Analysis Notes 131AH

OMKAR TASGAONKAR

Winter Quarter - January-March 2026

If you are reading this...

Hello! These are my notes for analysis, which I have typed up to save to github so I can reference them in the future. Things will be written in my own words and may not be fully correct, but this is just my attempt to fully internalize everything and write it down as precisely as possible.

Contents

1 Propositional and First-Order Logic

> **Definition 1.1.** *A proposition is a statement that takes a TRUE or FALSE value*

For example, "birds are mammals" is a valid proposition, which takes on the value of FALSE. We also define a "primitive proposition", which is one with no connectives or quantifiers.

> **Definition 1.2.** *A connective is a unary or binary operator allowing us to chain propositions to create new propositions*

The connectives we will work with are $\neg$ (logical not), $\wedge$ (and), $\vee$ (or), $\implies$ (implies), and $\iff$ (biconditional). The $\neg$ operator is the only unary operator, which switches the value of the proposition. The others are binary, requiring two propositions and outputting one value.

We view an interesting truth table for $\implies$:

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

When P is false but Q is true, the entire statement is "vacuously" true as the expected proposition Q is true no matter P. When both are false, the proposition is true as well. So in general when P is false, $P \implies Q$ is true.

> **Lemma 1.1.** $(P \implies Q) \iff \neg(P \wedge \neg Q) \iff \neg P \vee Q$

Here the biconditional means that the propositions are equivalent. We can use the biconditional to represent when two statements imply each other or are logically the same.

This lemma highlights proof by contradiction, where we assume P and $\neg Q$ and we show some contradiction (or show that the proposition is false). Then the logical not of the proposition is true. The second equivalence is an application of DeMorgan's Laws for logic.

> **Lemma 1.2.** $(P \implies Q) \iff \neg Q \implies \neg P$

This is the method of proof by contrapositive. This implication is called the **"only if"** direction, while the reverse is the **"if"** direction.

> **Definition 1.3.** *We propose two quantifiers: $\forall$ (for all/for each) and $\exists$ (there exists)*

A first order statement assumes the form: (quantifier)(variable): (proposition). This is the basis of First Order Logic, an extension of propositional logic. For example:

$$\forall x : P(x)$$

translates to: for all x such that $P(x)$ is true. The terminology "some" is equivalent to $\exists$.

Lemma 1.3. $\neg(\forall x : P(x)) \iff \exists x : \neg P(x)$ *and* $\neg(\exists x : P(x)) \iff \forall x : \neg P(x)$

Here we see how the logical not operator distributes over quantifier and proposition.

Theorem 1.1. $\forall x \forall y : P(x,y) \iff \forall y, \forall x : P(x,y)$ *and* $\exists x, \exists y : P(x,y) \iff \exists y, \exists x : P(x,y)$

If the quantifiers for two variables are the same, we can switch the order of the quantifiers. This will appear in proofs involving sets and supremums/infimums. However, if they are not the same, then we cannot suggest an equivalence.

Theorem 1.2. $\exists x \forall y : P(x,y) \implies \forall y \exists x : P(x,y)$

The reason why this is a one-way implication is because the left hand statement is more specific. We will encounter this when dealing with images and preimages. We take the following example:

Mathematical Example: If there exists one x greater than all y, then for each y there exists an $x > y$. In this case that x is the same. For the converse, if each y has an $x > y$, there may not be a singular x greater than all y.

Figurative Example: If there is one building to which all apartments belong, then all the apartments belong to that building. However, if all the apartments belong to a building, not necessarily they belong to one building. We can have the apartments in multiple buildings.

2 Naive Set Theory and Zermelo Fraenkel

2.1 Naive Set Theory

Naive Set Theory was the initial proposed set theory. A set being a container for some sort of objects. To construct a set, we have the following:

> **Definition 2.1.** *The Comprehension Principle:* $A := \{x : P(x)\}$

The symbol $:=$ means "defined as". If we use "$=$" when referring to sets, that is not the same as using $:=$.

The Comprehension Principle tells us that a set is some elements x that satisfy a proposition $P(x)$. We won't dwell on this theory too much, but it has pitfalls such as Russell's paradox:

$$A := \{x : x \notin x\}$$

Here if $x \notin x$, then we could say $A \notin A$, but the proposition tells us that $A \in A$, so we come across a contradiction.

2.2 Zermelo Fraenkel

> **Axiom 2.1.** *Axiom of Separation* Any set A can be defined as $\{x \in U : P(x)\}$.

We see that the elements $x \in A$ are being drawn from U which in this case can be a universe (all the possible sets, but this is not itself a set), or a bigger set. By restricting elements, we can prevent self references such as those in Russell's paradox.

> **Axiom 2.2.** *Axiom of Extensionality* $\forall A, B : A = B \iff \forall x : (x \in A \iff x \in B)$

We mentioned previously that $:=$ is not the same as $=$ for sets. Extensionality gives the definition of $=$ to be that the sets must have the same elements. The axiom itself is sometimes written as an implication rather than a biconditional.

> **Axiom 2.3.** *Empty Set:* $\exists \varnothing : \forall x \in U : x \notin \varnothing$

We postulate the existence of the empty set, which contains none of the possible elements in the universe. Now, how do we build the set of all subsets (aka the powerset)?

> **Definition 2.2.** $\subseteq$ *(subset relation):* $\forall A, B : A \subseteq B \implies \forall x \in A : x \in B$

Let's try to use the Axiom of Separation and this definition to construct the powerset. We can say maybe for a set A that $\mathcal{P}(A) := \{B \in C : B \subseteq A\}$. But this definition requires a set C to exist in which B belongs, so either this is the powerset or a set containing more elements than the powerset, and so we have a circular definition.

> **Axiom 2.4.** *Powerset:* $\forall A \exists \mathcal{P}(A) : (\forall B \subset A : B \in (P)(A))$

We have the same issue when we go to define a set such as $\{x, y\}$, which is called the pairset, as each element is also inherently a set.

Axiom 2.5. *Pairset:* $\forall x, y \exists A \forall Z \in A : x = z \lor y = z$

The pairset allows us to show the existence of a singleton set $\{x\}$ if both the elements given are the same as $\{x, x\} = \{x\}$. This can be shown by extensionality.

Axiom 2.6. *Infinite Set:* $I := \{x : x \in I \implies \{x\} \in I\}$

This is one possible construction of the infinite set we will use for the naturals; however, this axiom postulates the existence of infinite sets. Many of the constructions in analysis will hinge on this axiom; without it, we fail to construct the idea of "natural numbers" let alone "real numbers".

3   Ordered Pairs and Relations

We talk about pairs (or n-tuples) of numbers such as $(1, 2)$ or (x, y). However, we also distinguish the order unlike sets, saying that $(1, 2)$ is not the same as $(2, 1)$.

> **Definition 3.1. *Kuratowski Pair:*** $(x, y) \coloneqq \{\{x\}, \{x, y\}\}$

4 Isomorphisms of Dedekind Cuts and Ordered Fields

So far we have constructed the reals or $\mathbb{R}$ as the set of Dedekind cuts, which are the set of rationals less than that real number. By allowing for completeness, we can show that each cut admits an upper bound so it admits a supremum.

The supremum is the real number itself; however, quantifying the supremum will require the existence of an algebraic structure that encompasses the supremum. If we solely rely on $\mathbb{Q}$, then certain cuts may not admit supremums in $\mathbb{Q}$.

So we say there exists some complete ordered field $(F, +, 0, \cdot, 1, \leqslant)$ (completeness is important here, otherwise we cannot guarantee supremums). Each ordered field fundamentally has the $\mathbb{N}_F$. We examine a theorem:

Theorem 4.1. *If F is an ordered field, then it has characteristic 0*

We prove this by contrapositive:

Proof. Assume F does not have characteristic 0. Then $\sum_{k=1}^{n}(1) = 0$ for an arbitrary n. We then know that $0 < 1$, by ordered field properties:

$$0 < 1 \implies 0 < \sum_{k=1}^{n-1}(1) < \sum_{k=1}^{n-1}(1) + 1 = \sum_{k=1}^{n}(1)$$

Now assume the inequality to be true, then by mulitplicative property of ordered fields:

$$\implies \sum_{k=1}^{n-1}(1) \cdot \sum_{k=1}^{n-1}(1) < \sum_{k=1}^{n}(1) \cdot \sum_{k=1}^{n-1}(1)$$

$$\iff (n-1)^2 = n^2 - 2n + 1 < n(n-1) = n^2 - n \tag{1}$$

We know that n or 1 added n times repeatedly is equal to zero, so we get:

$$0 + 1 = 1 < 0$$

Which is a contradiction of a key property of ordered fields that $0 \leqslant 1$. $\qquad\square$

So now, the addition of of 1 to an element of the field becomes our sucessor operation, and 0 becomes our zero element. If we then consider the subset of elements greater than zero, then zero is not a sucessor of any number in the set. Thus we can construct the naturals in a field as:

$$\mathbb{N}_F = \bigcap_{\alpha \in I} \{A_\alpha \subseteq F : (\forall x \in A_\alpha : x + 1 \in A_\alpha)\}$$

Now that we have the naturals in every ordered field, by field properties, we require their additive and multiplicative inverses. Furthermore any combination of these must also remain in the field, giving rise to the rationals $\mathbb{Q}_F$ (considered the "minimal" ordered field).

So currently we have the set of Dedekind cuts $\mathbb{R}$ and an ordered field structure F. One is constructed from ground up, the other's existence is assumed respectively. We aim to show the two are isomorphic. We know the following theorem proved in prior chapters:

Theorem 4.2. *The rationals of different fields are isomorphic*

So there exists an isomorphism we call $\psi : \mathbb{Q} \to \mathbb{Q}_F$. Since the Dedekind cuts are merely sets of rationals, then we can construct cuts in F which we call $\mathbb{R}_F$. We can extend $\psi : \mathbb{R} \to \mathbb{R}_F$.

Theorem 4.3. $\forall A \in \mathbb{R}_F : (\exists b \in F, \forall a \in A : a \leqslant b) : \sup : A \to F$ *and* $\sup$ *is an isomorphism*

This lemma is essentially the restatement of the completeness of F. If we take any subset bounded above, then it admits a supremum. However, here we define the supremum as a map from a set to an element in the field. We prove it is an isomorphism for a cut specifically. We require the lemma:

Lemma 4.1. $\forall A \in \mathbb{R}_F : A = \{a \in \mathbb{Q}_F : a < \sup(A)\}$

For now we only show the inclusion $A \subseteq \{a \in \mathbb{Q}_F : a < \sup(A)\}$.

Proof. If $x \in A$, then $x \leqslant \sup(A)$. If $x = \sup(A) \implies \sup(A) \in A$. By the definition of Dedekind cut's, there cannot be a maximal element, so:

$$\exists b \in A : \sup(A) < b \wedge b \leqslant \sup(A)$$

The second inequality comes from the supremum definition. We see that we get a contradiction, so $\sup(A) \notin A$, thus $\forall a \in A : a < \sup(A)$. $\qquad\square$

We now return to proving the theorem at hand.

Proof. For the sake of brevity we assume sup is a homomorphism (which means that it respects the operations of sets and the field). We aim to show injectivity:

$$\forall A, B : \sup(A) = \sup(B) \implies A = B$$

If two cuts are not equal, then either $A \subset B \vee B \subset A$. Assume without loss of generality that $A \subset B$. Then:

$$\exists b \in B, \forall a \in A : a < b$$

This inequality we can derive from the ordering of F and the construction of the cuts. Then b is an upper bound for A and by definition of supremum $\sup(A) \leqslant b$. So by the lemma $a < \sup(A) \leqslant b < \sup(B)$. Thus $\sup(A) < \sup(B)$, proving injectivity of sup. $\qquad\square$

Proof. We now prove the surjectivity of sup:

$$\forall x \in F, \exists A \in \mathbb{R}_F : \sup(A) = x$$

Assume for a certain $x \in F$, that there does not exist an $A \in \mathbb{R}_F$ such that $\sup(A) = x$. Then all Dedekind cuts must admit a supremum in F as they are bounded above. (The following part may be a bit hand wavy :), but in essence by means of pigeonhole principle (whose proof we detailed before), two different cuts will then share the same supremum, which breaks injectivity. $\qquad\square$

Theorem 4.4. *For all complete ordered fields* $(F, +, 0, \cdot, 1, \leqslant)$, *we find that* $\mathbb{R} \cong F$.

The proof of this is simply a combination of the previous theorems and lemmas. We know that $\psi : \mathbb{R} \to \mathbb{R}_F$ is an isomorphism and $\sup : \mathbb{R}_F \to F$ is also an isomorphism. The composition is therefore an isomorphism:

$$\sup \circ \psi : \mathbb{R} \to F$$

So the Dedekind cut representation of the reals are isomorphic to any complete ordered field. So now instead of dealing with the reals as a set of Dedekind cuts, we can deal with it as an ordered field of the supremums of the Dedekind cuts. So $\sqrt{2}$ is a supremum of the cut, so we can explicitly refer to the cut as $\sqrt{2}$ itself even though the cut does not admit a supremum in $\mathbb{Q}$.

5 Properties of the Reals

We have determined that every complete ordered field is isomorphic to the real numbers. We now aim to show certain properties have correctly carried forward to this construction.

We dealt with rational roots and equalities such as $x^2 = 2$ that did not admit solutions in the rationals, but allowed us to construct $\sqrt{2}$ in the reals. We now see the reals hold for arbitrary roots as well:

> **Theorem 5.1.** $\forall a \in \mathbb{R}^+, \forall n \in \mathbb{N} \setminus \{0\} : (\exists x \in \mathbb{R}^+ : x^n = a) \wedge (x^n = y^n \implies x = y)$

In essence, there exists an *nth* root for every real number, which is unique. We use the following property:

$$\forall x, y \in \mathbb{R}, \forall n \in \mathbb{N} \setminus \{0\} : x^n - y^n = (x - y) \sum_{k=0}^{n-1} x^k y^{n-k-1}$$

Integer exponentiation can be defined by field properties as repeated multiplication of the element in the field.

HW 5 Problem: We prove the property below:

Proof. By the distributive property of multiplication in fields:

$$(x - y) \sum_{k=0}^{n-1} x^k y^{n-k-1} = x \sum_{k=0}^{n-1} x^k y^{n-k-1} - y \sum_{k=0}^{n-1} x^k y^{n-k-1}$$

$$= \sum_{k=0}^{n-1} x^{k+1} y^{n-k-1} - \sum_{k=0}^{n-1} x^k y^{n-k}$$

We can notice a pattern of cancellations here. The last term of the left term is x^n and the first term of the right term is y^n. We can then rewrite the sum as:

$$= x^n - y^n + \sum_{k=0}^{n-2} x^{k+1} y^{n-k-1} - \sum_{k=1}^{n-1} x^k y^{n-k}$$

We see that the two summations output the same terms, just that the starting and ending indices are up by 1. If our starting term is xy^{n-1}, then at $k = 0$ we sum $x^{k+1} y^{n-k-1}$, but at $k = 1$, we sum $x^k y^{n-k}$.

If our last term is $x^{n-1} y$, then the same applies. So the two summations are equal, thus their subtraction equals 0, and our resultant expression is $x^n - y^n$. $\square$

We return to the proof of the theorem. We define the set $A := \{y \in \mathbb{R}^+ : y^n \leqslant a\}$. Since this is a subset of the field, if it is bounded, then it admits a supremum:

$$1 + a > a \implies (1 + a)^n > (1 + a)^{n-1} a > a^n > a \quad \text{(Ordering Property of Multiplication)}$$

Thus $\forall y \in A : y \leqslant 1 + a$ so A is bounded above, thus it admits a supremum. We want to show that $\sup(A)^n = a$, thus defining the supremum of the set as the n-th root and showing its existence (as we know supremum exists).

Proof. We use the fact that $\forall y \in A : y \leqslant \sup(A)$, We first show $\sup(A)^n \leqslant A$. We also use the fact that:

$$\forall m \in \mathbb{N} \setminus \{0\} : \sup(A) - \frac{1}{m} < \sup(A) \implies \exists y \in A : \sup(A) - \frac{1}{m} \leqslant y$$

$$\sup(A)^n - a \leqslant \sup(A)^n - y^n = (\sup(A) - y) \sum_{k=0}^{n-1} \sup(A)^k y^{n-k-1}$$

9

Our previous implication that $\sup(A) - \frac{1}{m} \leqslant y \iff \sup(A) \leqslant y + \frac{1}{m} \implies \sup(A) \leqslant y$. Since $y \in A$, we know that $y \leqslant \sup(A)$, so $y = \sup(A)$. Our inequality becomes:

$$\sup(A)^n - a \leqslant \sup(A)^n - y^n = (\sup(A) - y)\sum_{k=0}^{n-1}\sup(A)^k y^{n-k-1} = 0 \iff \sup(A)^n \leqslant a$$

We now prove that $\sup(A)^n \geqslant a$. We note that $\forall m \in \mathbb{N} \setminus \{0\} : \sup(A) + \frac{1}{m} > \sup(A)$, so:

$$(\sup(A) + \frac{1}{m})^n > a \implies a - \sup(A)^n < (\sup(A) + \frac{1}{m})^n - \sup(A)^n$$

$$= \frac{1}{m}\sum_{k=0}^{n-1}(\sup(A) + \frac{1}{m})^k \sup(A)^{n-k-1} \leqslant \frac{n}{m}(\sup(A) + \frac{1}{m})^{n-1}$$

Thus, by dividing both sides (also the Archimedian Property is really a double implication), we see that:

$$\frac{m}{n}\frac{a - \sup(A)^n}{(\sup(A) + \frac{1}{m})^{n-1}} < 1 \implies a - \sup(A)^n \leqslant 0 \iff \sup(A)^n \geqslant a$$

So we have proven that $\sup(A)^n = a$, proving the existence of the nth root, which we will denote by $a^{1/n}$. $\square$

Proof. We now prove the uniqueness of the $n - th$ root. Assume there are two $n - th$ roots of a such that $y^n = a = \tilde{y}^n$. Then:
$$\sup(A)^n = y^n = \tilde{y}^n = a \implies a^{1/n} = \sup(A) = y = \tilde{y}$$

Since the supremum is unique, then if $y = \sup(A) = \tilde{y}$, then y must equal $\tilde{y}$, so the nth root is unique. $\square$

Now that we have the notion of natural powers and nth roots, we can extend this to rational roots. We know that any rational can be represented as the ratio of two integers p/q. Note the following:

> **Corollary 5.1.** $\forall n \in \mathbb{N}, \forall a \in \mathbb{R} : a^{-n} = (a^{-1})^n = (\frac{1}{a})^n \wedge a^0 = 1$

With this corollary, we are able to define the notion of integer powers (as integers have additive inverses not accounted for by natural powers). Another property of natural powers and by extension integer powers:

> **Corollary 5.2.** $\forall m, n \in \mathbb{N} : a^m \cdot a^n = a^{m+n} \wedge a^{m \cdot n} = (a^m)^n$

HW 5 Problem: Using this corollary, we can prove the following lemma. For this problem assume $a > 1$:

> **Lemma 5.1.** $\forall m, n, p, q \in \mathbb{Z} : n, q > 0 \wedge \frac{m}{n} = \frac{p}{q} \implies \forall a > 1 : (a^m)^{1/n} = (a^p)^{1/q}$

Proof. If $\frac{m}{n} = \frac{p}{q}$, then we know that $mq = np$ where the two are integers. If $m/n < 0$, then $p/n < 0$ and same for $m/n \geqslant 0$. Thus the two fractions take on the same sign. Therefore $mq = np$ will either be positive or negative.

If $mq = np < 0$, we simply define their additive inverses $m'q' = n'p' > 0$ and use them to construct an equality (eg $-a = -b \iff a = b$ by field properties). So without loss of generality we proceed with $mq = np > 0$:

$$a^{mq} = a^{qm} = (a^q)^m \wedge a^{np} = a^{pn} = (a^p)^n \tag{2}$$

This is true by properties of exponentiation with respect to naturals and field multiplication. Thus, by the existence of arbitary roots:

$$(a^q)^m = (a^p)^n \implies ((a^q)^m)^{1/n} = a^p \implies ((a^m)^q)^{1/n} = a^p$$

We need to show that we can exchange the places of q and $1/n$, so we prove $(a^m)^{1/n} = (a^{1/n})^m : n > 0$:

$$((a^m)^{1/n})^n = a^m \land ((a^{1/n})^m)^n = (a^{1/n})^{mn} = (a^{1/n})^{nm} = ((a^{1/n})^n)^m = a^m$$

Since both expressions to raised to the $n : n > 0$ are equal, then the nth roots are also equal so $(a^m)^{1/n} = (a^{1/n})^m$.

So returning to our initial statement:

$$((a^m)^q)^{1/n} = a^p \implies ((a^m)^{1/n})^q = a^p \implies (a^m)^{1/n} = (a^p)^{1/q}$$

$\square$

Corollary 5.3. $\forall p, q \in \mathbb{Z}, q \neq 0 : a^{\frac{p}{q}} = (a^p)^{1/q} = (a^{1/q})^p$

Lemma 5.2. $\forall r, s \in \mathbb{Q} : a^{r+s} = a^r \cdot a^s$

Proof. We let $r = \frac{m}{n}$ and $s = \frac{p}{q}$, where $m, n, p, q \in \mathbb{Z} \land p, q \neq 0$, then $r + s = \frac{m}{n} + \frac{p}{q} = \frac{mq+np}{pq}$. Thus by corollary 5.3:

$$a^{r+s} = a^{(mq+np)/pq} = (a^{1/pq})^{mq+np}$$

Now $mq + np$ is an integer expression, by corollary 5.2, we can rewrite the expression as:

$$= (a^{1/pq})^{mq} \cdot (a^{1/pq})^n p = (a^{mq/pq}) \cdot (a^{np/pq}) = a^{m/p} \cdot a^{n/q} = a^r \cdot a^s$$

$\square$

The idea of multiplicative inverses can now be brought about for rational powers. Let $p, q, m, n \in \mathbb{Z} : q, n \neq 0$, then:

$$a^{p/q} \cdot a^{m/n} = (a^{1/pq})^{mq+np} = 1$$

Using corollary 5.1, we account for the sign of $mq + np$. If the expression is greater than zero, than it is a natural number. If it is less than zero, we define its additive inverse $(mq + np)' > 0$ by field properties and change the base from a to a^{-1} as per the corollary.

Without loss of generality we proceed with the first case. Once again by corollary 5.1 stating that $a^0 = 1$:

$$mq + np = 0 \implies np = -mq$$

So by Lemma 5.2, $(a^{1/pq})^{mq+np} = (a^{1/pq})^{mq} \cdot (a^{1/pq})^{np} = (a^{1/pq})^{mq} \cdot (a^{1/pq})^{-mq}$. Since $mq \in \mathbb{Z}$, we again apply corollary 5.1 and 5.3:

$$(a^{1/pq})^{mq} \cdot (a^{1/pq})^{-mq} = a^{m/p} \cdot a^{-m/p} = 1$$

So we see that $(a^{m/p})^{-1} = a^{-m/p} = (a^{-1})^{m/p}$.

Now that we have defined and proven the properties of rational powers, we need to define real powers, which we do via supremums again leveraging completeness of the underlying ordered field:

Definition 5.1. $a^x := \begin{cases} \sup\{a^z : z \in \mathbb{Q} \land z \leqslant x\}, & a > 1 \\ \inf\{a^z : z \in \mathbb{Q} \land z \leqslant x\}, & a < 1 \\ 1, & a = 1 \end{cases}$

Proof. We prove the first case, where if $a > 1$ and $x \in \mathbb{Q}$, then $a^x = \sup\{a^z : z \in \mathbb{Q} \land z \leqslant x\}$. We know that $x < x + 1$. By density of the rationals:

$$\exists r \in \mathbb{Q} : x < r < x + 1$$

Since $z \leqslant q$, then $z < r$. We now need to prove the lemma:

Lemma 5.3. $\forall z, r \in \mathbb{Q} : z < r \implies a^z < a^r$.

If $z < r$, and $a > 1$, we see that $a^{r-z} > a > 1$ so by Lemma 5.2:

$$a^{r-z} = a^r \cdot a^{-z} = a^r \cdot (a^z)^{-1} > 1$$

Since $a > 1 > 0$, then $a^z > a > 0$. So if $(a^z)^{-1} < 0$. Thus:

$$a^r \cdot (a^z)^{-1} > 1 \implies 1 \cdot a^z < a^r \cdot (a^z)^{-1} \cdot a^z \iff a^z < a^r$$

Therefore a^r is an upper bound, allowing for the set to admit a supremum. We now want to show that a^x is the least upper bound. Consider any upper bound $a^p : p \in \mathbb{Q}$, then:

$$\forall z \leqslant x : a^z \leqslant a^p \implies a^x \leqslant a^p$$

If $a^p < a^x$, then $a^{x-p} > 1$. Since $a > 1$, then $a^{x-p} > a > 1$ only if $x > p$.

$$p < x \land z \leqslant x \implies \exists z \leqslant x : p < z \implies a^p < a^z$$

Thus, a^p is not an upper bound for all a^z, so we have proven the statement by contrapositive.

Lemma 5.4. $\forall x, y \in \mathbb{R} : a^{x+y} = a^x \cdot a^y$

We define $a^{x+y} = \sup\{a^z : z \in \mathbb{Q} \land z \leqslant x + y\}$.

Since $a^y = \sup\{a^p : p \in \mathbb{Q} \land p \leqslant y\}$ and $a^x = \sup\{a^q : q \in \mathbb{Q} \land q \leqslant x\}$:

$$a^x \cdot a^y = \sup\{a^q\} \cdot \sup\{a^p\} = \sup\{a^q \cdot a^p\} = \sup\{a^{q+p}\}$$

This follows from Lemma 5.2 and sup being a homomorphism. Multiplying the supremums combines the predicates in a logical $\land$ and results in a Minkowski product $\{a^{q+p} : q \leqslant x \land p \leqslant y\}$. By ordered field properties:

$$q \leqslant x \land p \leqslant y \implies q + p \leqslant x + y$$

By definition of the set of a^{x+y}, we see $\forall q \leqslant x, \forall p \leqslant y : a^{p+q} \in \{a^z : z \leqslant x + y\}$. Therefore:

$$\{a^{q+p}\} \subseteq \{a^z\} \implies \sup\{a^{q+p}\} \leqslant \sup\{a^z\} \iff a^x \cdot a^y \leqslant a^{x+y}$$

Thus $a^x \cdot a^y = \sup\{a^{q+p} : p \leqslant y \land q \leqslant x\} \leqslant \sup\{a^z : z \leqslant x + y\} = a^{x+y}$.

We now prove the opposite inequality, that $a^{x+y} \leqslant a^x \cdot a^y$. We choose a fixed $\tilde{q} \leqslant x$ and a fixed $\tilde{p} \leqslant y$, then:

$$a^{x+y} = \sup\{a^z : z \leqslant x + y\}$$

12

For every z, we can pick $\tilde{p}, \tilde{q}$ such that by Lemma 5.3:

$$(\forall z \in x + y : \exists \tilde{p} \leqslant y, \tilde{q} \leqslant x) : z \leqslant \tilde{p} + \tilde{q} \implies a^z \leqslant a^{\tilde{p}+\tilde{q}} = a^{\tilde{p}} \cdot a^{\tilde{q}}$$

Which means that since the choices of $\tilde{p}, \tilde{q}$ were arbitrary:

$$\forall z \leqslant x + y : a^z \leqslant a^{\tilde{p}} \cdot a^{\tilde{q}} \leqslant a^x \cdot a^y$$

By definition of supremum, a^{x+y} is the least upper bound for all a^z and we have shown that $a^x \cdot a^y$ is an upper bound, so:

$$a^z \leqslant a^x \cdot a^y \implies a^{x+y} \leqslant a^x \cdot a^y$$

Thus proving $a^{x+y} = a^x \cdot a^y$. This forces a multiplicative inverse to exist such that $a^x \cdot a^y = a^0 = 1 \implies x + y = 0 \implies y = -x$. Thus $(a^x)^{-1} = a^{-x}$. $\qquad \square$

Now that we have defined exponentiation with respect to a real valued base, along with elementary operations, we aim to define the inverse of exponentiation, which is the logarithm.

Theorem 5.2. *The exponential function: $f : \mathbb{R} \to \mathbb{R}^+ \setminus \{0\}$ is a bijection*

We define $f(x) = b^x$ for any $a \in \mathbb{R}^+ \setminus \{1\}$. The proof is the same if $b > 1$ or $b < 1$. Then:

Proof. f is injective if $\forall x, y \in \mathbb{R}^+ : b^x = b^y \implies x = y$. We can use Lemma 5.4 and multiplicative inverses:

$$b^x = b^y \implies b^x \cdot (b^y)^{-1} = 1 \implies b^x \cdot b^{-y} = b^{x-y} = 1$$

Since $b > 1$, then by the definition of exponentiation, $x - y = 0$, so $x = y$.

If f is surjective, then $\forall y \in \mathbb{R}^+ . \exists x \in \mathbb{R} : b^x = y$. Here x is called the logarithm of y base b.

We fix $a > 1$, then by the results of problem 1:

$$a^n - 1 = (a - 1) \sum_{k=0}^{n-1} a^k \cdot 1^{n-k-1} = (a - 1) \sum_{k=0}^{n-1} a^k \geqslant \sum_{k=0}^{n-1} 1^k = n(a - 1)$$

Now if we let a be the nth root of b such that $a = b^{1/n} > 1$, then:

$$a^n - 1 \geqslant n(a - 1) \implies b - 1 \geqslant n(b^{1/n} - 1)$$

If $t > 1$ and $n > (b-1)/(t-1)$, then $b^{1/n} < t$:

$$n(b^{1/n} - 1) \leqslant b - 1 < n(t - 1) \implies b^{1/n} - 1 < t - 1 \implies b^{1/n} < t$$

If w is such that $b^w < y$, then $b^{w+(1/n)} < y$ for sufficiently large n; to see this, apply part (c) to $t = y \cdot b^{-w}$.

$$b^{1/n} < t \implies b^{1/n} < y \cdot b^{-w} \implies b^{(1/n)+w} < y$$

If $b^w > y$, then $b^{w-(1/n)} > y$ for sufficiently large n. Let $t = y^{-1} \cdot b^w > 1$:

$$b^{1/n} < t \implies b^{1/n} < y^{-1} \cdot b^w \implies y < b^{w-(1/n)}$$

Let A be the set of all w such that $b^w < y$, and show that $x = \sup(A)$ satisfies $b^x = y$.

If $b^x > y$, then $b^{x-(1/n)} > y$, which means that $x - \frac{1}{n}$ is a smaller lower bound, so b^x is not greater than y. If $b^x < y$, then $b^{x+(1/n)} < y$, so $x + \frac{1}{n} \in A$, but $x < x + \frac{1}{n} \notin A$, showing a contradiction.

Thus $b^x = y$ and $x = \sup(A)$, proving the existence of a logarithm of y base b for any $y \in \mathbb{R}^+$, showing surjection. The uniqueness of x can be derived from the injection or also through supremums:

$$b^x = b^{x'} = y \implies x = \sup(A) \wedge x' = \sup(A)$$

Since the supremum is unique, then $x = x'$, showing injectivity, thus f is a bijection and therefore has an inverse which we will call log. $\qquad\square$

Corollary 5.4. $\log_b : \mathbb{R}^+ \setminus \{0\} \to \mathbb{R}$ *such that* $log_b(y) = x$

Then the composition of $\log_b(b^x) = x$ and $b^{\log_b(x)} = x$. Here we see the $\log_b$ acting as a left and right inverse, which is the expected behavior of the inverse of a bijective function. (Remember a surjection only has a right inverse, and an injection only has a left inverse).

6 Countability and Cardinality

The natural numbers or $\mathbb{N}$ are the backbone of the idea of "counting". When we say $0, 1, 2, 3...$ so on and so forth, we are iterating through the naturals. We define the following:

Definition 6.1. *The first n naturals are $[0, n) := \{k \in \mathbb{N} : k < n\}$*

We can now introduce the notion of a finite set.

Definition 6.2. *A finite set A is a set such that $\exists f, \exists n \in \mathbb{N} : f : [0, n) \to A$ and f is a bijection*

An infinite set is one that is not finite, or one for which we cannot find a bijection between an interval of the first n naturals for some $n \in \mathbb{N}$.

We also introduce the notion of Dedekind infinite. A set that is Dedekind infinite is one where we can find an injective but not surjective map between the set and itself. For example:

$$S : \mathbb{N} \to \mathbb{N} : 0 \notin \text{Ran}(S) \implies S(\mathbb{N}) \subseteq \mathbb{N}$$

So the naturals are a Dedekind infinite set under the successor operation. We can also come up with other injections such as $f(n) = n^2$. We say the following:

Lemma 6.1. *Given a set A: A is Dedekind infinite $\implies$ A is infinite*

The converse requires invoking axiom of choice as we need to assume the existence of a choice function that can generate a sequence of values to which we can apply an injective function.

If we can find a bijection f between an interval $[0, n)$ and a finite set A, then the value $n \in \mathbb{N}$ is called the cardinality of A, denoted by $|A|$. We see some inequalities related to cardinality:

Lemma 6.2. *For finite sets A and b:*

1. $B \subseteq A \implies |B| \leqslant |A|$

2. $|A \cup B| \leqslant |A| + |B|$

3. $|A \times B| = |A| \cdot |B|$

We prove Lemma 6.2.1:

Proof. We define $|A|$ as $n \in \mathbb{N}, \exists f : f : [0, n) \to A \wedge \text{Im}(f) = A$. The proof is as follows:

1. $B \subseteq A \iff \forall x \in B : x \in A$ (Subset Definition)

2. $\forall x \in B : (x \in A \implies x \in \text{Im}(f))$ (By definition of f)

3. $\implies \forall x \in B \wedge A, \exists m \in [0, n) : f(m) = x$ (Definition of Preimage)

4. Thus $f^{-1}(B) = \{m \in [0, n) : f(m) = x\}$

5. Since $\forall m \in f^{-1}(B) : m \in [0, n) \implies f^{-1}(B) \subseteq [0, n)$ (Definition of Subset)

6. Therefore either $f^{-1}(B) = [0, n) = \text{Dom}(f)$ or $f^{-1}(B) \subset [0, n)$ so it does not contain all first n naturals

7. Since f is a bijection, then $f|_{f^{-1}(B)} : f^{-1}(B) \to B$ is a bijection, so by line 6: $|B| \leqslant |A|$

$\square$

We prove Lemma 6.2.2:

Proof. The proof is as follows. Since A, B are finite, axiom of choice is not required:

1. $A \cup B \iff \forall x \in A \cup B : x \in A \lor x \in B$ (Definition of Union)

2. $f : [0, n) \to A \land g : [0, m) \to B$ such that f, g are bijections. Then $|A| = n, |B| = m$.

3. We define a function $p : [0, m + n) \to A \cup B$ such that $\text{Ran}(p|_{[0,n)}) = A \land \text{Ran}(p|_{[n,m+n)}) = B$

4. If $A \cap B \neq \varnothing$, then $\exists x : x \in A \land x \in B$

5. $\implies x \in \text{Im}(f) \land \text{Im}(g)$

6. $\implies \exists l \in [0, n) \land \exists k \in [m, m + n) : p|_{[0,n)}(l) = p|_{[n,m+n)} = x$, which means that p is not injective, so $|A \cup B| < m + n = |A| + |B|$

7. If $A \cap B = \varnothing$, then p is a bijection, so $|A \cup B| = m + n = |A| + |B|$

8. Thus $|A \cup B| \leqslant |A| + |B|$

$\square$

We prove Lemma 6.2.3:

Proof. The proof is as follows:

1. We define the cartesian product of A, B as $A \times B := \bigcup \{f : [0, 2) \to A \cup B : (f(0) \in A \land f(1) \in B)\}$

   (a) Remember that a function is fundamentally its graph, so each function is bijective to the set of pairs with the first element being from A and the seocnd being from B.

   (b) Let's restrict each f to a single, unique ordered pair so that all the graphs are disjoint.

2. Thus $|A \times B| = |\bigcup \{f : [0, 2) \to A \cup B : (f(0) \in A \land f(1) \in B)\}| = \#$ of f

3. Let $|A| = m \land |B| = n$

4. Fix an $a \in A : f(0) = a \land f(1) \in B \implies |\{f : [0, 2) \to A \cup B : f(0) = a \land f(1) \in B\}| = n$ as there are n elements in B

5. Since there are m elements in A, we have m sets $\{f : [0, 2) \to A \cup B : f(0) = a \land f(1) \in B\}$.

6. Thus in total we have m sets of n functions, so we have $m \cdot n$ functions in total, thus:

7. $|A \times B| = |A| \cdot |B|$

$\square$

We note the relation of Axiom of Choice and Cartesian Products/Powers. Take for example the cartesian power $\mathbb{R}^{\mathbb{N}}$ which is the set of real valued sequences. The question is the same as the problem proposed in Lemma 6.1. To construct the sequence, we encounter the following collection:

1. $x_0 \in \mathbb{R}$

2. $x_1 \in \mathbb{R} \setminus \{x_0\}$

3. $x_2 \in \mathbb{R} \setminus \{x_0, x_1\}$

4. $x_3 \in \mathbb{R} \setminus \{x_0, x_1, x_2\}$

5. The sequences continues countably infinitely (for as many naturals there are)

Here we see that we are required to select real numbers from an infinite family of infinite sets. We encounter the same problem for an infinite family of finite sets.

We now consider the cartesian power $\mathbb{N}^{\mathbb{N}}$ which is the set of all natural-valued sequences. We see the following family:

1. $x_0 \in \mathbb{N}$

2. $x_1 \in \mathbb{N} \setminus \{x_0\}$

3. The sequence continues for all of $\mathbb{N}$

This case does not require axiom of choice however, due to the well ordering principle. We know that every set of naturals admits a minimum (an infinimum that exists in the set). So we know that there exists one such possible sequence which is the naturals itself. We can also take different collections of subsets of $\mathbb{N}$, combined with the well-ordering, allowing us to define different sequences and their rules.

Question to Think About: What about $\mathbb{Q}^{\mathbb{N}}$, does such construction require Axiom of Choice?

Definition 6.3. *A countable set A is one such that $\exists f$ such that $f : \mathbb{N} \to A$ is a bijection*

We now exmaine countable sets. We will show results such as $\mathbb{Q}$ being countable, or the set of algebraic reals being countable, which on surface seem counterintuitive.

Theorem 6.1. *If A is countable, then $\forall B \subseteq A : B$ infinite $\implies B$ countable*

Proof. We first see that A is infinite as well as it has an infinite subset (How can we prove this?)

Since A is countable, then we can find some bijection $f : \mathbb{N} \to A$. We then define $f^{-1}(B) = \tilde{B} \subseteq \mathbb{N}$. If B is to be countable, then $\tilde{B}$ should also be countable. We prove by induction:

Let B_0 be our base case such that $B_0 = \tilde{B}$. We know that if B is infinite then $\tilde{B}$ must be infinite (why?). We then define the following:
$$B_{k+1} = \begin{cases} \varnothing : & B_k = \varnothing \\ B_k \setminus \{\inf(B_k)\} : & B_k \neq \varnothing \end{cases}$$

Since we are dealing with sets of naturals, we do not require Axiom of Choice for this infinite family. By the well ordering principle, we can pick out the infimum of each set.

If B_{k+1} is infinite, then we see that $B_k = B_{k+1} \cup \{\inf(B_k)\}$, so B_k is also infinite. What we have done is created the ocuntable sequence of sets:
$$\{B_k\}_{k \in \mathbb{N}} := (B_0, B_1, B_2, ..., B_k, B_{k+1}, ...) \text{ such that } B_k = B_{k+1} \cup \inf(B_k)$$

If we manipulate the defining proposition of the sequence, we see that $B_k \setminus B_{k+1} = \inf(B_k)$.

So if we define some function $h : \mathbb{N} \to \{B_k\}$, we get a bijection as there is a countable number of sets. We define the inf function:
$$\inf : \{B_k\} \to \tilde{B}$$

We first see that inf is injective. So $\forall B_x, B_y \in \{B_n\} : \inf(B_x) = \inf(B_y) \implies B_x = B_y$. Imagine $B_x \neq B_y$, then $B_x \subset B_y \vee B_y \subset B_x$. Assume without loss of generality that $B_x \subset B_y$:
$$\exists y \in B_y : y \notin B_x$$

17

Since each B_k is infinite (implying the sets are not bounded above) and $B_x \subset B_y$, then $\forall x \in B_x : y < x$ by the ordering of the naturals. Thus y is a lower bound for B_x but $\inf(B_x) \in B_x$, so $y \neq \inf(B_x)$. Then:

$$y \in B_y \wedge y \neq \inf(B_x) \implies \inf(B_y) \leqslant y < \inf(B_x) \tag{3}$$

The inequality $y < \inf(B_x)$ follows from the fact that $y \neq \inf(B_x)$ and if $y > \inf(B_x)$, then $y \in B_x$ which it is not by definition. Thus $\inf(B_y) < \inf(B_x)$ proving the statement by contrapositive.

We now see that inf is surjective. $\forall b \in \tilde{B}, \exists B_k : \inf(B_k) = b$. This follows from the well ordering property, that every subset of the naturals has an infimum. So if we pick any element of B which is a subset of the naturals, then through the recursive process, we can find a set:

$$\exists n \in \mathbb{N} : B_n := \{x \in \tilde{B} : b \leqslant x\}$$

The infimum of this set is b and B_n is a valid element in $\{B_k\}$ as it is either $\tilde{B} = B_0$ itself, or is a proper subset, which in that case $\exists B_m : B_m \subseteq B_0 \wedge B_n = B_m \setminus \inf(B_m)$.

Now that we know inf is also a bijection, we can construct a bijection $\inf \circ h : \mathbb{N} \to \tilde{B}$ which means that $\tilde{B}$ is countable. Now:

$$\exists g : g : \tilde{B} \to B \text{ such that g is a bijection}$$

Thus $g \circ \inf \circ h : \mathbb{N} \to B$, and the composition of two bijections is a bijection, showing that B is countable. $\qquad \square$

Corollary 6.1. *For any set A: A finite or countable $\iff \exists f : A \to \mathbb{N}$ such that f is an injection*

Proof. We examine the forward direction first. If A is finite, then by definition 6.2 it is bijective to the subset of the first n naturals, thus f is an injection. If it is countable and not finite, then by definition 6.3, f is a bijection so it is an injection.

For the converse, if f an injection, that means $\mathrm{Ran}(f) \subseteq \mathbb{N}$. If $\mathrm{Ran}(f) = \mathbb{N}$, then f is also a surjection, so A is infinite and countable.

If $\mathrm{Ran}(f) \subset \mathbb{N}$, then if the range is infinite by Thm 6.1, it is countable, so there exists a bijection between A and $\mathbb{N}$, so A is countable.

Finally, if $\mathrm{Ran}(f)$ is finite, then there exists a bijection between A and the first n naturals, so A is finite. $\qquad \square$

Lemma 6.3. *If A and B are countable, then $A \times B$ is countable*

Proof. We define bijections $f : \mathbb{N} \to A \wedge g : \mathbb{N} \to B$. We can define a function $h : A \times B \to \mathbb{N} \times \mathbb{N}$ such that $h(a, b) = (f^{-1}(a), g^{-1}(b))$.

We see that h is a surjection: This is supported by the bijections f, g.

$$\forall (m, n) \in \mathbb{N} \times \mathbb{N}, \exists (f(m), g(n)) \in A \times B : h((f(m), g(n))) = (m, n)$$

We now prove that h is an injection: $\forall (a, b), (a', b') \in A \times B : h(a, b) = h(a', b') \implies (a, b) = (a', b')$:

$$h(a, b) = h(a', b') \implies f^{-1}(a) = f^{-1}(a') \wedge g^{-1}(b) = g^{-1}(b')$$

Since f, g are bijections, their inverses are bijections, so they are also injections, thus $a = a' \wedge b = b'$.

We now define a map $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that $\psi(m,n) = \frac{1}{2}(m+n)(m+n+1) + m$. This is called the Cantor Pairing Function. We see that every natural number can be represented by this (why?).

We now prove it is an injection. Given two ordered pairs $(m,n), (m',n')$, if $m < m'$, then $\psi(m,n) < \psi(m',n')$. If $m = m. \wedge n < n'$, then we see again $\psi(m,n) < \psi(m',n')$. So ψ is strictly monotonic, therefore it is an injection as we have a total lexicographic ordering of the ordered pairs.

$\square$

Theorem 6.2. *The set $\mathbb{Q}$ is infinite and countable*

Proof. If $\mathbb{Q}$ were not infinite, then $\exists t \in \mathbb{Q}, \forall r \in \mathbb{Q} : r \leqslant t$. Every rational can be written as $\frac{p}{q} : p, q \in \mathbb{Z} \wedge q \neq 0$. The naturals are infinite, implying the integers are infinite, so:

$$\exists p, q \in \mathbb{Z} : t = \frac{p}{q} \wedge q > 0$$

However, since $\mathbb{Z}$ is infinite, $p + 1 > p$. So we can construct a rational $s \in \mathbb{Q}$, such that:

$$s = \frac{p+1}{q} = \frac{p}{q} + \frac{1}{q} = t + \frac{1}{q}$$

Since $q > 0$, then by field properties $q^{-1} = 1/q > 0$. If $q < 0$, then we simply use $p - 1$. So we can show there exists a larger rational, thus $\mathbb{Q}$ is infinite.

We can say that the set of positive rationals is countable through the Cantor "Snake" argument by arranging them in a table of naturals minus 0 as rows and as columns. A similar thing can be done with negative integers for the negative rationals. We let ψ be the function that takes in a positive or negative rational and outputs its "place".

Each positive and negative rational will appear in the same place by construction (it has the same natural number assigned to it). We formalize it as below:

$$\forall r \in \mathbb{Q}^{+} : \phi(r) = 2\psi(r) = 2n$$

Here n is the place of the positive rational in the Cantor Snake argument. For negative rationals:

$$\forall q \in \mathbb{Q}^{-} : \phi(q) = 2\psi(q) - 1 = 2n - 1$$

Finally $\phi(0) = 0$. So all the negative rationals map to the odd naturals, while the positive rationals map to the even naturals. We see that this is an injection as if two rationals have the same value of ϕ, then they must be both greater/less than 0.

Since the set of positive/negative rationals are bijective to the naturals, then the two rationals must be unique. Thus $\text{Ran}(\phi) \subseteq \mathbb{N}$, and because the rationals are infinite, the range is infinite. Thus by Thm 6.1, the range is countable.

$\square$

Lemma 6.4. *The set of algebraic reals: $\{x \in \mathbb{R} : (\exists n \in \mathbb{N}, \exists a_0, ..., a_n \in \mathbb{Z} : a_n \neq 0 \wedge a_n x^n + ... + a_1 x + a_0 = 0)\}$ is countable.*

We first prove that an n degree polynomial has at most n roots:

Proof. Our base case is when $n = 0$. This polynomial does not have roots trivially.

We assume that a polynomial $p(x)$ that is degree n with at most n roots.

Then if there is a $q(x) = (x - r)p(x)$ such that $\deg(q) = \deg(p) + 1 = n + 1$, if $p(r) \neq 0$, then $\# \text{ roots(q)} = \# \text{ roots(p)} \leqslant \deg(p) < \deg(q)$.

If $p(r) = 0$, then $\# \text{ roots(q)} = \# \text{ roots(p)} + 1 \leqslant \deg(p) + 1 = \deg(q)$. Therefore for any polynomial $\# \text{ roots(q)} \leqslant \deg(q)$

We can define each n degree polynomial by the $n + 1$-tuple of its coefficients. Thus, we can define the set of polynomials degree n:

$$P_n := \mathbb{Z}^n$$

By Thm 6.1, $\mathbb{Z} \subset \mathbb{Q} \implies Z$ countable. Thus the cartesian product $\mathbb{Z}^n$ is also countable as n is finite. (If the iterating set were infinite, then we would face an infinite cartesian product).

Since each polynomial $p \in P$ has at most $\deg(p)$ roots, then the set of roots for each P_n is also countable. If we define the set of roots for P_n to be T_n, then the set of all roots $\bigcup T_n$ is also countable.

Thus $\mathbb{R}_A := \bigcup T_n$, so the set of all algebraic roots is countable. $\qquad \square$

Theorem 6.3. *There exists non-algebraic reals.*

Since $\mathbb{R}_A$ is countable and $\mathbb{R}$ is not countable, then $\mathbb{R}_A \precnsim \mathbb{R}$, so we can define the nonalgebraic reals as:

$$\mathbb{R}_N := \mathbb{R} \setminus \mathbb{R}_A$$

We know this set is nonempty by the strict inequality for numerosity.

7   Uncountable Sets

Theorem 7.1. *The set of sequences $\{0,1\}^{\mathbb{N}}$ is uncountable*

We utilize Cantor's Diagonal argument for this theorem.

Proof. We let $f : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$. We construct a sequence $\{x_n\}_{n \in \mathbb{N}} : x_n = 1 - f(n)_n$.

Since every nth term in this new sequence is different from the nth sequence, the sequence does not belong in $\mathrm{Im}(f)$, so f is not surjective, meaning $\{0,1\}^{\mathbb{N}}$ is uncountable. $\qquad\square$

Lemma 7.1. *The set of functions $\{0,1\}^{\mathbb{R}}$ is uncountable*

We once again utilize Cantor's Diagonal argument for this lemma.

Proof. This time we deal with graphs of functions and not sequences. Each function $f : \mathbb{R} \to \{0,1\}$. We also have $\phi : \mathbb{N} \to \{f\}$.

We define $h : \mathbb{R} \to \{0,1\}$ as the function where the value of n of nth function is $(n, 1 - f_n(n))$.

We see that h does not fit in any of the functions, so the set of such functions are uncountable. $\qquad\square$

Definition 7.1. *Equinumerosity of two sets A, B: $A \simeq B \implies \exists f : (f : A \to B) \wedge f \text{ bijection}$*

The Equinumerosity class of a set $[A]$ is the set of all sets equinumerous to A.

Theorem 7.2. *For all A, $A \notin [\mathcal{P}(A)]$*

Proof. We want to prove that there exists no bijection $f : A \to \mathcal{P}(A)$. Since $\forall x \in A, \{x\} \subset A$, then $\forall x \in A : x \in \mathcal{P}(A)$ by definition of powerset. If two such singletons are equal, then by Axiom of Exensionality, their elements must be equal. Thus f is an injection.

We prove that f is not a surjection. We know that $A \subseteq A$, so $A \in \mathcal{P}(A)$, but $A \notin A$, so f is not a surjection.

If $A \in A$, then we would have a loop of infinite containment, which is not allowed by Axiom of Regularity. $\quad\square$

Definition 7.2. *We define nonequinumerosity: $A \lesssim B \implies \exists f : (f : A \to B) \wedge f \text{ injection}$*

Cardinality is limited to countable or finite sets, so this inequality allows us to compare numerosity or size for uncountable sets as well.

We observe that $A \subseteq B \implies A \lesssim B$. However $A \lesssim B \wedge B \lesssim A$ does not imply $A = B$. For example:

$$\mathbb{Q} \lesssim \mathbb{N} \wedge \mathbb{N} \lesssim \mathbb{Q}$$

Yet, $\mathbb{Q} \neq \mathbb{N}$. We however do formulate the Cantor-Bernstein theorem as follows:

Theorem 7.3. $A \lesssim B \wedge B \lesssim A \implies A \simeq B$

Proof. Proof is left as an excercise to the reader. (I will complete it later, I am the reader) □

However, a powerful result of this theorem is that we only need to show two injections exist between both sets to show they are equinumerous.

> **Definition 7.3.** *The equinumerosity class* $[\mathbb{R}]$ *is called the continuum.*

We can show that $\mathcal{P}(\mathbb{N}) \in [\mathbb{R}]$.

Proof. $\mathcal{P}(\mathbb{N}) \simeq \{0,1\}^{\mathbb{N}}$ via the following bijection:

$$\forall A \in \mathcal{P}(\mathbb{N}) : f_n(A) := \begin{cases} 0 : & n \notin A \\ 1 : & n \in A \end{cases}$$

If two sequences are equal, then we know that their preimages must be the same, so f is injective. We also see f is surjective, since every sequence is the size of $\mathbb{N}$, it obtains an equivalent expression as a subset of $\mathbb{N}$.

Before we derived an injection from $\{0,1\}^{\mathbb{N}} \to [0,1]$, thus we can derive an injection from $\mathcal{P}(\mathbb{N}) \to [0,1]$. Since $[0,1] \subset \mathbb{R}$, the injection follows, so $\mathcal{P}(\mathbb{N}) \lesssim \mathbb{R}$.

Now $\mathbb{R}$ is the set of Dedekind cuts to $\mathbb{R} \lesssim \mathcal{P}(\mathbb{Q})$.

Since $\mathbb{Q} \simeq \mathbb{N}$ then $\mathcal{P}(\mathbb{Q}) = \mathcal{P}(\mathbb{N})$ (why? Technically by 2^n rule but not proven). Thus $\mathbb{R} \lesssim \mathcal{P}(\mathbb{N})$, so:

$$\mathcal{P}(\mathbb{N}) \simeq \{0,1\}^{\mathbb{N}} \simeq [0,1] \simeq \mathbb{R}$$

□

> **Lemma 7.2.** $\mathbb{R} \times \mathbb{R} \simeq \mathbb{R}$ *and* $\forall n \in \mathbb{N} : \mathbb{R}^n \simeq \mathbb{R}$

We first prove that $\mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{N})$. We see trivially that $\mathcal{P}(\mathbb{N}) \lesssim \mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N})$.

$$\mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$$

We can construct an injection $f : \{0,1\}^{\mathbb{N}} \times \{0.1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$. If we take a pair of sequences, simply concatenating them provides a new, unique sequence. Therefore:

$$\mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \lesssim \{0,1\}^{\mathbb{N}} \simeq \mathcal{P}(\mathbb{N})$$

Thus $\mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{N})$.

Proof. We note that $\mathbb{R} \lesssim \mathbb{R} \times \mathbb{R}$ trivially, as the real numbers are embedded in a 2D Euclidean space.

$$\mathbb{R} \times \mathbb{R} \lesssim \mathcal{P}(\mathbb{Q}) \times \mathcal{P}(\mathbb{Q}) \simeq \mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{N}) \simeq \mathbb{R}$$

Thus $\mathbb{R} \times \mathbb{R} \lesssim \mathbb{R}$, so $\mathbb{R} \times \mathbb{R} \simeq \mathbb{R}$. □

Proof. We now prove that $\forall n \geqslant 1 : \mathbb{R}^n \simeq \mathbb{R}$. The inequality $\mathbb{R} \lesssim \mathbb{R}^n$ is trivial. We prove the other direction:

$$\mathbb{R}^n \lesssim \mathcal{P}(\mathbb{Q})^n \simeq \mathcal{P}(\mathbb{N})^n$$

Since $n \in \mathbb{N}$, if n is even, then we create $n/2$ number of pairs of $\mathcal{P}(\mathbb{N})$. Using the property initially proved, we get:

$$\mathcal{P}(\mathbb{N})^n \simeq \mathcal{P}(\mathbb{N})^{n/2} \simeq \mathcal{P}(\mathbb{N})^{n/4} \simeq ...$$

22

We repeat this process till we get $\mathcal{P}(\mathbb{N})$. If n is odd, then we apply this process to $n-1$. We then get:

$$\mathcal{P}(\mathbb{N})^n \simeq \mathcal{P}(\mathbb{N})^{n-1} \times \mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{N})$$

Therefore we complete our previouos inequality:

$$\mathbb{R}^n \lesssim \mathcal{P}(\mathbb{Q})^n \simeq \mathcal{P}(\mathbb{N})^n \simeq \mathcal{P}(\mathbb{N}) \simeq \mathbb{R}$$

Now tha we have both inequalities, we have proven that $\mathbb{R}^n \simeq \mathbb{R}$. $\qquad\square$

Lemma 7.3. $\mathbb{R}^{\mathbb{N}} \simeq \mathbb{R}$

We know that $\mathbb{R} \simeq \{0,1\}^{\mathbb{N}}$, so by Lemma 6.3 and 7.2:

$$\mathbb{R}^{\mathbb{N}} \simeq \{0,1\}^{\mathbb{N} \times \mathbb{N}} \simeq \{0,1\}^{\mathbb{N}} \simeq \mathcal{P}(\mathbb{N}) \simeq \mathbb{R}$$

Lemma 7.4. $\mathbb{R} \lesssim \mathbb{R}^{\mathbb{R}}$

We first prove that $\mathbb{N} \times \mathbb{R} \simeq \mathbb{R}$. The inequality $\mathbb{R} \lesssim \mathbb{N} \times \mathbb{R}$ is trivial. We prove the other direction by Lemma 7.2:

$$\mathbb{N} \times \mathbb{R} \simeq \mathbb{N} \times \mathcal{P}(\mathbb{N}) \lesssim \mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{N}) \simeq \mathbb{R}$$

Proof. We can now track the following inequality:

$$\mathbb{R} \lesssim \mathcal{P}(\mathbb{R}) \simeq \{0,1\}^{\mathbb{R}} \simeq \{0,1\}^{\mathbb{N} \times \mathbb{R}} \simeq \mathcal{P}(\mathbb{N})^{\mathbb{R}} \simeq \mathbb{R}^{\mathbb{R}}$$

$\qquad\square$

8 Metric Spaces

We earlier defined a sequence as the graph of a function $f : \mathbb{N} \to A$. While this is the underlying machinery, we consider a sequence as a set of elements indexed by the naturals, denoted as $\{x_n\}_{n \in \mathbb{N}}$.

First we deal with "real-valued sequences", which means that all $x_n \in \mathbb{R}$. We define two important notions, in the context of the reals:

Definition 8.1. *A Cauchy sequence:* $\forall k \in \mathbb{N}, \exists n_0 \in \mathbb{N}, \forall m, n \geqslant n_0 : |x_m - x_n| < \frac{1}{k+1}$

The reason we set the bound as $\frac{1}{k+1}$ is that the bound will always exist in the reals and as we incresae k, the bound will get smaller, which is the behavior we want to model. We could set the bound to $\frac{1}{r+1}$, where r is a real, but the natural numbers allos us to create a sequence of bounds unlike the reals which are uncountable.

Definition 8.2. *A convergent sequence:* $\exists L \in \mathbb{R}, \forall k \in \mathbb{N}, \exists n_0 \in \mathbb{N}, \forall n \geqslant n_0 : |x_n - L| < \frac{1}{k+1}$

This is different from a Cauchy sequence as we are now fixing a set "limit point". As we decrease the bound, we can always find some element in the sequence arbitrarily close. So for a very large k we see that x_n tends to L.

Example: We show the following is Cauchy and convergent:

$$a_0 = 1 \wedge a_{n+1} = 3 - \frac{1}{a_n}$$

We show that $\forall n \in \mathbb{N} : 1 \leqslant a_{n+1} \leqslant 3$ and $a_n < a_{n+1}$:

$$1 \leqslant a_n \leqslant 3 \implies -1 \leqslant -\frac{1}{a_n} \leqslant -\frac{1}{3}$$

$$a_{n+1} \leqslant 3 - \frac{1}{a_n} \leqslant 3 - \frac{1}{3} < 3$$

Also $a_n \geqslant 1 \implies a_{n+1} > a_n \geqslant 1$, so $1 \leqslant a_{n+1} \leqslant 3 \wedge a_n < a_{n+1}$. Since each consecutive term is greater, we call this monotone (monotonically increasing).

We now prove the sequence is Cauchy:

Proof. We need to show we can find some n_0 such that $\forall k \in \mathbb{N}$ the proposition holds true. We apply the triangle inequality:

$$|a_{n+1} - a_n| \leqslant |a_{n+1}| + |a_n|$$

$\square$

8.1 Metrics

We now generalize the notion of distance to the concept of a "metric":

Definition 8.3. *A metric space is a set X equipped with a metric ϱ such that for (X, ϱ):*

1. $\forall x, y \in X : \varrho(x, y) \geqslant 0 \wedge \varrho(x, x) = 0$ (Positivity)

2. $\forall x, y \in X : \varrho(x, y) = \varrho(y, x)$ (Symmetry)

3. $\forall x, y, z \in X : \varrho(x, z) \leqslant \varrho(x, y) + \varrho(y, z)$ (Triangle Inequality)

We will now look at various metrics. The typical Euclidean metric (distance between two points in $\mathbb{R}^n$) is defined as:

$$\left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{\frac{1}{2}}$$

We can in fact generalize this notion to create p-metrics for $\updownarrow^p$ spaces (which are spaces of X valued sequences with a p-norm. Think, vector spaces with a fixed basis such that the elements can be represented as finite sequences, or infinite with Axiom of Choice):

$$\text{p-metric} := \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{\frac{1}{p}}$$

The ∞-metric is then defined as the following and is referred to as the "box metric":

$$\max_{i=1,2,\ldots,n} |x_i - y_i|$$

We address the "shapes" of these metrics. This construct stems from the idea of a norm, attached to a vector space (which creates a normed vector space):

Definition 8.4. *A norm $||\cdot||$ is a metric that gives the magnitude of a vector from the "origin" or 0 element. Let V be a vector space:*

1. $\forall v \in V : ||v|| \geqslant 0$ (Positivity)

2. $\forall v \in V \wedge c \in F : ||cv|| = |c|||v||$ (Scalar Multiplication)

3. $\forall v, w \in V : ||v + w|| \leqslant ||v|| + ||w||$ (Triangle Inequality)

The norm pretty closely mirrors the properties of a metric space (that is because it is, or it induces one). We can derive the remaining properties of the metric space (which is symmetry) by the following:

$$\forall v, w \in V : ||v - w|| = || - (w - v)|| = | - 1|||w - v|| = ||w - v||$$

This completes all the properties needed for the vector space to be a metric space. We can now draw out the p-norm unit circles (which are the set of points/vectors a distance 1 away from the origin): We will return
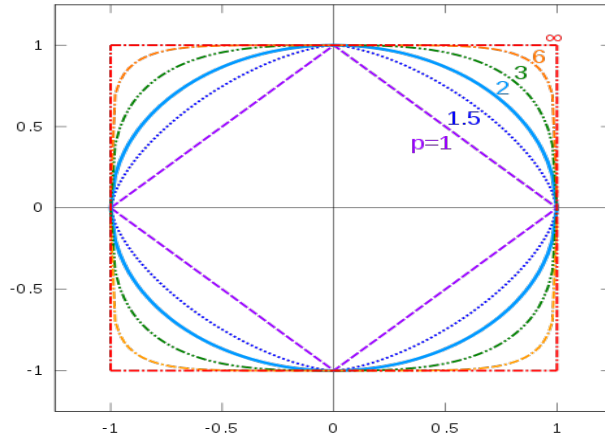

Figure 1: Unit circles of p-norms

to verifying the p-metrics to be valid metrics via the norms. The p-norm now becomes the distance between the origin and the vector:

$$\forall v \in V : ||v||_p = \left(\sum_{k=1}^{n} (|v_k|^p)\right)^{\frac{1}{p}}$$

With the introduction of metrics, we can now generalize the notion of Cauchy and convergent sequences for an X valued sequence $\{x_n\}_{n\in\mathbb{N}}$ with a metric (X,ϱ):

Definition 8.5. *A Cauchy sequence:* $\forall \epsilon > 0, \exists n_0 \geqslant 0, \forall m, n \geqslant n_0 : \varrho(x_m, x_n) < \epsilon$
A convergent sequence: $\exists L \in X, \forall \epsilon > 0, \exists n_0 \geqslant 0, \forall n \geqslant n_0 : \varrho(x_n, L) < \epsilon$

The definition is largely similar, with the difference being an arbitrary upper bound ϵ as it is not specified what the set X is and what the metric ϱ is.

Lemma 8.1. *A convergent sequence is Cauchy.*

Proof. Given (X, ϱ) and a sequence $\{x_n\}_{n\in\mathbb{N}}$, we let the limit point be L. Then:

$$\forall \epsilon > 0, \exists n_0 \geqslant 0, \forall n \geqslant n_0 : \varrho(x_n, L) < \epsilon/2$$

We choose an $m \geqslant n_0$ such that $\varrho(x_m, L) < \epsilon/2$. Then by the triangle inequality:

$$\varrho(x_n, x_m) \leqslant \varrho(x_n, L) + \varrho(L, x_m) = \varrho(x_n, L) + \varrho(x_m, L) = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus $\forall \epsilon > 0$, we have shown there exists an n_0, which is the same as our initial one in the limit defintion, such that $\forall m, n \geqslant n_0 : \varrho(x_m, x_n) < \epsilon$. $\qquad\square$

Lemma 8.2. *For (X, ϱ), an X valued, convergent sequence $\{x_n\}_{n_k \in \mathbb{N}}$ has at most one limit L.*

Proof. We assume there to be two limits L and L' such that $x_n \to L \wedge x_n \to L'$. Since $L \neq L'$, then we can find an ϵ such that $\varrho(L, L') > \epsilon$. Then we define an n_0, such that:

$$\forall n \geqslant n_0 : \varrho(x_n, L) < \frac{\epsilon}{2} \wedge \varrho(x_n, L') < \frac{\epsilon}{2} \tag{4}$$

By positivity we know that $\varrho(x, x) = 0$, so by the triangle inequality:

$$\varrho(L, L') \leqslant \varrho(x_n, L) + \varrho(x_n, L') = \epsilon < \varrho(L, L')$$

We get that $\varrho(L, L') < \varrho(L, L')$ which is a contradiction by metric properties, so $L = L'$. $\qquad\square$

Definition 8.6. *A subsequence is defined by $\{x_{n_k}\}_{n_k \in \mathbb{N}}$ where $\{n_k\}_{k \in \mathbb{N}}$ is a strictly increasing/monotonically increasing sequence*

Typically we can define the n_k sequence by some rule, such as even naturals or odd naturals.

Lemma 8.3. *For a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$, if there is a convergent subsequence such that $x_{n_k} \to z \implies x_n \to z$ where z is a limit point*

Proof. The sequence being Cauchy means $\forall \epsilon > 0, \exists n_0 \geqslant 0, \forall m, n \geqslant n_0 : \varrho(x_m, x_n) < \epsilon/2$.

For the convergent subsequence, we get that $\exists z \in X, \forall \epsilon > 0, \exists n_1 \geqslant 0, \forall n \geqslant n_1 : \varrho(x_n, z) < \epsilon/2$.

Then we choose an $m \geqslant n_1 \wedge m \geqslant n_0 \wedge n \geqslant n_0 \wedge n \geqslant n_1$. By triangle inequality:

$$\varrho(x_m, z) \leqslant \varrho(x_m, x_n) + \varrho(x_n, z) = \epsilon$$

Thus the sequence for any $m \geqslant n_1 \wedge m \geqslant n_0$ converges to z. We can simply set the $n_2 = \max\{n_0, n_1\}$ such that:

$$\exists z \in X, \forall \epsilon > 0, \exists n_2 \geqslant 0, \forall n \geqslant n_2 : \varrho(x_n, z) < \epsilon$$

<div align="right">□</div>

Intuition Note: We use $\epsilon > 0$, as if $\epsilon = 0$, then that would mean the two values being compared are equal. However if you take the *nth* term for a very vary large n in a sequence, not necessarily you will reach the limit, but you may be ever so close, which is why there will be some small difference ϵ. However, if we say it equals ϵ and then choose the next term of the sequence, then the difference will not necessarily be equal to. However, if it is converging, then it will likely be less, so we say $< \epsilon$. If we choose the ϵ then we can guarantee a closer term to the limit always.

What if the sequence does actually reach the limit? We view the discrete metric for a point y:

$$\varrho(x, y) := \begin{cases} 0 : & x = y \\ 1 : & x \neq y \end{cases}$$

> **Corollary 8.1.** *If a sequence in (X, ϱ) is Cauchy, then it is convergent and eventually constant.*

Proof. If it is Cauchy, then $\forall \epsilon > 0, \exists n_0 \geqslant 0, \forall m, n \geqslant n_0 : \varrho(x_m, x_n) < \epsilon$. For any epsilon (eg $\epsilon = 1$), the metric being less than epsilon implies that $x_m = x_n$. So we can define a limit L such that $x_m = x_n = L$.

We can see that the sequence does meet and stagnates at the limit under this metric. □

> **Lemma 8.4.** *Urysohn's Subsequence Principle says the following are equivalent for a sequence $\{x_n\}_{n \in \mathbb{N}}$ and a limit point z:*
>
> 1. *$x_n \to z$*
>
> 2. *Every subsequence $\{x_{n_k}\}_{n_k \in \mathbb{N}}$ contains a convergent subsubsequence such that $x_{n_{k_j}} \to z$*

Proof. We prove equivalence of the statements by showing an if and only if relation:

$\implies$: By lemma 8.1, we get that the sequence is Cauchy. We apply Lemma 8.3 twice to see that there exists such a subsubsequence converging to z.

$\impliedby$: We proceed by contrapositive. Assume $\{x_n\}_{n \in \mathbb{N}}$ does not converge at all, then $\forall \epsilon > 0, \exists n_0 \geqslant 0, \forall n \geqslant n_0 : \varrho(x_n, z) \geqslant \epsilon$. If we then set:

$$n_0 = 1 \wedge n_k := \inf\{m > n_k \wedge \varrho(x_m, z) > \epsilon\}$$

Essentially each n_k of the subsequence is the smallest value of m so that the sequence does not converge, then the subsequence does not converge. So by the contrapositive of lemma 8.3, the subsubsequence does not converge. □