

Since f and g are differentiable at x , there exists the differential $df(x)(h) = f'(x)h$ and $dg(x)(h) = g'(x)h$ where

$$\begin{aligned}f(x+h) &= f(x) + df(x)(h) + o(h) = f(x) + f'(x)h + o(h) \\ g(x+h) &= g(x) + dg(x)(h) + o(h) = g(x) + g'(x)h + o(h)\end{aligned}$$

From this relation, we can clearly see that a certain property of the differential automatically implies the same property of the derivative. (Remember that $f'(x)$ and $g'(x)$ are not functions! They are scalars defined on fixed point x .)

1. Even though this derivation may be a bit long, every step is included to minimize ambiguity.

$$\begin{aligned}
 (f + g)(x + h) - (f + g)(x) &= (f(x + h) + g(x + h)) - (f(x) + g(x)) \\
 &= (f(x + h) - f(x)) + (g(x + h) - g(x)) \\
 &= (df(x)(h) + o(h)) + (dg(x)(h) + o(h)) \\
 &= (f'(x)h + o(h)) + (g'(x)h + o(h)) \\
 &= (f'(x) + g'(x))h + o(h) \\
 &= (f' + g')(x)(h) + o(h) \\
 &= d(f + g)(x)h + o(h)
 \end{aligned}$$

2. For the product rule, we have

$$\begin{aligned}
 (f \cdot g)(x + h) - (f \cdot g)(x) &= f(x + h)g(x + h) - f(x)g(x) \\
 &= (f(x) + df(x)(h) + o(h))(g(x) + dg(x)(h) + o(h)) - f(x)g(x) \\
 &= (f(x) + f'(x)h + o(h))(g(x) + g'(x)h + o(h)) - f(x)g(x)
 \end{aligned}$$

Expanding this gives

$$\begin{aligned}
 &(f'(x)g(x) + f(x)g'(x))h + (f(x) + g(x))o(h) + \\
 &\quad f'(x)g'(x)h^2 + (f'(x) + g'(x))ho(h) + (o(h))^2
 \end{aligned}$$

but note that since $f(x), g(x), f'(x), g'(x)$ are constants, we see that

(a) $(f(x) + g(x))o(h) = o(h)$ because

$$\lim_{h \rightarrow 0} \frac{(f(x) + g(x))o(h)}{h} = (f(x) + g(x)) \lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$$

(b) $f'(x)g'(x)h^2 = o(h)$ since

$$\lim_{h \rightarrow 0} \frac{f'(x)g'(x)h^2}{h} = f'(x)g'(x) \lim_{h \rightarrow 0} h = 0$$

(c) $(f'(x) + g'(x))ho(h) = o(h)$ because

$$\lim_{h \rightarrow 0} \frac{(f'(x) + g'(x))ho(h)}{h} = (f'(x) + g'(x)) \lim_{h \rightarrow 0} o(h) = 0$$

In fact, this term is of $o(h^2)$.

(d) We can see that $(o(h))^2 = o(h)$ since

$$\lim_{h \rightarrow 0} \frac{(o(h))^2}{h} = \lim_{h \rightarrow 0} \frac{o(h)}{h} \cdot \lim_{h \rightarrow 0} o(h) = 0 \cdot 0 = 0$$

In fact, $(o(h))^2 = o(h^2)$.

Therefore, the above simplifies to

$$(f \cdot g)(x + h) - (f \cdot g)(x) = (f'(x)g(x) + f(x)g'(x))h + o(h)$$

But this means that the differential (best approximation) $d(f \cdot g)(x)$ must be

$$(f \cdot g)'(x)(h) = (f \cdot g)'(x)h = (f'(x)g(x) + f(x)g'(x))h$$

3. Since the function $g(x) \neq 0$ at point x , then by continuity we can assume that there exists a neighborhood $U(x)$ where the image of that neighborhood does not vanish. That is, we can guarantee that $g(x+h) \neq 0$ for sufficiently small values of h . We assume h is small in the following computations.

$$\begin{aligned} \left(\frac{f}{g}\right)(x+h) - \left(\frac{f}{g}\right)(x) &= \frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)} \\ &= \frac{1}{g(x)g(x+h)}(f(x+h)g(x) - f(x)g(x+h)) \\ &= \left(\frac{1}{g^2(x)} + o(1)\right) \left((f(x) + f'(x)h + o(h))g(x) \right. \\ &\quad \left. - f(x)(g(x) + g'(x)h + o(h)) \right) \\ &= \left(\frac{1}{g^2(x)} + o(1)\right) \left((f'(x)g(x) - f(x)g'(x))h + o(h) \right) \\ &= \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}h + o(h) \end{aligned}$$

Note that here we have used the continuity of g at the point x and the fact that $g(x) \neq 0$ to deduce that

$$\lim_{h \rightarrow 0} \frac{1}{g(x)g(x+h)} = \frac{1}{g^2(x)} \iff \frac{1}{g(x) + g(x+h)} = \frac{1}{g^2(x)} + o(1)$$

where $o(1)$ is infinitesimal as $h \rightarrow 0$.

Theorem 6.2 (Chain Rule for Composite Functions over \mathbb{R})

Let there be functions $f : E_1 \subset \mathbb{R} \rightarrow E_2 \subset \mathbb{R}$ is differentiable at a point $x \in E_1$ and the function $g : E_2 \subset \mathbb{R} \rightarrow \mathbb{R}$ is differentiable at point $y = f(x) \in E_2$, with respective differentials

$$\begin{aligned} df(x) : T_x \mathbb{R} &\longrightarrow T_y \mathbb{R} \\ dg(y) : T_y \mathbb{R} &\longrightarrow T_{g(y)} \mathbb{R} \end{aligned}$$

Then the composite function $g \circ f : E_1 \rightarrow \mathbb{R}$ is differentiable at x , and $d(g \circ f)(x) : T_x \mathbb{R} \rightarrow T_{g \circ f(x)} \mathbb{R}$ is

$$d(g \circ f)(x) = dg(y) \circ df(x) \iff (g \circ f)'(x) = g'(f(x)) \circ f'(x)$$

Proof.

We will denote the increment of the argument with the variables h and t . Then, by differentiability of f and g , we have

$$\begin{aligned} f(x+h) - f(x) &= f'(x)h + o(h) \text{ as } h \rightarrow 0 \\ g(y+t) - g(y) &= g'(y)t + o(t) \text{ as } t \rightarrow 0 \end{aligned}$$

Since the function $o(t)$ can be represented as $o(t) = \gamma(t)t$, where $\gamma = o(1)$ and hence is infinitesimal as $t \rightarrow 0$, meaning that we can assume $\gamma(0) = 0$ (since $o(t)$ is defined for $t = 0$).

We can think of the displacement of x as like a chain reaction: $As x \mapsto x+h, f(x) \mapsto f(x+h)$, which we could interpret as $y \mapsto y+t$ and hence means that $g(y) \mapsto g(y+t)$. So, setting $f(x) = y$ and $f(x+h) = y+t$, by differentiability and hence continuity of f at point x , we can conclude that $t \rightarrow 0$

as $h \rightarrow 0$. So, we have

$$\gamma(f(x+h) - f(x)) = \gamma((y+t) - y) = \gamma(t) = \alpha(h) \rightarrow 0 \text{ as } h \rightarrow 0$$

Thus, we get

$$\begin{aligned} o(t) &= \gamma(t)t = \gamma(f(x+h) - f(x))(f(x+h) - f(x)) \\ &= \alpha(h)(f'(x)h + o(h)) \\ &= \alpha(h)f'(x)h + \alpha(h)o(h) \\ &= o(h) + o(h) = o(h) \text{ as } h \rightarrow 0 \\ (g \circ f)(x+h) - (g \circ f)(x) &= g(f(x+h)) - g(f(x)) \\ &= g(y+t) - g(y) \\ &= g'(y)t + o(t) \\ &= g'(f(x))(f(x+h) - f(x)) + o(f(x+h) - f(x)) \\ &= g'(f(x))(f'(x)h + o(h)) + o(f(x+h) - f(x)) \\ &= g'(f(x))(f'(x)h) + g'(f(x))(o(h)) + o(f(x+h) - f(x)) \end{aligned}$$

Since $g'(f(x))(o(h))$ is really just a constant multiplied by a function that is $o(h)$, it is $o(h)$. $o(f(x+h) - f(x))$. As for $o(f(x+h) - f(x))$, we see that since $f(x+h) - f(x) = t$, a function that is $o(f(x+h) - f(x))$ becomes infinitesimal compared to t as $t \rightarrow 0$. As already stated before, we have

$$o(f(x+h) - f(x)) = o(h) \text{ as } h \rightarrow 0$$

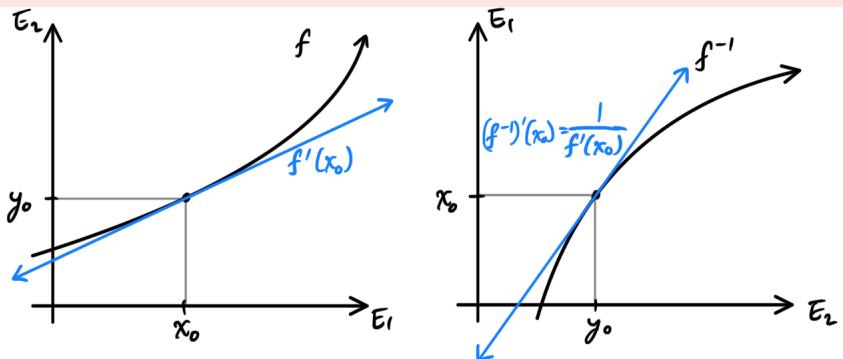
and thus, we proved that

$$\begin{aligned} (g \circ f)(x+h) - (g \circ f)(x) &= g'(y)f'(x)h + o(h) \\ &= (dg(y) \circ df(x))(h) + o(h) \end{aligned}$$

Theorem 6.3 (Differentiation of Inverse Functions over \mathbb{R})

Let $E_1, E_2 \subset \mathbb{R}$, and $f : E_1 \rightarrow E_2$ and $f^{-1} : E_2 \rightarrow E_1$ be mutually inverse and continuous at points $x_0 \in E_1$ and $f(x_0) = y_0 \in E_2$. If f is differentiable at x_0 and $f'(x_0) \neq 0$, then f^{-1} also differentiable at the point y_0 , and

$$(f^{-1})^{-1}(y_0) = (f'(x_0))^{-1} \iff df^{-1}(y_0) = (df(x_0))^{-1}$$



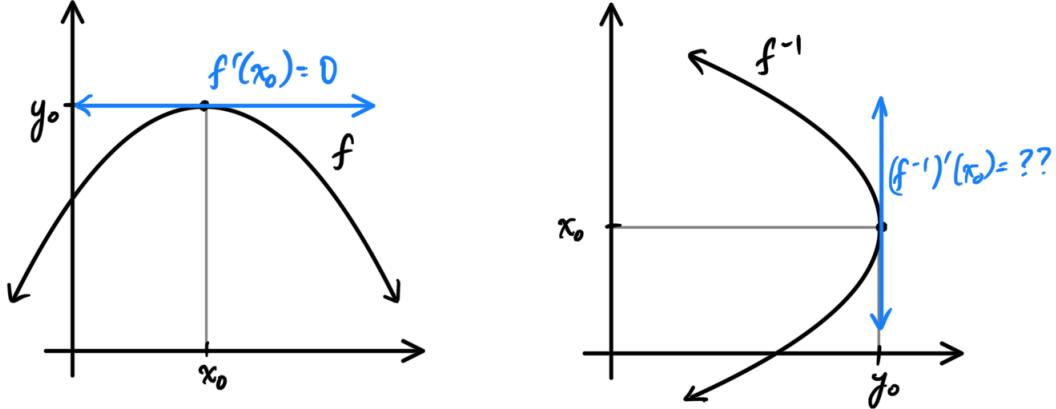
Note that if we knew in advance that f^{-1} was differentiable at y_0 (which is a stronger hypothesis), we can find immediately by the identity

$$(f^{-1} \circ f)(x) = x$$

and the theorem on the differentiation of a composite function that

$$(f^{-1})'(y_0) \cdot f'(x_0) = 1$$

Note that if the hypothesis was satisfied, but $f'(x_0) = 0$, then f^{-1} would not be differentiable since it would have an undefined differential.



6.3.2 Higher-Order Derivatives

Definition 6.6 (Global Derivative Function)

If function $f : E \rightarrow \mathbb{R}$ is differentiable at every point $x \in E$, then a new function $f' : E \rightarrow \mathbb{R}$, whose value at a point $x \in E$ equals the derivative $f'(x)$ of the function f at the point.

Definition 6.7 (Second, Nth Derivative)

This function f' may itself have a derivative $(f')' : E \rightarrow \mathbb{R}$, called the **second derivative** of the original function f , denoted

$$f''(x), \quad \frac{d^2 f(x)}{dx^2}$$

By induction, if the derivative $f^{(n-1)}(x)$ of order $n - 1$ of f has been defined, then the **derivative of order n** is defined by the formula

$$f^{(n)}(x) \equiv (f^{(n-1)})'(x)$$

The set of function $f : E \rightarrow \mathbb{R}$ having continuous derivatives up to order n inclusive is denoted $C^n(E, \mathbb{R})$.

Lemma 6.1 (Leibniz' Formula)

Let $u(x)$ and $v(x)$ be functions having derivatives up to order n inclusive on a common set E . Then,

$$(uv)^{(n)} = \sum_{m=0}^n \binom{n}{m} u^{(n-m)} v^{(m)}$$

This means that given a polynomial $P_n(x) = c_0 + c_1(x - x_0) + \dots + c_n(x - x_0)^n$, then

$$\begin{aligned} P_n(x_0) &= 0 \\ P'_n(x_0) &= 1!c_1 \\ P''_n(x_0) &= 2!c_2 \\ &\dots = \dots \\ P_n^{(n)}(x_0) &= n!c_n \\ P_n^{(k)}(x_0) &= 0 \text{ for } k > n \end{aligned}$$

and thus the polynomial $P_n(x)$ can be written as

$$P_n(x) = P_n^{(0)}(x_0) + \frac{1}{1!}P_n^{(1)}(x_0)(x - x_0) + \frac{1}{2!}P_n^{(2)}(x_0)(x - x_0)^2 + \dots + \frac{1}{n!}P_n^{(n)}(x_0)(x - x_0)^n$$

6.4 Theorems of Differential Calculus

6.4.1 Fermat's Lemma, Rolle's Theorem

Definition 6.8 (Local Extrema)

A point $x_0 \in E \subset \mathbb{R}$ is called a **local maximum** (resp. **local minimum**) and the value of a function $f : E \rightarrow \mathbb{R}$ at that point a **local maximum value** (resp. **local minimum value**) if there exists a neighborhood $U_E(x_0)$ of x_0 in E such that at any point $x \in U_E(x_0)$ we have

$$f(x) \leq f(x_0), (\text{resp. } f(x) \geq f(x_0))$$

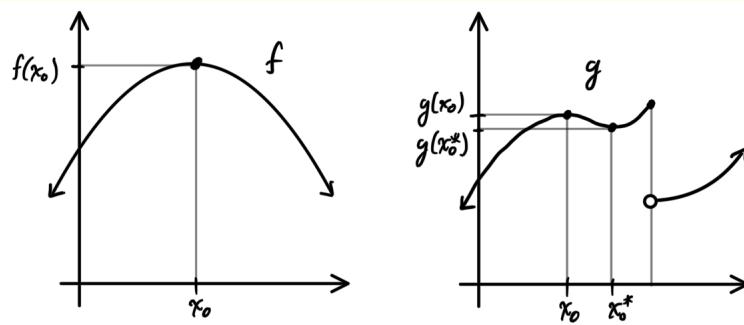
If this is a strict inequality

$$f(x) < f(x_0), (\text{resp. } f(x) > f(x_0))$$

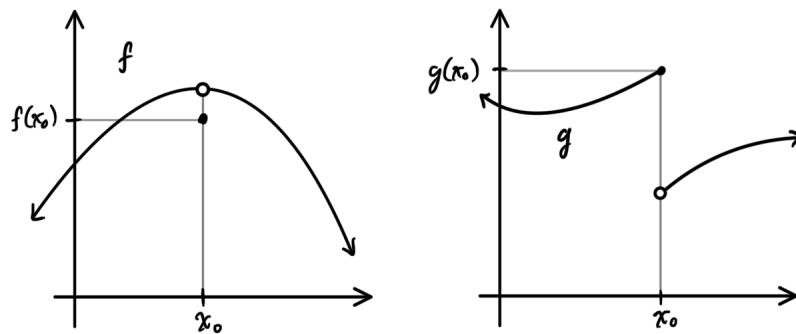
then x_0 is called a **strict local maximum** (resp. **strict local minimum**).

Definition 6.9 (Interior Extrema)

An extremum $x_0 \in E$ of the function $f : E \rightarrow \mathbb{R}$ is called an **interior extremum** if x_0 is not on the boundary of E , or more rigorously, x_0 is a limit point of both sets $E_- = \{x \in E \mid x < x_0\}$ and $E_+ = \{x \in E \mid x > x_0\}$. For example, the graphs below are interior extrema at x_0, x_0^* .



But the following graphs show extrema (at x_0) that are not interior extrema.

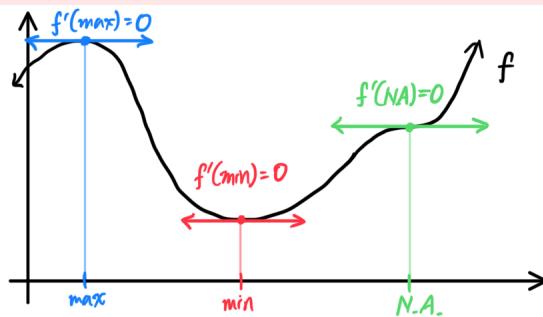


Lemma 6.2 (Fermat)

If a function $f : E \rightarrow \mathbb{R}$ is differentiable at an interior extremum $x_0 \in E$, then its derivative at x_0 is 0. That is,

$$f'(x_0) = 0$$

Thus, this lemma gives a necessary condition for an interior extremum of a differentiable function. But for non-interior extrema, it is generally not true that $f'(x_0) = 0$ and so the converse does not hold (labeled in green).



Proof.

By definition of differentiability at x_0 we get

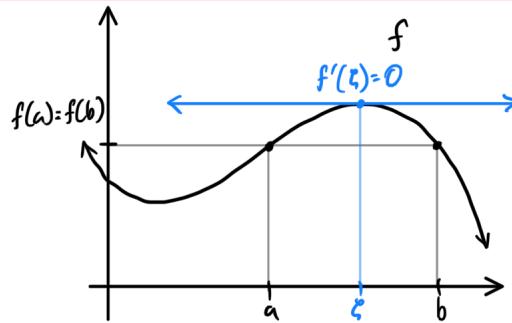
$$\begin{aligned} f(x_0 + h) - f(x_0) &= f'(x_0)h + o(h) \\ &= f'(x_0)h + \alpha(x_0; h)h \\ &= (f'(x_0) + \alpha(x_0; h))h \end{aligned}$$

where we know that $o(h)$ can be written as $o(1)h$ for some infinitesimal $o(1)$ as $h \rightarrow 0$. If $f'(x_0) \neq 0$, then for h sufficiently close to 0 the quantity $f'(x_0) + \alpha(x_0; h)$ would have the same sign as $f'(x_0)$, since $\alpha(x_0; h) \rightarrow 0$ as $h \rightarrow 0$. But the value of h can be both positive or negative, given that x_0 is an interior extremum. This contradiction must imply that $f'(x_0) = 0$.

Geometrically, Fermat's lemma is obvious, since it asserts that at an extremum of a differentiable function, the tangent to its graph is horizontal. Physically, this lemma means that in motion along a line the velocity must be zero at the instant then the direction reverses.

Theorem 6.4 (Rolle's Theorem)

If a function $f : [a, b] \rightarrow \mathbb{R}$ is continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) and $f(a) = f(b)$, then there exists a point $\zeta \in (a, b)$ such that $f'(\zeta) = 0$.

**6.4.2 Mean Value Theorem, Cauchy's Finite-Increment Theorem**

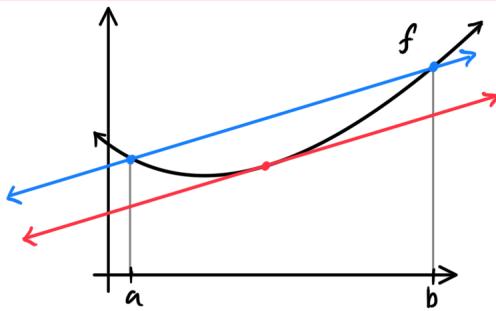
The following theorem is extremely useful in studying numerical valued functions.

Theorem 6.5 (Mean Value Theorem)

If a function $f : [a, b] \rightarrow \mathbb{R}$ is continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) , there exists a point $\zeta \in (a, b)$ such that

$$f'(\zeta) = \frac{f(b) - f(a)}{b - a} \iff f(b) - f(a) = f'(\zeta)(b - a)$$

Geometrically, this means that there exists a tangent line somewhere at $\zeta \in (a, b)$ that is parallel the secant line connecting the two points $(a, f(a))$ and $(b, f(b))$.



Some remarks:

1. Physically, if x is interpreted as time and $f(b) - f(a)$ as the amount of displacement over the time $b - a$ of a particle moving along the line, this theorem says that the velocity $f'(x)$ of the particle at some time $\zeta \in (a, b)$ is such that if the particle had moved with constant velocity $f'(\zeta)$ over the whole time interval, it would have been displaced by the same amount $f(b) - f(a)$. We call $f'(\zeta)$ the **average velocity** over the time interval $[a, b]$.
2. Note that the Mean Value Theorem is important in that it connects the increment of a function over a finite interval with the derivative of the function on that interval. Up to now, we have characterized only the local (infinitesimal) increment of a function in terms of the derivative or differential at a given point. MVT connects the increment of a function over a **finite** interval with the derivative of the function.

The MVT actually leads to multiple useful corollaries.

Corollary 6.1 (Derivative of a Monotonic Function)

If the derivative of a function is nonnegative (resp. positive) at every point of an open interval, then the function is nondecreasing (resp. increasing) on that interval.

Proof.

If $x_1 < x_2$ are two points of the interval, then the MVT

$$f(x_2) - f(x_1) = f'(\zeta)(x_2 - x_1)$$

shows that the sign of the left hand side must equal that of the right.

Corollary 6.2 (Derivative of a Constant Function)

A function that is continuous on a closed interval $[a, b]$ is constant on it if and only if its derivative equals 0 at every point of the interval $[a, b]$ or the open interval (a, b) .

Therefore, if the derivatives $f'_1(x)$ and $f'_2(x)$ of two functions $f_1(x)$ and $f_2(x)$ are equal on some interval (that is, $f'_1(x) = f'_2(x)$ on the interval), then the difference

$$(f_1 - f_2)(x) = f_1(x) - f_2(x)$$

is constant.

Proof.

Given constant function f , the MVT equation

$$0 = f(x_2) - f(x_1) = f'(\zeta)(x_2 - x_1)$$

implies that $f'(\zeta) = 0$ for all $x_1, x_2 \in E$. It follows that by the arithmetic properties of the derivative, given two functions f_1, f_2 with the same derivative on an interval, the derivative of their difference $(f_1 - f_2)' = 0$, and therefore must be constant on that interval.

The following proposition is a useful generalization of Lagrange's theorem.

Theorem 6.6 (Cauchy's Finite-Increment Theorem)

Let $x = x(t)$ and $y = y(t)$ be functions that are continuous on a closed interval $[\alpha, \beta]$ and differentiable on the open interval (α, β) . Then, there exists a point $\tau \in [\alpha, \beta]$ such that

$$x'(\tau)(y(\beta) - y(\alpha)) = y'(\tau)(x(\beta) - x(\alpha))$$

If in addition $x'(t) \neq 0$ for each $t \in (\alpha, \beta)$, then $x(\alpha) \neq x(\beta)$ and we have the equality

$$\frac{y(\beta) - y(\alpha)}{x(\beta) - x(\alpha)} = \frac{y'(\tau)}{x'(\tau)}$$

6.4.3 Taylor's Formula

From the following results one may deduce that the more derivatives of two functions coincide (including the derivative of the 0th order) at a point, the better these functions approximate each other in a neighborhood of that point. Using Leibniz's rule, approximations up to a certain degree at a point can be expressed as a polynomial

$$P_n(x_0; x) = P_n(x_0) + \frac{P'_n(x_0)}{1!}(x - x_0) + \dots + \frac{P_n^{(n)}(x_0)}{n!}(x - x_0)^n$$

where each coefficient of the polynomial

Definition 6.10 (Taylor Polynomial)

If a function $f : E \rightarrow \mathbb{R}$ has derivatives of all orders $n \in \mathbb{N}$ at a point x_0 , the unique series

$$P_n(x_0; x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

is the **Taylor polynomial of order n of $f(x)$ at x_0** . We can see that the derivatives of f and P_n coincide up to order n .

Definition 6.11 (Analytic Functions)

We cannot assume that the Taylor series of an infinitely differentiable function converges to the function f within a neighborhood $U(x_0)$, nor can we assume that it converges at all! These types of "nice" functions that have a Taylor approximation within the neighborhood of x_0 are called **analytic functions** and can be written in the form

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r_n(x_0; x)$$

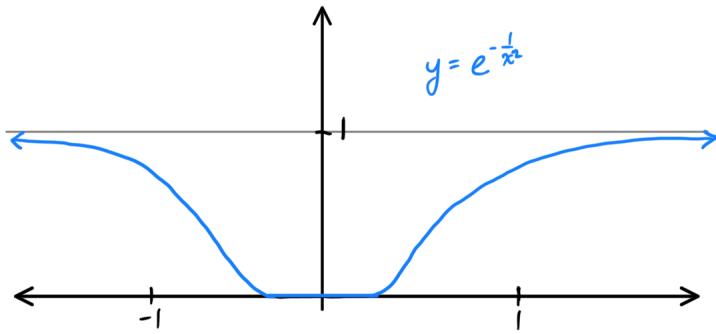
where r is called the **remainder term**.

Example 6.3 (Infinitely Differentiable, Non-Analytic Function)

A example of a non-analytic function is

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

which looks like



One can verify that the derivative $f^{(k)}(0) = 0$ for all k and hence the Taylor series is identically equal to 0, while $f(x) \neq 0$ if $x \neq 0$.

The relationship between these different conditions is nicely summarized in the figure.

f infinitely differentiable at $x_0 \iff$ Taylor series of f exists at x_0

Taylor series converges at x_0

Taylor series converges to f at $x_0 \iff f$ is analytic

The following lemma proves why Taylor Polynomials are considered a "good" approximations to analytic functions.

Lemma 6.3 (Infinitesimality of Functions with Vanishing Derivative up to Order n)

Given a function $\varphi : E \rightarrow \mathbb{R}$ defined on a closed interval E with endpoint x_0 , let its derivatives vanish up to order n at x_0 . That is

$$\varphi(x_0) = \varphi'(x_0) = \dots = \varphi^{(n)}(x_0) = 0$$

Then, $\varphi = o((x - x_0)^n)$ as $x \rightarrow x_0$.

Proof.

We prove by induction. For $n = 1$, the definition of differentiability states that

$$\varphi(x) = \varphi(x_0) + \varphi'(x - x_0) + o(x - x_0) \text{ as } x \rightarrow x_0$$

and so we have proved that

$$\varphi(x_0) = \varphi'(x_0) = 0 \implies \varphi(x) = o(x - x_0) \text{ as } x \rightarrow x_0$$

Now, suppose this assertion has been proved for order $n = k - 1 \geq 1$. That is, we have shown that

$$\varphi(x_0) = \dots = \varphi^{(k-1)}(x_0) = 0 \implies \varphi = o((x - x_0)^{k-1}) \text{ as } x \rightarrow x_0$$

Then we must show that this is valid for order $n = k \geq 2$. Assume that

$$\varphi(x_0) = \varphi'(x_0) = \dots = \varphi^{(k)}(x_0) = 0$$

We can see that this is equivalent to

$$(\varphi')'(x_0) = (\varphi')^{(2)}(x_0) = \dots = (\varphi')^{(k-1)}(x_0) = 0$$

and therefore by the induction assumption, we have

$$\varphi' = o((x - x_0)^{k-1}) \text{ as } x \rightarrow x_0$$

which means that we can put it in form

$$\varphi(x) = \alpha(x)(x - x_0)^{k-1} \text{ so that } \lim_{x \rightarrow x_0} \varphi(x) = \lim_{x \rightarrow x_0} \alpha(x) = 0$$

From the mean value theorem and substituting what we have above, we get

$$\begin{aligned} \varphi(x) &= \varphi(x) - \varphi(x_0) = \varphi'(\zeta)(x - x_0) \\ &= \varphi(\zeta)(\zeta - x_0)^{k-1}(x - x_0) \end{aligned}$$

where $\zeta \in (x_0, x)$. However, this implies that $|\zeta - x_0| < |x - x_0|$, and thus, as $x \rightarrow x_0$, $\zeta \rightarrow x_0$, which then makes $\alpha(\zeta) \rightarrow 0$. Since

$$|\varphi(x)| \leq |\alpha(\zeta)| |x - x_0|^{k-1} |x - x_0| = |\alpha(\zeta)| |x - x_0|^k$$

This means that $\varphi(x)$ is bounded by function $|\alpha(\zeta)| |x - x_0|^k$, which is $o((x - x_0)^k)$, and so

$$\varphi = o((x - x_0)^k) \text{ as } x \rightarrow x_0$$

By induction, this works for all orders n .

Theorem 6.7 (Peano's Form of the Remainder)

Given analytic function $f : E \rightarrow \mathbb{R}$, a point $x_0 \in E$, and its n th order Taylor polynomial $P_n(x_0; x)$ around x_0 , P_n is a "good" approximation of f in the fact that its error term is $o((x - x_0)^n)$. That is,

$$f(x) = P_n(x_0; x) + o((x - x_0)^n) \text{ as } x \rightarrow x_0$$

This equation where $r_n(x; x_0) = o((x - x_0)^n)$ is called the **Peano's form of the remainder**.

Proof.

Since the Taylor polynomial $P_n(x_0; x)$ is constructed from the requirement that its derivatives up to order n inclusive must coincide with the corresponding derivatives of f at x_0 , it follows that

$$r_n(x_0; x_0) \equiv f^{(k)}(x_0) - P_n^{(k)}(x_0; x_0) = 0 \text{ for } k = 0, 1, \dots, n$$

Using the previous lemma, a this means that $r_n(x; x_0) = o((x - x_0)^n)$ as $x \rightarrow x_0$.

Theorem 6.8 (Lagrange Form of the Remainder)

If $f : E \rightarrow \mathbb{R}$ has derivatives of order $n + 1$ on the open interval with endpoints x_0 and x , then

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r_n(x; x_0)$$

where

$$r_n(x; x_0) = \frac{f^{(n+1)}(\zeta)}{(n+1)!}(x - x_0)^{n+1}$$

This form is called **Taylor's formula with the Lagrange form of the remainder**. Furthermore, this form says that if $f^{(n+1)}(x)$ is bounded in a neighborhood of x_0 , it also implies the formula

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + O((x - x_0)^{n+1})$$

Therefore, we can use this boundedness of $f^{(n+1)}$ to find the maximum error bound

$$|r_n(x; x_0)|$$

of $P_n(x; x_0)$.

Proof.

It is a direct result from the lemma. This is actually a generalization of the mean value theorem but for higher orders.

Corollary 6.3 (Table of Asymptotic Formulas for Elementary Functions)

We write the Maclaurin series (Taylor series around $x = 0$) for elementary functions. Note that these error terms are $O(x^{n+1})$ (bounded compared to x^{n+1}) and $o(x^n)$ (infinitesimal compared to x^n).

$$\begin{aligned} e^x &= 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \dots + \frac{1}{n!}x^n + O(x^{n+1}) \\ \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \dots + \frac{(-1)^n}{(2n)!}x^{2n} + O(x^{2n+2}) \\ \sin x &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \dots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + O(x^{2n+3}) \\ \cosh x &= 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots + \frac{1}{(2n)!}x^{2n} + O(x^{2n+2}) \\ \sinh x &= x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \dots + \frac{1}{(2n+1)!}x^{2n+1} + O(x^{2n+3}) \\ \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots + \frac{(-1)^n}{n}x^n + O(x^{n+1}) \\ (1+x)^\alpha &= 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n + O(x^{n+1}) \end{aligned}$$

6.5 The Study of Functions using Differential Calculus

6.5.1 Conditions for Monotonicity of Functions

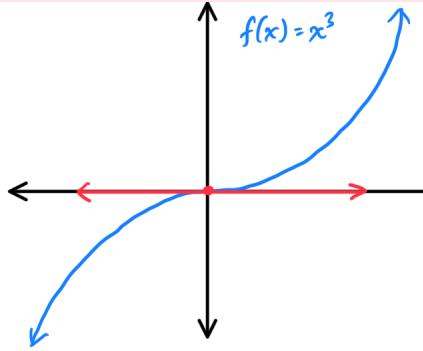
We can now connect the concepts of derivatives and monotonicity.

Theorem 6.9 (Derivative \Rightarrow Monotonicity)

Given function $f : E \rightarrow \mathbb{R}$ that is differentiable on an open interval $(a, b) = E$,

- $f'(x) > 0 \implies f$ is increasing
- $f'(x) \geq 0 \iff f$ is nondecreasing
- $f'(x) \equiv 0 \iff f$ is constant
- $f'(x) \leq 0 \iff f$ is nonincreasing
- $f'(x) < 0 \implies f$ is decreasing

Note that if f is strictly increasing (resp. decreasing), we cannot determine that $f'(x) \geq 0$ (resp. $f'(x) \leq 0$). For example, take the function $f(x) = x^3$, which is strictly increasing, but has derivative $f'(0) = 0$ at $x = 0$.



It is clearly strictly increasing within a neighborhood $U(0)$, so we can see that

$$\begin{aligned} f \text{ is increasing} &\implies f'(x) \geq 0 \\ f \text{ is decreasing} &\implies f'(x) \leq 0 \end{aligned}$$

6.5.2 Conditions for Extrema of Functions

Similarly, we can connect the concepts of extrema and derivatives.

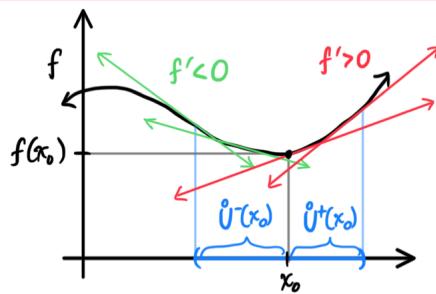
Theorem 6.10 (First Derivative Test)

Let function $f : E \rightarrow \mathbb{R}$ be defined in a neighborhood $U(x_0)$ of point x_0 , which is continuous at x_0 and differentiable in $\dot{U}(x_0)$, a deleted neighborhood of x_0 . (Note that this is broader hypothesis than just assuming that f be differentiable at x_0 .) Let

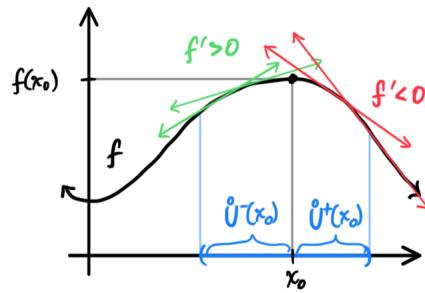
$$\dot{U}^-(x_0) \equiv \{x \in U(x_0) \mid x < x_0\}, \quad \dot{U}^+(x_0) \equiv \{x \in U(x_0) \mid x > x_0\}$$

That is, $\dot{U}^-(x_0)$ is the left portion of $\dot{U}(x_0)$ and $\dot{U}^+(x_0)$ is the right portion of $\dot{U}(x_0)$. Then,

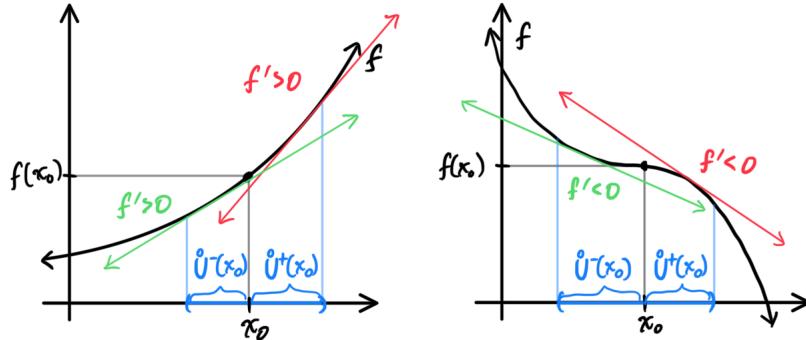
1. $(x_0, f(x_0))$ is strict local minimum if $f'(x) < 0$ in $\dot{U}^-(x_0)$ and $f'(x) > 0$ in $\dot{U}^+(x_0)$.



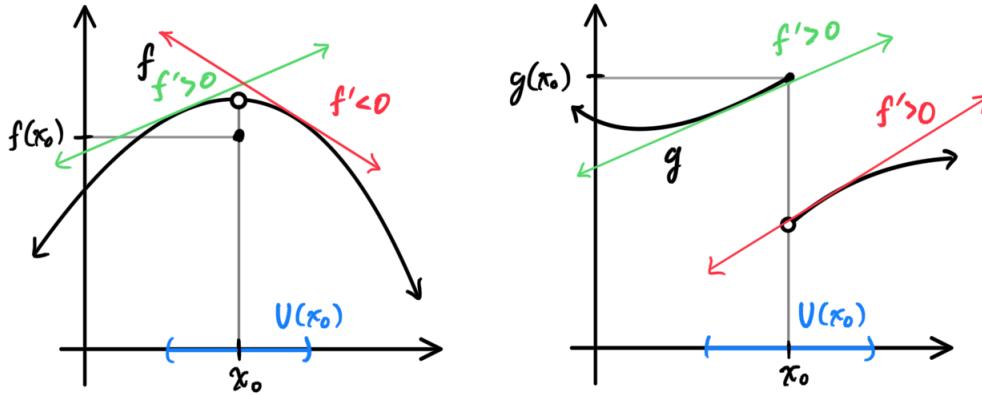
2. $(x_0, f(x_0))$ is strict local maximum if $f'(x) > 0$ in $\dot{U}^-(x_0)$ and $f'(x) < 0$ in $\dot{U}^+(x_0)$.



3. $(x_0, f(x_0))$ has no extremum at x_0 if $f'(x) > 0$ in both $\dot{U}^-(x_0), \dot{U}^+(x_0)$, or if $f'(x) < 0$ in both $\dot{U}^-(x_0), \dot{U}^+(x_0)$.



Note that if there is a discontinuity at a point x_0 , then this theorem does not apply. For example, $(x_0, f(x_0))$ in the graph below is a local minimum, even though the derivatives to the left of x_0 are positive and those to the right of x_0 are negative (within neighborhood $U(x_0)$). Similarly, $(x_0, g(x_0))$ is a local maximum, even though the derivative to the left and to the right of x_0 are both positive.



Proposition 6.1 (2nd, n th Derivative Test)

Let function $f : E \rightarrow \mathbb{R}$ be defined on a neighborhood $U(x_0)$ of x_0 has derivatives of order up to n inclusive at x_0 . If its derivatives up to the $(n-1)$ th order vanishes

$$f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0$$

but the n th derivative at x_0 does **not** vanish

$$f^{(n)}(x_0) \neq 0$$

then

1. n is odd \implies there is no local extremum at x_0
2. n is even \implies there is a local extremum at x_0
 - (a) $f^{(n)}(x_0) > 0 \implies$ it is a strict local minimum
 - (b) $f^{(n)}(x_0) < 0 \implies$ it is a strict local maximum

6.5.3 Important Algebraic Inequalities

We also introduce various inequalities that may be useful for producing future results. The following lemmas can be proved with elementary algebra.

Lemma 6.4 (Young's Inequalities)

If $a > 0$ and $b > 0$, and the numbers p and q are such that $p \neq 0, 1$ and $q \neq 0, 1$, and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{1}{p}a + \frac{1}{q}b \text{ if } p > 1$$

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \geq \frac{1}{p}a + \frac{1}{q}b \text{ if } p < 1$$

and equality holds in both cases if and only if $a = b$.

Lemma 6.5 (Holder's Inequalities)

Let $x_i \geq 0, y_i \geq 0$ for $i = 1, 2, \dots, n$, and let $\frac{1}{p} + \frac{1}{q} = 1$. Then,

$$\sum_{i=1}^n x_i y_i \leq \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q \right)^{\frac{1}{q}} \text{ for } p > 1$$

$$\sum_{i=1}^n x_i y_i \geq \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q \right)^{\frac{1}{q}} \text{ for } p < 1, p \neq 0$$

Lemma 6.6 (Minkowski's Inequalities)

Let $x_i \geq 0, y_i \geq 0$ for $i = 1, 2, \dots, n$. Then,

$$\left(\sum_{i=1}^n (x_i + y_i)^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n y_i^p \right)^{\frac{1}{p}} \text{ when } p > 1$$

$$\left(\sum_{i=1}^n (x_i + y_i)^p \right)^{\frac{1}{p}} \geq \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n y_i^p \right)^{\frac{1}{p}} \text{ when } p < 1, p \neq 0$$

6.5.4 Conditions for a Function to be Convex**Definition 6.12 (Convex, Concave Functions)**

A function $f : (a, b) \rightarrow \mathbb{R}$ defined on an open interval $(a, b) \subset \mathbb{R}$ is **convex** if the inequality

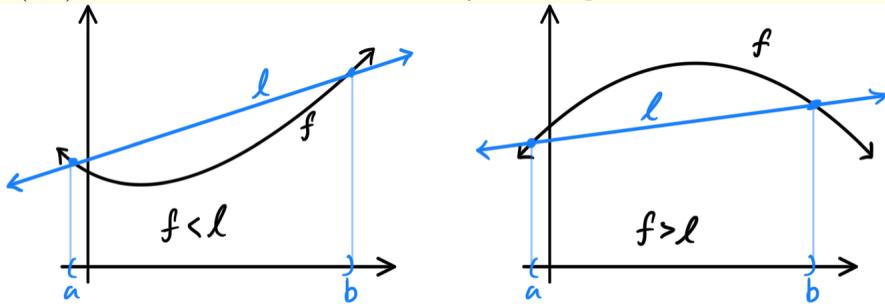
$$f(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2)$$

holds and **concave**, or **convex upward**, if the inequality

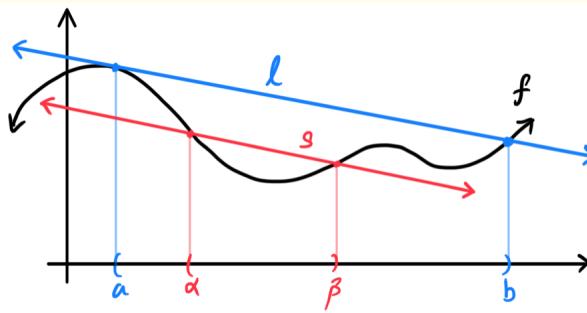
$$f(\alpha_1 x_1 + \alpha_2 x_2) \geq \alpha_1 f(x_1) + \alpha_2 f(x_2)$$

holds for all pairs of points $x_1, x_2 \in (a, b)$ and any numbers $\alpha_1, \alpha_2 \geq 0$ such that $\alpha_1 + \alpha_2 = 1$. If this inequality is strict whenever $x_1 \neq x_2$ and $\alpha_1 \alpha_2 \neq 0$, the function is said to be **strictly convex** and **strictly concave**, respectively.

Visually, this just means that given any two points a, b , the graph of a convex function (left) in (a, b) always lies underneath the secant line formed by the two points and the graph of a concave function (right) in (a, b) lies over the secant line formed by the two points.



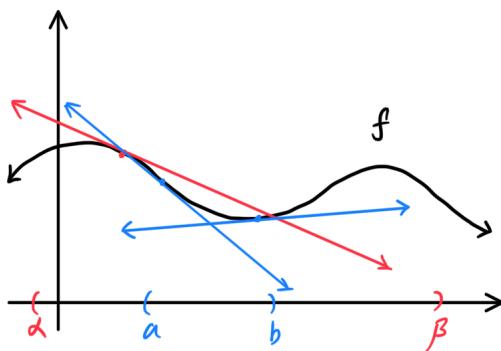
However, this is only a visual aid. In reality, it is actually not only the secant line formed by the two endpoints, but every pairs of points within that interval. For example, even though the secant line l formed by the endpoints a, b is above the whole graph in (a, b) , f is not convex over (a, b) since the secant line formed by points α, β do not lie completely underneath f .



The following is also another equivalent condition for a function to be convex over (a, b) .

Proposition 6.2 ()

A function $f : (a, b) \rightarrow \mathbb{R}$ that is differentiable on the open interval (a, b) is convex on (a, b) if and only if its graph contains no points below any tangent drawn to it.



Theorem 6.11 (2nd Derivatives of Convex Functions)

Given a function $f : (a, b) \rightarrow \mathbb{R}$ that is differentiable in its domain,

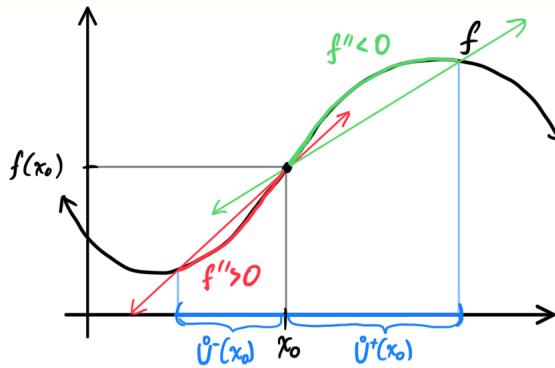
1. f is convex $\iff f'$ is nondecreasing on $(a, b) \iff f'' \geq 0$ on (a, b)
2. f is strictly convex $\iff f'$ is increasing on $(a, b) \iff f'' > 0$ on (a, b)
3. f is concave $\iff f'$ is nonincreasing on $(a, b) \iff f'' \leq 0$ on (a, b)
4. f is strictly concave $\iff f'$ is decreasing on $(a, b) \iff f'' < 0$ on (a, b)

Definition 6.13 (Inflection Point)

Let $f : E \rightarrow \mathbb{R}$ be a function defined and differentiable on a neighborhood $U(x_0)$. If the function is convex downward (resp. upward) on the set $\mathring{U}^-(x_0) = \{x \in U(x_0) \mid x < x_0\}$ and convex upward (resp. downward) on $\mathring{U}^+(x_0) = \{x \in U(x_0) \mid x > x_0\}$, then the point

$$(x_0, f(x_0))$$

is called a **inflection point of the graph**.



Proposition 6.3 (Jensen's Inequality)

If $f : (a, b) \rightarrow \mathbb{R}$ is a convex function, x_1, \dots, x_n are points of (a, b) , and $\alpha_1, \dots, \alpha_n$ are nonnegative numbers such that $\alpha_1 + \dots + \alpha_n = 1$, then

$$f(\alpha_1 x_1 + \dots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \dots + \alpha_n f(x_n)$$

6.5.5 L'Hopital's Rule

Theorem 6.12 (L'Hopital's Rule)

Let c be an extended real number (i.e. $c \in \mathbb{R} \cup \{+\infty, -\infty\}$) and let (a, b) be an open interval containing c (for a two-sided limit) or an open interval with endpoint c (for a one-sided limit, or a limit at infinity if c is infinite). Assume that f and g are assumed to be differentiable on $(a, b) \setminus c$, and additionally $g'(x) \neq 0$ on $(a, b) \setminus c$. If either

$$\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0 \text{ or } \lim_{x \rightarrow c} |f(x)| = \lim_{x \rightarrow c} |g(x)| = \infty$$

then

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}$$

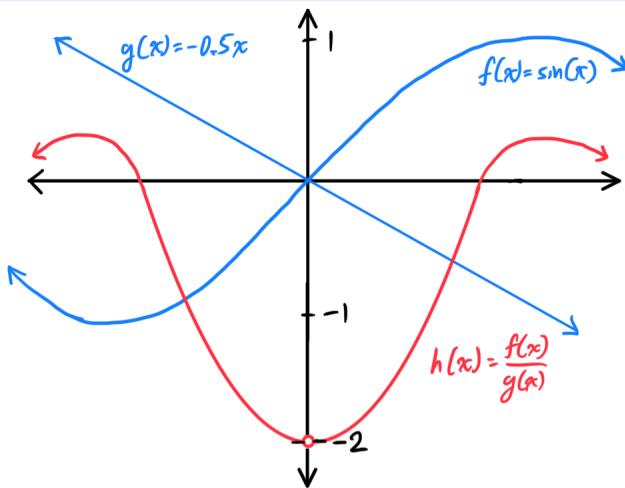
L'Hopital's rule can be stated colloquially, but not quite accurately, as follows: **The limit of a ratio of functions equals the limit of the ratio of their derivatives if their derivatives exist.**

Example 6.4 ()

Let $f(x) = \sin x$ and $g(x) = -0.5x$. Then, the function

$$h(x) = \frac{f(x)}{g(x)} = \frac{\sin x}{-0.5x}$$

is clearly undefined at $x = 0$.

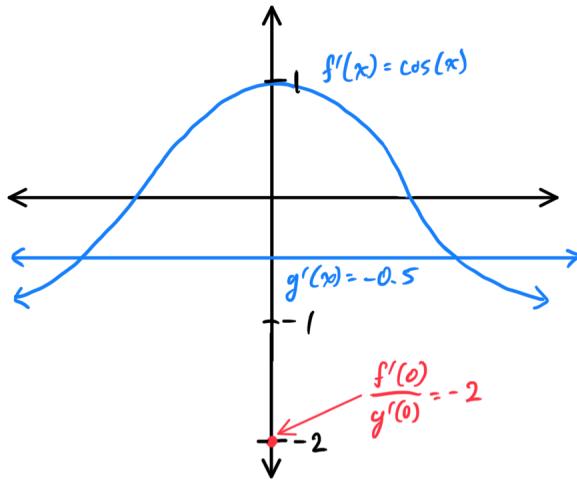


However, we can solve the limit using L'Hopital's rule to get

$$\lim_{x \rightarrow 0} \frac{\sin x}{-0.5x} = \lim_{x \rightarrow 0} \frac{\cos x}{-0.5} = -2$$

Therefore, $h : \mathbb{R} \setminus 0 \rightarrow \mathbb{R}$ can be completed to continuous function on all of \mathbb{R} by defining the extension:

$$H(x) \equiv \begin{cases} h(x), & x \neq 0 \\ -2, & x = 0 \end{cases}$$



6.6 Complex Analysis: An Introduction

Just as the equation $x^2 = 2$ has no solutions in the domain \mathbb{Q} of rational numbers, the equation $x^2 = -1$ has no solutions in the domain \mathbb{R} of real numbers. Just as we adjoin the symbol $\sqrt{2}$ as a solution of $x^2 = 2$ and connect it with rational numbers to get new numbers of the form

$$r_1 + r_2\sqrt{2}, \quad r_1, r_2 \in \mathbb{Q}$$

we introduce the symbol i as a solution of $x^2 = -1$ and attach this number to real numbers.

One feature of this enlargement of the field \mathbb{R} of real numbers into the resulting field \mathbb{C} of complex numbers, every algebraic equation with real or complex coefficients now has a solution.

6.6.1 Algebraic Extension of Field \mathbb{R}

We introduce the number i , called the **imaginary unit**, such that $i^2 = -1$. We may multiply real numbers y to i to get yi , and we can add real numbers to such numbers, to get numbers of the form

$$x + yi, \quad x, y \in \mathbb{R}$$

We then define all objects of the form $x + iy$ as the **complex numbers**, with addition defined

$$(x_1 + iy_1) + (x_2 + iy_2) \equiv (x_1 + x_2) + i(y_1 + y_2)$$

and multiplication defined

$$(x_1 + iy_1) \cdot (x_2 + iy_2) \equiv (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

As expected, this makes $+$ and \cdot commutative operations. Furthermore, two complex numbers $z = x_1 + iy_1$ and $w = x_2 + iy_2$ are equal if and only if $x_1 = x_2$ and $y_1 = y_2$.

One nontrivial property of field \mathbb{C} is that every element $z \in \mathbb{C}$ has a multiplicative inverse z^{-1} . To find this, we must define the following.

Definition 6.14 (Complex Conjugate)

Given complex number $z = x + iy$, its **complex conjugate** is

$$\bar{z} = \overline{x + iy} = x - iy$$

Note that

$$z \cdot \bar{z} = x^2 + y^2 \neq 0 \text{ iff } z \neq 0$$

Thus, given z ,

$$z^{-1} = \frac{1}{z \cdot \bar{z}} \cdot \bar{z} \iff (x + yi)^{-1} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}$$

6.6.2 Geometric Interpretation of \mathbb{C}

Once the algebraic operations $+$ and \cdot has been introduced, the symbol i is no longer needed. That is, we can define a new set $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ with the operations $+_{\mathbb{R}}, \cdot_{\mathbb{R}} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined

$$\begin{aligned} (x_1, y_1) +_{\mathbb{R}} (x_2, y_2) &\equiv (x_1 + x_2, y_1 + y_2) \\ (x_1, y_1) \cdot_{\mathbb{R}} (x_2, y_2) &\equiv (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) \end{aligned}$$

We can check that this new set $(\mathbb{R}^2, +_{\mathbb{R}}, \cdot_{\mathbb{R}})$ is isomorphic to $(\mathbb{C}, +, \cdot)$ as fields, and therefore one can identify complex numbers with vectors $z = (x, y)$ of the plane \mathbb{R}^2 , where $x = \operatorname{Re} z$ is called the **real part** and $y = \operatorname{Im} z$ is called the **imaginary part**.

Definition 6.15 (Norm, Metric of \mathbb{C})

Moreover, the isomorphism

$$\gamma : \mathbb{C} \rightarrow \mathbb{R}^2, \quad \gamma(x + yi) = (x, y)$$

induces additional structures on \mathbb{C} , such as the norm and metric.

1. The norm of $z = x + yi \in \mathbb{C}$ is defined as the norm of $\gamma(z) = (x, y) \in \mathbb{R}^2$. That is,

$$|z| = |x + yi| = |(x, y)| = \sqrt{x^2 + y^2}$$

Or more simply,

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

2. The metric of two complex numbers $z_1, z_2 \in \mathbb{C}$ is defined

$$|z_1 - z_2| = |(x_1, y_1) - (x_2, y_2)| = |(x_1 - x_2, y_1 - y_2)| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Or more simply,

$$|z_1 - z_2| = (z_1 - z_2) \cdot \overline{(z_1 - z_2)}$$

Definition 6.16 (Polar Coordinates of \mathbb{C})

Given the basis transformation of polar coordinates $(r, \varphi) \mapsto p(r, \varphi) = (x, y)$ where

$$p \begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

the isomorphism $\mathbb{C} \simeq \mathbb{R}^2$ induces a similar polar transformation in \mathbb{C}

$$\rho = \gamma^{-1} \circ p \circ \gamma : \mathbb{C}_{(r, \theta)} \longrightarrow \mathbb{C}_{(x, y)}, \quad \rho(r + \theta i) = r \cos \theta + r \sin \theta i = x + yi$$

as shown in the commutative diagram.

$$\begin{array}{ccc} \mathbb{C}_{(r, \theta)} & \xrightarrow{\rho} & \mathbb{C}_{(x, y)} \\ \downarrow \gamma & & \downarrow \gamma \\ \mathbb{R}_{(r, \theta)}^2 & \xrightarrow{p} & \mathbb{R}_{(x, y)}^2 \end{array}$$

Therefore, we can write

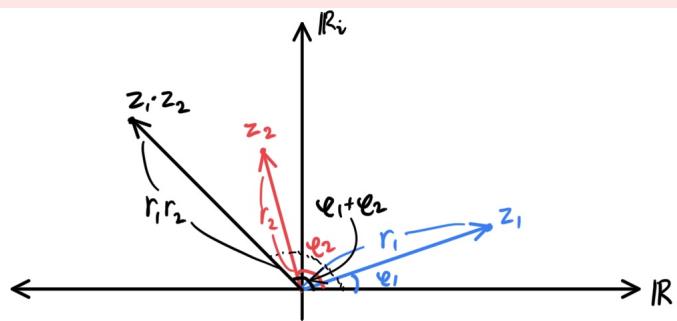
$$z = r(\cos \varphi + i \sin \varphi)$$

where $r = |z|$ is called the **magnitude** of z , and $\varphi = \text{Arg } z$ is called the **argument** of z .

Lemma 6.7 (Multiplication of Complex Numbers in Polar Form)

It turns out that multiplication is a lot easier in polar coordinates than in rectangular ones:

$$\begin{aligned} z_1 \cdot z_2 &= (r_1 \cos \varphi_1 + ir_1 \sin \varphi_1)(r_2 \cos \varphi_2 + ir_2 \sin \varphi_2) \\ &= \dots \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)) \end{aligned}$$



Theorem 6.13 (De Moivre's Formula)

By induction using the previous lemma, we get

$$z = r(\cos \varphi + i \sin \varphi) \implies z^n = r^n(\cos n\varphi + i \sin n\varphi)$$

Corollary 6.4 (Roots of Unity)

The n complex solutions of the equation

$$z^n = a$$

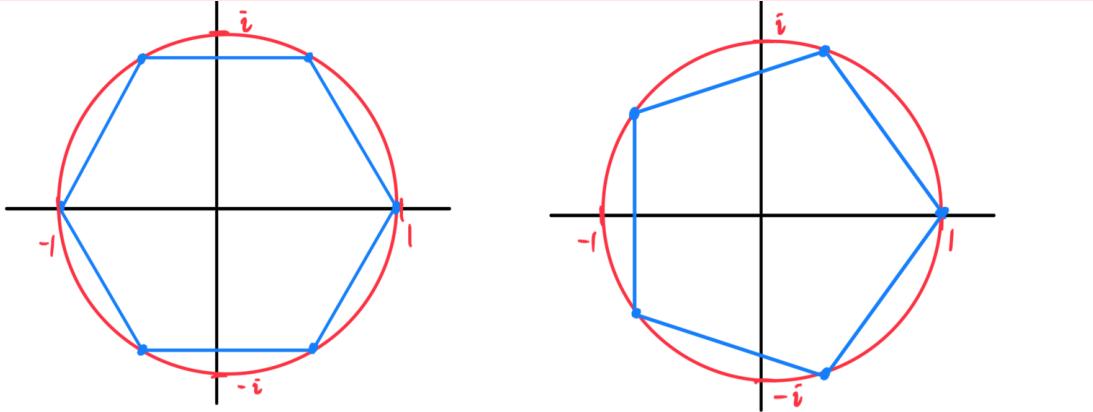
where $a = \rho(\cos \psi + i \sin \psi)$ is

$$z_k = \sqrt[n]{\rho} \left(\cos \left(\frac{\psi + 2\pi k}{n} \right) + i \sin \left(\frac{\psi + 2\pi k}{n} \right) \right), \quad k = 0, 1, 2, \dots, n-1$$

Moreover, if $a = 1$, then the n complex solutions are called the **n th roots of unity**, defined

$$z_k = \cos \left(\frac{2\pi k}{n} \right) + i \sin \left(\frac{2\pi k}{n} \right), \quad k = 0, 1, 2, \dots, n-1$$

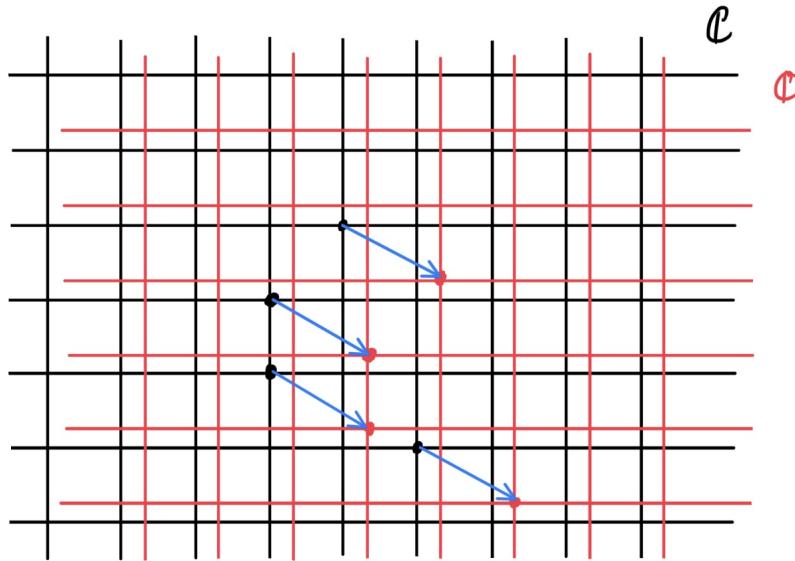
which shows that the n th roots of unity are at the vertices of a regular n -sided polygon inscribed in the unit circle, with one vertex at 1, within the complex plane. The 5th and 6th roots of unity are shown below.



Finally, we can visualize certain transformations in \mathbb{C} . For a fixed $b \in \mathbb{C}$, the sum $z + b$ can be interpreted as the mapping of \mathbb{C} onto itself given by the formula

$$z \mapsto z + b$$

This mapping is a translation of the plane by the vector b .



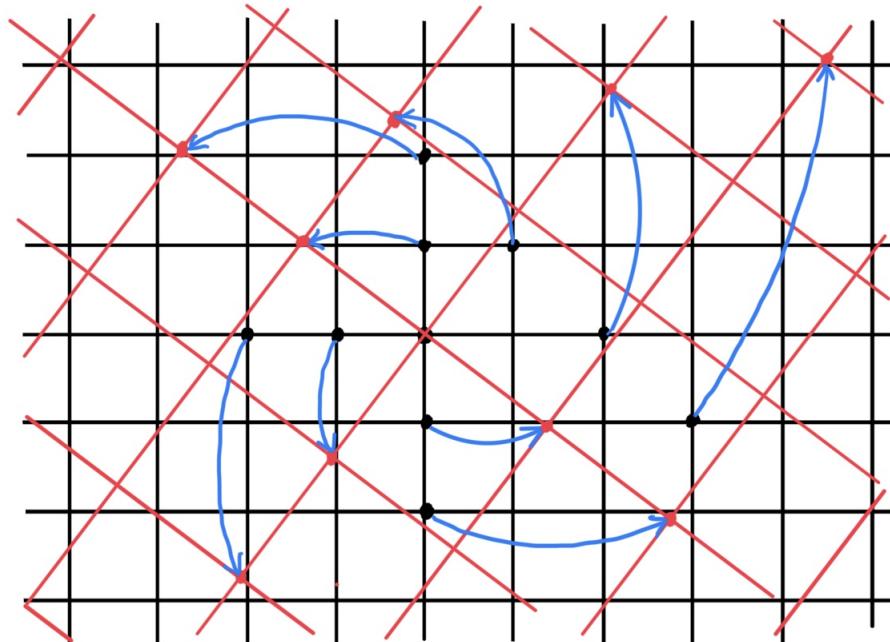
Visualizing multiplication is a bit harder. Given a

$$a = |a|(\cos \varphi + i \sin \varphi) \neq 0$$

the product az can be interpreted as the mapping of \mathbb{C} onto itself given by the formula

$$z \mapsto az$$

which is the composition of a dilation by a factor of $|a|$ and a rotation through the angle $\varphi \in \text{Arg } a$.

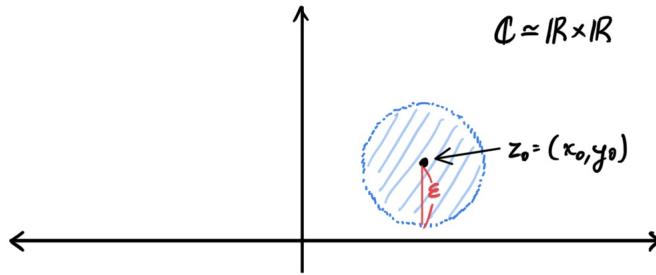


6.6.3 Sequences and Series in \mathbb{C}

Our previous construction of a metric within \mathbb{C} enables to define the ϵ -neighborhood of a number $z_0 \in \mathbb{C}$ as the set

$$U_\epsilon(z_0) \equiv \{z \in \mathbb{C} \mid |z - z_0| < \epsilon\}$$

which can be visualized as an open disk of radius ϵ in \mathbb{R}^2 centered at point (x_0, y_0) if $z_0 = x_0 + iy_0$.



Definition 6.17 (Convergence of a Sequence in \mathbb{C})

A sequence $\{z_n\}$ of complex numbers **converges** to $z_0 \in \mathbb{C}$ if and only if

$$\lim_{n \rightarrow \infty} |z_n - z_0| = 0$$

It is clear from the inequality

$$\max\{|x_n - x_0|, |y_n - y_0|\} \leq |z_n - z_0| \leq |x_n - x_0| + |y_n - y_0|$$

that a sequence of complex numbers converges if and only if the sequences of its real and imaginary parts of the terms of the sequence both converge. That is,

$$\{z_n\} \text{ converges} \iff \{\operatorname{Re} z\} \text{ and } \{\operatorname{Im} z\} \text{ converges}$$

Lemma 6.8 (Convergence of Cauchy Sequences over \mathbb{C})

A sequence of complex numbers $\{z_n\}$ is called a **Cauchy sequence** if for every $\epsilon > 0$ there exists an index $N \in \mathbb{N}$ such that

$$|z_n - z_m| < \epsilon \text{ for all } n, m > N$$

It is also clear that

$$\{z_n\} \text{ is Cauchy} \iff \{\operatorname{Re} z\} \text{ and } \{\operatorname{Im} z\} \text{ is Cauchy}$$

and using the Cauchy criterion for sequences of real numbers, we can easily see that a sequence of complex numbers converges if and only if it is a Cauchy sequence.

Lemma 6.9 (Convergence of Cauchy Series over \mathbb{C})

Interpreting the sum of a series of complex numbers

$$z_1 + z_2 + \dots + z_n + \dots$$

as the limit of the sequence its partial sums $\{s_n\}$, where $s_n = z_1 + \dots + z_n$ as $n \rightarrow \infty$, we can see that the series converges if and only if for every $\epsilon > 0$ there exists a $N \in \mathbb{N}$ such that

$$|z_m + \dots + z_n| < \epsilon$$

for any natural numbers $n \geq m > N$.

Definition 6.18 (Absolute Convergence of \mathbb{C})

A series $z_1 + \dots + z_n + \dots$ of complex numbers is **absolutely convergent** if the series

$$|z_1| + |z_2| + \dots + |z_n| + \dots$$

converges. Clearly, if a series converges absolutely, then it converges due to the inequality

$$|z_m + \dots + z_n| \leq |z_m| + \dots + |z_n|$$

Example 6.5 ()

The following complex series converges because they converges absolutely. That is,

$$\begin{aligned} 1 + \frac{1}{1!}|z| + \frac{1}{2!}|z^2| + \dots &\text{ converges } \forall z \implies 1 + \frac{1}{1!}z + \frac{1}{2!}z^2 + \dots \text{ converges } \forall z \\ |z| + \frac{1}{3!}|z|^3 + \frac{1}{5!}|z|^5 + \dots &\text{ converges } \forall z \implies z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 + \dots \text{ converges } \forall z \\ 1 + \frac{1}{2!}|z|^2 + \frac{1}{4!}|z|^4 + \dots &\text{ converges } \forall z \implies 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \dots \text{ converges } \forall z \end{aligned}$$

Definition 6.19 (Complex Power Series)

Series of the form

$$\sum_{n=0}^{\infty} c_n(z - z_0)^n = c_0 + c_1(z - z_0) + \dots + c_n(z - z_0) + \dots$$

are called **complex power series**, or **power series over \mathbb{C}** .

But a power series is quite useless unless we know the domain in which it converges (again, note that it is not always guaranteed to converge onto the function f if its power series expansion does converge at all). To develop more sophisticated tests of convergence of a complex power series, we introduce the complex analogue of the root test for real power series.

Theorem 6.14 (Cauchy-Hadamard Theorem)

The complex power series

$$c_0 + c_1(z - z_0) + \dots + c_n(z - z_0) + \dots$$

converges inside the disk $|z - z_0| < R$ with center at z_0 and radius given by the formula

$$R = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|}} = \frac{1}{\limsup_{n \rightarrow \infty} \sup \sqrt[n]{|c_n|}}$$

Where $\overline{\lim}$ denotes the superior limit. Furthermore,

1. the power series diverges at any point exterior to the disk.
2. the power series converges absolutely at any point interior to the disk.
3. the power series is indeterminate at any point on the boundary of the disk.

Note that in the degenerate case when $R = 0$, the series converges only at the point $z = z_0$.

Corollary 6.5 (Abel's First Theorem on Power Series)

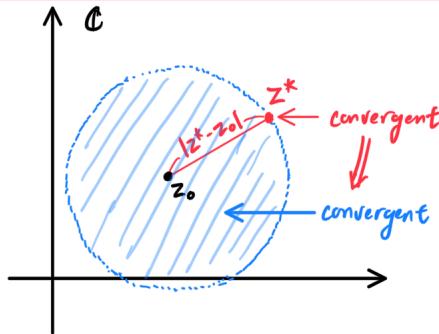
If the power series

$$c_0 + c_1(z - z_0) + \dots + c_n(z - z_0) + \dots$$

converges at some value z^* , then it converges absolutely for any value of z satisfying

$$|z - z_0| < |z^* - z_0|$$

The values of z satisfying the inequality above can be intuitively visualized as the following region.

**Theorem 6.15 (Product of Absolutely Convergent Series)**

Let $a_1 + a_2 + \dots$ and $b_1 + b_2 + \dots$ be an absolutely convergent series such that

$$\sum_{i=1}^{\infty} a_i = A \text{ and } \sum_{j=1}^{\infty} b_j = B$$

Then, the Cauchy product of the two series

$$\left(\sum_{i=1}^{\infty} a_i \right) \cdot \left(\sum_{j=1}^{\infty} b_j \right) = \sum_{k=0}^{\infty} c_k = AB, \text{ where } c_k = \sum_{l=0}^k a_l b_{k-l}$$

$a_1 b_1 + a_2 b_2 + \dots$ is absolutely convergent and

$$\sum_{i=1}^{\infty} a_i b_i = AB$$

Proof.

To be done.

Example 6.6 (Convergence of the Cauchy Product of Absolutely Convergent Complex Series)

The two series

$$\sum_{n=0}^{\infty} \frac{1}{n!} a^n \text{ and } \sum_{m=0}^{\infty} \frac{1}{m!} b^m$$

converges absolutely. Therefore, we can see that their Cauchy product can be nicely represented by

grouping together all monomials of the form $a^n b^m$ having the same total degree $m + n = k$.

$$\left(\sum_{n=0}^{\infty} \frac{1}{n!} a^n \right) \cdot \left(\sum_{m=0}^{\infty} \frac{1}{m!} b^m \right) = \sum_{k=0}^{\infty} \left(\sum_{n+m=k} \frac{1}{n!} a^n \frac{1}{m!} b^m \right)$$

But we can simplify

$$\sum_{m+n=k} \frac{1}{n!m!} a^n b^m = \frac{1}{k!} \sum_{n=0}^k \frac{k!}{n!(k-n)!} a^n b^{k-n} = \frac{1}{k!} (a+b)^k$$

and therefore we find that

$$\left(\sum_{n=0}^{\infty} \frac{1}{n!} a^n \right) \cdot \left(\sum_{m=0}^{\infty} \frac{1}{m!} b^m \right) = \sum_{k=0}^{\infty} \frac{1}{k!} (a+b)^k$$

6.6.4 Euler's Formula

Definition 6.20 (Complex Taylor Expansions of Transcendental Functions)

Since we have determined absolute convergence, and therefore convergence, of all these series in all of \mathbb{C} , it is natural to extend the definitions of

$$\exp, \cos, \sin : \mathbb{R} \longrightarrow \mathbb{R}$$

to the complex field

$$\exp, \cos, \sin : \mathbb{C} \longrightarrow \mathbb{C}$$

by defining them as

$$\begin{aligned} e^z &\equiv 1 + \frac{1}{1!} z + \frac{1}{2!} z^2 + \frac{1}{3!} z^3 + \dots \\ \cos z &\equiv 1 - \frac{1}{2!} z^2 + \frac{1}{4!} z^4 - \frac{1}{6!} z^6 + \dots \\ \sin z &\equiv z - \frac{1}{3!} z^3 + \frac{1}{5!} z^5 - \frac{1}{7!} z^7 + \dots \end{aligned}$$

Notice that even in the complex field, $\cos z$ is an even function and $\sin z$ is an odd function.

$$\cos(-z) = \cos(z)$$

$$\sin(-z) = -\sin(z)$$

In fact, the last example in the previous subsection just proves the following.

Lemma 6.10 (Exponential Map as a Group Homomorphism)

The exponential map $\exp : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \setminus \{0\}$ satisfies the following

$$\exp(z_1 + z_2) = \exp(z_1) \cdot \exp(z_2)$$

That is, \exp is a group homomorphism from $(\mathbb{C}, +)$ to $(\mathbb{C} \setminus \{0\}, \cdot)$.

Definition 6.21 (Euler's Formula)

By making the substitution $z = yi$ in the series expansion of e^z (where y is an arbitrary complex number), we get

$$\begin{aligned} e^{iy} &= 1 + \frac{1}{1!}(iy) + \frac{1}{2!}(iy)^2 + \frac{1}{3!}(iy)^3 + \frac{1}{4!}(iy)^4 + \dots \\ &= \left(1 - \frac{1}{2}y^2 + \frac{1}{4!}y^4 - \dots\right) + i\left(\frac{1}{1!}y - \frac{1}{3!}y^3 + \frac{1}{5!}y^5 - \dots\right) \end{aligned}$$

which brings us the identity

$$e^{iy} = \cos y + i \sin y$$

Since \cos is even and \sin is odd, we can add the two identities

$$\begin{aligned} e^{iz} &= \cos z + i \sin z \\ e^{-iz} &= \cos z - i \sin z \end{aligned}$$

to get

$$\begin{aligned} \cos z &= \frac{1}{2}(e^{iz} + e^{-iz}) \\ \sin z &= \frac{1}{2i}(e^{iz} - e^{-iz}) \end{aligned}$$

This gives us a very elegant connection between these three transcendental functions.

Definition 6.22 (Hyperbolic Functions)

Likewise, the following series are convergent (since they are absolutely convergent) and therefore we can define the extension of \cosh and \sinh into the complex field as

$$\begin{aligned} \cosh z &\equiv 1 + \frac{1}{2!}z^2 + \frac{1}{4!}z^4 + \frac{1}{6!}z^6 + \dots \\ \sinh z &\equiv z + \frac{1}{3!}z^3 + \frac{1}{5!}z^5 + \frac{1}{7!}z^7 + \dots \end{aligned}$$

The following identities immediately follow

$$\begin{aligned} \cosh z &= \frac{1}{2}(e^z + e^{-z}) \\ \sinh z &= \frac{1}{2}(e^z - e^{-z}) \end{aligned}$$

Lemma 6.11 (Trigonometric, Hyperbolic Identities over \mathbb{C})

Common identities, which are exactly the same as their real analogues, are listed.

1. $\cos^2 z + \sin^2 z = 1$
2. $\cosh^2 z - \sinh^2 z = 1$
3. $e^{i(z_1+z_2)} = (\cos z_1 \cos z_2 - \sin z_1 \sin z_2) + i(\sin z_1 \cos z_2 + \cos z_1 \sin z_2)$
4. $\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2$
5. $\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2$
6. $\cosh z = \cos iz$
7. $\sinh z = -i \sin iz$

However, to obtain even such geometrically obvious facts as the equality

$$\sin \pi = 0 \text{ or } \cos z + 2\pi = \cos z$$

from the power series definitions of \cos and \sin is extremely difficult. What the properties actually do is present the remarkable unity of these seemingly different trigonometric and hyperbolic functions, which would have been impossible to detect without going into the domain of complex numbers.

If we just take the following identities

$$\begin{aligned}\cos x &= \cos(x + 2\pi) \\ \sin x &= \sin(x + 2\pi) \\ \cos 0 &= 1 \\ \sin 0 &= 0\end{aligned}$$

then we get the following identity.

Theorem 6.16 (Euler's Identity)

The following relation is true.

$$e^{i\pi} + 1 = 0$$

which immediately implies

$$\exp(z + 2\pi i) = \exp z$$

That is, the exponential function is a periodic function on \mathbb{C} with the purely imaginary period $T = 2\pi i$.

Corollary 6.6 (Trigonometric Notation of Complex Number)

With Euler's formula and the periodic relation of $\exp z$, the trigonometric form of a complex number can be presented as

$$z = r(\cos \varphi + i \sin \varphi) = re^{i\varphi}$$

We can rewrite DeMoivre's formula as

$$z^n = r^n e^{n\varphi i}$$

6.6.5 Visualizing Complex Functions

TBD

6.6.6 Continuity, Differentiability, Analyticity of Complex Functions

The definitions of continuity and differentiability are the same, just under a different field.

Definition 6.23 (Limit of a Complex Function)

The function $f : E \subset \mathbb{C} \rightarrow \mathbb{C}$ tends to $A \in \mathbb{C}$ as $z \rightarrow a$, or that

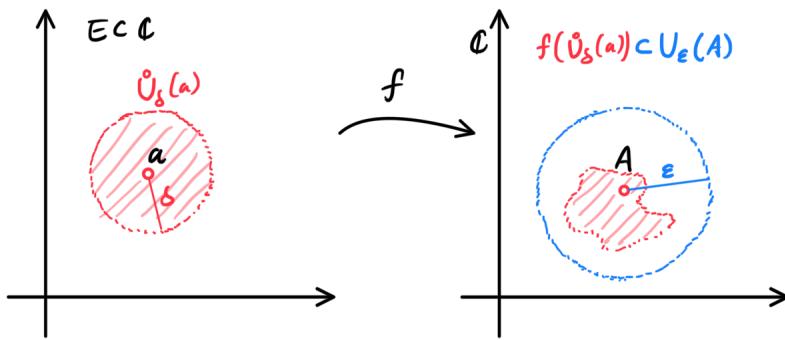
$$\lim_{z \rightarrow a} f(z) = A$$

if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$0 < |z - a| < \delta \implies |f(z) - A| < \epsilon$$

Note that we set $0 < |z - a|$ to ensure that $z \neq a$.

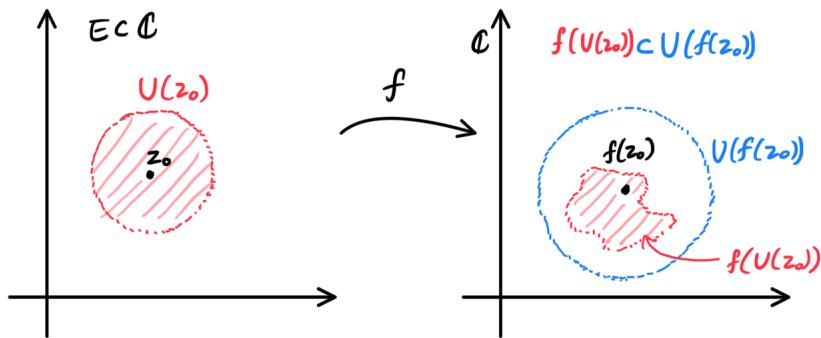
Therefore, in other words, for any arbitrarily small $\epsilon > 0$, we can find a $\delta > 0$ such that the image of the deleted δ -neighborhood of a , denoted $\tilde{U}_\delta(a)$, is completely within the ϵ -neighborhood $U_\epsilon(A)$.



Definition 6.24 (Continuity of a Complex Function)

A function $f : E \subset \mathbb{C} \rightarrow \mathbb{C}$ is **continuous** at a point $z_0 \in E$ if for any neighborhood $U(f(z_0))$ there exists a neighborhood $U(z_0)$ such that its image is contained in $U(f(z_0))$. In short,

$$\lim_{z \rightarrow z_0} f(z) = f(z_0)$$



Definition 6.25 (Differentiability of a Complex Function)

The **derivative** of a function $f : E \subset \mathbb{C} \rightarrow \mathbb{C}$ is defined

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

if this limit exists. f **differentiable** at z_0 means that a differential function

$$df(z_0) : T_{z_0}\mathbb{C} \rightarrow T_{f(z_0)}\mathbb{C}, \quad h \mapsto df(z_0)(h)$$

exists such that

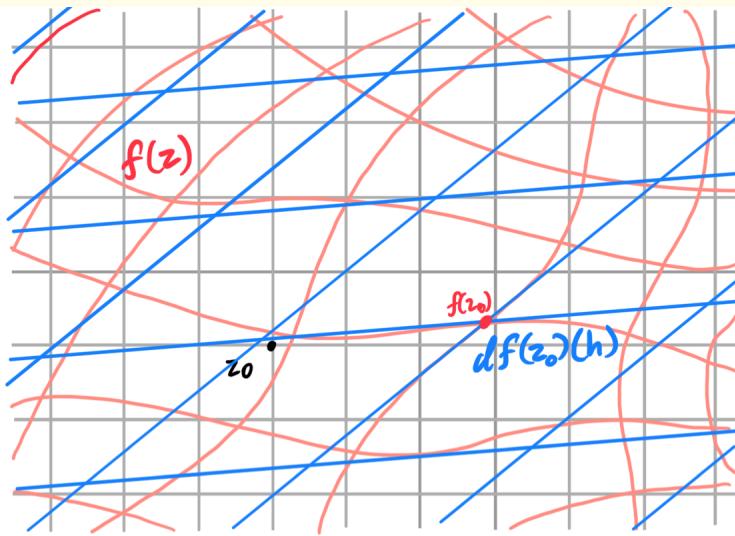
$$f(z) = f(z_0) + df(z_0)(h) + o(h)$$

where $h = z - z_0$ is the increment of the argument. Just like the real case, it turns out that $df(z_0)(h) = f'(z_0)h$, and

$$f(z) - f(z_0) = f'(z_0)(z - z_0) + o(z - z_0)$$

which elegantly weaves together the two concepts of differentiability and the derivative.

Visualizing this, we can see that for whatever function $f : \mathbb{C} \rightarrow \mathbb{C}$ there is a linear function that transforms the entire space as such at z_0 (along with a given point $z_0 \in \mathbb{C}$),



The differential $df(z_0)$ at the point z_0 is a linear mapping that "best" approximates f , with an error of $o(h) = o(z - z_0)$.

Lemma 6.12 (Arithmetic Properties of Differentiation over \mathbb{C})

If functions $f, g : E \subset \mathbb{C} \rightarrow \mathbb{C}$ are differentiable at a point $z \in E$, then

1. their sum is differentiable at z , and

$$d(f + g)(z) = df(z) + dg(z) \iff (f + g)'(z) = (f' + g')(z)$$

2. their product is differentiable at z , and

$$d(f \cdot g)(z) = g(z)df(z) + f(z)dg(z) \iff (f \cdot g)'(z) = f'(z)g(z) + f(z) \cdot g'(z)$$

3. their quotient is differentiable at z if $g(z) \neq 0$, and

$$d\left(\frac{f}{g}\right)(z) = \frac{g(z)df(z) - f(z)dg(z)}{g^2(z)} \iff \left(\frac{f}{g}\right)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{g^2(z)}$$

Just like the real case, the operation of taking the derivative is a linear operator.

Lemma 6.13 (Chain Rule for Composite Functions over \mathbb{C})

Let there be functions $f : E_1 \subset \mathbb{C} \rightarrow \mathbb{C}$ differentiable at point $z \in E_1$ and $g : E_2 \subset \mathbb{C} \rightarrow \mathbb{C}$ differentiable at point $w = f(z) \in E_2$, with respective differentials

$$\begin{aligned} df(z) &: T_z \mathbb{C} \rightarrow T_w \mathbb{C} \\ dg(w) &: T_w \mathbb{C} \rightarrow T_{g(w)} \mathbb{C} \end{aligned}$$

Then, the composite function $g \circ f : E_1 \rightarrow \mathbb{C}$ is differentiable at z , and $d(g \circ f)(z) : T_z \mathbb{C} \rightarrow T_{g \circ f(z)} \mathbb{C}$ is

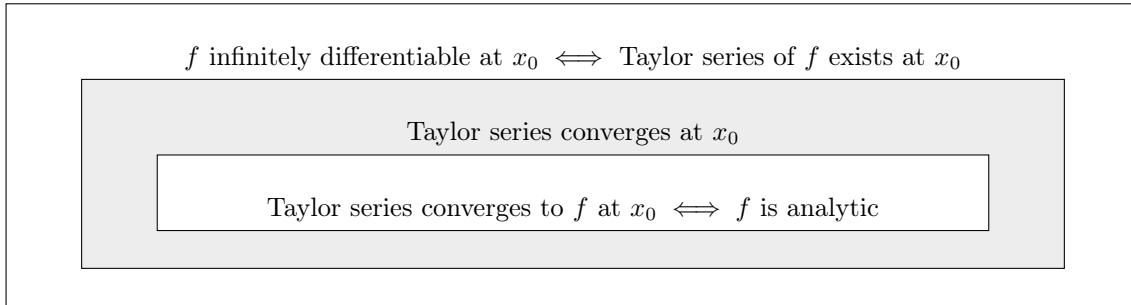
$$d(g \circ f)(z) = dg(w) \circ df(z) \iff (g \circ f)'(z) = g'(f(z)) \circ f'(z)$$

6.6.7 Power Series Representation of a Function

Definition 6.26 (Holomorphic Function)

If function $f : E \subset \mathbb{C} \rightarrow \mathbb{C}$ is (complex) differentiable at a point $z_0 \in E$, then f is said to be **holomorphic at z_0** .

We recall the diagram that summarizes the conditions of differentiability and analyticity of a function f over the field \mathbb{R} .



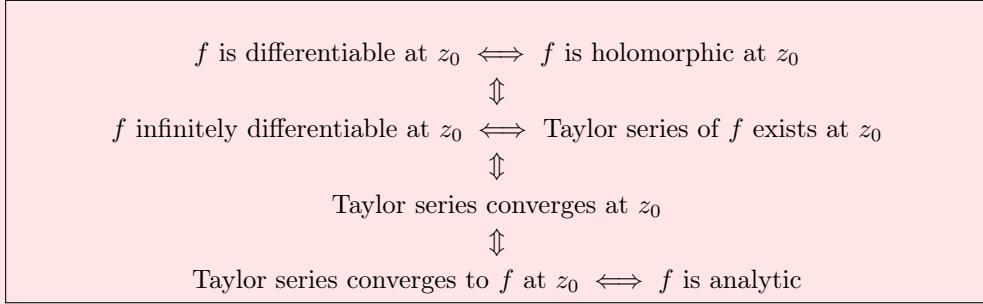
In the theory of functions of a complex variable we actually have a remarkable theorem that does not have an analogue for functions over \mathbb{R} .

Theorem 6.17 (Analyticity of Differentiable Functions over \mathbb{C})

If a function $f : E \subset \mathbb{C} \rightarrow \mathbb{C}$ is differentiable in a neighborhood of a point $z_0 \in E$, then it is analytic at that point. In other words,

$$f \text{ is holomorphic at } z_0 \implies f \text{ is analytic at } z_0$$

This means that the conditions in the diagram above all are equivalent! Visually,



This is certainly an amazing fact, since it then follows from the theorem that if a function $f(z)$ has one derivative $f'(z)$ in a neighborhood of a point, it also has derivatives of all orders in that neighborhood.

6.6.8 Algebraic Closedness of the Field \mathbb{C}

Definition 6.27 (Algebraically Closed Field)

A field \mathbb{F} is **algebraically closed** if every nonconstant polynomial in $\mathbb{F}[x]$ (the polynomial ring with coefficients in \mathbb{F}) has a root in \mathbb{F} .

Theorem 6.18 (Fundamental Theorem of Algebra)

\mathbb{C} is algebraically closed. That is, every polynomial

$$P(z) \equiv c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n$$

of degree $n \geq 1$ with complex coefficients $c_i \in \mathbb{C}$ ($i = 0, 1, \dots, n$) has a root in \mathbb{C} . This immediately implies that every polynomial $P(z)$ admits a representation (unique up to the order of the factors) in the form

$$P(z) = c_n(z - z_1)(z - z_2) \dots (z - z_n)$$

where $z_1, \dots, z_n \in \mathbb{C}$ not necessarily all distinct.

We can also prove the interesting property about zeroes of polynomials in $\mathbb{R}[x]$.

Corollary 6.7 (Complex Conjugate Roots of Real Polynomials)

Given a polynomial with real coefficients

$$P(z) \equiv a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n$$

P , as we know, does not always have real roots (e.g. $P(x) = x^2 + 1$). However, we state that

$$\text{if } P(z_0) = 0, \text{ then } P(\bar{z}_0) = 0$$

Therefore, every polynomial P with real coefficients can be expanded as a product of linear and quadratic polynomial with real coefficients.

Proof.

We can see from the properties of complex numbers that

$$\begin{aligned} \overline{(z_1 + z_2)} &= \overline{z_1} + \overline{z_2} \\ \overline{(z_1 \cdot z_2)} &= \overline{(r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2})} \\ &= \overline{r_1 r_2 e^{i(\varphi_1 + \varphi_2)}} = r_1 r_2 e^{-i(\varphi_1 + \varphi_2)} \\ &= r_1 e^{-i\varphi_1} \cdot r_2 e^{-i\varphi_2} = \bar{z}_1 \cdot \bar{z}_2 \end{aligned}$$

Thus, if $P(z_0) = 0$, then

$$0 = \overline{P(z_0)} = \overline{a_0 + \dots + a_n z_0^n} = \bar{a}_0 + \dots + \bar{a}_n \bar{z}_0^n = a_0 + \dots + a_n \bar{z}_0^n = P(\bar{z}_0)$$

and thus $P(\bar{z}_0) = 0$.

6.7 Primitives

Definition 6.28 (Primitive)

A function $F(x)$ is a **primitive** of a function $f(x)$ on an interval if F is differentiable on the interval and satisfies the equation

$$F'(x) = f(x)$$

or equivalently, if their respective differentials satisfy

$$dF(x) = f(x) dx$$

Lemma 6.14 ()

If $F_1(x)$ and $F_2(x)$ are two primitives of $f(x)$ on the same interval, then the difference $(F_1 - F_2)(x)$ is constant on that interval.

Example 6.7 ()

Both

$$F_1(x) \equiv \arctan x \text{ and } F_2(x) \equiv \operatorname{arccot} \frac{1}{x}$$

are primitives of $f(x) = \frac{1}{1+x^2}$. Indeed, we can see by direct calculation that in the domain $\mathbb{R} \setminus 0$,

$$F_1(x) - F_2(x) = \arctan x - \operatorname{arccot} \frac{1}{x} = \begin{cases} 0, & x > 0 \\ -\pi, & x < 0 \end{cases}$$

which is supported by the lemma.

Notice how given a function $f(x)$, the operation of finding its differential, denoted with d , gives us a new function of h , called the differential

$$df(x)(h)$$

Similarly, the operation of finding a primitive of function $f(x)$, denoted with the symbol \int , gives us a new function.

Definition 6.29 (Indefinite Integration)

The operation of finding a primitive of a certain function $f(x)$ is called **indefinite integration**, and the mathematical notation

$$\int f(x) dx$$

is called the **indefinite integral** of $f(x)$ on a given interval (f called the **integrand** and $f(x) dx$ called the **differential form**).

1. It immediately follows from the lemma that if $F(x)$ is any particular primitive of $f(x)$ on the interval, then on that interval

$$\int f(x) dx = F(x) + C$$

2. If $F'(x) = f(x)$ (that is, F is a primitive of f on some interval), then we have

$$d \int f(x) dx = dF(x) = F'(x) dx$$

3. It also follows that

$$\int dF(x) = \int F'(x) dx = F(x) + C$$

Theorem 6.19 (Basic Methods of Indefinite Integration)

The definition of the indefinite integral has three basic properties that can be used to solve indefinite integrals.

1. Linearity of the indefinite integral.

$$\int (\alpha u(x) + \beta v(x)) dx = \alpha \int u(x) dx + \beta \int v(x) dx + C$$

2. Integration by parts.

$$\int (uv)' \, dx = \int u'(x)v(x) \, dx + \int u(x)v'(x) \, dx + C$$

3. Change of Variable, or U -substitution. Given that $F'(x) = f(x)$ on an interval I_x and $\varphi : I_t \rightarrow I_x$ is a C^1 mapping of interval I_t into I_x , then

$$\int (f \circ \varphi)(t) \varphi'(t) \, dt = (F \circ \varphi)(t) + C$$

7 Integration

7.1 Construction of the Riemann Integral

We shall first define the integral using the familiar notation of Riemann sums.

Definition 7.1 (Partitions with Distinguished Points)

A **partition** P of a closed interval $[a, b]$, $a < b$, is a finite system of points x_0, \dots, x_n of the interval such that

$$a = x_0 < x_1 < x_2 < \dots < x_n = b$$

The intervals $[x_{i-1}, x_i]$, $i = 1, 2, \dots, n$, are called the **intervals** of the partition P . The largest of the lengths of the intervals of the partition P , denoted $\lambda(P)$, is called the **mesh** of the partition.

A **partition with distinguished points** (P, ξ) on the closed interval $[a, b]$ is a partition P of $[a, b]$ along with the set of n points

$$\xi_1 \in [x_0, x_1], \xi_2 \in [x_1, x_2], \dots, \xi_n \in [x_{n-1}, x_n]$$

The n -tuple of ξ_i 's is denoted by the single letter ξ

$$\xi = (\xi_1, \xi_2, \dots, \xi_n)$$

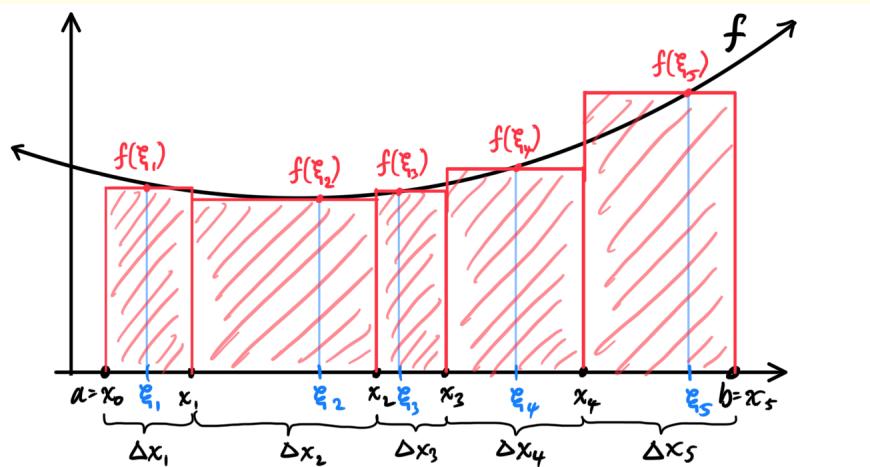
This naturally leads to the following construction.

Definition 7.2 (Riemann Sums)

If a function f is defined on a closed interval $[a, b]$ and (P, ξ) is a partition with distinguished points on this closed interval, the sum

$$\sigma(f; P, \xi) \equiv \sum_{i=1}^n f(\xi_i) \Delta x_i, \text{ where } \Delta x_i = x_i - x_{i-1},$$

is the **Riemann sum** of the function f corresponding to the partition (P, ξ) with distinguished points on $[a, b]$.



Thus, when a function f is fixed, the Riemann sum $\sigma(f; P, \xi)$ is a mapping that takes in a partition with distinguished points $p = (P, \xi)$ on the closed interval $[a, b]$ and outputs a number representing the total area of the Riemann sums. That is, for a fixed f and some input $p = (P, \xi)$, we can define

the function

$$\Phi : \mathcal{P} \longrightarrow \mathbb{R}, \quad \Phi(p) \equiv \sigma(f; p) \equiv \sigma(f; (P, \xi))$$

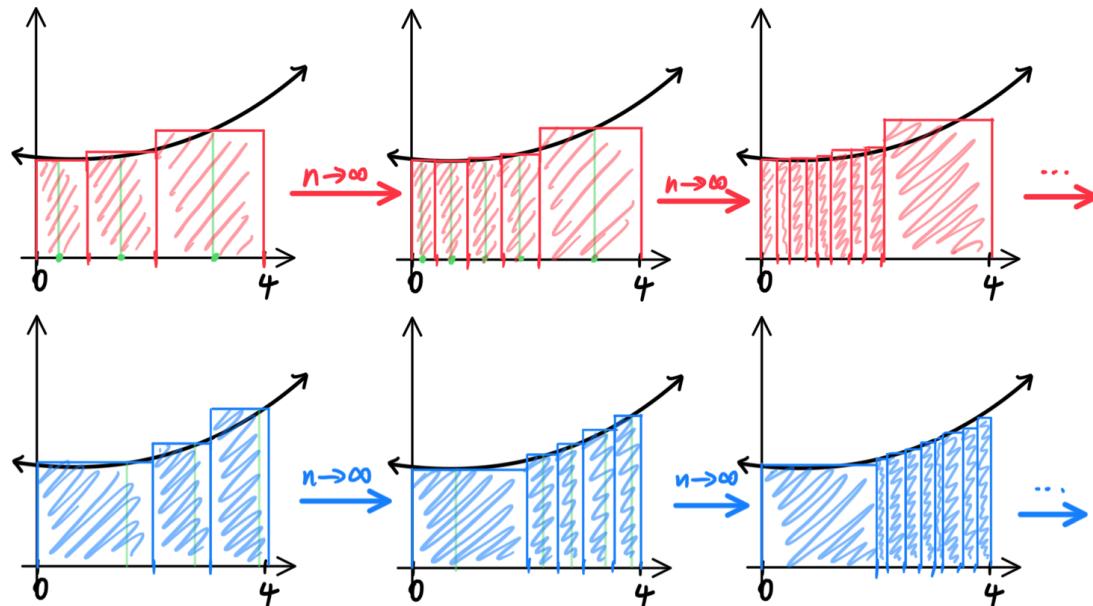
that takes in a partition with distinguished points on $[a, b]$ and outputs the corresponding Riemann sum for that fixed f .

Definition 7.3 (Riemann Integral)

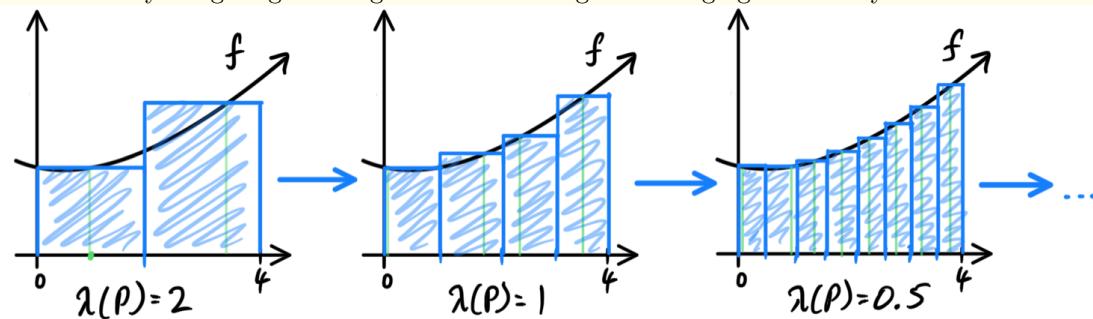
The number $\int_a^b f(x) dx$ is the **Riemann integral** of the function f on the closed interval $[a, b]$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i) \Delta x_i \right| < \epsilon$$

for any partition (P, ξ) with distinguished points on $[a, b]$ whose mesh $\lambda(P)$ is less than δ . We can view this as a limit where $n \rightarrow \infty$, but there is a problem since we can increase the partition within different subsets of $[a, b]$, leading to multiple values of convergence.



Rather, we can set the mesh $\lambda(P)$ to approach 0, which would take care of the problems. We can visualize this by imagining the lengths of the rectangles converging "uniformly."



Therefore, we can culminate by defining the Riemann integral of $f(x)$ over $[a, b]$ as

$$\int_a^b f(x) dx \equiv \lim_{\lambda(P) \rightarrow 0} \sum_{i=1}^n f(\xi_i) \lambda x_i$$

7.1.1 Conditions for Integrability

Definition 7.4 (Riemann Integrable Functions)

A function f is **Riemann integrable** on the closed interval $[a, b]$ if

$$\int_a^b f(x) dx \equiv \lim_{\lambda(P) \rightarrow 0} \sum_{i=1}^n f(\xi_i) \lambda x_i$$

is defined, i.e. if the limit of the right-hand side of Riemann sums exists as $\lambda(P) \rightarrow 0$ (that is, the Riemann integral of f is defined).

Furthermore, the set of Riemann-integrable functions on a closed interval $[a, b]$ is denoted $\mathcal{R}[a, b]$.

Remember that the Riemann integral, as complicated as the formula is, is still a limit of a function. That means that we can apply the Cauchy criterion to it to determine convergence.

Lemma 7.1 (Cauchy Criterion on Existence of Riemann Integral)

Given a function f , the integral of f over $[a, b]$, defined

$$\int_a^b f(x) dx \equiv \lim_{\lambda(P) \rightarrow 0} \sum_{i=1}^n f(\xi_i) \lambda x_i$$

exists if and only if for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|\sigma(f; P', \xi') - \sigma(f; P'', \xi'')| < \epsilon$$

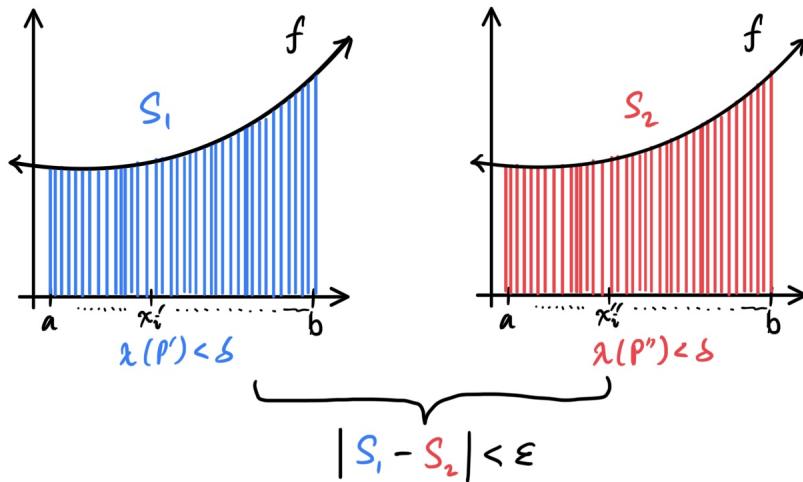
or, what is the same,

$$\left| \sum_{i=1}^{n'} f(\xi'_i) \Delta x'_i - \sum_{i=1}^{n''} f(\xi''_i) \Delta x''_i \right| < \epsilon$$

for any partitions (P', ξ') and (P'', ξ'') with distinguished points on the interval $[a, b]$ with

$$\lambda(P'), \lambda(P'') < \delta$$

In words, this means that for any $\epsilon > 0$ that we choose, there always exists a $\delta > 0$ such that **any** two Riemann sums with mesh size **both** smaller than δ will have an error difference of less than ϵ .



Theorem 7.1 (Necessary Condition for Integrability)

A necessary condition for f defined on a closed interval $[a, b]$ to be Riemann integrable on $[a, b]$ is that f be bounded on $[a, b]$. That is,

$$f \in \mathcal{R}[a, b] \implies f \text{ is bounded on } [a, b]$$

We can clearly see the necessity of f being bounded by looking at the contrapositive of the following statement.

Theorem 7.2 (Refinement)

Given a partition P on interval $[a, b]$, recall that we have points x_0, \dots, x_n such that

$$a = x_0 < x_1 < \dots < x_n = b$$

Here we introduce new notation:

1. Δ_i denotes the interval $[x_{i-1}, x_i]$
2. Δx_i denotes the difference $x_i - x_{i-1}$, i.e. the length of Δ_i

If a partition \tilde{P} of the closed interval $[a, b]$ is obtained from the partition P by the addition of new points to P , we call \tilde{P} a **refinement** of P .

When a refinement \tilde{P} of a partition P is constructed, some (perhaps all) of the closed intervals $\Delta_i = [x_{i-1}, x_i]$ of the partition P themselves undergo partitioning.

$$x_{i-1} = x_{i0} < x_{i1} < \dots < x_{in_i} = x_i$$

In that connection, it will be useful to label the points of \tilde{P} by double indices, where in the notation x_{ij} the first index i means that

$$x_{ij} \in \Delta_i = [x_{i-1}, x_i]$$

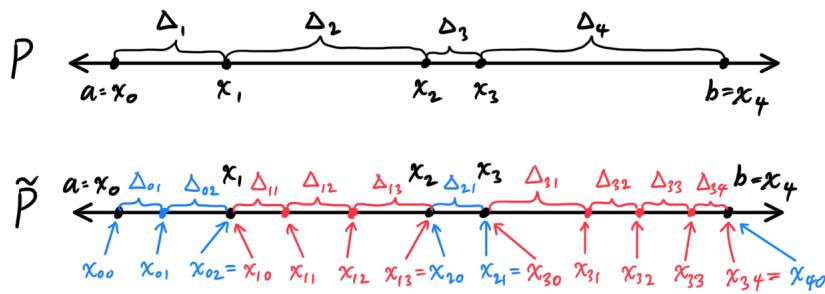
and the second index j is the ordinal number of the point on the closed interval $\Delta_i = [x_{i-1}, x_i]$. Therefore, it is natural to set the notations

1. $\Delta_{ij} = [x_{ij-1}, x_{ij}]$
2. $\Delta x_{ij} = x_{ij} - x_{ij-1}$

This means that

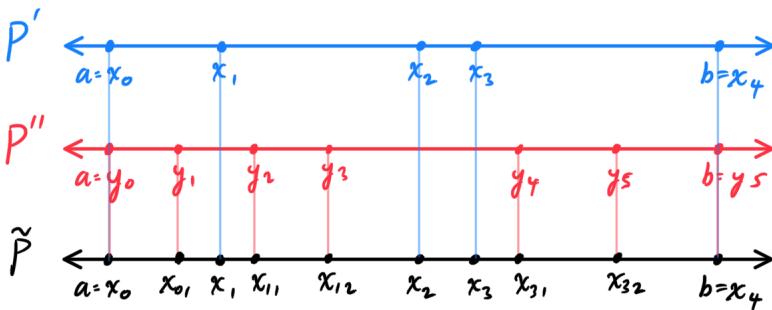
$$\Delta x_i = \Delta x_{i1} + \Delta x_{i2} + \dots + \Delta x_{in_i}$$

which can be visualized below



Example 7.1 (Union of Partitions as a Refinement)

For some interval $[a, b]$, given partitions P' ($a = x_0 < \dots < x_n = b$) and P'' ($a = y_0 < \dots < y_n = b$), the union of the two partitions $\tilde{P} = P' \cup P''$ is a refinement of both P' and P'' .



Recall that $\omega(f; E)$ denotes the oscillation of the function f on the set E ; that is,

$$\omega(f; E) \equiv \sup_{x', x'' \in E} |f(x') - f(x'')|$$

In particular, $\omega(f; \Delta_i)$ is the oscillation of f on the closed interval Δ_i .

Theorem 7.3 (Sufficient Condition for Integrability)

Let f be a bounded on a closed interval $[a, b]$ such that for every $\epsilon > 0$ there exists a number $\delta > 0$ such that

$$\sum_{i=1}^n \omega(f; \Delta_i) \Delta x_i < \epsilon$$

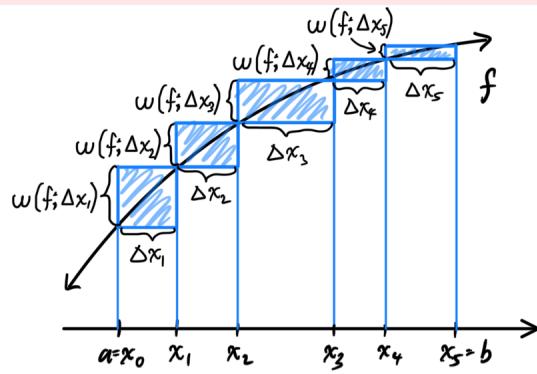
for any partition P of $[a, b]$ with mesh $\lambda(P) < \delta$. This is equivalent to saying that

$$\lim_{\lambda(P) \rightarrow 0} \sum_{i=1}^n \omega(f; \Delta_i) \Delta x_i = 0$$

Then, f is integrable. We can visualize

$$\sum_{i=1}^n \omega(f; \Delta_i) \Delta x_i$$

as the following sum of rectangles below.



What the theorem states, visually, is that as we make all the rectangles smaller and smaller (by putting a limit on the mesh $\lambda(P) < \delta$), we can make the sum of all these rectangles also arbitrarily small.

Corollary 7.1 (Integrability of Continuous Functions)

Every continuous function on a closed interval is integrable on that closed interval. That is,

$$f \in C[a, b] \implies f \in \mathcal{R}[a, b]$$

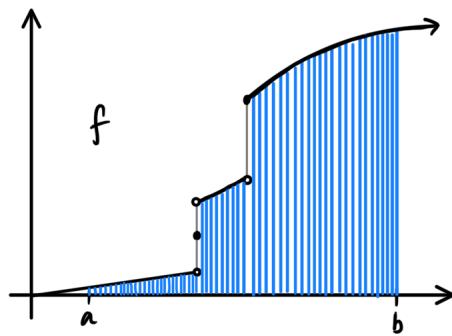
We can actually make a stronger claim.

Corollary 7.2 (Integrability of Discontinuous Functions)

If a bounded function f on a closed interval $[a, b]$ is continuous everywhere except at a finite set of points, then $f \in \mathcal{R}[a, b]$.

Corollary 7.3 (Integrability of Monotonic Functions)

A bounded monotonic function on a closed interval is integrable on that interval.



Definition 7.5 (Upper, Lower Riemann Sums)

Let $f : [a, b] \rightarrow \mathbb{R}$ be a real-valued function that is defined and bounded on the closed interval $[a, b]$,

and let P be a partition of $[a, b]$, and let Δ_i ($i = 1, 2, \dots, n$) be the intervals of the partition P . Let

$$m_i = \inf_{x \in \Delta_i} f(x)$$

$$M_i = \sup_{x \in \Delta_i} f(x)$$

be the infimum and supremum of f over Δx_i . Then, the sums

$$s(f; P) \equiv \sum_{i=1}^n m_i \Delta x_i$$

$$S(f; P) \equiv \sum_{i=1}^n M_i \Delta x_i$$

are respectively called the **lower** and **upper Riemann sums** of the function f on the interval $[a, b]$ corresponding to the partition P of that interval.

Given an arbitrary partition (P, ξ) with distinguished points on $[a, b]$, it is clear that

$$s(f; P) = \inf_{\xi} \sigma(f; P, \xi) \leq \sigma(f; P, \xi) \leq \sup_{\xi} \sigma(f; P, \xi) = S(f; P)$$

Theorem 7.4 ()

A bounded real-valued function $f : [a, b] \rightarrow \mathbb{R}$ is Riemann integrable on $[a, b]$ if and only if the following limits exist and are equal to each other.

$$\underline{I} \equiv \lim_{\lambda(P) \rightarrow 0} s(f; P) = \lim_{\lambda(P) \rightarrow 0} S(f; P) \equiv \bar{I}$$

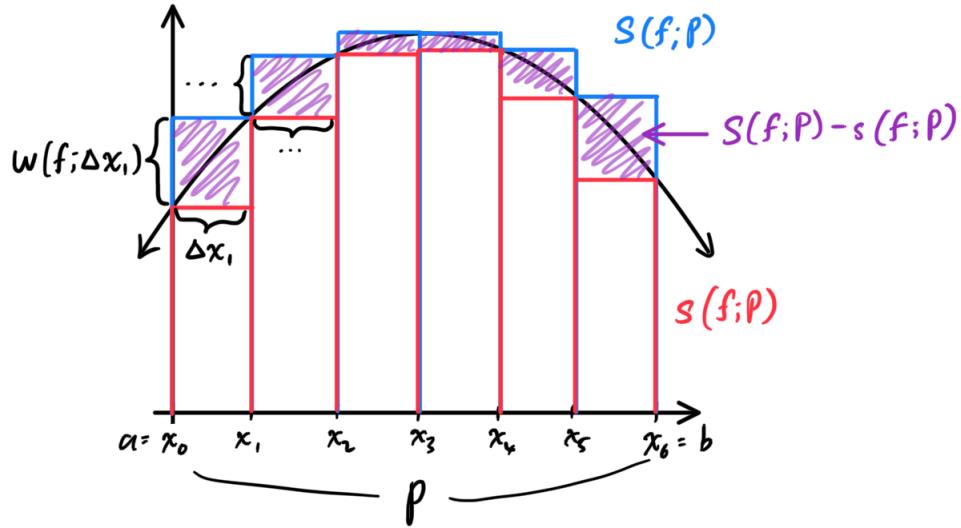
When the relation is true, then the integral is this common value.

$$\int_a^b f(x) dx = \underline{I} = \bar{I}$$

Note that this condition of the upper and lower Riemann sums converging to the same value and the condition that

$$\lim_{\lambda(P) \rightarrow 0} \sum_{i=1}^n \omega(f; \Delta_i) \Delta x_i = 0$$

are the same. For we can see that the rectangles visualized from the equation above are the exact same rectangles formed by $S(f; P) - s(f; P)$!



7.1.2 The Vector Space of Riemann Integrable Functions

Theorem 7.5 (The Vector Space of Integrable Functions)

The set of Riemann integrable functions $\mathcal{R}[a, b]$ over closed interval $[a, b]$ is a vector space. That is, given $f, g \in \mathcal{R}[a, b]$ and $\alpha \in \mathbb{R}$, then

1. $(f + g) \in \mathcal{R}[a, b]$
2. $(\alpha f) \in \mathcal{R}[a, b]$

Furthermore,

1. $|f| \in \mathcal{R}[a, b]$
2. The restriction of f in any $[c, d] \subset [a, b]$, denoted $f|_{[c, d]}$, is in $\mathcal{R}[c, d]$
3. $(f \cdot g) \in \mathcal{R}[a, b]$

Proof.

7.1.3 Lebesgue's Criterion for Riemann Integrability

We give Lebesgue's version of an intrinsic description of a Riemann integrable function.

Definition 7.6 (Measure)

A set $E \subset \mathbb{R}$ has (**Lebesgue**) **measure zero** if for every number $\epsilon > 0$ there exists a covering of the set E by an at most countable system $\{I_k\}$ of intervals, the sum of whose lengths

$$\sum_{k=1}^{\infty} |I_k| \leq \epsilon$$

This means that the above series summing up the lengths of the intervals is an absolutely convergent series.

Lemma 7.2 ()

We can deduce measures of basic sets.

1. A finite number of points are sets of measure zero.
2. The union of a finite or countable number of sets of measure zero is a set of measure zero.
3. A subset of a set of measure zero is itself a set of measure zero.
4. A closed interval $[a, b]$ with $a < b$ is not a set of measure zero.

Definition 7.7 ()

If a property holds at all points of a set X except possibly the points of a set of measure zero, we say that this property holds **almost everywhere on X** or at **almost every point of X** .

Now, we can state Lebesgue's criterion for integrability, which nicely summarizes what we have so far.

Theorem 7.6 (Lebesgue's Criterion for Integrability)

A function defined on a closed interval is Riemann integrable on that interval if and only if it is bounded and continuous at almost every point.

Example 7.2 (Non-Integrability of the Dirichlet Function)

The Dirichlet function

$$\mathcal{D}(x) \equiv \begin{cases} 1, & \text{for } x \in \mathbb{Q} \\ 0, & \text{for } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

on the interval $[0, 1]$ is not integrable on that interval. We state two different reasons why.

1. For any partition P of $[0, 1]$ we can find in each interval Δ_i both a rational point ξ'_i and an irrational point ξ''_i . Then, we can see that the lower and upper Riemann sums do not necessarily converge to each other since

$$\sigma(f; P, \xi') = \sum_{i=1}^n 1 \cdot \Delta x_i = 1 \text{ while } \sigma(f; P, \xi'') = \sum_{i=1}^n 0 \cdot \Delta x_i = 0$$

as $\lambda(P) \rightarrow 0$.

2. From the point of view of the Lebesgue criterion the nonintegrability of the Dirichlet function is obvious since $\mathcal{D}(x)$ is discontinuous at every point of $[0, 1]$, which is not a set of measure zero.

Notice that by the Lebesgue criterion, integrability is a weaker condition than continuity. That is,

$$f \text{ continuous} \implies f \text{ Riemann integrable}$$

but not necessarily the other way around. It turns out that this has consequences when determining the composition of functions.

Proposition 7.1 (Integrable + Continuous Composition)

Let $f : I_1 = [a, b] \rightarrow \mathbb{R}$ be a function that is integrable on $[a, b]$, with $\text{Im } f = [c, d] = I_2$. Define a continuous (remember, continuity is stronger than integrability) function $g : [c, d] \rightarrow \mathbb{R}$. Then the composition

$$g \circ f : [a, b] \rightarrow \mathbb{R}$$

is clearly defined and continuous at all the points of $[a, b]$ where f is continuous. But since f is integrable, the union of all the discontinuities in $[a, b]$ must have measure zero, and so it follows that

since $[a, b]$ is the same

$$g \circ f \in \mathcal{R}[a, b]$$

Therefore, we can find out that

$$f \text{ integrable and } g \text{ continuous} \implies g \circ f \text{ integrable}$$

as visualized in the commutative diagram below.

$$\begin{array}{ccc} & g \circ f & \\ I_1 & \xrightarrow{f} & I_2 \xrightarrow{g} \mathbb{R} \end{array}$$

However, contrary to intuition,

$$f \text{ integrable and } g \text{ integrable} \not\implies g \circ f \text{ integrable}$$

We present a counterexample.

Example 7.3 ()

Consider the functions

$$|sgn|(x) \equiv \begin{cases} 1 & x \neq 0 \\ 0 & x = 0 \end{cases}$$

and the Riemann function

$$\mathcal{R}(x) \equiv \begin{cases} \frac{1}{n} & x = \frac{m}{n} \in \mathbb{Q}, \gcd(m, n) = 1 \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

We can see that \mathcal{R} is continuous at all irrational points and discontinuous at all rational points except 0, meaning that it is integrable (\mathbb{Q} has measure zero). Then, the composition of these two functions is precisely the Dirichlet function

$$\mathcal{D}(x) = |sgn| \circ \mathcal{R}$$

which is not integrable.

7.2 Basic Properties of the Integral

One of the most basic properties of the integral is that it is a linear map.

Lemma 7.3 (Linearity of the Integral)

Given closed interval $[a, b] \subset \mathbb{R}$, the Riemann integration function

$$\int_a^b : \mathcal{R}[a, b] \longrightarrow \mathbb{R}$$

is a linear functional living within the dual space $\mathbb{R}^*[a, b]$. That is, given $f, g \in \mathcal{R}[a, b]$, a linear combination of them $\alpha f + \beta g$ is also integrable on $[a, b]$, and

$$\int_a^b (\alpha f + \beta g)(x) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx$$

Proof.

It is clear from basic algebraic transformation that the Riemann sums for the integral expressions on both sides are equal.

$$\sum_{i=1}^n (\alpha f + \beta g)(\xi_i) \Delta x_i = \alpha \sum_{i=1}^n f(\xi_i) \Delta x_i + \beta \sum_{i=1}^n g(\xi_i) \Delta x_i$$

Taking the limit as $\lambda(P) \rightarrow 0$ on both sides leads to the respective Riemann integrals.

The next property of the Riemann integral is its additive property **on the interval of integration**. Note that the value of the integral

$$\int_a^b f(x) dx \equiv \lim_{\lambda(P) \rightarrow 0} \sigma(f; P, \xi)$$

depends on both the integrand and the closed interval over which the integral is taken.

Lemma 7.4 (Properties of the Interval of Integration)

If $a < b < c$ and $f \in \mathcal{R}[a, c]$, then $f|_{[a, b]} \in \mathcal{R}[a, b]$, $f|_{[b, c]} \in \mathcal{R}[b, c]$, and the following equality holds

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$$

From these we set

$$\int_a^b f(x) dx \equiv - \int_b^a f(x) dx$$

and

$$\int_a^a f(x) dx \equiv 0$$

Theorem 7.7 (Symmetry of the Riemann Integral)

Let $a, b, c \in \mathbb{R}$ and let f be integrable over the largest closed interval having two of these points as endpoints. Then, the restriction of f to each of the other closed intervals is also integrable over those intervals and the following equality holds.

$$\int_a^b f(x) dx + \int_b^c f(x) dx + \int_c^a f(x) dx = 0$$

This property can be abstractified to those of additive interval functions, which will be shown soon.

We finally end with an important property of the integral which, as seen later, allows us to define inner products on function spaces.

Theorem 7.8 ()

If $a \leq b$ and $f \in \mathcal{R}[a, b]$, then $|f| \in \mathcal{R}[a, b]$, and

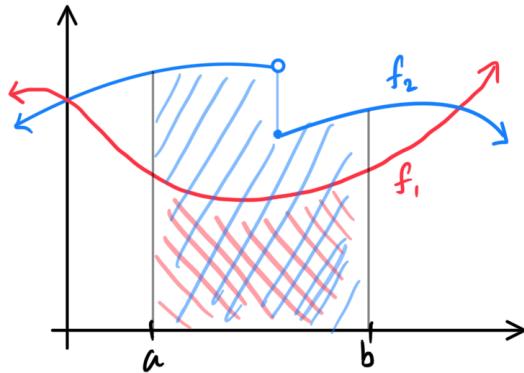
$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f|(x) dx$$

7.2.1 Mean Value Theorem of the Integral

Lemma 7.5 (Monotonicity of the Integral)

If $a \leq b$, $f_1, f_2 \in \mathcal{R}[a, b]$, and $f_1(x) \leq f_2(x)$ for every $x \in [a, b]$, then

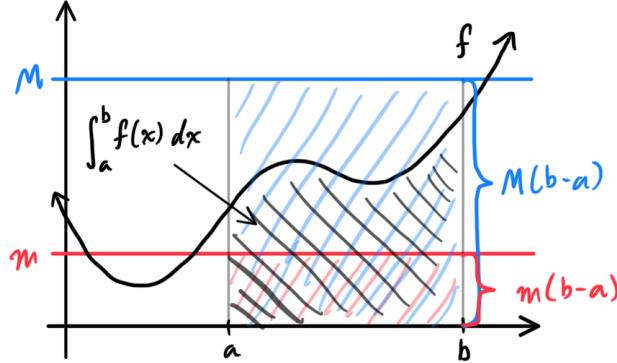
$$\int_a^b f_1(x) dx \leq \int_a^b f_2(x) dx$$



This immediately implies that given constants m, M such that $m \leq f(x) \leq M$ at each $x \in [a, b]$, we have

$$m \cdot (b - a) \leq \int_a^b f(x) dx \leq M \cdot (b - a)$$

This is very easily visualized below.



In particular, if $0 \leq f(x)$ on $[a, b]$, then

$$0 \leq \int_a^b f(x) dx$$

Theorem 7.9 (Mean Value Theorem of the Integral)

Given $f \in \mathcal{R}[a, b]$, with

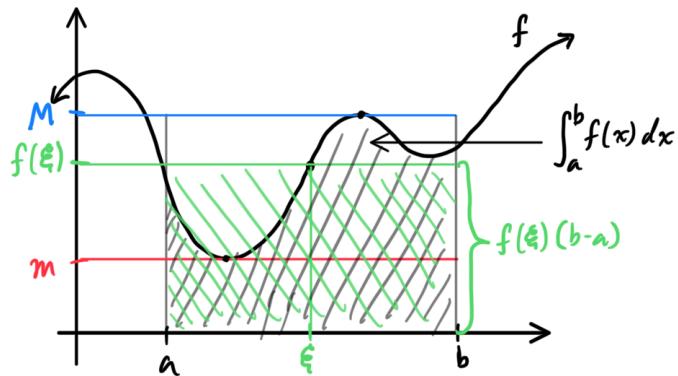
$$m = \inf_{x \in [a, b]} f(x) \text{ and } M = \sup_{x \in [a, b]} f(x)$$

then there exists a number $\mu \in [m, M]$ such that

$$\int_a^b f(x) dx = \mu \cdot (b - a)$$

Furthermore, if $f \in C[a, b]$ (that is, continuous on $[a, b]$), it immediately follows by the intermediate value theorem that there exists a point $\xi \in [a, b]$ such that

$$\int_a^b f(x) dx = f(\xi)(b - a)$$



Due to the length of the proof, we ask the reader to take it for granted the following theorem.

Theorem 7.10 (Bonnet's Formula)

If $f, g \in \mathcal{R}[a, b]$ and g is a monotonic function on $[a, b]$, then there exists a point $\xi \in [a, b]$ such that

$$\int_a^b (f \cdot g)(x) dx = g(a) \int_a^\xi f(x) dx + g(b) \int_\xi^b f(x) dx$$

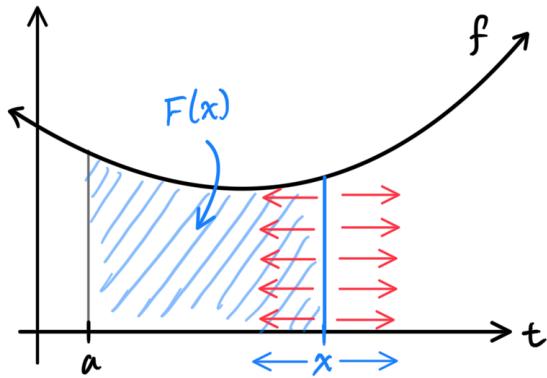
7.3 Connections between Integrals, Primitives, Derivatives

Definition 7.8 (Integral with Variable Upper Limit)

Let $f \in \mathcal{R}[a, b]$, and let us choose an $x \in [a, b]$ in order to construct the function

$$F(x) \equiv \int_a^x f(t) dt$$

which is called an **integral with variable upper limit**. Note that since $[a, x] \subset [a, b]$, it follows that $f|_{[a, x]} \in \mathcal{R}[a, x]$ and therefore the function $x \mapsto F(x)$ is unambiguously defined for $x \in [a, b]$.



Furthermore, $F(x)$ is continuous on $[a, b]$. Since f is integrable on $[a, b]$, it is bounded by a constant C such that

$$|f(t)| \leq C \text{ on } [a, b]$$

It follows from the additive properties of the integral and boundedness theorem that

$$|F(x + h) - F(x)| \leq C|h|$$

if $x, x + h \in [a, b]$, as visualized. This means that for any δ -neighborhood of $F(x)$, we can find an arbitrary small h such that the $C|h|$ -neighborhood of $F(x)$ is completely contained in the δ -neighborhood. But by the inequality above, this means that there exists an $\epsilon = h$ -neighborhood of x such that its entire image is contained within the $C|h|$ -neighborhood, which itself is contained within the δ -neighborhood. This shows that F is continuous.

Theorem 7.11 (First Fundamental Theorem of Calculus)

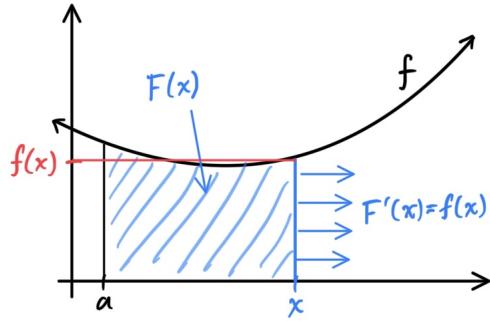
Let $f \in \mathcal{R}[a, b]$ be continuous at point $x \in [a, b]$ (resp. continuous on closed interval $[a, b]$). Let F be the function, defined for all $x \in [a, b]$ by

$$F(x) \equiv \int_a^x f(t) dt$$

Then, f is continuous and differentiable at x (resp. uniformly continuous on $[a, b]$ and differentiable on (a, b)),

$$F'(x) = f(x)$$

at x (resp. for all $x \in [a, b]$). This is an amazing fact, because visually, it tells us that the rate at which the integral F is increasing at x (represented by the increasing area under the curve of f) is equal to the value of f at the point x itself!



Proof.

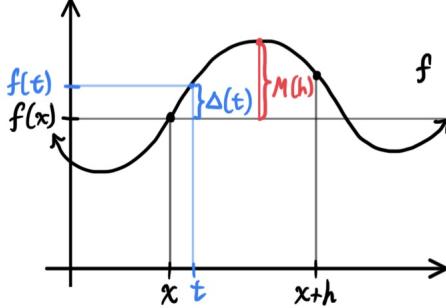
Let $x, x+h \in [a, b]$, and let us estimate the difference $F(x+h) - F(x)$. It follows from the continuity of f at x that $f(t) = f(x) + \Delta(t)$, where $\Delta(t) \rightarrow 0$ as $t \rightarrow x$. If point x is held fixed, the function

$$\Delta(t) = f(t) - f(x)$$

is integrable on $[a, b]$, being the difference of the integrable function $t \mapsto f(t)$ and the constant $f(x)$. Let us denote

$$M(h) \equiv \sup_{t \in [x, x+h]} |\Delta(t)|$$

which means that $M(h)$ is the largest difference between $f(x)$ and $f(t)$ in the interval $[x, x+h]$.



Clearly $M(h) \rightarrow 0$ as $h \rightarrow 0$. We can now find

$$\begin{aligned} F(x+h) - F(x) &= \int_a^{x+h} f(t) dt - \int_a^x f(t) dt \\ &= \int_x^{x+h} f(t) dt \\ &= \int_x^{x+h} (f(x) + \Delta(t)) dt \\ &= \int_x^{x+h} f(x) dt + \int_x^{x+h} \Delta(t) dt \\ &= f(x)h + \alpha(h)h \end{aligned}$$

where we have set

$$\int_x^{x+h} \Delta(t) dt = \alpha(h)h$$

where α is infinitesimal as $h \rightarrow 0$, since

$$\left| \int_x^{x+h} \Delta(t) dt \right| \leq \left| \int_x^{x+h} |\Delta(t)| dt \right| \leq \left| \int_x^{x+h} M(h) dt \right| = M(h)|h| = \alpha(h)|h|$$

Therefore, we have shown that if the function f is continuous at a point $x \in [a, b]$, then for displacements h from x such that $x+h \in [a, b]$, the following equality holds.

$$F(x+h) - F(x) = f(x)h + \alpha(h)h$$

where $\alpha(h) \rightarrow 0$ as $h \rightarrow 0$, and by definition, this means that $F(x)$ is differentiable on $[a, b]$ at the point $x \in [a, b]$ and that $F'(x) = f(x)$.

Corollary 7.4 ()

Every bounded function $f : [a, b] \rightarrow \mathbb{R}$ on the closed interval $[a, b]$ and has only a finite number of points of discontinuity has a primitive, and every primitive of f on $[a, b]$ has the form

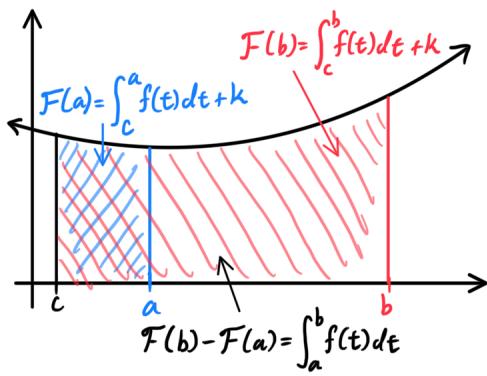
$$\mathcal{F}(x) \equiv \int_a^x f(t) dt + c$$

where c is a constant.

Theorem 7.12 (Second Fundamental Theorem of Calculus)

Let f be a real-valued function on a closed interval $[a, b]$ with \mathcal{F} any primitive of f on $[a, b]$. If f is Riemann-integrable (i.e. f bounded with finite points of Lebesgue measure zero) on $[a, b]$, then

$$\int_a^b f(x) dx = \mathcal{F}|_a^b \equiv \mathcal{F}(b) - \mathcal{F}(a)$$

**Proof.**

We already know that a bounded function on a closed interval having a finite number of discontinuities is integrable, and by the corollary, we are guaranteed an existence of a primitive $\mathcal{F}(x)$ of the function f on $[a, b]$ with the form

$$\mathcal{F}(x) \equiv \int_a^x f(t) dt + c$$

Setting $x = a$, we find that $c = \mathcal{F}(a)$, and so

$$\mathcal{F}(x) \equiv \int_a^x f(t) dt + \mathcal{F}(a)$$

Evaluating \mathcal{F} at $x = b$ gives

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

7.3.1 Integration by Parts and Taylor's Formula

Theorem 7.13 (Definite Integration by Parts)

If the functions $u(x)$ and $v(x)$ are continuously differentiable on a closed interval with endpoints a and b , then

$$\int_a^b (u \cdot v')(x) dx = (u \cdot v)|_a^b - \int_a^b (v \cdot u')(x) dx$$

which is customarily written in the form as

$$\int_a^b u dv = u \cdot v|_a^b - \int_a^b v du$$

Proof.

By the product rule of differentiation, we have

$$(u \cdot v)'(x) = (u' \cdot v)(x) + (u \cdot v')(x)$$

where by hypothesis, $u' \cdot v, u \cdot v'$ are continuous and hence integrable on $[a, b]$. Using the linearity of the integral and the 2nd fundamental theorem of calculus, we get

$$(u \cdot v)(x)|_a^b = \int_a^b (u' \cdot v)(x) dx + \int_a^b (u \cdot v')(x) dx$$

Theorem 7.14 (Integral Form of the Remainder)

If $f : E \rightarrow \mathbb{R}$ has continuous derivatives up to order n on the closed interval $[a, x]$, then Taylor's formula holds

$$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \dots + \frac{f^{(n-1)}(a)}{(n-1)!}(x - a)^{n-1} + r_{n-1}(a; x)$$

where

$$r_{n-1}(a; x) = \frac{1}{(n-1)!} \int_a^x f^{(n)}(t)(x - t)^{n-1} dt$$

This form is called **Taylor's formula with the integral form of the remainder**.

Proof.

Using the 2nd fundamental theorem and the definite integration by parts formula, we can carry out

the following chain of transformations, assuming continuity and differentiability when needed.

$$\begin{aligned}
 f(x) - f(a) &= \int_a^x f'(t) dt \\
 &= - \int_a^x f'(t)(x-t)' dt \\
 &= -f'(t)(x-t)|_a^x + \int_a^x f''(t)(x-t) dt \\
 &= f'(a)(x-a) - \frac{1}{2} \int_a^x f''(t)((x-t)^2)' dt \\
 &= f'(x-a) - \frac{1}{2} f''(t)(x-t)^2|_a^x + \frac{1}{2} \int_a^x f'''(t)(x-t)^2 dt \\
 &= f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2 - \frac{1}{2 \cdot 3} \int_a^x f'''(t)((x-t)^3)' dt \\
 &= \dots \\
 &= f'(a)(x-a) + \dots + \frac{1}{(n-1)!} f^{(n-1)}(a)(x-a)^{n-1} + r_{n-1}(a; x)
 \end{aligned}$$

where $r_{n-1}(a; x)$ is given by the integral formula mentioned.

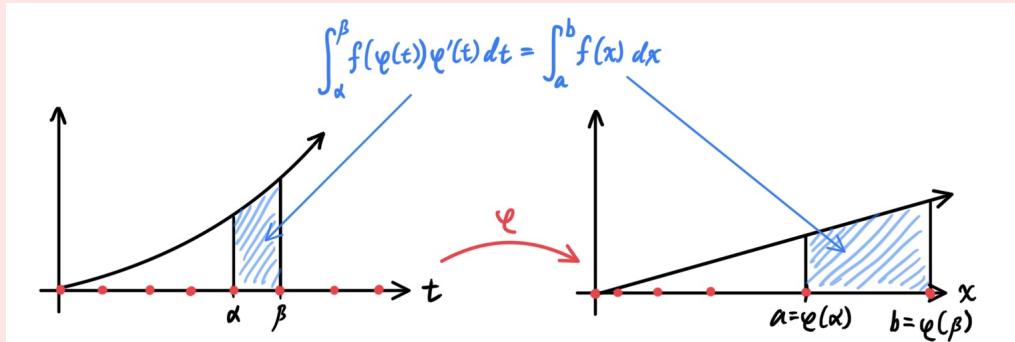
7.3.2 Change of Variables in Integration

We now show and prove the method what we call "u-substitution" for definite integration.

Theorem 7.15 (Change of Variable)

If $\varphi : [\alpha, \beta] \rightarrow [a, b]$ is a continuously differentiable mapping such that $\varphi(\alpha) = a$ and $\varphi(\beta) = b$, then for any continuous function $f(x)$ on $[a, b]$ the function $f(\varphi(t))\varphi'(t)$ is continuous on the closed interval $[\alpha, \beta]$ and

$$\int_a^b f(x) dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t) dt$$



Proof.

We prove a slightly weaker form of the theorem with the additional hypothesis that φ is strictly monotonic.

7.3.3 Additive Interval Functions and the Integral

In this section we take a step back and construct the integral in a more abstract sense, using the concepts of an additive interval function.

Definition 7.9 (Additive Interval Function)

An **additive (oriented) interval function** is a function

$$(\alpha, \beta) \mapsto I(\alpha, \beta) \in \mathbb{R}$$

that assigns a number $I(\alpha, \beta)$ to each ordered pair of points (α, β) of a fixed closed interval $[a, b]$ in such a way that the following equality holds for any triple of points $\alpha, \beta, \gamma \in [a, b]$.

$$I(\alpha, \gamma) = I(\alpha, \beta) + I(\beta, \gamma)$$

Notice that the integral holds this property, shown in the theorem on the symmetric property of the integral. It follows that all additive interval functions are anticommutative:

$$I(\alpha, \beta) + I(\beta, \alpha) = 0$$

which immediately results in

$$I(\alpha, \alpha) = 0$$

Lemma 7.6 (Generating Functions of Additive Interval Functions)

For any function $x \mapsto \mathcal{F}(x)$ that maps points on the interval $[a, b]$ to \mathbb{R} , we set

$$\mathcal{F}(x) \equiv I(a, x)$$

and by additivity we have

$$I(\alpha, \beta) = I(\alpha, \beta) - I(a, \alpha) = \mathcal{F}(\beta) - \mathcal{F}(\alpha)$$

and thus, every additive oriented interval function has the form

$$I(\alpha, \beta) = \mathcal{F}(\beta) - \mathcal{F}(\alpha)$$

By constructing I in this manner, we say that **the function \mathcal{F} generates the additive function I** .

Example 7.4 ()

If $f \in \mathcal{R}[a, b]$, the function $\mathcal{F} = \int_a^x f(t) dt$ generates the additive function

$$I(\alpha, \beta) = \mathcal{F}(\beta) - \mathcal{F}(\alpha) = \int_a^\beta f(t) dt - \int_a^\alpha f(t) dt = \int_\alpha^\beta f(t) dt$$

We conclude by stating a sufficient condition for an additive interval function to be generated by an integral.

Theorem 7.16 ()

Suppose the additive function $I(\alpha, \beta)$ defined for points $\alpha, \beta \in [a, b]$ has the property that, for some

known function $f \in \mathcal{R}[a, b]$,

$$\inf_{x \in [\alpha, \beta]} f(x)(\beta - \alpha) \leq I(\alpha, \beta) \leq \sup_{x \in [\alpha, \beta]} f(x)(\beta - \alpha)$$

holds for any closed interval $[\alpha, \beta] \subset [a, b]$ ($\alpha \leq \beta$). Then, the additive function I must be the definite integral

$$I(a, b) = \int_a^b f(x) dx$$

This theorem is extremely useful. It says that if we have any abstract additive interval function $I(\alpha, \beta)$ that satisfies the properties above, then it **must** be generated by an integral with variable upper limit, meaning that (by the previous example) I itself must be a definite integral!

7.3.4 Arc Length

When modeling systems in physics, one of the most fundamental tools we use are path functions that models the movement of a particle in \mathbb{R}^3 .

Definition 7.10 (Path)

A **path** in \mathbb{R}^3 is a continuous mapping $r : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^3$ defined

$$t \mapsto (x(t), y(t), z(t))$$

of an interval of the real line into \mathbb{R}^3 defined by the (continuous) scalar functions x, y, z . The endpoints

$$A = (x(a), y(a), z(a)) \text{ and } B = (x(b), y(b), z(b))$$

in \mathbb{R}^3 are called the **initial point** and **terminal point** of the path. Furthermore, a path is **closed** if its initial and terminal points coincide.

Definition 7.11 (Support)

If $\Gamma : I \rightarrow \mathbb{R}^3$ is a path, the image $\Gamma(I) \subset \mathbb{R}^3$ is called the **support** of the path.

Definition 7.12 (Simple Paths)

A path $\Gamma : I \rightarrow \mathbb{R}^3$ that is injective is called a **simple path**, or a **paramaterized curve**, and its support is called a **curve** in \mathbb{R}^3 .

A closed path $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ is called a **simple closed path/curve** if the path $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ is simple.

Definition 7.13 (Smooth Paths)

A path $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ is C^k smooth if the functions $x(t), y(t), z(t)$ are C^k smooth. Γ is **piecewise smooth** if the closed interval $[a, b]$ can be partitioned into a finite number of closed intervals on each of which the corresponding restriction of Γ is smooth.

Now, we are ready to construct the length of a smooth path $\Gamma : [a, b] \rightarrow \mathbb{R}^3$. Our initial ideas about the length $l[a, b]$ of the path traversed during the time interval $\alpha \leq t \leq \beta$ are as follows:

1. If $\alpha < \beta < \gamma$, then l is an additive interval function.

$$l[\alpha, \gamma] = l[\alpha, \beta] + l[\beta, \gamma]$$

2. If $v(t) = (x'(t), y'(t), z'(t))$ is the velocity of the point at time t , then

$$\int_{x \in [\alpha, \beta]} |v(t)|(\beta - \alpha) \leq l[\alpha, \beta] \leq \sup_{x \in [\alpha, \beta]} |v(t)|(\beta - \alpha)$$

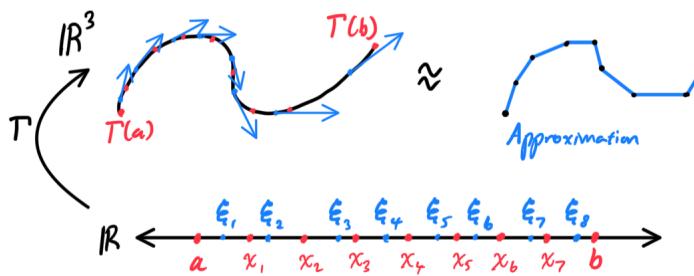
Thus, if the functions x, y, z are continuously differentiable on $[a, b]$, this is sufficient condition (by the theorem in the previous subsection) that the additive function l is an integral.

Definition 7.14 (Arc Length Integral)

The length of a smooth path $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ is defined by

$$l[a, b] \equiv \int_a^b |\Gamma'(t)| dt \equiv \int_a^b \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} dt$$

We can visualize this by partitioning the interval $[a, b]$ into the intervals Δ_i , each with point $\xi_i \in \Delta_i$. This would partition the path to $\Gamma(\Delta_i)$, each with points $\Gamma(\xi_i)$, and at each point $\Gamma(\xi_i)$, we can imagine the velocity vector of the curve. By taking the magnitude of this vector $\Gamma'(\xi_i)$, we multiply it by the length of the interval Δx_i to get one rectangle, creating an approximation for one partition of the path.



An immediate result of this formula is the formula for the length of a graph of a function $f : [a, b] \rightarrow \mathbb{R}$ in \mathbb{R}^2 , by looking at the paramaterization $t \mapsto \Gamma(t) = (t, f(t))$.

$$l[a, b] \equiv \int_a^b \sqrt{1 + (f'(t))^2} dt$$

The question on the effect of paramaterization on the integral now arises.

Definition 7.15 (Admissible Change of Parameter)

The path $\tilde{\Gamma} : [\alpha, \beta] \rightarrow \mathbb{R}^3$ is obtained from $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ by an **admissible change of parameter** if there exists a smooth mapping

$$T : [\alpha, \beta] \rightarrow [a, b]$$

such that $T(\alpha) = a, T(\beta) = b, T'(\tau) > 0$ (that is, the reparamaterization T is monotonic) on $[\alpha, \beta]$, and

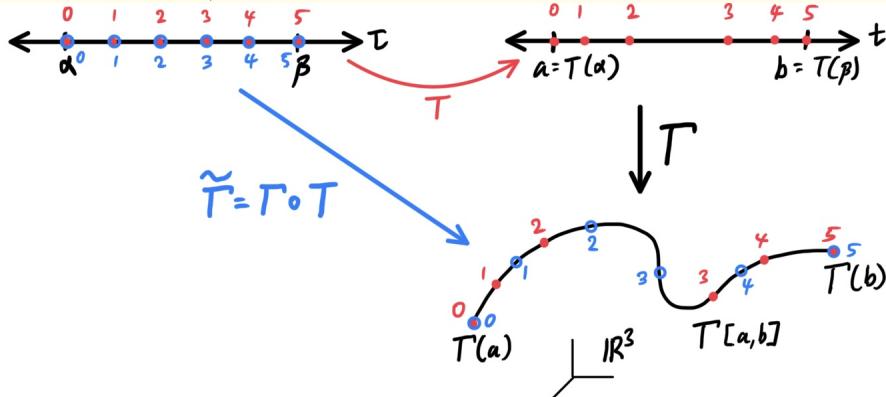
$$\tilde{\Gamma} = \Gamma \circ T$$

The series of mappings can be represented with the following commutative diagram, where $I_{\alpha, \beta} =$

$[\alpha, \beta] \subset \mathbb{R}$ and $I_{\alpha, \beta} = [a, b] \subset \mathbb{R}$.

$$\begin{array}{ccc} I_{\alpha, \beta} & \xrightarrow{T} & I_{a, b} \\ & \searrow \tilde{\Gamma} & \downarrow \Gamma \\ & & \mathbb{R}^3 \end{array}$$

or with the more detailed visual below (Note that the points are labeled 0, 1, 2, 3, 4, 5 do not represent numerical values, but rather the order in which the points are parameterized. We can see from this ordering that T is monotonic.)



Theorem 7.17 (Invariance of Arclength Integral under Admissible Change of Parameters)

If a smooth path $\tilde{\Gamma} : [\alpha, \beta] \rightarrow \mathbb{R}^3$ is obtained from a smooth path $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ by an admissible change of parameter, then the lengths of the two paths are equal. That is,

$$\int_a^b |\Gamma'(t)| dt = \int_\alpha^\beta |\tilde{\Gamma}'(t)| dt \equiv \int_\alpha^\beta |(\Gamma \circ T)'(t)| dt$$

7.4 Improper Integrals

Due to some limitations of the Riemann integral, we cannot integrate over "singularities" where either the interval or the function is unbounded. We develop the tools of improper integration to deal with this problem; there are two types of improper integrals.

Definition 7.16 (Improper Integral of Unbounded Interval)

Suppose the function $x \mapsto f(x)$ is defined on the interval $[a, +\infty)$ and is integrable on every closed interval $[a, b]$ contained in that interval. Then, we call the following term

$$\int_a^{+\infty} f(x) dx \equiv \lim_{b \rightarrow +\infty} \int_a^b f(x) dx$$

the **improper Riemann integral of f over the interval $[a, +\infty)$** and

$$\int_{-\infty}^b f(x) dx \equiv \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

the **improper Riemann integral of f over the interval $(-\infty, b]$** . If the limit exists, then we say that the integral **converges** and **diverges** otherwise.

Definition 7.17 (Improper Integral of Unbounded Function)

Suppose the function $x \mapsto f(x)$ is defined on the interval $[a, B)$ and integrable on any closed interval $[a, b] \subset [a, B)$. Then, we call the following term

$$\int_a^B f(x) dx \equiv \lim_{b \rightarrow B^-} \int_a^b f(x) dx$$

the **improper Riemann integral of f over interval $[a, B)$** and

$$\int_A^b f(x) dx \equiv \lim_{a \rightarrow A^+} \int_a^b f(x) dx$$

the **improper Riemann integral of f over interval $(A, b]$** .

For cohesiveness, we can combine these two definitions of improper integrals into the following one.

Definition 7.18 (Improper Integrals)

Let $[a, \omega)$ be a finite or infinite interval and $x \mapsto f(x)$ a function defined on that interval and integrable over every closed interval $[a, b] \subset [a, \omega)$. Then, by definition

$$\int_a^\omega f(x) dx \equiv \lim_{b \rightarrow \omega} \int_a^b f(x) dx$$

if this limit exists as $b \rightarrow \omega, b \in [a, \omega)$. Similarly, given the finite or infinite interval $(\omega, b]$ with f integrable over every closed interval $[a, b] \subset (\omega, b]$, we have

$$\int_\omega^b f(x) dx \equiv \lim_{a \rightarrow \omega} \int_a^b f(x) dx$$

Note that if $\omega \in \mathbb{R}$ and $f \in \mathcal{R}[a, \omega]$, the improper integral is equivalent to the regular Riemann integral.

$$\int_a^\omega f(x) dx = \lim_{b \rightarrow \omega} \int_a^b f(x) dx$$

Lemma 7.7 (Properties of the Improper Integral)

Suppose f, g are functions defined on interval $[a, \omega)$ (without loss of generality, we let ω be the upper limit of integration) and integrable on every closed interval $[a, b] \subset [a, \omega)$. Suppose the improper integrals

$$\int_a^\omega f(x) dx \text{ and } \int_a^\omega g(x) dx$$

are well-defined.

1. For any $\lambda_1, \lambda_2 \in \mathbb{R}$ the function $(\lambda_1 f + \lambda_2 g)(x)$ is integrable in the improper sense on $[a, \omega)$ and

$$\int_a^\omega (\lambda_1 f + \lambda_2 g)(x) dx = \lambda_1 \int_a^\omega f(x) dx + \lambda_2 \int_a^\omega g(x) dx$$

2. For any $c \in [a, \omega)$,

$$\int_a^\omega f(x) dx = \int_a^c f(x) dx + \int_c^\omega f(x) dx$$

3. If $\varphi : [\alpha, \gamma) \rightarrow [a, \omega)$ is a smooth strictly monotonic mapping with $\varphi(\alpha) = a$ and $\varphi(\beta) \rightarrow \omega$ as

$\beta \rightarrow \gamma^-$, then the improper integral of the function $t \mapsto (f \circ \varphi)(t)\varphi'(t)$ over $[\alpha, \gamma)$ exists and

$$\int_a^\omega f(x) dx = \int_\alpha^\gamma (f \circ \varphi)(t)\varphi'(t) dt$$

7.4.1 Convergence of an Improper Integral

Note that by definition, an improper integral

$$\int_a^\omega f(x) dx \equiv \lim_{b \rightarrow \omega} \int_a^b f(x) dx$$

is a limit of the function

$$\mathcal{F}(b) \equiv \int_a^b f(x) dx$$

as $b \rightarrow \omega$. This means that we can use the Cauchy criterion to determine the convergence of this limit, and hence, existence of this improper integral.

Theorem 7.18 (Cauchy Criterion for Convergence of an Improper Integral)

If the function $x \mapsto f(x)$ is defined on the interval $[a, \omega)$ and integrable on every closed interval $[a, b] \subset [a, \omega)$, then the integral

$$\int_a^\omega f(x) dx$$

converges if and only if for every $\epsilon > 0$ there exists $B \in [a, \omega)$ such that the relation

$$\left| \int_{b_1}^{b_2} f(x) dx \right| < \epsilon$$

holds for any $b_1, b_2 \in [a, \omega)$ satisfying $B < b_1$ and $B < b_2$.

Proof.

We have

$$\int_{b_1}^{b_2} f(x) dx = \int_a^{b_2} f(x) dx - \int_a^{b_1} f(x) dx = \mathcal{F}(b_2) - \mathcal{F}(b_1)$$

and therefore the condition is simply the Cauchy criterion for the existence of a limit for the function $\mathcal{F}(b)$ as $b \rightarrow \omega$.

Definition 7.19 (Absolute Convergence of an Improper Integral)

The improper integral

$$\int_a^\omega f(x) dx$$

converges absolutely if the integral

$$\int_a^\omega |f|(x) dx$$

converges. Clearly, the inequality

$$\left| \int_{b_1}^{b_2} f(x) dx \right| \leq \left| \int_{b_1}^{b_2} |f|(x) dx \right|$$

implies that if an improper integral converges absolutely, then it converges.

This study of absolute convergence reduces to the study of convergence of integrals of nonnegative functions. The following lemma is useful in determining convergence of such functions.

Lemma 7.8 ()

Let there be a function f defined on interval $[a, \omega)$ that is also integrable over every closed interval $[a, b] \subset [a, \omega)$. If $f(x) \geq 0$ on $[a, \omega)$, then the improper integral

$$\int_a^\omega f(x) dx$$

exists if and only if the function

$$\mathcal{F}(b) = \int_a^b f(x) dx$$

is bounded on $[a, \omega)$.

Proof.

It is clear that

$$\int_a^\omega f(x) dx = \lim_{b \rightarrow \omega} \mathcal{F}(b)$$

If $f(x) \geq 0$, then the function $\mathcal{F}(b)$ is nondecreasing on $[a, \omega)$ and therefore has a limit as $b \rightarrow \omega$ only if it is bounded (since every monotonically increasing sequence that is bounded always converges).

This leads to the familiar integral test for convergence of a series.

Theorem 7.19 (Integral Test for Convergence of a Series)

If the function $x \mapsto f(x)$ is defined on the interval $[1, +\infty)$, nonnegative, nonincreasing, and integrable on each closed interval $[1, b] \subset [1, +\infty)$, then the series

$$\sum_{n=1}^{\infty} f(n) = f(1) + f(2) + \dots$$

and the integral

$$\int_a^{+\infty} f(x) dx$$

either both converge or both diverge.

We can use the comparison test analogue to determine convergence of improper integrals.

Theorem 7.20 (Comparison Test for Convergence of Improper Integrals)

Suppose the functions $f(x), g(x)$ are defined on the interval $[a, \omega)$ and integrable on any closed interval $[a, b] \subset [a, \omega)$. If

$$0 \leq f(x) \leq g(x)$$

on $[a, \omega)$, then

$$\int_a^\omega g(x) dx \text{ converges} \implies \int_a^\omega f(x) dx \text{ converges}$$

and the inequality

$$\int_a^\omega f(x) dx \leq \int_a^\omega g(x) dx$$

holds. Also,

$$\int_a^\omega f(x) dx \text{ diverges} \implies \int_a^\omega g(x) dx \text{ diverges}$$

7.4.2 Improper Integrals with Multiple Singularities

Definition 7.20 (Improper Integral with Both Limits as Singularities)

Given singularities ω_1, ω_2 , the improper integral is defined

$$\int_{\omega_1}^{\omega_2} f(x) dx \equiv \int_{\omega_1}^c f(x) dx + \int_c^{\omega_2} f(x) dx$$

where c is an arbitrary point in (ω_1, ω_2) .

Example 7.5 (Gaussian Integral)

The integral

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$