Introduction to Analysis for Computer Scientists

Lecture Notes

March 21, 2002

Contents

Chapter 1

Introduction

Classical mathematics can be coarsely divided into two separate sub–disciplines:

Algebra that deals with operators acting on objects.

Analysis that deals with objects and their relations to each other.

Modern mathematics can be viewed as dividing into these same two sub–disciplines, but it tries to provide further abstractions to the same objects and relations that the classical mathematics describe. This course is an introduction to some of the fields of modern mathematical analysis.

The most basic field of modern mathematical analysis is *set theory*. Sets provide only one type of relation between objects: $\in$, or the test whether one object is a member of another. The second field that will be discussed here is *topology*. Topology adds to the basic notions of set theory the notion of neighborhoods or proximity, allowing us to generalize classical concepts such as continuity of functions. The last part of this course introduces *measure theory* which is an attempt to generalize the classical notions of probability, area, volume, integrals, etc.

Chapter 2

Set Theory

Set theory is divided into two parts: Naive Set Theory and Axiomatic Set Theory. The more interesting part of Naive Set Theory is infinite combinatorics while the most significant results of Axiomatic Set Theory are those of independence. In this course we only deal with Naive Set Theory.

2.1 Basic Definitions

We begin with some basic definitions of sizes of sets:

Definition 2.1: *Two sets A, B are* equipotent *(written as $A \sim B$) if there exist a function $f : A \mapsto B$ that is both one to one and onto B. Note that this defined an equivalence relation.*

Definition 2.2: *A set is called* countable *if and only if it is equipotent to $I\!N$.*

Definition 2.3: *We define the* cardinality *of a set as the equivalence class under the equivalence relation $\sim$ to which the set belongs. We denote the cardinality of a set A as $|A|$.*

We now move on to define *infinite* sets. We give three alternate definitions that are equivalent.

Definition 2.4: *A set A is infinite if and only if the following holds (all three are equivalent):*

1. *There exists a function $f : A \overset{1-1}{\longmapsto} A$ that is* not *onto A.*
2. *There exists a function $f : I\!N \overset{1-1}{\longmapsto} A$ (or in other words, the naturals are equipotent to a subset of A).*
3. *A is not equipotent to any initial segment of $I\!N$.*

Note that definitions 2 and 3 of infinity are based on the definition of the set of all natural numbers $I\!N$.

Theorem 2.1 *(Kantor's theorem):* *There exists infinite sets that are* not *equipotent. Specifically, for all sets A, it holds that $A \not\sim \{0, 1\}^A$.*

Corollary 2.2: *We can define an ordering of the cardinality of infinite sets.*

We write $A \preceq B$ if there is a function $f : A \overset{1-1}{\longmapsto} B$, or alternatively, if there is $C \subseteq B$ s.t. $A \sim C$. Note that we still have to prove that the relation $\preceq$ is indeed an ordering. Kantor's theorem shows that $A \prec \{0,1\}^A$. Note that it is trivial that $A \preceq \{0,1\}^A$.

We denote by $\aleph_0$ the cardinality of the naturals and by $\aleph_1$ the cardinality of the first set that has cardinality greater than that of the set of all naturals (we will later prove that such a set indeed exists).

While we do not repeat the full proof of Kantor's theorem, we demonstrate the use of the technique it introduced. This technique is called *diagonalization*. For this demonstration we consider functions from $I\!N$ to $I\!N$. For any two functions $f, g \in I\!N^{I\!N}$ we say that f dominates g if there exists n_0 s.t. for all $k > n_0$ it holds that $g(k) < f(k)$.

Theorem 2.3 ***(Demonstration of diagonalization):*** *For any countable series of functions from $I\!N$ to $I\!N$, $f_1, f_2, f_3, \ldots, f_n, \ldots$ there exists $g \in I\!N^{I\!N}$ that dominates all the f_is.*

Proof: For every $k \in I\!N$ define $g(k) = \max\{f_1(k), f_2(k), \ldots, f_k(k)\} + 3$. It is easy to see that this g dominates all the f_is, as for any $i \in I\!N$, $g(k) > f_i(k)$ for any $k > i$. □

We still have not demonstrated that the relation $\preceq$ is indeed an ordering. To do this we need to show three properties of the relation $\preceq$:

Transitivity Easy to prove, by simply taking the composition of the two functions that are used in the two instances of the relation $\preceq$.

Antisymmetry This part is the Kantor–Bernstein theorem.

Well defined We have to show that the relation is independent of the choice of sets from the equivalence classes by which we define cardinality. This immediately follows from the definition of the equivalence relation $\sim$.

Theorem 2.4 ***(Kantor–Bernstein):*** *If there are two functions $f : A \overset{1-1}{\longmapsto} B$ and $g : B \overset{1-1}{\longmapsto} A$ then there is a function $h : A \overset{1-1}{\longmapsto} B$ that is onto B.*

We will now demonstrate the use of the Kantor–Bernstein theorem to prove the following claim:

Claim 2.1*:* *The reals have the same cardinality as the power set of the naturals, or in other words:*

$$\Re \sim P(I\!N)$$

Proof: It is easy to see that $\Re \sim (0,1)$, so we will only prove that $(0,1) \sim \{0,1\}^{I\!N}$. We do that by showing a function $f : (0,1) \overset{1-1}{\longmapsto} \{0,1\}^{I\!N}$. We define $f(r)$ to be the series that represents r as a (possibly infinite) binary fraction. Note that numbers that are finite binary fractions can be written in two ways (either with a trailing series of zeros or a trailing series of ones). We choose one of these representations arbitrarily. For example we may define $f(\frac{1}{2}) = 1000\ldots$. Note that while f is indeed one on one, it is not onto $\{0,1\}^{I\!N}$. To complete the proof we will show a function $g : \{0,1\}^{I\!N} \overset{1-1}{\longmapsto} (0,1)$. We will use a function g that transforms into decimal notation the series of binary digits provided on its input. For example, the result of $g(f(\frac{1}{2}))$ we saw before would be 0.1 (i.e. $\frac{1}{10}$). This will make sure that inputs which end with an infinite series of zeros or ones will be mapped to unique real numbers. By the Kantor–Bernstein, the existence of these two functions implies that $(0,1) \sim \{0,1\}^{I\!N}$. □

We show yet another use of the Kantor–Bernstein theorem. Now we will use it to demonstrate that $\Re \sim \Re^2$ (which is also true for any other finite dimension, i.e. for any natural n $\Re \sim \Re^n$). To do that we will, again, show two 1–1 functions. By transitivity it is enough to show such functions between $(0,1)$ and $(0,1)^2$.

Claim 2.2: $(0,1) \sim (0,1) \times (0,1)$

Proof: First, we construct $f : (0,1) \stackrel{1-1}{\longmapsto} (0,1) \times (0,1)$. We may define $f(x) \triangleq (x, \frac{1}{2})$. Now we have to construct another function $g : (0,1) \times (0,1) \stackrel{1-1}{\longmapsto} (0,1)$. Given a point $(r,s) \in (0,1) \times (0,1)$ we will look at the representation of r and s as binary fractions without an infinite suffix that is all ones. Denote this representation as $r_1 r_2 r_3 \ldots$ and $s_1 s_2 s_3 \ldots$. Now define $g(r,s) \triangleq 0.r_1 s_1 r_2 s_2 r_3 s_3 \ldots$. It is easy to see that g is indeed one to one. By the Kantor–Bernstein theorem it follows that $(0,1) \sim (0,1) \times (0,1)$. $\square$

The following claim demonstrates that cardinalities do not grasp the full meaning of what we intuitvly see as sizes of sets (at least in some contexts). We know that rationals are dense within the reals (i.e., any real can be approximated arbitrarily well by a rational), and the rationals are countable while the reals are not, yet the following holds:

Claim 2.3: *For any countable $A \subseteq [0,1]$ there exists a countable set of segments $\{I_i : i \in I\!N\}$ such that:*

- *For all i, I_i is a segment of length > 0.*
- $\sum_i |I_i| < \frac{1}{16}$.
- $A \subseteq \bigcup_{i \in I\!N} I_i$.

Proof: Let $f : I\!N \stackrel{1-1}{\longmapsto} A$ be onto (there exists such f as A is countable). Define the segments to be:

$$I_i \triangleq (f(i) - \frac{1}{2^{i+6}}, f(i) + \frac{1}{2^{i+6}})$$

Note that the length of each I_i is $\frac{1}{32 \cdot 2^i}$ and therefor $\sum_{i \in I\!N} |I_i| = \frac{1}{16}$ but they still cover A (as every point $a \in A$ is the center of a segment $I_{f(a)}$). $\square$

2.2 Cardinality Arithmetics

We present some extensions of arithmetics on finite numbers to infinite cardinalities of sets.

Definition 2.5:

- *If $A \cap B = \emptyset$ then $|A| + |B| \triangleq |A \cup B|$.*
- *For any A, B, $|A| \cdot |B| \triangleq |A \times B|$.*
- *For any A, B, $|A|^{|B|} \triangleq |A^B|$.*

Some basic rules that are extensions of the standard rules we know for finite numbers follow immediately from these definitions:

- $2^\lambda > \lambda$.

- $\lambda^\mu \cdot \lambda'^\mu = (\lambda \cdot \lambda')^\mu$.
- $(\lambda^\mu)^\tau = \lambda^{\mu \cdot \tau}$.

We now move further to investigate the properties in which the arithmetics of infinite cardinalities differ from arithmetic of finite numbers.

Claim 2.4: *For any $\lambda \geq \aleph_0$, it holds that $\lambda + \aleph_0 = \lambda$.*

Proof: Given two sets A, B such that $A \cap B = \emptyset$ and $|A| = \lambda$ and $|B| = \aleph_0$ we need to calculate the cardinality of the union, $|A \cup B|$, and demonstrate that $A \cup B \sim A$. By definition 2.4, A has a subset of cardinality $\aleph_0$, and by the definition of cardinality it follows that there exists $f : I\!N \overset{1-1}{\longmapsto} A$. We now define a new function $h : A \cup B \overset{1-1}{\longmapsto} A$ as follows:

$$h(x) \triangleq \begin{cases} x & \text{if } x \notin \text{Range}(f),\ x \notin B \\ f(2f^{-1}(x)) & x \in \text{Range}(f) \\ f(g(x)) & x \in B \end{cases}$$

Where g is any 1–1 function from B to the set of odd naturals. $\square$

Corollary 2.5: *Denote by Ir the set of irrational reals, then $|Ir| = 2^{\aleph_0}$, as $|\Re| = 2^{\aleph_0}$ while $|\mathcal{Q}| = \aleph_0$ and $\Re = \mathcal{Q} \cup Ir$.*

To summarize, for addition and multiplication of infinite cardinalities, the result is only effected by the greater of the two operands. The power operation, on the other hand, has similar behavior to that it exhibits on finite numbers.

2.3 Orderings and Zorn's Lemma

While most of our discussion is of naive set theory, as opposed to axiomatic set theory, and as such is not based on a formal set of axioms, we do show one axiom of set theory explicitly, as it will prove to be a valuable tool.

Definition 2.6:

Ordering of a set *An ordering $R \subseteq X \times X$ of a set X is a relation which is transitive and antisymmetric.*

Chain of an ordering *A set $A \subseteq X$ is a chain for the ordering $<$ if for every $x \neq y \in A$ either $x < y$ or $y < x$.*

Anti-chain of an ordering *A set $A \subseteq X$ is an anti—chain for the ordering $<$ if for every $x \neq y \in A$ both $x \not< y$ and $y \not< x$.*

For example, on the set of all finite binary series $X \triangleq \{0,1\}^*$ we may define the order $\sigma < \tau \iff \sigma$ is a prefix of τ. Under this order, the set 0^* will be an infinite chain, while the set 1^*0 will be an infinite anti-chain.

Exercise 2.1: *Show that under the ordering $\subseteq$ over $P(I\!N)$ there are both a chain and an anti-chain of cardinality $2^{\aleph_0}$.*

Before stating Zorn's Lemma, we need one more definition:

Definition 2.7: *Given a subset $A \subseteq X$, a member $y \in X$ is called an* upper bound *for A if for every $a \in A$ either $y = a$ or $a < y$. An element $y \in X$ is called* maximal *if there is no $a \in X$ s.t. $y < a$.*

We are now ready to state Zorn's Lemma.

Lemma 2.1: *(Zorn) For every partial order $(X, <)$, if $X \neq \emptyset$ and every chain $A \subseteq X$ has an upper bound in X, then there is a maximal element in $(X, <)$.*

To demonstrate the power of Zorn's Lemma we prove the following:

Claim 2.5: *For every two sets A, B either there exists $f : A \overset{1-1}{\longmapsto} B$ or there exists $g : B \overset{1-1}{\longmapsto} A$. In other words, either $|A| \leq |B|$ or $|B| \leq |A|$.*

Proof: Examine all one to one mappings from any subset of A to any subset of B:

$$X = \{f : f \subseteq A \times B \text{ and } f \text{ is 1–1}\}$$

Define the following ordering on X: We say that $f \prec g$ if $f \subset g$ (as sets of pairs from $A \times B$). We need to check that the defined $(X, \prec)$ satisfies the conditions of the lemma. There are two conditions to check:

1. $X \neq \emptyset$ because $\emptyset \in X$. This means that the first condition required by the lemma holds.
2. Let $A \subseteq X$ be a chain for the ordering $\prec$. We claim that $\bigcup A$ is an upper bound for A. It is obvious that $\bigcup A \in X$ as the elements of A are functions extending each other, so their union is still a function. It is also clear that $\bigcup A$ is an upper bound as for any $f \in A$ it holds that $f \subseteq \bigcup A$.

Now the lemma guarantees us a maximal element in the ordering $\prec$ on X. We still have to show that for this maximal $f \subseteq A \times B$ it holds that either $\text{Range}(f) = B$ or that $\text{Dom}(f) = A$. Assume that neither is true, then there is a pair $(a, b) \in (A \times B) \setminus f$ s.t. $a \notin \text{Dom}(f)$ and $b \notin \text{Range}(f)$. But then, $g \triangleq f \cup \{(a, b)\}$ is also a function in X, and $f \prec g$, in contradiction with the maximality of f in $(X, \prec)$. □

Zorn's Lemma also applies in other "classical" areas of mathematics such as linear algebra. For example:

Theorem 2.6: *For every vector space F^V and for every set of linearly independent elements $A \subseteq V$ there is an extension $B \supseteq A$ of A such that B is a base for F^V.*

Proof: Define the set

$$X \triangleq \{S \mid A \subseteq S \subseteq V \text{ and } S \text{ is linearly independent}\}$$

We use the ordering $\subset$ for X. Again, we check the conditions for Zorn's Lemma:

1. $X \neq \emptyset$ as $A \in X$.

2. Given a chain C for $(X, \subset)$, we again examine the union $\bigcup C$. It is obvious that for any $B \in C$, $B \subseteq \bigcup C$, and so $B \preceq \bigcup C$. All that is left is to verify that $\bigcup C$ is indeed a member of X. Note that it is clear that $A \subseteq \bigcup C$ because the elements of C are all in X. So, if we could show that $\bigcup C$ is linearly independent we would get $\bigcup C \in X$. Assume that $\bigcup C$ is dependent. This means that there are $\beta_1, \ldots, \beta_k \in \bigcup C$ such that $\sum_{i=1}^{k} x_i \beta_i = 0$ and not all the x_is are zero. But for every i there must be $B_i \in C$ s.t. $\beta_i \in B_i$ and because C is a chain, *all* these β_is are in the maximal such B_i. But this contradicts the assumption that this B_i is itself l.i.

By the Zorn Lemma, $(X, \prec)$ has a maximal element. Denote that element S^*. This S^* is a set of linearly independent vectors, $A \subseteq S^*$ and S^* is maximal under the ordering of containment. This implies that S^* is a base. $\square$

2.4 Linear Orderings

The concepts we have so far described such as chains, anti-chains and maximal elements mostly apply to *partial orderings*. We now focus on *linear* or *complete* orderings.

Definition 2.8: *An ordering is called* linear *or* complete *if for every* $(x, y) \in X \times Y$ *either* $x \leq y$ *or* $y \leq x$.

Definition 2.9: *An* initial segment *of an ordering* $(X, \leq)$ *is a set* $A \subseteq X$ *for which it holds that* $\forall x \in A, y \leq x \rightarrow y \in A$. *$A$ is a* proper initial segment *if A is an initial segment and* $A \neq X$. *We may also define the converse, i.e. a final segment of the ordering.*
An interval *of an ordering* $(X, \leq)$ *is a set* $A \subseteq X$ *for which it holds that* $\forall x, y \in A, x \leq z \leq y \rightarrow z \in A$.

When we discussed the cardinality of sets we defined an equivalence relation between sets. The following are analogous definitions for orderings on sets:

Definition 2.10: *Given a pair of orderings* $(X, <_X)$ *and* $(Y, <_Y)$, *a function* $f : X \mapsto Y$ *is called an embedding if:*

1. *f is one to one.*
2. *f preserves the ordering, that is* $\forall a, b \in X, a <_X b \rightarrow f(a) <_Y f(b)$.

If f is onto it will be called an isomorphism.

For example, the function $f(n) = 2n$ is an isomorphism between $I\!N$ and $\{2n : n \in I\!N\}$ with the standard ordering of the natural numbers.

On the other hand, while it's easy to find f that is an embedding of $I\!N$ in Q with the standard ordering, no such function is an isomorphism. This follows from the fact that the rationals are dense, that is for any $x, y \in Q$ there is $x < z < y, z \in Q$. If f was an isomorphism then a number z for which it holds that $f(1) < z < f(2)$ would not be in Range(f), or otherwise f would not have preserved the ordering.

Claim 2.6: *Isomorphism on orderings is an equivalence relation.*

Proof:

1. The identity embedding is an isomorphism from any ordering onto itself.

2. Isomorphism is symmetric. If $f : X \overset{1-1}{\longmapsto} Y$ is onto and preserves ordering, then $g \triangleq f^{-1}$ is $g : Y \overset{1-1}{\longmapsto} X$, onto and preserves the ordering.

3. Isomorphism is transitive because the composition of two isomorphisms is itself an isomorphism.

□

Recall that for cardinalities of sets we had not only an equivalence relation but also an ordering of the cardinalities. Such a definition on orderings of sets gives a relation that is not necessarily antisymmetric. We denote $A \hookrightarrow B$ the fact that there is an embedding of A in B, and $A \simeq B$ the fact that A and B are isomorphic. The next example shows that it is possible that $A \hookrightarrow B$ and $B \hookrightarrow A$ and still $A \not\simeq B$:

Example 2.1: *Define $X = (0,1)$ and $Y = [0,1]$. The identity function is an embedding of X in Y, while the function $g(x) = \frac{1}{2}x + \frac{1}{4}$ is an embedding of Y in X. However, we claim that $X \not\simeq Y$, as Y has a first element while X does not. Given a function $f : Y \mapsto X$ examine $f(0) \in (0,1)$. There is an element $z < f(0)$ in $(0,1)$, but because f preserves the ordering this z cannot be in Range(f) and therefor f is not onto.*

Just like cardinalities, which are the equivalence classes of sets under the relation $\sim$, had well–known notations, the *ordering types*, which are the equivalence classes of orderings under the relation $\simeq$ have well–known notations.

Notation:

- Denote by ω the ordering type of $(I\!N, <)$.
- Denote by η the ordering type of $(\mathcal{Q}, <)$.
- Denote by λ the ordering type of $(\Re, <)$.
- For any ordering θ denote by θ^* the reverse ordering of θ. That is, if $(A, \leq)$ is a representative of θ then $(A, \leq^*)$ would be a representative of θ^* where $x \leq^* y$ if and only if $y \leq x$.

Note that for example ω^* (the ordering on the negative integers) is different from ω because while ω has a first element and does not have a last element, ω^* has a last element but does not have a first element. On the other hand $\eta^* = \eta$ and $\lambda^* = \lambda$, as the function $f(x) = -x$ is an isomorphism between η and η^* or between λ and λ^*.

Just like we defined a calculus of the infinite cardinalities of sets we may also define operators on the types of orderings. We thus define the following operators:

Definition 2.11: *Given two ordering types θ and θ' define the ordering $\theta + \theta'$ by the following: Let $(A, <_A)$ be an ordering of the type θ and $(B, <_B)$ be an ordering of the type θ' such that A and B are disjoint. We define the following ordering on $A \cup B$:*

$$x <_{A\cup B} y \ \text{iff} \begin{pmatrix} x, y \in A \text{ and } x <_A y & \text{or} \\ x, y \in B \text{ and } x <_B y & \text{or} \\ x \in A \text{ and } y \in B & \end{pmatrix}$$

We then define $\theta + \theta'$ to be the type of the ordering $<_{A\cup B}$.

For example, the ordering $\omega^* + \omega$ is the standard ordering of the integers $\mathbb{Z}$. The ordering type $\omega + \omega^*$ is just the opposite ordering, with all the positive integers preceding the negative integers. It is easy to verify that $\omega^* + \omega \neq \omega + \omega^*$ showing that ordering type addition is not commutative.

Having seen more than one ordering type on the naturals, the question of how many countable ordering types there are pops into mind. It is almost immediate to verify that there can be no more than $2^{\aleph_0}$ ordering types on countable sets. Given an ordering $(A, <_A)$ we define an ordering on $\mathbb{N}$ by choosing $f : A \overset{1-1}{\longmapsto} \mathbb{N}$ that is onto, but does not necessarily preserve the ordering and defining for every $l, k \in \mathbb{N}$ the ordering $l <'_A k$ if and only if $f^{-1}(l) <_A f^{-1}(k)$. For any two orderings $(A, <_A) \not\cong (B, <_B)$ it must hold that $\{(l,k) : l <'_A k\} \neq \{(l,k) : l <'_B k\}$ or else we would have obtained an isomorphism between $(A, <_A)$ and $(B, <_B)$.

The last argument shows that there is a one to one mapping of ordering types on countable sets to subsets of $\mathbb{N} \times \mathbb{N}$ so the amount of countable ordering types can be no more than $|P(\mathbb{N} \times \mathbb{N})| = 2^{\aleph_0}$.

Claim 2.7: *There are* $2^{\aleph_0}$ *types of orderings on countable sets.*

Proof: We define a one to one mapping in the opposite direction from the one we used in our previous argument. Given a set of indexed orderings $\{(A_i, <_{A_i}) : i \in I\}$ where all the A_is are disjoint and an ordering $<_I$ on I we define the *generalized sum*

$$\sum_{i \in I} A_i$$

as an ordering on $\bigcup_{i \in I} A_i$ and given $x, y \in \bigcup_{i \in I} A_i$ if both x and y are in the same A_i, make $x < y \iff x <_{A_i} y$. Otherwise, $x \in A_i$ and $y \in A_j$ for $i \neq j$ and define $x < y \iff i <_I j$. If for example, each of the A_is is a pair of elements $\{0_i, 1_i\}$ with the ordering $0_i <_i 1_i$ and $I = \mathbb{Q}$ then the ordering $\sum_{i \in \mathbb{Q}} A_i$ is not dense as there is no element between 0_i and 1_i for all i. Every element in this ordering will have an immediate successor or an immediate predecessor, but not both.

One possible mapping from $\{0,1\}^{\mathbb{N}}$ to the set of countable ordering types is defined by transforming $f \in \{0,1\}^{\mathbb{N}}$ into the ordering $\sum_{i \in \mathbb{N}} A_{f(i)}$ where $A_{f(i)}$ is defined by:

$$A_{f(i)} \triangleq \begin{cases} \omega & \text{if } f(i) = 0 \\ \omega^* & \text{if } f(i) = 1 \end{cases}$$

□

Having defined a summation operator on ordering types we now move forward to define the multiplication of ordering types.

Definition 2.12: *Given the ordering types θ, θ' define the ordering type $\theta \cdot \theta'$ to be the type of the ordering $\sum_{i \in I} A_i$ where $(I, <_I)$ is of the type θ' and the $(A_i, <_{A_i})$s are disjoint orderings of the type θ.*

Note that like ordering types addition, the multiplication operator is not commutative either. For example, it is immediate to see that $2 \cdot \omega = \omega$ while $\omega \cdot 2 = \omega + \omega$ (where 2 is an ordering on a set of cardinality 2).

Definition 2.13: *An ordering is called* dense *if for any $x < y$ in the ordering there is a z in the ordering such that $x < z < y$.*

It turns out that there are just 4 types of countable orderings that are dense. Possible representatives for them are $\mathcal{Q}$, $\mathcal{Q} \cap [0,1)$, $\mathcal{Q} \cap (0,1]$ and $\mathcal{Q} \cap [0,1]$. It is easy to see that these four ordering types are not isomorphic to each other (the first has neither a first nor a last element, the second has a first element but doesn't have a last element, etc.), and the following theorem demonstrate (part) of the evidence that these four ordering types are the only four countable dense orderings.

Theorem 2.7 ***(Kantor):*** *Every two countable dense orderings that have neither a first nor a last element are isomorphic to each other.*

Proof: Assume we have two countable dense orderings $(A, <_A)$, $(B, <_B)$. We look at two enumerating functions for the orderings: $f : I\!N \stackrel{1-1}{\longmapsto} A$, $g : I\!N \stackrel{1-1}{\longmapsto} B$, both are onto. We define a series of mappings $h_0 \subseteq h_1 \subseteq h_2 \ldots$ by induction on $I\!N$ such that every mapping is an extension of its predecessors and every mapping is one to one from A onto B and preserves the ordering. We will build the h_is so that the even functions will gradually cover the set A, and the odd functions will gradually cover the set B. We start with:

$$\begin{aligned} h_0 = &\quad \{(f(0), g(0))\} \\ h_1 = &\quad \{(f(0), g(0)), (f(1), ?)\} \end{aligned}$$

and use the following induction step:

- If i is even then if $f(\frac{i}{2})$ is in Range(h_{i-1}) we take $h_i = h_{i-1}$ otherwise we set $h_i = h_{i-1} \cup \{f(\frac{i}{2}, b_i)\}$ where b_i is chosen so that h_i will preserve the ordering. In other words, if $f(\frac{i}{2})$ is greater (under $<_A$) from all the elements of A for which h_{i-1} was defined then b_i will be an element of B that is greater (under the ordering $B_<$) from any element in Range(h_{i-1}), and similarly if $f(\frac{i}{2})$ is smaller from all the elements in A for which h_{i-1} was defined. Such a b_i exists in B as the ordering $(B, <_B)$ has neither a first nor a last element.

  If none of the above conditions hold, then there is an immediate predecessor and an immediate successor of $f(\frac{i}{2})$ in the domain of h_{i-1}. Let us denote them by $y <_A f(\frac{i}{2}) <_A x$. We pick any b_i such that $h_{i-1}(y) <_B b_i <_B h_{i-1}(x)$. Such a b_i must exist in B as B is dense.

- If i is odd we use the same method to add $g(\frac{i-1}{2})$ to Range(h_i) while preserving the order.

It is immediate that $\bigcup_{i \in I\!N} h_i$ is an isomorphism between $(A, <_A)$ and $(B, <_B)$. □

Definition 2.14*:* *An ordering is called* complete *if for every proper initial segment there is a minimal upper bound.*

For example, the ordering ω on $I\!N$ is complete. On the other end, $\omega + \omega^*$ is not complete.

Claim 2.8*:* *The ordering λ on $\Re$ is complete. Note that the reals are sometimes defined as the extension of the rationals so that the extension forms a complete ordering. Here we look at reals as numbers being represented by an infinite binary series.*

Proof: Let X be a non-empty proper initial segment of $\Re$ (when we treat $\Re$ as an ordering we mean the standard ordering λ). We will find a minimal upper bound for X. Denote by n the largest integer s.t. $n \in X$ but $n+1 \notin X$. We define by induction on i a series $\sigma(i)$ of 0 and 1:

$$\sigma(1) \triangleq \begin{cases} 0 & n + \frac{1}{2} \in X \\ 1 & n + \frac{1}{2} \notin X \end{cases}$$

With the induction step being:

$$\sigma(i+1) \triangleq \begin{cases} 0 & \frac{1}{2^{i+1}} + n + \sum_{j \le i} \frac{\sigma(j)}{2^j} \notin X \\ 1 & otherwise \end{cases}$$

We have thus defined a real number $r_\sigma = n.\sigma(1)\sigma(2)\ldots$ such that for all $x \in X$ it holds that $x \le r_\sigma$ while for all $x \notin X$, $x \ge r_\sigma$ and so r_σ is the required minimal upper bound on X. □

2.5 Well Orderings

Definition 2.15: *A linearly ordered set $(X, <_X)$ is called* well–ordered *if for all nonempty $A \subseteq X$ there is a minimal element in A.*

For example, $I\!N$ under the ordering ω is well ordered. On the other hand, $Z\!\!\!Z$ is not well ordered under the standard ordering as the subset $A = Z\!\!\!Z$ has no minimal element. The real interval $[0,1]$ is not well ordered under the ordering λ either, as the subset $(0,1)$ does not have a minimal element.

Theorem 2.8: *A set is well ordered if and only if the ordering ω^* cannot be embedded in it (or in other words, if and only if it does not have an infinite descending series).*

Proof: For the first direction, assume we are given an ordering $(X, <_X)$ and an order–preserving function $f : \omega^* \overset{1-1}{\longmapsto} (X, <_X)$. It is immediate that $A = \text{Range}(f)$ has no minimal element so X is not well ordered by $<_X$.

For the second direction, we assume that X is not well ordered by $<_X$, so there exists $\emptyset \ne A \subseteq X$ that has no minimal element. We define by induction on $I\!N$ a function $f : \omega^* \overset{1-1}{\longmapsto} A$ so that f preserves the ordering. We pick any $a_0 \in A$ (such an element is guaranteed to exist as A is known not to be empty). We set $f(0) \triangleq a_0$. Now, given any $f(-n) \in A$, we know that it is not a minimal element in A (since A has no minimal element), so we can pick an element $b \in A$ s.t. $b <_X a$ and set $f(-n-1) \triangleq b$. □

One of the main reasons we are interested in well–ordered sets is that these orderings have the *induction property.*

Definition 2.16: *We say that an ordering $(X, <_X)$ has the induction property if for all subsets $A \subseteq X$ it holds that:*

$$\forall y(\{x : x < y\} \subseteq A \rightarrow y \in A) \rightarrow A = X$$

or in other words: the only initial segment that "consumes" its upper bound is the whole set X.

Theorem 2.9: *A set is well–ordered if and only if it has the induction property.*

Proof: For the first direction of the proof assume that X is well ordered and that A is a consuming set. We have to prove that $A = X$. Assume, by contradiction, that $A \ne X$, so there is $\emptyset \ne B = X \setminus A$ and since X is well ordered, B has a minimal element b. Since b is minimal in B, $\{x : x < b\} \subseteq A$ and since A consumes, $b \in A$ in contradiction with the definition of B, and so it must hold that $B = \emptyset$.

For the second direction we assume that the induction property holds, and we prove that X is well ordered. If X is not well ordered there is a subset $\emptyset \ne B \subseteq X$ that does not have a minimal element. Define

$$A \triangleq \{y \in X : \forall b \in B \; y < b\}$$

We show that A is a consuming set. Given some element a, if $\{x : x < a\} \subseteq A$ we have to show that $a \in A$. Assume that $a \notin A$, so there exists $b \in B$ s.t. $b \leq a$. It can't be that $b < a$ as all elements that are less than a are in A, and A and B are disjoint, and so $a = b$ and $a \in B$. But note that this element is an element of B whose predecessors are not elements of B, so, in other words, it's a minimal element of B, in contradiction to our assumption that B has no minimal element. Therefore, if A is consuming then B is empty and $A = X$. $\square$

It is easy to see that the operators we've defined on orderings (i.e. addition and multiplication of ordering types) preserve well–ordering, or, simply put, the sum of two well–orderings is a well–ordering and the multiplication of two well–orderings is a well–ordering.

So far we've only seen countable well–orderings. A natural question would be: Are there non–countable well–orderings? The following theorem answers that in the affirmative:

Theorem 2.10 ***(Well–Ordering Theorem):*** *For every set X there exists a well–ordering on X (or there is a set $Y \subseteq X \times X$ s.t. Y is a linear well–ordering of X).*

Proof: We use Zorn's Lemma. We define the set:

$$F \triangleq \{Y' \subseteq X \times X : \mathrm{Dom}(Y') = \mathrm{Range}(Y') \text{ and } Y' \text{ is a well–ordering on its domain}\}$$

Note that F is non–empty as $Y' = \emptyset$ is a well–ordering on $\emptyset$. We now move to define an ordering on F itself. We say that $Y' \leq Y''$ if Y'' is an end–extension of Y', or $Y' \subseteq Y''$ and for every $a \in \mathrm{Dom}(Y'') \setminus \mathrm{Dom}(Y')$ and every $b \in \mathrm{Dom}(Y')$, $b <_{Y''} a$.

We will first have to show that if S is a chain of well–orderings under the ordering of F then $\bigcup S$ is a well–ordering. Given $\emptyset \neq A \subseteq \mathrm{Dom}(S)$ we pick $Y' \in S$ s.t. $\emptyset \neq A \cap \mathrm{Dom}(Y') \subseteq Y'$ and Y' is a well–ordering, so that $A \cap \mathrm{Dom}(Y')$ has a minimal element under the ordering Y'. Such a minimal element a is also a minimal element of the whole A as for every $b \in A$ there is a $Y'' \in S$ s.t. $b \in \mathrm{Dom}(Y'')$. If this Y'' satisfies $Y'' < Y'$ then $Y'' \subseteq Y'$ and so $b \in \mathrm{Dom}(Y)'$ and from a's minimality in Y' it holds that $a \leq b$. Otherwise, $Y'' \geq Y'$ and then either $b \in \mathrm{Dom}(Y')$ or that b is greater than any element in $\mathrm{Dom}(Y')$. In any case, $a \leq b$.

Zorn's Lemma now guarantees the existence of a maximal element $Y^* \in F$. As we have just seen, this Y^* is a well ordering on $\mathrm{Dom}(Y^*) \subseteq X$. If it is not defined for all of X then we pick an element $t \in X \setminus \mathrm{Dom}(Y^*)$ and define the ordering $Y^{**} \triangleq Y^* \cup \{(y,t) : y \in Y^*\}$, or in other words, we concatenate t at the end of the ordering Y^*. By definition $Y^* < Y^{**}$, in contradiction with the maximality of Y^*. $\square$

The following claims points out some of the amusing properties of well–orderings.

Claim 2.9*:* *If $(X, <_X)$ is a well ordering then for every order–preserving $f : X \overset{1-1}{\longmapsto} X$ it holds for all $a \in X$ that $f(a) \geq a$.*

Proof: By contradiction. Assume that the claim does not hold, so the set $B = \{x : f(x) < x\}$ is non–empty. Pick a minimal element $b \in B$. It satisfies $f(b) < b$. We apply f again, and since f preserves the ordering we get $f(f(b)) < f(b)$, implying that $f(b) \in B$ in contradiction with the minimality of b in B. $\square$

We can attempt to define a linear ordering on well–orderings types. We say that $(X, <_X) \leq (Y, <_Y)$ if there is an embedding of Y as an initial segment of X. It is clear that this relation is reflexive and transitive. How do we prove it to be antisymmetric? We must show that if $(X, <_X)$ can be embedded in an initial segment of $(Y, <_Y)$ and $(Y, <_Y)$ can be embedded in an initial segment of $(X, <_X)$ then it must be that $(X, <_X) \simeq (Y, <_Y)$.

Claim 2.10: *If in the previous setting we have two functions $f : X \overset{1-1}{\longmapsto} Y$ and $g : Y \overset{1-1}{\longmapsto} X$ both preserving the orderings then f must be onto Y.*

Proof: If f is not onto Y then $f \circ g$ is an order–preserving one to one mapping of X to a proper initial segment of itself and so there must be $a \in X$ s.t. $f \circ g(a) < a$ in contradiction with the previous claim. $\square$ We leave to the reader to prove that this ordering is linear (i.e. that given any two well–orderings either the first is embeddable in an initial segment of the second or the second is embeddable in an initial segment of the first).

Claim 2.11: *The ordering $\leq$ of well–ordering types is a well–ordering.*

The proof of this claim is also left to the reader.

Claim 2.12: *For any set A there is an ordering that is:*

1. *A well–ordering (we've already proved that).*
2. *Every proper initial segment of the ordering has cardinality strictly less than that of A.*

Proof: Examine the set of all well orderings of A under the ordering $\leq$ on well orderings we've just mentioned. Let $<^*$ be the minimal element in that set. If $B \subset A$ is a proper initial segment under $<^*$ of A then $|B| < |A|$ or else, the ordering on B can be used to define an ordering on A that can be embedded in a proper initial prefix of $(A, <^*)$, in contradiction with the minimality of $<^*$. $\square$

To sum up the section on well orderings, here is a list of the important properties we've proved for them so far:

1. A well ordering can be defined by any one of the following three properties:
    - Every subset has a minimal element.
    - It has no infinite descending series.
    - It has the induction property: The complete set is the only consuming set.
2. Well–orderings are preserved by addition and multiplication.
3. The well–ordering theorem: A well–ordering can be defined for any set.
4. We defined the relation "Embeddable as an initial prefix" on ordering types.
5. We have shown that this relation is an ordering of well ordering types.
6. We proved that any order preserving function f from a well ordered set to itself satisfies for any x $x \leq f(x)$.
7. The ordering we defined on well–ordering types is a linear ordering.
8. This ordering on well–ordering types is itself a well–ordering on any set of well–ordering types.
9. For any set there is a well–ordering of that set for which every proper initial segment has strictly smaller cardinality than the whole set.

We call an element x in a well ordering a successor if it has an immediate predecessor and a limit if it doesn't. Note that every element has a successor or else the group of all elements greater than it will not have a minimal element. In good orderings we many times mark by an element the set of all elements less than itself.

Definition 2.17: *An element in a well ordering (or, equivalently, an initial segment of the ordering) is called an* ordinal *if the cardinality of the initial segment is greater than the cardinality of any proper initial segment of it.*

Corollary 2.11: *Any infinite ordinal has no last element (otherwise, the initial segment that contains all elements but the last would have the same cardinality).*

Claim 2.13: *If λ and μ are two ordinals with the same cardinality then λ and μ are isomorphic.*

Proof: By property (7) of well orderings, for any two well orderings there is an embedding of one as an initial segment of the other. Assume, w.l.o.g. that λ can be embedded as an initial segment of μ. This embedding is 1 to 1 and preserves the ordering. We have to show it is onto. If it wasn't, this embedding would have been an embedding of λ in a proper initial segment of μ. But since μ is an ordinal, all its initial segments have cardinality strictly less than that of μ in contradiction with the assumption that $|\lambda| = |\mu|$. $\square$

From the last claim it follows that we may use ordinals to represent cardinalities. This means that the cardinalities are well–ordered under the ordering $\preceq$, as it is a sub-ordering of the ordering $\hookrightarrow$ on well–ordering types. This means that any cardinality has an immediate successor. For any ordinal λ denote by λ^+ its immediate successor. The use of ordinals to represent cardinalities has other benefits as demonstrated by the proof of the following claim:

Claim 2.14: *For any infinite cardinality λ it holds that $\lambda \cdot \lambda = \lambda$, or, in other words, for any infinite A, $A \times A \sim A$.*

Proof: Take any set B whose cardinality is greater than λ and define a well ordering on it. We use lower Greek letters to denote elements of B under this ordering. We prove by induction on B's ordering that for any $\alpha \in B$ it holds that $|\alpha \times \alpha| = |\alpha|$ (note that we use α to denote the initial segment of B that contains all elements less than α). Assume this holds for $\{\beta : \beta < \gamma\}$ and prove that it must hold for γ. There are two cases:

1. All $\beta < \gamma$ is finite. In this case, either γ is finite or $\gamma = \omega$. In both cases, the claim holds.
2. There is an infinite $\beta < \gamma$. Assume for contradiction that $|\gamma \times \gamma| \succ |\gamma|$. Note that both $\gamma \times \gamma$ and γ are well orderings. Therefore, there is an embedding of γ as an initial segment of $\gamma \times \gamma$, and from our contradiction assumption, this must be a proper initial segment of $\gamma \times \gamma$. Therefore, there is an element $(\delta, \eta) \in \gamma \times \gamma$ that is greater than the image of γ. Both δ and η are less than γ. W.l.o.g. assume that $\delta > \eta$, so under the ordering on $\gamma \times \gamma$ $(\delta, \eta) \leq (\delta, \delta)$. By the induction hypothesis $|\delta \times \delta| = |\delta|$, but $|\gamma| < |\delta \times \delta| = |\delta| \leq |\gamma|$ which implies that $|\gamma| = |\delta|$.

$\square$

Chapter 3

Point Set Topology

Topology is an attempt to add to the structures and objects we got familiar with in the previous chapter properties such as: proximity of points, continuity, open and closed intervals, etc.

3.1 Basic Definitions

Definition 3.1: *A* Topology Space *is a pair* (X, U) *where* $U \subseteq P(X)$. *U represents the set of "neighborhoods" in* X. *It must satisfy:*

1. $\emptyset, X \in U$.
2. *U is closed under* finite *intersections.*
3. *U is closed under* any *unions.*

A set $b \subseteq X$ *is called* open *if and only if* $b \in U$. *It is called* closed *if and only if* $X \setminus b \in U$.

A few natural examples of topology spaces include:

1. Set $X = \Re$ and U be the set of "open" sets on the reals - the set of any unions of open–ended intervals over the reals. Natural generalizations of this topology space include the topology $U = \{A : A \text{ is a union of sets of the form } (\alpha, \beta) \text{ in } X\}$ where X is ordered by any linear ordering (the ordering topology) or the topology $U = \{A : A \text{ is a union of open spheres in } X\}$ where X is any metrical space (the metric topology).

2. Let X be any infinite set and define U as $U \triangleq \{A \subseteq X : X \setminus A \text{ is finite}\} \cup \{\emptyset\}$ (the co–finite topology). It is easy to show that this is indeed a topology, as the finite intersection of two sets from this U gives a set whose complement in X is finite.

We will usually find it easier to define topology by the sets from which it is generated.

Definition 3.2: *A set* $B \subseteq P(X)$ *is called a* base for a topology *if it satisfies:*

1. *For any* $x \in X$ *there is a set* $b \in B$ *such that* $x \in b$.
2. *For any* $x \in X$ *and any* $b_1, b_2 \in B$, *if* $x \in b_1 \cap b_2$ *then there exists* $b_3 \in B$ *s.t.* $x \in b_3 \subseteq b_1 \cap b_2$.

An example for such a base may be the set of open intervals on the reals.

Definition 3.3: *Given a set B that is a base for a topology, we define the topology space that is based on B as:*

$$U(B) \triangleq \{A : A \text{ is a union of elements from } B\}$$

Claim 3.1: *If B is a base for a topology then $U(B)$ is indeed a topology.*

Proof:

1. $X \in U(B)$ as for any $x \in X$ there is a set $b \in B$ s.t. $x \in b$, and their union is in $U(B)$. $\emptyset \in U(B)$ because by the definition of unions $\emptyset = \bigcup \emptyset$.

2. $U(B)$ is closed under any unions since a union of elements in $U(B)$ is a union of unions of elements in B which is a union of elements in B itself.

3. We have to show that $U(B)$ is closed under finite intersection. For that purpose, we first show that for any B it holds that $A \in U(B)$ if and only if for any $x \in A$ there exists $b \in B$ s.t. $x \in b \subseteq A$.

   If $A \in U(B)$ then A is a union of elements from B. Since $x \in A$, at least one of the sets in that union contains x. Use that for the required b (note that it indeed complies with all the conditions we required of b). If, on the other hand, for any element in A there is $b \in B$ that contains it and is a subset of A we can construct A as a union of elements from B in the following way: For any $x \in A$ choose one such set in $b_x \in B$, and define A to be $\bigcup_{x \in A} b_x$.

   Now if $A_1, A_2 \in U(B)$ we have to show that $A_1 \cap A_2 \in U(B)$. We show that by showing that for any $x \in A_1 \cap A_2$ there is $b \in B$ s.t. $x \in b \subseteq A_1 \cap A_2$. Given such an x, we know that both A_1 and A_2 have the required property, so we have $b_{A_1} \in B$ s.t. $x \in b_{A_1} \subseteq A_1$ and $b_{A_2} \in B$ s.t. $x \in b_{A_2} \subseteq A_2$. But, because $x \in b_{A_1} \cap b_{A_2}$, by the second property of B we have $b_3 \in B$ s.t. $x \in b_3 \subseteq b_{A_1} \cap b_{A_2}$ and so $b_3 \subseteq A_1 \cap A_2$ and the required property holds.

□

We call the topology $U = \{X, \emptyset\}$ the *trivial topology* and the topology $U = P(X)$ the *discrete topology*. The discrete topology may also be defined as the topology based on $B = \{\{x\} : x \in X\}$.

We leave as an exercise the proof of the following (very useful) lemma:

Lemma 3.1: *The set of open–ended intervals with rational endpoints is a base for the standard topology on $\Re$.*

From this lemma, the following is immediate:

Corollary 3.1:

1. *The set of open sets on $\Re$ has cardinality $2^{\aleph_0}$.*

2. *The set of closed sets on $\Re$ has cardinality $2^{\aleph_0}$ (the mapping of open sets to closed sets is 1–1).*

3. *There are only $2^{\aleph_0}$ sets that are either closed or open.*

4. *The vast majority of subsets of $\Re$ are neither closed nor open.*

Definition 3.4: *We say a set $A \subseteq X$ is* dense *in a topology (X, U) if for any $b \in U$ s.t. $b \neq \emptyset$, $A \cap b \neq \emptyset$.*

For example, in the trivial topology, any set $A \neq \emptyset$ is dense. In the discrete topology, on the other hand, X is the only dense set. In an ordering topology a set A is dense in the topology if and only if it is dense in the ordering upon which the topology is based.

Definition 3.5: *A topology space is called* separable *if it has a countable dense set.*

We also introduce a concept which is, in some sense, the opposite of dense sets:

Definition 3.6: *A set is* Nowhere Dense (ND) *if it is a set $A \subseteq X$ s.t. for any non–empty $b \in U$ there is $\emptyset \neq b' \subseteq b$ s.t. $A \cap b' = \emptyset$ and $b' \in U$.*

As an example of NDs, recall the intervals $\{I_i\}$ of claim 2.3 we used to cover all the rationals by a small interval. We will show that $A = (\Re \setminus \bigcup_i I_i)$ is ND for any choice of such open intervals I_i under the standard ordering topology on the reals. Given any $b \in U$ we can find a rational number q and a neighborhood of it $b' = I_i$ which is an open set and $A \cap b' = \emptyset$.

Claim 3.2: *If $<$ is a complete and dense ordering on X then X cannot be covered completely by a countable set of NDs.*

Definition 3.7: *A* set of the first category *is a set that may be constructed by a* countable *union of NDs.*

For example, we may look on the branches of an infinite binary tree as our domain. These branches are labeled by words from $\{0,1\}^{\boldsymbol{N}}$. We define the "cauliflower" topology to be the topology based on the set $B_{\text{cf}} \triangleq \{b_\sigma : \sigma \in \{0,1\}^*\}$ where each b_σ is defined by

$$b_\sigma \triangleq \left\{ f \in \{0,1\}^{\boldsymbol{N}} : \sigma \text{ is a prefix of } f \right\}$$

Lemma 3.2: *The set B_{cf} satisfies the requirements for a base for a topology, namely:*

1. *Every function $f \in \{0,1\}^{\boldsymbol{N}}$ has an element $b_\sigma \in B_{cf}$ s.t. $f \in b_\sigma$.*
2. *For all $b_1, b_2 \in B_{cf}$, and for all $x \in b_1 \cap b_2$ there is $b_3 \in B_{cf}$ s.t. $x \in b_3$ and $b_3 \subseteq b_a \cap b_2$.*

Proof: It is immediate to see that both properties hold:

1. We have $b_\epsilon \in B_{\text{cf}}$ where every f is in b_ϵ.
2. If there is $f \in b_\sigma \cap b_\tau$ then it must be that one of b_σ, b_τ contains the other, and, w.l.o.g. we may choose b_τ as the required b_3.

□

Note that all the base elements (i.e. all b_σ) are closed sets as well. We have:

$$X \setminus b_\sigma = \bigcup_{\tau \neq \sigma, |\tau| = |\sigma|} b_\tau$$

Definition 3.8: *An* Oracle Turing Machine *(OTM for short) is a TM that has a basic operation of "querying an oracle". The machine may ask $Q = n$ and receives as an answer $Ans = f(n) \in \{0,1\}$. Denote by $L(M^{[f]})$ the language a machine M accepts where the queries are answered by the function f.*

Definition 3.9: *We say* $L_1 <_p L_2$ *(*L_1 polynomially reduces *to* L_2*) if there is a polynomial function* $h : \Sigma^* \to \Sigma^*$ *such that for every* ω*,* $\omega \in L_1$ *if and only if* $h(\omega) \in L_2$*.*

Claim 3.3: *If* $L_1 <_p L_2$ *then there is a polynomial OTM s.t.* $L(M^{[L_2]}) = L_1$*.*

Definition 3.10: *We define a* complexity class relative to an oracle *as follows: Given the oracle* f *define:*

$$P^{[f]} \triangleq \left\{L : \textit{There exists a polynomial OTM } m, \textit{ s.t. } L(m^{[f]}) = L\right\}$$

It immediately follows that for any oracle f that is in P itself the class $P^{[f]}$ is exactly P.

3.2 Applications to Game Theory

We can use topological definitions to define games, for example on the reals.

The model we concern ourselves with is the following: A game is played by two players who together construct a real number (as a series of binary digits). At the ith stage, player I picks $\sigma_i^{\mathrm{I}} \in \{0,1\}^*$ and after player II sees player I's choice it picks $\sigma_i^{\mathrm{II}} \in \{0,1\}^*$. Define a subset $A \subseteq \Re$ of the reals. Player I's goal is to ensure the constructed real number $\sigma_1^{\mathrm{I}}\sigma_1^{\mathrm{II}}\sigma_2^{\mathrm{I}}\sigma_2^{\mathrm{II}}\ldots$ is in A.

Definition 3.11: *A* strategy *is a function mapping the history of a game to the next move a player will take. In our case, it maps a finite series of the form:* $\sigma_1^{\mathrm{I}}\sigma_1^{\mathrm{II}}\ldots\sigma_n^{\mathrm{I}}\sigma_n^{\mathrm{II}}$ *(or one with only player I's last move in a strategy for player II) to a single move* $\sigma_{n+1}^{\mathrm{I}}$*. A strategy* H *is called a* winning strategy *for the player* i *if any series (or game) where player* i *plays by* H *end with* i*'s victory.*

Note that by our definition, given a set A, there can be a winning strategy for *at most* one of the players. From counting arguments it can be easily shown that there are goal sets A for whom no winning strategy exist. A game is called *open* if the set A is open in the cauliflower topology.

Theorem 3.2: *In any open game there is a winning strategy for one of the players.*

Theorem 3.3 *(Bank–Mazur)*: *Player II has a winning strategy if and only if* A *is a set of the first category in the cauliflower topology.*

It is immediate to find a strategy for player II if A is a set of the first category. In that case, A may be written as $\bigcup I_i$ where each I_i is ND. All player II has to do is choose its next move in the ith stage so that it arrives in the subset of the current neighborhood which is disjoint to A_i, and such a set can always be found for all A_i.

3.3 Further Topological Definitions

Definition 3.12: *Given a set* $A \subseteq X$ *in a space* (X,U) *we define the* closure *of the set* A *to be:*

$$\overline{A} \triangleq \bigcap \{B : A \subseteq B, B \textit{ is closed}\}$$

It may be worthwhile to note the following about the closure of sets:

1. $\overline{A}$ is always a closed set, as the set of closed sets is closed under any intersections (since the set of open sets is closed under any unions).

2. The closure of a set A is the minimal closed set (under containment) that contains the set A.

3. $x \in \overline{A}$ iff for any open neighborhood of x (i.e. for every $b \in U$ s.t. $x \in b$) it holds that $A \cap b \neq \emptyset$.

The proof for the third claim on the list above is:

Proof: If there is a $x \in b \in U$ s.t. $b \cap A = \emptyset$ then $X \setminus b$ is closed and $A \subseteq X \setminus b$, and so $\overline{A} \subseteq X \setminus b$ but $x \notin X \setminus b$ and so it must be that $x \notin \overline{A}$.

For the other direction of the proof, assume that $x \notin \overline{A}$. Then $b = (X \setminus \overline{A})$ is an open set (as it is the complement of a closed set) and $x \notin b$ but $A \subseteq b$ as $A \subseteq \overline{A}$. □

Definition 3.13: *Given $A \subseteq X$ we define the induced topology (from the space (X, U)) on A to be $(A, \{b \cap A : b \in U\})$.*

It is easy to verify that this is indeed a definition of a topology space. As an example for this definition we may take the standard topology on $\Re$ and pick $A = [0, 1)$. The induced topology is based on $\{(a, b) : 0 \leq a \leq b \leq 1\}$ and $\{[0, b) : 0 \leq b \leq 1\}$. Another subset based on the standard topology on the reals may be the set $A = \{0\} \cup \{\frac{1}{n} : n \in I\!N\}$. In this subspace all singletons, except for $\{0\}$ are open. The neighborhoods of $\{0\}$ are of the form: $\{0\} \cup \{\frac{1}{n} : n > k\}$ for any natural number k.

Note that if A is open in (X, U) then a set $b \subseteq A$ is open in the induced topology if and only if b is open in (X, U). Analogously, if A is closed in (X, U) then a set $b \subseteq A$ is closed in the induced topology if and only if it is closed in (X, U).

3.4 Functions on Topology Spaces

Definition 3.14: *Given two topology spaces $(X, U_X), (Y, U_Y)$, a function $f : X \mapsto Y$ is called* continuous *if for all $b \in U_Y$ it holds that $f^{-1}(b) \in U_X$ where*

$$f^{-1}(b) \triangleq \{x : f(x) \in b\}$$

Recall that the standard definition for continuous functions from $\Re$ to $\Re$ used in calculus is: A function $f : \Re \mapsto \Re$ is called continuous if for all $\epsilon > 0$ and all $x \in \Re$ there exists $\delta > 0$ s.t.

$$|x' - x| < \delta \rightarrow |f(x') - f(x)| < \epsilon$$

Claim 3.4: *A function $f : \Re \mapsto \Re$ is continuous under the calculus definition if and only if it is continuous in the topological sense under the standard topology on $\Re$.*

Proof: First we assume that we are given a function that is known to be continuous in the topological sense and we prove it to be continuous under the calculus definition. Given ϵ, x we have to find the required δ. We define the set $b = (f(x) - \epsilon, f(x) + \epsilon)$. This set b is open and so its source under f must also be open. We also know that $x \in f^{-1}(b)$, and so there is a basic open set (i.e. an open set that is in the base of the topology) that is contained in b and contains x. This basic set is an interval of the form (s, t). We define the required δ as follows:

$$\delta \triangleq \min(x - s, t - x)$$

The interval $(x - \delta, x + \delta) \subseteq f^{-1}(b)$ so for any $x' \in (x - \delta, x + \delta)$ it must hold that $f(x') \in (f(x) - \epsilon, f(x) + \epsilon)$.

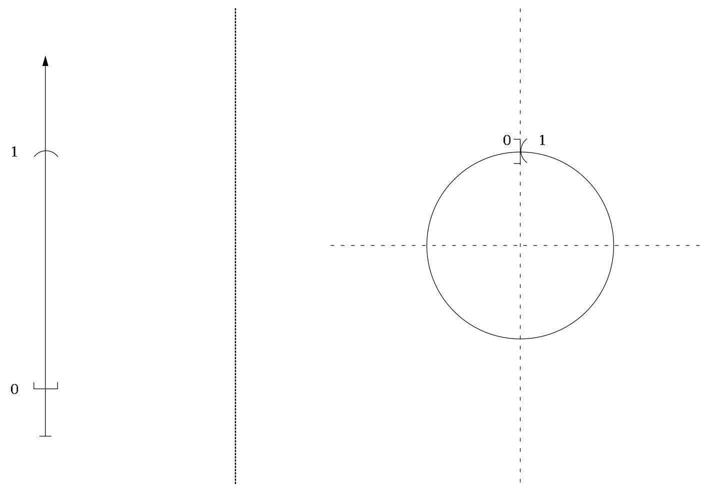

Figure 3.1: Mapping $[0,1)$ to a circle

For the second direction, we assume we are given a function that is continuous in the calculus sense. We now pick an open $b \in \Re$ and show that its source under f is open as well. It is enough to show that for any $x \in b$ there is an open interval (s,t) s.t. $x \in (s,t)$ and $(s,t) \subseteq f^{-1}(b)$. If b is the empty set then its source is the empty set, which is open in any topology. Assume there exists $x \in f^{-1}(b)$. It follows that $f(x) \in b$ and since b is an open set there is an interval $(u,v) \subseteq b$ s.t. $f(x) \in (u,v)$. Define $\epsilon \triangleq \min(v - f(x), f(x) - u)$. It is clear that $\epsilon > 0$ and that $(f(x) - \epsilon, f(x) + \epsilon) \subseteq b$. By the continuity of f under the calculus sense it follows that there exists δ s.t. if $|x' - x| < \delta$ then $|f(x') - f(x)| < \epsilon$, or in other words, if $x' \in (x - \delta, x + \delta)$ then $f(x') \in (f(x) - \epsilon, f(x) + \epsilon)$, which implies that $f(x') \in b$ and so $(x - \delta, x + \delta) \subseteq f^{-1}(b)$ and $x \in (x - \delta, x + \delta)$ so for any element in $f^{-1}(b)$ there is an open neighborhood of x in $f^{-1}(b)$, so $f^{-1}(b)$ must be an open set. $\square$

Analogously to that last claim it may be proven that a continuous function $f : \Re^n \mapsto \Re^n$ under the calculus sense is simply a function that is continuous in the topological sense under the open spheres topology on $\Re^n$.

For example, consider the two topologies X, the standard ordering topology on $\Re$ and $Y = (\Re, U(B))$ where B is the set of all intervals of the form $[a,b)$. The identity function from X to Y is not continuous, but the identity function *is* continuous from Y from X.

As another example, consider the mapping that transforms the interval $[0,1)$ to the circumference of a circle in $\Re^2$ (see figure 3.1) where the topology on that ring in $\Re^2$ is the topology induced by the open spheres topology on $\Re^2$ and the topology on the interval $[0,1)$ is the topology induced by the standard topology on the reals. This transformation is continuous from the interval to the ring in $\Re^2$ but its inverse is not, because the source of the open set of the form $[0,a)$ in $\Re^2$ will not be an open set.

We now show an alternative definition of continuity that may relate more easily to the intuitive notion of continuity:

Claim 3.5: *$f : X \mapsto Y$ is continuous if and only if for all $A \subseteq X$, $f(\overline{A}) \subseteq \overline{f(A)}$.*

Note that the alternative definition offered by the last claim means that the values given by f for elements at the limit of A are contained in the limit of the values of f itself.

Proof: Assume that f is continuous. Given $y \in \overline{A}$ it is known that for any $b \in U$ s.t. $y \in b$, $b \cap A \neq \emptyset$. Given an open set b' s.t. $f(y) \in b'$ we have to show that $b' \cap f(A) \neq \emptyset$. Consider $f^{-1}(b')$. This is an open set in X since f is continuous and $y \inf^{-1}(b')$ because $f(y) \in b'$, and so $f^{-1}(b') \cap A \neq \emptyset$ (as y is surely in that intersection) which means that $f(f^{-1}(b')) \cap f(A) \neq \emptyset$, or, simply put, $b' \cap f(A) \neq \emptyset$.

For the other direction of the proof, assume that $\forall A f(\overline{A}) \subseteq \overline{f(A)}$ and we will show that f is continuous. Given an open set $b \subseteq Y$ we consider $Y \setminus b$ (which is a closed set). We show that $f^{-1}(Y \setminus b)$ is a closed set in X. This will suffice as $f^{-1}(Y \setminus b) = f^{-1}(Y) = f^{-1}(b) = X \setminus f^{-1}(b)$.

We begin by showing that $\overline{f^{-1}(Y \setminus b)} \subseteq f^{-1}(Y \setminus b)$. Note that the following is immediate from the assumption on f and from the fact that b is an open set:

$$f(\overline{f^{-1}(Y \setminus b)}) \subseteq \overline{f(f^{-1}(Y \setminus b))} = \overline{Y \setminus b} = Y \setminus b$$

Now we apply f^{-1} to both sides of the last equation to get:

$$f^{-1}(f(\overline{f^{-1}(Y \setminus b)})) \subseteq f^{-1}(Y \setminus b)$$

and again, f and f^{-1} cancel each other on the left hand side to get $\overline{f^{-1}(Y \setminus b)} \subseteq f^{-1}(Y \setminus b)$. □

We can now use the definitions we've just seen and their equivalence to the classic definitions from calculus to derive some basic properties of continuous functions much more easily:

Corollary 3.4: *The composition of two continuous functions is a continuous function.*

Proof: By the topological definition of continuity this is immediate. If there is a set $b \in Z$ then $g^{-1}(b) \in Y$ is an open set for any open set b and so $(f \circ g)^{-1}(b) = f^{-1}(g^{-1}(b))$ is also an open set. □

Definition 3.15: *We define* Homeomorphism *between topology spaces. Two topology spaces X, Y are homeomorphic if there exists a function $f : X \overset{1-1}{\longmapsto} Y$ that is onto, continuous and f^{-1} is also continuous.*

For example, the real segment $(0, 1)$ is homeomorphic to $\Re$, when both are considered with respect to the standard ordering topology on the reals (it can be shown by a function based on tan, for example).

3.5 Connectedness

So far we demonstrated the topological generalizations of two major properties of the real numbers:

- The reals are an ordering topology.
- That ordering is complete and dense.

We now move on to consider the property of connectivity.

Definition 3.16: *A topology space X is called* connected *if for any two open sets $A, B \subseteq X$ s.t. $A, B \neq \emptyset$ and $A \cup B = X$ it holds that $A \cap B \neq \emptyset$.*

For example, $\Re \setminus \{0\}$ under the standard topology is not connected because we may have $A = \{x : x < 0\}$ and $B = \{x : x > 0\}$. It is also easy to see that $\mathcal{Q}$ with the standard topology is not connected. Here we may pick $A = \{q : q < 0 \vee q^2 < 2\}$ and $B = \{q : q^2 > 2 \wedge q > 0\}$.

Theorem 3.5: *If X is a topology space defined by an ordering topology induced by a complete and dense ordering then X is connected (and so is any interval in X).*

Note that any complete ordering that has a dense countable subset is isomorphic to $\Re$ and so its induced ordering topology is homeomorphic to the standard ordering topology on $\Re$. As a different ordering, consider the lexicographic ordering of $\Re^2$. This ordering is not separable because the set of intervals (under this ordering) $\{[(x,0),(x,1)] : x \in \Re\}$ are pairwise disjoint and of cardinality $2^{\aleph_0}$ so any dense set must intersect each one of them.

Proof: Let X be the topology space induced by a complete and dense ordering and let $A, B \subseteq X$ two open nonempty sets that cover X. Pick two elements $a \in A$ and $b \in B$. W.l.o.g. assume that $a < b$. We prove by contradiction that $A \cap B \neq \emptyset$. Assume that $A \cap B = \emptyset$.

Consider the set $\{x \in A : x < b\}$. This is a bounded subset of X and so (as the ordering is complete) there must be a minimal upper bound for it c. We now prove that $c \notin A \cup B$. First, assume that $c \in B$. Since B is an open set there must be a basic open set $u = (s,t)$ s.t. $c \in (s,t)$. This means that $s < c < t$ and because the ordering is dense there must be a point w s.t. $s < w < c$. But because $(s,t) \in B$ this means that this w is an upper bound on $\{x \in A : x < b\}$ because $A \cap B = \emptyset$ and $[w,c] \subseteq B$, in contradiction with the choice of c as the minimal upper bound. So, it must hold that $c \notin B$.

In a similar way, if $c \in A$ there are s', t' s.t. $c \in (s',t') \subseteq A$ and because the ordering is dense we may pick $c < w' < t'$ and we get $[c,w'] \subseteq A$ and because A and B are disjoint $b \notin [c,w']$ so it must be that $w' < b$ which means that c is not greater than all of $\{x \in A : x < b\}$.

All in all, we have that if $A \cap B = \emptyset$ then $A \cup B \neq X$. □

Lemma 3.3: *If X is connected and $f : X \mapsto Y$ is continuous then Range(f) is connected under the topology induced on it by Y's topology.*

Proof: Immediate. Assume that Range(f) is not connected, then there are disjoint sets A, B that cover Range(f) and because f is continuous, $f^{-1}(A)$ and $f^{-1}(B)$ are also open, clearly disjoint and cover X, in contradiction to the fact that X is itself connected. □

This last lemma's proof actually describes the intuition that a continuous function over a connected domain (such as $\Re$) can be drawn without lifting the pen from the paper.

Theorem 3.6 *(Mid–Value)*: *If $f : \Re \mapsto \Re$ is a continuous function, then for every $(s,t) \in \Re$, if $f(s) < f(t)$ then for all z s.t. $f(s) \le z \le f(t)$ there exists w between s and t such that $f(w) = z$.*

Proof: If there is a point z that contradicts the theorem then Range$(f_{|[s,t]})$ can be written as:

$$\text{Range}(f_{|[s,t]}) = (\text{Range}(f_{|[s,t]}) \cap \{y : y < z\}) \cup (\text{Range}(f_{|[s,t]}) \cap \{y : y > z\})$$

and by applying f to both sides we get:

$$[s,t] = \{s \le x \le t : f(x) < z\} \cup \{s \le x \le t : f(x) > z\}$$

which is a union of two disjoint open set, and that is, by the last lemma we proved, a contradiction on the connectivity of $[s,t]$. □

This theorem has many uses. One easy use is demonstrated by the proof of the following claim:

Claim 3.6: *For every $f : [0,1] \mapsto [0,1]$ that is continuous there is a fix point (a point x s.t. $f(x) = x$).*

Proof: If f is continuous then so is $g(x) \triangleq x - f(x)$ (because it is a composition of two continuous functions). If 0 is a fix point (i.e. $f(0) = 0$) then we are done. Otherwise, $g(0) = 0 - f(0) < 0$. If 1 is a fix point we are again, done. Otherwise, $g(1) = 1 - f(1) > 0$. By the mid–value theorem there is $z \in [0,1]$ s.t. $g(z) = 0$ which means exactly that $f(z) = z$. □

We now define a slightly different notion of connectedness:

Definition 3.17: *A topology space X is called* path–connected *if for any $x, y \in X$ there exists a continuous function $f : [0,1] \mapsto X$ s.t. $f(0) = x$ and $f(1) = y$.*

Claim 3.7: *If a topology space X is path–connected then X is connected.*

Proof: Assume that X is not connected. Pick two disjoint nonempty open sets A, B that cover X and two elements $a \in A$ and $b \in B$. Be f the function guaranteed for a and b by the definition of path–connectiveness. We have that $A \cap \text{Range}(f)$ and $B \cap \text{Range}(f)$ are disjoint, open in the topology space induced on Range(f), cover Range(f) and are nonempty (they contain the points a and b respectively) and so we have that Range(f) is not connected, in contradiction with the definition. □

This last claim gives a tool that's useful in showing the connectedness of some topology spaces where it might be easier to show path connectedness rather than simple connectivity.

Corollary 3.7: *For any positive integer n, the space $\Re^n$ under the open spheres topology is connected.*

Proof: We show this by demonstrating that this space is path–connected. Given two points $x, y \in \Re^n$ we define $f(t) : [0,1] \mapsto \Re^n$ as:

$$f(t) \triangleq (1-t)\overline{x} + t\overline{y}$$

which is clearly a continuous function (because multiplication and addition are continuous functions). □

The natural question that now pops into mind is whether all connected topology spaces are also path connected. We show that there are two reasons why a connected space will not be path connected. Intuitively, one can think of these as either the space is "too long" so that any continuous function from the unit interval cannot reach both its ends or the structure of the topology space may such that there is no specific line along which a point may be reached.

For the first type of counterexamples, consider the topology defined on $\Re^2$ by the lexicographic ordering. This is a complete and dense ordering so $\Re^2$ under this ordering must be a connected topology space. However, the interval between $(0,0)$ and $(1,1)$ in this topology covers a set of continuum cardinality of disjoint open intervals (i.e. all intervals of the form $\{((x,0),(x,1)) : 0 < x < 1\}$). Assume now that there was a continuous function $f : [0,1] \mapsto \Re^2$ such that $f(0) = (0,0)$ and $f(1) = (1,1)$. By the mid–value theorem, every point in the interval $[(0,0),(1,1)]$ is in f's range so for all $0 < x < 1$, $f^{-1}((x,0),(x,1)) \neq \emptyset$ is an open set on $[0,1]$ and for different values of x those are disjoint sets, so we have $2^{\aleph_0}$ disjoint open subsets of $[0,1]$ in contradiction with the separability of $[0,1]$. Note that this same argument holds for any non-separable topology space:

Corollary 3.8: *If a topology space X is path connected it is also separable.*

We now introduce a different type of topology spaces that have some valuable properties.

Definition 3.18: *A* metric space *is a set X and a function $d : X^2 \mapsto \Re^+$ that satisfies:*

1. $x = y \iff d(x, y) = 0$.
2. *The triangle inequality:* $\forall x, y, z \in X$, $d(x, y) + d(y, z) \geq d(x, z)$.
3. $\forall x, y$, $d(x, y) = d(y, x)$.

As an example of a metric space we may consider the set $\Re$ with the function $d(x, y) \triangleq |x - y|$, the set $\Re^n$ with the function $d(x, y) \triangleq \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$ or the set $\Re^n$ with the metric function $d(x, y) \triangleq \sum_{i=1}^{n} |x_i - y_i|$.

Definition 3.19: *For any set X, define the* discrete metric *to be:*

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

3.6 Compactness

So far we have concerned ourselves with the generalization of the following attributes of the real line:

- The reals form a complete dense and separable ordering.
- The reals are connected.

Another characterization of the reals that we will generalize is the attribute that implies properties such as: "All continuous functions over a closed interval have a maximum and minimum value within that interval".

Definition 3.20: *A topology space X is called* compact *if for any set of open sets S that cover X (i.e. $\bigcup S = X$), there is a finite sub–cover, or a subset $A \subseteq S$ that is finite and $\bigcup A = X$.*

As an easy example, consider the topology space induced over the domain $X = \{0\} \cup \{\frac{1}{n} : n \in I\!N\}$ by the standard topology on $\Re$. Given an open cover S for X, pick any $a \in S$ s.t. $0 \in a$. a is an open set in this topology, so it must be of the form $a = b \cap X$ where b is an open set in the standard topology on the reals. It follows that $0 \in b$ and so there is a basic open set s.t. $0 \in (s, t) \subseteq b$ and $(s, t) \cap X \subseteq a$ and $0 < t$. This means that $X \setminus a$ has only finitely many elements. For any $a \in X \setminus a$ pick a set $a_i \in S$ that covers it and we have $A = \{a, a_1, \ldots, a_k\}$ as a finite sub-cover of X.

On the other hand, note that the same space without the zero point (i.e. the space induced on $X = \{\frac{1}{n} : n \in I\!N\}$) is not compact, because all singletons are open sets and they form an infinite open cover that has no finite sub-cover for X.

It is immediate to see that all finite spaces are compact, and that the discrete topology over any infinite set is not compact (by an argument similar to the one given for the last example above).

Having already demonstrated the relation between our new topological definitions and classical calculus definitions and results we go on to demonstrate the connection between the compactness on topology spaces we've just defined and the notion of compactness used in sentential logic.

To do so, we define a topology space for sentential logic: we set

$$X = \{V : V \text{ is an assignment of truth values for all atoms}\}$$

and for any sentence α set $b_\alpha \triangleq \left\{V : \overline{V}(\alpha) = T\right\} \subseteq X$ and use the set of all such sets $B \triangleq \{b_\alpha :\ \alpha \text{ is a sentence}\}$ as the base for the topology space on X.

We have to verify that B is indeed a base for a topology space. If $b_\alpha \cap b_\beta \neq \emptyset$ and V is some assignment in $b_\alpha \cap b_\beta$ then $b_{(\alpha\wedge\beta)} \subseteq b_\alpha \cap b_\beta$ and clearly, $v \in b_{(\alpha\wedge\beta)}$ and so B satisfies the conditions for a base set for a topology space. We now consider the topology space $(X, U(B))$. Recall the compactness theorem from logic:

Theorem 3.9 ***(Sentential Logic Compactness):*** *A set of sentences Σ is satisfiable if and only if every finite subset of it is satisfiable.*

Note that it is immediate that if Σ is satisfiable then every finite subset of it is satisfiable. We will thus only prove the "interesting" direction of the theorem.

Proof: Assume Σ is not satisfiable, so that for any $V \in X$ there exists $\alpha \in \Sigma$ such that $\overline{V}(\alpha) = F$ or $\overline{V}(\neg\alpha) = T$ and so $S_\Sigma = \left\{b_{(\neg\alpha)} : \alpha \in \Sigma\right\}$ is a cover of X. Therefor, if the topology is compact there exists a finite sub–cover and so there exists a finite subset $A \subseteq \Sigma$ that is not satisfiable. This demonstrates that the compactness theorem from logic is equivalent to the compactness of the space $(X, U(B))$ we've just defined. □

Theorem 3.10*:* *If X is compact and there is a linear ordering defined on the space Y and the function $f : X \mapsto Y$ is continuous for the ordering topology on Y then f has a maximum, i.e. there exists $x_0 \in X$ such that for all $x \in X$, $f(x) \leq f(x_0)$.*

Proof: Assume for contradiction that there is no maximum in Range(f). Define the set $b_a \triangleq \{y : y < a\}$. We now examine the set $b_a : a \in \text{Range}(f)$. If there is no maximum in Range(f) then this set is a covering of Y by open sets. Because f is continuous, the set $S = \{f^{-1}(b_a) : a \in \text{Range}(f)\}$ is an open–set cover of X.

By the compactness of X there iss a finite subset $A \subset S$ that covers X. This means that there are elements $a_1 \ldots a_k \in Y$ such that $\{f^{-1}(a_1) \ldots f^{-1}(a_k)\}$ is a cover of Y, but then $\{b_{a_1} \ldots b_{a_k}\}$ is a cover of Range(f). Because the ordering on Y is linear, we can pick a maximal a_i among $a_1 \ldots a_k$ and then

$$\text{Range}(f) = \bigcup_{j=1}^{k} b_{a_j} \subseteq b_{a_i}$$

and this contradicts the assumption because $a_i \notin b_{a_i}$ while $a_i \in \text{Range}(f)$. □

We now introduce another simple definition of compactness:

Theorem 3.11*:* *X is compact if and only if for any set W of closed sets from X if for every finite subset $A \subseteq W$, $\bigcap A \neq \emptyset$ then $\bigcap W \neq \emptyset$.*

Proof: We use the complements of those sets, and De-Morgan's laws, noting the following immediate condition:

$$\emptyset \neq \bigcap\{t : t \in A\} \iff \bigcup\{X \setminus t : t \in A\} \neq X$$

□

Recall that our motivation for defining compactness for topological spaces was to generalize some of the attributes of the real interval. The following theorem justifies that:

Theorem 3.12*:* *The real interval $[0, 1]$ (or more generally, any closed interval in any topology space induced by a complete ordering) is compact.*

Proof: Given an open covering S for the interval $[0,1]$ define the set

$$k \triangleq \{0 \leq a \leq 1 : \text{ there exists a finite sub-cover for } [0,a] \text{ in } S\}$$

The set k is an initial segment of the interval $[0,1]$ because for all $s < t$ if $t \in k$ then $s \in k$ (as the cover for $[0,t]$ will also be a cover for $[0,s]$). Also note that $0 \in k$ and so $k \neq \emptyset$. Therefore, since the ordering is complete, there is a least upper bound t_0 for k. We now prove two attributes of this t_0 that will apply the theorem:

$t_0 \in k$ Pick any $b_0 \in S$ such that $t_0 \in b_0$. b_0 is an open set so there must be an interval $(r,s) \subseteq b_0$ s.t. $r < t_0 < s$. Since t_0 is a least upper bound for k it follows that $r \in k$ and so $[0,r]$ has a finite sub-cover. Add the set b_0 to that finite sub-cover to get a (still) finite sub-cover that covers $[0,t_0]$ as well.

$t_0 = 1$ Here we assume for simplicity that the topology is dense (though the proof can be modified to hold for non–dense spaces as well). This means that if $t_0 < 1$ there is s' so that $t < s' < s$ (where s is the one we used in the previous item to prove that $t_0 \in k$). This means that the finite sub-cover we've described in the previous paragraph is a cover for $[0,s']$ as well, contradicting the choice of t_0 as an upper bound on k.

□

We now show some nice properties of compactness for topology spaces:

Claim 3.8: *If X is a compact topology space and $f : X \mapsto Y$ is continuous then Range(f) is compact in the topology induced by the topology on Y.*

Proof: Let S be an open cover of Range(f). Consider $\{f^{-1}(b) : b \in S\}$ which is an open cover of X. Since S has a finite sub-cover A it follows that $\{b \in S : f^{-1}(b) \in A\}$ is a finite sub-cover of Range(f). □

Claim 3.9: *If X is a compact topology space and $A \subseteq X$ is closed then A is compact in the topology space induced by the topology space X.*

Proof: Given S, an open covering of A, the set $S' \cup \{X \setminus A\}$ is an open covering if X where S' is the expansion into open sets in X of the open sets in S. Because X is compact, there is a finite $K \subseteq \cup\{X \setminus A\}$ sub-cover for X. The set $\{b' \cap A : b' \in K\}$ is a finite open sub-covering of S. Its elements are open on A because they are the intersection of A with sets that are open in X, it is finite because K is finite, it covers A because it is an intersection of sets that cover X with A, and it is a sub-covering of S because the elements that went into K are expansions of S to open sets in X. □

Let us demonstrate that the converse of the last claim is false. Consider the following topology space over some infinite set X:

$$U = \{b \subseteq X : X \setminus b \text{ is finite}\} \cup \{\emptyset\}$$

It is easy to verify that this is indeed a topology space (known as the co–finite topology). We claim that in this topology any $A \subseteq X$ is compact. Given a set A, we pick some element $x_0 \in A$. Given a covering S for A, pick any $b \in S$ s.t. $x_0 \in b$. All that is left is to cover $A \setminus b$, which is a finite set. For each $x_i \in A \setminus b$ pick some $b_i \in S$ s.t. $x_i \in b_i$. The set $\{b, b_1, \ldots, b_k\}$ is a finite covering of A. On the other hand, any nonempty set whose complement is infinite is not open, so its complement is not closed so compactness does not imply closedness.

Theorem 3.13*:* *In metric spaces (where the topology is induced by a metric function), for any compact* X*, and any subset* $A \subseteq X$*,* A *is compact if and only if it is closed.*

Corollary 3.14*:* *There is no metric function for an infinite* X *that induces the co–finite topology.*

Proof: For any $x_0 \notin A$ we show that there exists an open set b s.t. $x_0 \in b$ and $b \cap A = \emptyset$. How does one find such a b? Given $x_0 \notin A$ for any $y \in A$ define:

$$U_y \triangleq \left\{ z : d(z,y) < \frac{d(x_0,y)}{2} \right\}$$

and

$$V_y \triangleq \left\{ x : d(x,x_0) < \frac{d(x_0,y)}{2} \right\}$$

The set $S = \{U_y \cap A : y \in A\}$ is an open covering of A. From the assumption that A is compact we have a finite sub-covering $k \subseteq S$ for A. The set $\bigcap \{V_y : U_y \in K\}$ is a finite intersection of open sets, so it is itself an open set. This intersection contains x_0, because for any y, $x_0 \in V_y$. This intersection is disjoint to A because it is disjoint to $\bigcup K$ which contains A, therefore $X \setminus A$ is an open set (because any x_0 in it has an open neighborhood in it) and so A is closed. Note that the proof holds for any compact A, even if X is not compact. □

Corollary 3.15*:* *Any set* A *that is compact in any metric space is closed in that space.*

Corollary 3.16*:* *A set* $A \subseteq \Re$ *is compact if and only if* A *is closed and bounded (i.e. there exists a finite* m *such that for any* $x, y \in A$*,* $d(x,y) < m$*).*

Proof: If A is compact then it is bounded because we can build an open covering for A by taking the set of all open segments of length $\frac{1}{8}$ around a point in A. Because this (possibly infinite) covering has a finite sub-covering it must follow that A is bounded. On the other hand if A is bounded then there is a closed interval $[s,t]$ such that $A \subseteq [s,t]$ therefore A is compact as a closed subset of $[s,t]$ which is compact. □

In calculus we used to say that y is an accumulation point of a series $\{a_n : n \in I\!N\}$ if for any $\epsilon > 0$ the set $\{n : |a_n - y| < \epsilon\}$ is infinite.

We say that $\{a_n\}$ converges to y if for any $\epsilon > 0$ the set $\{n : |a_n - y| < \epsilon\}$ is co–finite.

We wish to generalize those definitions using the tools we have seen so far for topology spaces. We replace the condition $\epsilon > 0$ by an empty set b such that $y \in b$ and the condition $|a_n - y| < \epsilon$ with the condition $a_n \in b$. We would now like to find an analogous notion for the notion of "most" point are closed used in the above definitions from calculus. We would like, given a set A, to check whether "most" elements in A satisfy some condition.

Definition 3.21*:* *A set of subsets* $F \subseteq 2^A$ *is called a filter if:*

1. $\emptyset \notin F$ *and* $F \neq \emptyset$.
2. $(K \in F \wedge K \subseteq B) \rightarrow B \in F$.
3. *For any* $K, K' \in F$, $K \cap K' \in F$.

F will represent the sets that are considered "majority" in A. For example, for an infinite A we may define

$$F_{co-fin} = \{K \subseteq A : A \setminus K \text{ is finite}\}$$

Let $(A, <)$ be a linear ordering. For any $x \in A$ be $K_x = \{y : x \leq y\}$, define:

$$F_{\leq} = \{B \subseteq A : \exists x K_x \subseteq B\}$$

For a non–countable A we may also define

$$F_{\aleph_0} = \{K \subseteq A : |A \setminus K| \leq \aleph_0\}$$

we may also define for any cardinality $\lambda \geq \aleph_0$ the filter

$$F_\lambda = \{K \subseteq A : |A \setminus K| \leq \lambda\}$$

And given a probability space (A, Pr) we may define the filter

$$F_{p=1} = \{K \subseteq A : \Pr(K) = 1\}$$

Following definitions from calculus, we use filters to define the following:

Definition 3.22: *Given a topology space (X, U) and a filter F over X we say that a point $y \in X$ is an accumulation point for F if for any open neighborhood b of y and for any $K \in F$, $K \cap b \neq \emptyset$.*

Definition 3.23: *The filter F converges to y if for* any *open neighborhood b of y it holds that $b \in F$.*

Note that if F_A is a filter on A where $A \subseteq X$ it has a natural extension to a filter on X by defining $F_X \triangleq \{K : K \cap A \in F_A\}$.

Those definitions coincide with the calculus definitions we've seen earlier for countable series of the reals when using the filter F_{co-fin}.

Theorem 3.17: *For a topology space X it holds that X is compact if and only if all filters F on X have an accumulation point.*

Proof: Recall that X is compact if and only if for any set S of closed sets of X, if for any finite $S' \subseteq S$ $\emptyset \neq \cap S'$ then $\emptyset \neq \cap S$.

Assume that X is compact. Given a filter F for X, define $S_F = \left\{\overline{K} : K \in F\right\}$. Note that if $y \in \bigcap S_F$ then y is an accumulation point of F (because if $y \in \bigcap S_F$ then for any $K \in F$ $y \in \overline{K}$ which means that for any neighborhood b of y, $b \cap K \neq \emptyset$). Therefore all that is left is to prove that $\cap S_F$ is not empty. Because of compactness, it is enough to prove that for any finite $S' \subseteq S_F$, $\bigcap S' \neq \emptyset$. The elements of S' are all in F (because each one of them is of the form $\overline{K}$ and $K \subseteq \overline{K}$ where $K \in F$ so $\overline{K} \in F$ as well). By property 3 of filters, $\cap S' \in F$ (because S' is finite). Therefore $\bigcap S' \neq \emptyset$.

For the second direction of the proof, assume that any filter F on X has an accumulation point, and let S denote a set of closed subsets of X for which any finite intersection is nonempty. Define the filter F_S:

$$F_S \triangleq \{K \subseteq X : \exists n, B_1, \ldots, B_n \text{ s.t. } B_1 \cap \ldots, \cap B_n \subseteq K\}$$

$F_S \neq \emptyset$ because any finite intersection of elements of S is nonempty. By the assumption, F_S has an accumulation point $y \in X$. We now show that $y \in \bigcap S$ and therefor $\bigcap S \neq \emptyset$. Given $B \in S$ we show that $y \in B$. $B \in F_S$ and so any neighborhood of y intersects B so $y \in \overline{B}$, but B is closed and so $y \in B$. □

We have seen various types of filters, including filters defined by cardinality (such as the filters F_{co-fin}, $F_{\aleph_0}$, etc.) and filters defined by ordering (the filter $F_{\leq}$). We may also define a filter based on a topology space:

$$F_{\text{cat}} \triangleq \{A \subseteq X : X \setminus A \text{ is a set of the first category}\}$$

In order for F_{cat} to be a filter, we have to make sure that X is not itself a set of the first category. Recall the definition: we say a set $B \subseteq X$ is nowhere dense if for any open set U there is $V \subseteq U$ that is open and $B \cap V = \emptyset$. A set is said to be of the first cateegory if it is a union of countably many sets that are nowhere dense.

Theorem 3.18: *Let X be a compact metric space, then X is not of the first category.*

Proof: Given a countable sequence $\{A_i\}_{i \in \boldsymbol{N}}$ of sets that are nowhere dense, we have to find an element $y \in X \setminus (\bigcup_i A_i)$. We construct a sequence of sets $\{K_i\}$ that will have a nonempty intersection to show the existence of such an element. Note that A is nowhere dense if an only if $\overline{A}$ does not contain any open set.

We take $K_0 = \overline{A_1}$. $X \setminus \overline{A_1}$ is open and nonempty, and so we may pick $x_0 \in X \setminus \overline{A_1}$, and because this is an open set, there must be a basic open neighborhood b of x_0 such that $b \subseteq X \setminus \overline{A_1}$. Let b be:

$$b = B_{(x_0, \epsilon)} = \{y : d(x_0, y) < \epsilon\}$$

From this it follows that $\overline{B_{(x_0, \frac{\epsilon}{2})}} \subseteq X \setminus \overline{A_1}$. Define $K_1 \triangleq \overline{B_{(x_0, \frac{\epsilon}{2})}}$. Similarly, we may take (immediate by induction) each $K_{i+1} \subseteq K_i$ closed and disjoint with $\overline{A_{i+1}}$.

From compactness we get that $\emptyset \neq \bigcap_{i \in \boldsymbol{N}} K_i$ (as any finite intersection of them is nonempty, and they are all closed sets). However, by definition $\forall i (\bigcap_{i \in \boldsymbol{N}} K_i) \cap A_i = \emptyset$ and since the intersection of the K_is is not empty it must be that $\bigcap A_i \neq X$. □

We now show that F_{cat} is different from all other filters we have seen so far. Actually, we will show a specific case: We show that in the standard topology over the unit interval $[0, 1]$ $F_{\text{cat}} \neq F_{\aleph_0}$. We do that by constructing a set K that is of the first category but not countable. This means that $[0, 1] \setminus K \in F_{\text{cat}}$ while $[0, 1] \setminus K \notin F_{\aleph_0}$.

We take K to be Kantor's Set. K is defined by using a sequence of subsets of the unit interval defined as follows:

$$\begin{aligned} K_0 = &\ [0, 1] \\ K_1 = &\ [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, 1] \\ K_2 = &\ [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{3}{9}] \cup [\tfrac{6}{9}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1] \\ &\ \vdots \end{aligned}$$

That is, in every step we remove the middle third of each interval we had in the previous step. We define Kantor's Set to be $K \triangleq \bigcap_{i \in \boldsymbol{N}} K_i$. It is immediate that K is nowhere dense, because any interval will have a subinterval missing. On the other hand $|K| = 2^{\aleph_0}$

Chapter 4

Measure Theory

4.1 Basic Definitions

We can view all the tools and techniques we have introduced so far as different ways to characterize the "size" of sets. We have seen the following tools that may be used for that purpose:

1. The cardinality of a set.
2. The topological properties of a set, or more specifically, saying that a set is "small" if it is of the first category.

We now wish to demonstrate another such tool. This tools is a generalization of the classic way of measuring areas and probability by integration.

What are the properties that we would like a "natural" measure function to have? Such a function $f : P(X) \mapsto \Re^+$ should probably satisfy the following:

1. f is monotone, so that $A \subseteq B \rightarrow f(A) \leq f(B)$.
2. For all sets A and B such that $A \cap B = \emptyset$, $f(A \cup B) = f(A) + f(B)$ (the triangle inequality).
3. $f(\emptyset) = 0$;

This intuitive "definition" is actually redundant:

Corollary 4.1: *In this definition* $2 \rightarrow 1$. *Note that for all* $A \subseteq B$, $B = A \cup (B \setminus A)$ *and thus* $f(B) = f(B \setminus A) + f(A) \geq f(A)$.

Corollary 4.2: *In this definition, also* $2 \rightarrow 3$. *For any* A, $f(A) = f(A \cup \emptyset) = f(A) + f(\emptyset)$ *and it must be that* $f(\emptyset) = 0$.

So it seems that our intuition boiled down to just the second requirement from the above list. We will now define a version of that property that is slightly more powerful than our original intuition:

Definition 4.1: *A function* $f : P(A) \mapsto \Re^+$ *is* σ-additive *if for any countable series* $\{A_i\}_{i=0}^{\infty}$, *if for all* $i \neq j$, $A_i \cap A_j = \emptyset$ *then*

$$f(\bigcup_{i \in N} A_i) = \sum_{i \in N} f(A_i)$$

Recall that for a sequence of real numbers $\{r_i\}_{i\in \boldsymbol{N}}$, the sum of the sequence is defined to be

$$\sum_{i\in \boldsymbol{N}} r_i \triangleq \lim_{i\to\infty} \sum_{n\le i} r_n$$

. We justify our requirements for σ-additive functions with the following claim:

Claim 4.1: *Let $f : P(X) \mapsto \Re^+$ be an additive function, f is σ-additive if and only if for any sequence $\{A_i\}_{i\in \boldsymbol{N}}$ where for all i it holds that $A_{i+1} \subseteq A_i$ and $\bigcap_{i\in \boldsymbol{N}} A_i = \emptyset$ it holds that* $\lim_{i\to\infty} f(A_i) = 0$.

Proof: Assume σ-additiveness of f. Note that all the sets $B_i = A_i \setminus A_{i+1}$ are pairwise disjoint sets, and for all n we have $A_n = \bigcup_{n\le i} B_i$. By the σ-additiveness of f we have $f(A_i) = \sum_{i\ge 1} f(B_i)$. But $f(A_1)$ is a finite number, and so the tail of the sequence $f(B_i)$ converges to 0. That is: $\lim_{n\to\infty}(\sum_{n\le i} f(B_i)) = 0$. We also have $\sum_{n\le i} f(B_i) = f(A_n)$ and from this it must be that the limit of $f(A_n)$ is 0.

For the other direction we have that the value of f on a telescopic sequence that converges to an empty set is 0 and we have to show that f is σ-additive. That is we have to show that given a set B_i of pairwise disjoint sets $f(\bigcup B_i) = \sum f(B_i)$. We then simply define $A_n \triangleq A \setminus \bigcup_{i\le n} B_i$ and show that the result follows from the assumption applied to these A_is. □

The following are all σ-additive functions:

1. $f \triangleq 0$
2. The counting measure:
$$f(A) = \begin{cases} |A| & A \text{ is finite} \\ \infty & \text{otherwise} \end{cases}$$
Note that in order for the last claim to remain true when we define the measure function to map some values to ∞, we need to add the following conditions to it: $f(A_1) < \infty$, $A_{i+1} \subseteq A_i$ and $\bigcap A_i = \emptyset$.
3. If we have a σ-additive function f over a domain $P(A)$ and $A \subseteq B$ we can define f' over $P(B)$ by setting $f'(K) \triangleq f(K \cap A)$.

We now try to demonstrate some of the difficulties of constructing "interesting" measure functions, and some of the ways to overcome these difficulties. For this, we first elaborate a little on the concept of filters introduced in the previous chapter. We define the following:

Definition 4.2: *A subset $F \subseteq 2^X$ is an* ultra–filter *over a set X if the following conditions hold:*

1. *F is a filter over X.*
2. *For any $A \subseteq X$ either $A \in F$ or $X \setminus A \in F$.*

Note that an ultra–filter is a maximal subset of 2^X that has the "finite– intersection" property (that is, the intersection of any finite number of elements of F is nonempty).

Definition 4.3: *A set $I \subseteq 2^X$ is called an* ideal *over X if the following holds:*

1. *$\emptyset \in I$, $X \notin I$.*

2. For any set $A \in I$ and a subset $B \subseteq A$ of it, $B \in I$.

3. For any two sets $A, B \in I$, $A \cup B \in I$.

This definition is, in a sense, dual to the definition of a filter.

Definition 4.4: *A set $I \subseteq 2^X$ is called a σ–*ideal *over X if:*

1. I is an ideal over X.

2. I is closed under unions of a countable number of sets.

Let F be a filter over X. Define:

$$f(A) \triangleq \begin{cases} 1 & A \in F \\ 0 & \text{otherwise} \end{cases}$$

Let us check that f is additive. Given $A, B \in F$, their intersection is nonempty, so the additiveness of f poses no condition on the values of $f(A)$ or $f(B)$. If $A \in F$ and $B \notin F$, since F is a filter it holds that $A \cup B \in F$ (since it contains A) and indeed it holds that $f(A \cup B) = 1 = 1 + 0 = f(A) + f(B)$. The problem arises when $A, B \notin F$.

One of the ways to overcome this problem is to require F to be an ultra–filter. If F is an ultra–filter and we have $A, B \notin F$ it follows that $X \setminus A, X \setminus B \in F$ and so $A \cap B \notin F$. Alternatively we may define a slightly different definition:

$$f'(A) \triangleq \begin{cases} 1 & A \in F \\ 0 & (X \setminus A) \in F \end{cases}$$

This f' is additive for any filter F. The problem here is that if F is not an ultra–filter then f' will be undefined for some subsets of X. Note that f is σ–additive if and only if $\{X \setminus A : A \in F\}$ is a σ–ideal.

Note that for this function to be both σ–additive and defined for any element in $P(X)$ we must have:

- F is an ultra–filter
- $\{X \setminus A : A \in F\}$ is a σ–ideal.

We leave as a simple exercise to the reader the proof of the following claim:

Claim 4.2: *If $f : P(X) \mapsto \{0, 1\}$ is σ–additive then it must be that $\{A : f(A) = 1\}$ is an ultra–filter and $\{A : f(A) = 0\}$ is a σ–ideal.*

One of the ways to construct such an ultra–filter is to pick any $x_0 \in X$ and define $F \triangleq \{A \subseteq X : x_0 \in A\}$. Such an F is called a principal ultra–filter. The question whether there exists a σ–ultra–filter that is not a principal ultra–filter is independent of set theory.

4.2 Measures of the Reals

We now ask ourselves what we would like a measure function over the reals to look like. The following two requirements seem to be intuitive:

1. For any interval (a, b), $f((a, b)) = b - a$.

2. For any real r and $A \subseteq \Re$, $f(A + r) = f(A)$ (where the set $A + r$ is defined as $A + r \triangleq \{x \in A : x - r \in A\}$).

Theorem 4.3 ***(Vitali):*** *There is no σ-additive function $f : P(\Re) \mapsto \Re^+$ that satisfies these last two requirements.*

For the proof of this theorem we need to introduce two more definitions. We will also restrict the proof to the interval $[0, 1]$.

Definition 4.5*:* *We define the following equivalence relation on $[0, 1]$: We say that $r \sim s$ if and only if $(r - s) \in \mathcal{Q}$.*

Definition 4.6*:* *A set $V \in [0, 1]$ is called a* Vitali Set *if and only if for any equivalence class $[s]$, $|V \cap [s]| = 1$.*

Proof: First, note that a Vitali set indeed exists (look at all the sets whose intersection with $[s]$ is of size at most 1, ordered under containment, and use Zorn's lemma).

Let V be a Vitali set. For any two rationals $q \neq q'$, $(V + q) \cap (V + q') = \emptyset$. Assume by contradiction that this is not the case, then there are $s, r \in V$ such that $x = s + q, x' = r + q'$ but then $r - s = q - q' \in \mathcal{Q}$ and so $[r] = [s]$, in contradiction with the definition of a Vitali set.

We now claim that given $y \in [0, 1]$ there is $s \in V$ such that $s \sim y$, as the set V has a representative from *every* equivalence class. Therefore (by the definition of the relation $\sim$), $y - s \in \mathcal{Q}$. Since y and s are both in the interval $[0, 1]$, $y - s \in [-1, 1]$ and since $y = s + (y - s)$ and so $y \in V + (y - s)$. Note that $\bigcup_{q \in [-1,1] \cap \mathcal{Q}} (V + q) \subseteq [-1, 2]$. Now assume that f is a measure function as stated in the theorem and observe $f(\bigcup_{q \in [-1,1]} (V + q))$. This is a countable union of disjoint sets and so, since f is σ–additive,

$$f\left(\bigcup_{q \in [-1,1] \cap \mathcal{Q}} (V + q)\right) = \sum_{q \in [-1,1] \cap \mathcal{Q}} f(V + q) = \sum_{q \in \mathcal{Q}} f(V) \leq 3$$

and therefore it must be that $f(V) = 0$. On the other hand, the set of all rational shifts of V covers the interval $[0, 1]$, and hence $\sum_{q \in [-1,1] \cap \mathcal{Q}} f(V + q) \geq 1$ - a contradiction. □

Corollary 4.4*:* *Any σ–additive function that is invariant to shifts and maps intervals to their lengths cannot be complete.*

We will now use a different strategy: instead of defining measure functions for all subsets of the reals we define specific subsets $B \subseteq \Re$ and limit ourselves to functions defined for sets in this B.

Definition 4.7*:* *B is called an* algebra *if B is closed under finite unions and set differences and $X, \emptyset \in B$. If B is also closed to countable unions we say that B is a σ–algebra.*

Theorem 4.5: *For any algebra $B \subseteq P(X)$ there exists a σ–algebra $\overline{B} \supseteq B$ that is minimal under containment among all the σ–algebras that contain B.*

One of the main uses for this last theorem is the following: Given a topology space (X, U), let Bor be the σ–algebra induced (i.e. the minimal σ–algebra that contains U). Its elements are Borell sets. It is easy to see that the cardinality of the set of all Borell sets on the real line is $2^{\aleph_0}$.

Definition 4.8: *A* measure space *is (X, B, μ) where X is the domain, $B \subseteq P(X)$ is a σ–algebra and $\mu : B \mapsto \Re^+ \cup \{\infty\}$ is a σ–additive function.*

A set $A \subseteq X$ is called μ–measurable *if $A \in B$.*

Theorem 4.6: *Given an algebra $\mathcal{A}$ of subsets of X and a σ–additive function $\mu : \mathcal{A} \mapsto \Re^+ \cup \{\infty\}$ there is an extension of μ to a σ–additive function $\overline{\mu}$ that is defined for all the elements of the induced σ–algebra $\overline{\mathcal{A}}$.*

We give here just the outline of the proof. Fiest we extend μ to a function μ^* defined over all the elements of $P(X)$:

$$\mu^*(E) \triangleq \inf \left\{ \sum_{i \in \mathbb{N}} \mu(A_i) : A_i \in \mathcal{A} \wedge E \subseteq \bigcup A_i \right\}$$

The problem with the function we've just defined is that it may no longer be σ–additive. The second stage of the proof indeed proves that over $\overline{\mathcal{A}}$, this μ^* is indeed σ–additive. For this purpose we pick a σ–algebra $B \supseteq \overline{\mathcal{A}}$ over which we will prove that μ^* is σ–additive. Note that this B will not necessary be maximal under containment.

For example, we may pick the following B:

$$B_1 = \{E \subseteq X : \forall k \subseteq X, \mu^*(k) = \mu^*(k \cap E) + \mu^*(k \setminus E)\}$$

We will concern ourselves with the special case where $\mathcal{A}$ is the algebra induced by the set of open sets in a topology space over X. We will denote this B_2 by L (for Lesbegue). A set $E \subseteq X$ is in L if for any $\epsilon > 0$ there is an open set F and a closed set G where $F \subseteq E \subseteq G$ and $\mu(G \setminus F) < \epsilon$. Such an $E \in L$ is called a Lesbegue set relative to (X, U) (if the topology space is omitted, it is meant to refer to open ended sphere topology over $\Re^n$).

We would like this L to have the following properties:

1. L should be a σ–algebra (i.e. closed to countable unions and set differences).
2. L should contain $\mathcal{A}$ (together with the first property this also implies that $L \supseteq \overline{\mathcal{A}}$).
3. μ^* should be σ–additive over L.
4. μ^* should equal μ for all elements in $\mathcal{A}$.

The proof that L indeed have these properties is omitted.

Definition 4.9: *Given a function μ that is σ–additive over an algebra $\mathcal{A}$, a set $E \subseteq X$ will be called a* zero set *if $\mu^*(E) = 0$ (or, in other words, for any $\epsilon > 0$ there exists a sequence of sets $k_1, k_2, \ldots \in \mathcal{A}$ such that $E \subseteq \bigcup k_i$ and $\sum_i \mu(k_i) < \epsilon$).*

Note that in both definitions of measurable sets we introduced (i.e. B_1 and L), every zero set is measurable. For example, in the definition of B_1, if E is a zero set then for any k $\mu^*(k) = \mu^*(k \cap E) + \mu^*(k \setminus E)$ because it is always true that $\mu^*(k) \le \mu^*(k \cap E) + \mu^*(k \setminus E)$ but $\mu^*(k \cap E) = 0$ and $\mu^*(k \setminus E) \le \mu^*(k)$ and therefor $\mu^*(k) = \mu^*(k \cap E) + \mu^*(k \setminus E)$.

Claim 4.3: *Denote by L_0 the set of all zero sets in L. If $X = \Re$ and $B = Bor$, $L_0 \nsubseteq B$ (with μ induced by lengths of segments).*

Proof: Recall that $|\text{Bor}| = 2^{\aleph_0}$. We will show that $|L_0| = 2^{2^{\aleph_0}}$. In order to show this it is enough to find one zero set of cardinality $2^{\aleph_0}$, because all its subsets must also be zero sets. Kantor's set is such a zero set, as it is the intersection of decreasing sets $\mu(k_0) = 1, \mu(k_1) = \frac{2}{3}, \dots, \mu(k_n) = \frac{2^n}{3^n}, \dots$ and so it can be covered by arbitrarily small unions of open sets. □

Let us recall our motivation of trying to define what "small" sets are. We've just added another criterion to the two we had before (countable sets and sets of the first category): Sets of measure zero - L_0. We list a few immediate properties of sets in L_0:

1. $|L_0| = 2^{2^{\aleph_0}}$.
2. There are sets in L_0 that are not countable (i.e. $L_0 \neq C_0$).
3. Every countable set has measure 0, or $C_0 \subset L_0$.

To prove the last property above we note the following: Given a countable set $E \subset \Re$, we pick an enumerating function $f : I\!N \overset{1-1}{\longmapsto} E$ onto it. Given ϵ we define for any n:

$$K_n^\epsilon = \left(f(n) - \frac{\epsilon}{2^{n+2}}, f(n) + \frac{\epsilon}{2^{n+2}}\right)$$

For any n, $f(n) \in K_n$ and so $E \subseteq \bigcup_{n \in I\!N} K_n$ while for any ϵ, $\sum_{n \in I\!N} \mu(K_n^\epsilon) \le \epsilon$, so E itself must have zero measure.

4.3 Relation to Topology and the Duality Theorem

Claim 4.3 shows that L_0 actually extends our previous notion of smallness induced by the cardinality of sets. The following claim indicates that this is not the case for the notion of smallness induced by sets of the first category.

Claim 4.4: *The interval $[0,1]$ can be covered by two disjoint subsets, one is of the first category, and the other is of measure zero.*

Proof: We describe the construction of such sets. As in the proof of property 3 above, define, for any natural number i, the set K_n^i to be K_n^ϵ defined above where the countable set is taken to be $\mathcal{Q} \cap [0,1]$ and $\epsilon = \frac{1}{i}$.

Now we define $E \triangleq \bigcap_i \bigcup_n K_n^i$. Since for all n and i, $f(n) \in K_n^i$ we have for all n, $f(n) \in \bigcap_i K_n^i$ and therefor $[0,1] \cap \mathcal{Q} \subseteq E$. Thus E has measure zero.

Now we claim that $[0,1] \setminus E$ is of the first category. Note that for all i, $[0,1] \setminus \bigcup_n K_n^i$ is nowhere dense (since $\bigcup_i K_n^i$ is an open set and contains $\mathcal{Q}$, so it contains an interval around any rational number, and its complement will have a "hole" within any open interval). But $[0,1] \setminus E = \bigcup_i ([0,1] \setminus \bigcup_n K_n^i)$ so it is the union of a countable number of sets that are N.D. and so it is a set of the first category. □

Let us summarize the construction of the Lesbeuge measure on the reals:

1. We defined μ over the algebra defined on segments by their length.

2. We extend μ to $\overline{\mu}$, defined over Borell sets.

3. We get a measure μ defined over the larger set L.

It can be shown that $L = \{\text{closure of } B \cup L_0 \text{ under } \sigma\text{–algebra}\}$. Moreover, every measurable set E is of the form $E = G \cup N$ where $G \in G_\delta$ and N is of measure zero.

Analogously, a similar definition in topological terms may be formulated:

Definition 4.10: *A set E is called a* Baire set *if there exists an open set U and a set of the first category P such that $E = U \triangle P$.*

Claim 4.5: *A set E is a Baire set if and only if there exists a set $G \in G_\delta$ and a set of the first category P such that $E = G \cup P$.*

Proof: Note that the closure of a N.D. set is still N.D. In fact, if $P = \bigcup A_i$ (where each A_i is N.D.), we are interested in $P' = \bigcup \overline{A_i}$. Note that $P' \in F_\sigma$ (a countable union of closed sets) and $P \subseteq P'$. We can thus express E as $E = (U \setminus P') \cup K$ where $K \subseteq P'$ is a set of the first category and $U \setminus P' \in G_\delta$.

The other direction follows from the claim that a Baire set is a σ–algebra and it is the σ–algebra that is created from Borell sets and sets of the first category (i.e. is based on sets of the form $U \cup C_I$). □

Theorem 4.7 *(Shifting):* *If A is a measurable set and $\mu(A) > 0$ then there exists $\delta > 0$ such that for any ϵ where $|\epsilon| < \delta$, $(A + \epsilon) \cap A \neq \emptyset$.*

Note that from this theorem it immediately follows that Vitali sets is not measurable as for any $q \in \mathcal{Q}$, $(V + q) \cap V = \emptyset$.

We show a similar theorem for topology (concerning Baire sets):

Theorem 4.8: *For any Baire set $A \subseteq \Re$ that is not of the first category there exists $\delta > 0$ such that for any x that satisfies $|x| < \delta$, $(A + x) \cap A \neq \emptyset$.*

Proof: A is a Baire set and therefor $A = U \triangle N$ where U is an open set and N is of the first category. It must be that $U \neq \emptyset$ or else A was of the first category itself. Pick a nonempty open interval I such that $I \subseteq U$, and set $\delta = \frac{\mu(I)}{2}$. We now have:

$$I \cap (I + x) \setminus [N \cap (N + x)] \subseteq A \cap (A + x)$$

But since $I \cap (I + x)$ is an interval and $N \cap (N + x)$ is a set of the first category, this is nonempty (a set of the first category cannot cover an open interval). □

Theorem 4.9 *(Bernstein):* *There exists a set $B_e \subseteq [0, 1]$ such that:*

1. *For any closed set $F \subseteq B_e$, $\mu(F) = 0$ (under the Lesbeuge measure).*

2. *For any closed set $F \subseteq [0, 1] \setminus B_e$, $\mu(F) = 0$.*

3. *For any Baire set $E \subseteq B_e$, E is a set of the first category.*

4. *For any Baire set $E \subseteq [0, 1] \setminus B_e$, E is a set of the first category.*

Proof: We will assume the Continuum Hypothesis for the sake of this proof (i.e., we assume $\aleph_1 = 2^{\aleph_0}$).

First we show that $|\{F \subseteq [0,1] : \text{ F is closed}\}| = 2^{\aleph_0}$. We have already seen that there are only $2^{\aleph_0}$ open sets, as any open set is made up of the union of intervals with rational end points, and there are only $\aleph_0$ such intervals. The mapping $H(U) \triangleq [0,1] \setminus U$ is one on one from the set of open sets onto the set of closed sets, and so there are only $2^{\aleph_0}$ closed sets as well.

By the well–ordering theorem (and the Continuum Hypothesis), there exists a well–ordering of the set $\{F : F \text{ is closed}\}$ where every initial segment of the ordering is of cardinality $\aleph_0$. Let $\{F_\alpha : \alpha < 2^{\aleph_0}\}$ be such a list of the closed sets. We build by induction on this ordering (which is actually ω_1) a series of pairs of points $\{(a_\alpha, b_\alpha)\}$: At the α stage we have so far have defined the series $\{(a_\beta, b_\beta) : \beta < \alpha\}$ which is a set of cardinality strictly less than $2^{\aleph_0}$. Now we examine the set F_α. If $|F_\alpha| < 2^{\aleph_0}$ we do nothing. Otherwise,

$$|F_\alpha \setminus \{a_\beta, b_\beta : \beta < \alpha\}| = 2^{\aleph_0}$$

and we can pick two arbitrary points a_α, b_α in this set.

Now define B_e to be

$$B_e \triangleq \{b_\alpha : \alpha < 2^{\aleph_0}\}$$

Note that so far we have not yet taken advantage of the Continuum Hypothesis (as we could have repeated the same argument with $\aleph_1$ instead of $2^{\aleph_0}$). However, the Continuum Hypothesis is what gives the first two required properties of B_e because for any closed set F of cardinality $2^{\aleph_0}$, $B_e \cap F \neq \emptyset$ because F appeared in the series we have defined at a certain point α and the corresponding b_α is in B_e. Because of a similar argument, for every closed F of cardinality $2^{\aleph_0}$, $(F \setminus B_e) \neq \emptyset$ because the corresponding a_α from F is not in B_e.

To show that for any Baire set E such that $E \subseteq B_e$ or $E \subseteq [0,1] \setminus B_e$, E is of the first category, we use the following claim (whose proof is left as an exercise):

Every set G_δ of cardinality $2^{\aleph_0}$ has a closed subset of cardinality $2^{\aleph_0}$.

Now, since for every Baire set E we have $E = G \cup N$ where $G \in G_\delta$ then if G is not of the first category, it is not countable. Therefor it is a countable union of sets that are not all countable and the Baire set has a closed subset that is not countable. Therefor E cannot be contained in B_e. □

Given an irrational number $r \in \Re \setminus \mathcal{Q}$ the question is "how well" r can be approximated by rational numbers.

Theorem 4.10: *If r is a root of a polynomial of degree n with integer coefficients then there exists $M > 1$ such that for any $p, q \in \mathbb{Z}$, $\left|r - \frac{p}{q}\right| \geq M \cdot \frac{1}{q^n}$.*

Definition 4.11: *A real number r is called a* Liouville number *if for every n there exists p, q such that $\left|r - \frac{p}{q}\right| < \frac{1}{q^n}$ and r is irrational.*

As a corollary from the last theorem we have:

Corollary 4.11: *Liouville numbers are transcendental (non algebraic numbers).*

A Liouville number can be easily constructed. Start with "0.1" and at the i'th stage, assuming the last non-zero digit was the k'th digit, add $i \cdot k$ zeros followed by a one. This construction also demonstrates that there are exactly $2^{\aleph_0}$ Liouville numbers as we could have replaced some of the 1s in the above construction with 2s.

Claim 4.6: *The set of all numbers that are* not *Liouville numbers is of the first category.*

Proof: Denote by G_n the set of numbers r that satisfy that for any $p, q \in \mathbb{Z}$, $\left|r - \frac{p}{q}\right| > \frac{1}{q^n}$. The set of numbers that are not Liouville numbers is exactly $\bigcup_{n\in\mathbb{N}} G_n$. On the other hand every G_n is nowhere dense, as it does not contain all the intervals of the form $(\frac{p}{q} - \frac{1}{q^n}, \frac{p}{q} + \frac{1}{q^n})$, so G_n is missing an interval around every rational number. Therefore the set of numbers that are not Liouville numbers is a countable union of sets that are nowhere dense, and hence, is a set of the first category. □

It is left as an exercise to verify that the set of Liouville numbers has zero measure.

Theorem 4.12 ***(Duality Theorem of Erdös and Serpinsky):*** *If P is a claim that is formulated in terms of "Set of the first category", "Set of measure zero" and cardinalities on sets of real numbers and P^* is the claim formed from P by replacing every instance of "Set of the first category" by "Set of measure zero" and vice versa then P holds if and only if P^* holds.*

The duality theorem's proof is immediate from the following lemma:

Lemma 4.1: *There exists a function $f : \Re \overset{1-1}{\longmapsto} \Re$ that is onto $\Re$ and for every $E \subseteq \Re$:*

1. *$\mu(E) = 0$ if and only if $f(E)$ is of the first category.*
2. *E is of the first category if and only if $\mu(f(E)) = 0$.*

To present the proof of this lemma we will first have to establish some background:

Definition 4.12: *A σ–ideal I over a set X will be called* well–behaved *if the following holds:*

1. *There exists $G \subseteq I$ such that $|G| = 2^{\aleph_0}$ and for all $A \in I$ there exists $B \in G$ such that $A \subseteq B$.*
2. *$\bigcup I = X$.*
3. *For all $A \in I$ there exists $B \subseteq X \setminus A$ such that $B \in I$ and $|B| = 2^{\aleph_0}$.*

We can easily verify that both L_0 and C_I are well–behaved σ–ideals. The first property holds for L_0 because for every A of measure 0 there's a set $B \in G_\delta$ such that $A \subseteq B$ and $\mu(B) = 0$, and $|G_\delta| = 2^{\aleph_0}$. The same holds for C_I because for any set A of the first category $A = \bigcup_{i\in\mathbb{N}} k_i$ where every k_i is N.D., in which case, as we've already seen, $\overline{k_i}$ is also N.D. and $A \subseteq \bigcup_{i\in\mathbb{N}} \overline{k_i}$ but $\bigcup_{i\in\mathbb{N}} \overline{k_i} \in F_\sigma$, while $|F_\sigma| = 2^{\aleph_0}$.

Claim 4.7: *Assuming the Continuum Hypothesis, if I is a well–behaved ideal then there is a decomposition of X into pairwise disjoint sets: $X = \bigcup_{\alpha<\omega_1} X_\alpha$ such that for all $A \subseteq X$, $A \in I$ if and only if there exists $\beta < \omega_1$ such that $A \subseteq \bigcup_{\alpha<\beta} X_\alpha$, or in other words, A is contained in a countable union of X_αs.*

Note that from this last claim it immediately follows that for all α $X_\alpha \in I$ (as $X_\alpha \subseteq \bigcup_{i\leq\alpha+1} X_i$).

Proof: Let $h : G \overset{1-1}{\longmapsto} \omega_1$ onto ω_1. Denote by G_α the set $h^{-1}(\alpha)$. We claim that $X = \bigcup_{\alpha\in\omega_1} G_\alpha$. Define:

$$H_\alpha \triangleq \bigcup_{\beta\leq\alpha} G_\beta \setminus \bigcup_{\beta<\alpha} G_\beta$$

We first show that $|\{\alpha : |H_\alpha| = \aleph_1\}| = \aleph_1$. If this was not true then there must have been a $\beta < \aleph_1$ such that

$$\{\alpha : |H_\alpha| = \aleph_1\} \subseteq \{\alpha : \alpha < \beta\}$$

because every countable subset of ω_1 is bounded.
$\bigcup_{\alpha \leq \beta} H_\alpha \in I$, and therefor, by the definition of well–behaved ideals there exists

$$A \subseteq X \setminus \bigcup_{\alpha \leq \beta} H_\alpha$$

and $A \in I$, $|A| = \aleph_1$. Pick a value γ such that $A \subseteq G_\gamma$ (property 1 of well–behaved ideals guarantees us the existence of such a set). This implies that A is a set of cardinality $\aleph_1$ covered by a countable number of countable sets. This means that our claim $|\{\alpha : |H_\alpha| = \aleph_1\}| = \aleph_1$ is indeed true. $\square$

Now we go on to define:

$$s : \omega_1 \overset{1-1}{\longmapsto} \{\alpha : |H_\alpha| = \aleph_1\}$$

and also:

$$X_\alpha \triangleq \bigcup_{\beta \leq s(\alpha)} G_\beta \setminus \bigcup_{\beta < s(\alpha)} G_\beta$$

We now show that $X = \bigcup X_\alpha$ where X_α are pairwise disjoint and each X_α has cardinality $\aleph_1$ and $A \in I$ if and only if $A \subseteq \bigcup_{n \in \boldsymbol{N}} X_n$.

Claim 4.8: *Let K and L be two well–behaved σ–ideals such that there is a decomposition of $\Re$, $\Re = A \cup B$, $A \cap B = \emptyset$ where $A \in K$ and $B \in L$, then there exists a function $f : \Re \overset{1-1}{\longmapsto} \Re$ onto, and $f^{-1} = f$ such that for all $E \subseteq \Re$, $E \in K$ if and only if $f(E) \in L$.*

Proof: Let X_α be the decomposition of the set X for the ideal K that was specified in claim 4.7, where X_0 is set to A (it is immediate from the construction in the proof of claim 4.7 that such a decomposition can indeed be constructed). Similarly, let Y_α be a decomposition of the Y constructed for the ideal L where $Y_0 = B$.

For all α let $f_\alpha : X_\alpha \overset{1-1}{\longmapsto} Y_\alpha$ onto and define the function f required in the statement of the claim to be:

$$f(x) \triangleq \begin{cases} f_\alpha(x) & \text{there exists } \alpha > 0 \text{ s.t. } x \in X_\alpha \\ f_\alpha^{-1}(x) & x \in X_0 \text{ or, in other words, there exists } \alpha > 0 \text{ s.t. } x \in Y_\alpha \end{cases}$$

Note that the constrcution is well defined, as X_0 and Y_0 form a decomposition of $\Re$, and so does X_α and $Y\alpha$, so that if $x \notin X_0$ it must be in Y_0 and vice versa.

It is now easy to see that this function f we've just constructed satisfies the conditions required in the claim's statement. Given $E \in K$, there exists $\beta < \omega_1$ such that $E \subseteq \bigcup_{\alpha<\beta} X_\alpha$ and thus $f(E) \subseteq \bigcup_{\alpha<\beta} Y_\alpha$ and so $f(E) \in L$. Given $f(E) \in L$ it holds that $f(E) \subseteq \bigcup_{\alpha<\beta} Y_\alpha$ which means that $E \subseteq \bigcup_{\alpha<\beta} X_\alpha$ and so $E \in K$. $\square$