General Topology

Pete L. Clark

Thanks to Mikhail Botchkarev, Kaj Hansen, Scott Higinbotham, Ishfaaq Imtiyas, William Krinsman and Maddie Locus for pointing out typos in these notes. Thanks to Alex Eldridge for helpful discussions related to Example 3.21. I am grateful to David Galban for informing me about the relevance of ultrafilters to social choice theory.

Contents

Preface

1. Why Another General Topology Textbook?

Apart from linear algebra it is hard to think of a subject that plays a broader foundational role in theoretical mathematics than general topology: most mathematicians use the concepts of topological space, continuous function and convergence regularly in their work. However the shrift given to general topology in contemporary American university curricula is rather short: typically only a single course is offered, usually the advanced undergraduate / very basic graduate level.

Textbooks for advanced undergraduate level mathematics classes seem to have long half-lives compared to other academic texts. The first edition of the most widely used undergraduate analysis textbook [**Rud**] was published in 1953; its third (and last) edition was published in 1976. The most popular general topology text for American students is [**Mu**], for which the first edition was published in 1974. This was my topology textbook in 1996, and the second (and last) edition was published in 2000. Other general topology textbooks in contemporary use include a 1966 work of Dugundji [**Du**], a 1977 work of Engelking [**En**], a 1955 text of Kelley [**Ke**], and a 1970 text of Willard [**Wi**].

Each of the above texts is, in fact, excellent. Moreover, though general topology remains active as a research field, it would be difficult or impossible to incorporate contemporary research developments into an introductory course. So one should ask: why not just continue making use of these excellent texts of yore? What would or could be gained from using a more contemporary text?

Here are some possible answers:

• None of the above texts are freely and legitimately available online.

This is a practical point, but an important one. I write this introduction in July 2020, in the middle of the COVID-19 pandemic and preparing for a Fall 2020 topology course at UGA. Having online course materials and reducing economic burdens on students has never seemed more important than right now – but this was important before the pandemic and (I sure hope!) will continue to be important after it. There is simply no need for anyone to pay hundreds of dollars for textbooks – or dozens of dollars, or dozens of cents – when these resources can be made freely available to all. You can't access food, clothing and shelter freely online, but you can access mathematics texts.

• Although recent research in general topology has had (and presumably should have) little effect on the teaching of introductory courses, in recent years many expository papers in general topology have been published. Some of these papers give alternate and/or simpler approaches to core material. Examples include

[**Bl19**], [**Ch92**], [**CHN74**], [**Cl19**], [**DF92**], [**Gr12**], [**Ha11**], [**Ka07**], [**Ma93**], [**Mi70**], [**NP88**], [**Ro76**], [**Sh72**], [**Sh74**], [**Ul03**].

Others introduce or go into more depth on topics that are accessible to undergraduate topology students. Examples include

[**Az18**], [**BC10**], [**Be98**], [**Be06**], [**Cl73**], [**CdC**], [**Di05**], [**DS14**], [**Do00**], [**DS84**], [**Ed62**], [**El09**], [**Fi91**], [**Fu55**], [**Ga71**], [**GMGS11**], [**GS79**], [**Go59**], [**Ha78**], [**HM09**], [**Hu81**], [**JH90**], [**Ki99**], [**Ko88**], [**Ma88**], [**NZ07**], [**Os14**], [**Pe90**], [**Si08**], [**Ve97**].

A contemporary text can take advantage of the advances, simplifications and fresh ideas from (some of) these articles.

• A contemporary text can respond to changes in the undergraduate mathematics curriculum and/or the needs of a particular cohort of students.

One recent curricular change is the reduction of undergraduate analysis courses. When I was an undergradate at Chicago in the 1990's, I took three undergraduate analysis courses before I took a topology course. All of these made use of metric spaces. At UGA there is a required undergraduate course on real sequences and series, which therefore gives a kind of introduction to the topology of $\mathbb{R}$, and no undergraduate analysis course is prerequisite to topology and indeed only one such course is offered. Thus a UGA student beginning a topology course may well have seen the definition of a metric space but probably will not have studied such spaces in much generality or depth: for instance, one should not assume mastery of completeness and compactness in metric spaces. In [**Mu**], metric spaces do not appear until near the end of the first long chapter on topological spaces. In [**Ke**] metric spaces do not appear until the discussion of metrization theorems, almost halfway through. In [**Wi**] metric spaces appear in Chapter 1 but only for eight pages, most of which consists of (very nice, but not very introductory) exercises.

To be honest about it, for many mathematicians who use topology, most or all of the spaces they do business with are, if not metric, then at least metrizable – that is, the topology is induced from a metric, though not necessarily in a canonical way. I am not one of these mathematicians: much of my work involves algebraic geometry, and in that subject some of the spaces one works with do not even satisfy the Hausdorff separation axiom. However, from a pedagogical perspective it must be better to do too much with metric spaces than too little: metric spaces provide valuable topological intuition, whereas contemplating a topological space without any separation axioms is more likely to develop an appreciation for pathology. Though in a somewhat tautological sense metric space theory is not *quite* general topology, nevertheless it is a very important subject and it uses and applications are similar to those of much of general topology.

2. Some Features of this Text

Let me now explain some features of this text, in a way that responds to the issues raised in the previous section.

First, a **bug**: this text is as yet unfinished. There is more material here than could be covered in a one semester course in general topology, but when one compares to the classic texts like [**Ke**], [**Mu**], [**Wi**] mentioned above, one finds that they go on to develop significantly more material. Here are some standard topics that are missing here:

- Paracompactness
- General metrization theorems (necessary and sufficient criteria for metrizability)
- Uniform spaces

The absence of paracompactness is surely the most distressing, as this is a general topological concept that plays an important role in most other parts of topology (e.g. differential and algebraic). This ought to be remedied in the future.

The lack of treatment of metrization theorems beyond Urysohn somewhat scopes and dates the text: the works of Nagata, Smirnoff and Bing, each giving general metrization theorems, were published in 1950 and 1951. In so doing they resolved what had been the outstanding problem in general topology for about half a century. It could in fact be argued that these "book-closing results" contributed to the *decline* of general topology as a research field in the ensuing seventy-something years; since then, the field has lacked outsanding, unifying open problems. This is not a text on general topology *for prospective general topologists*: this is a text on the general topology that a mathematician in another field may either need or want to know. General metrization theorems seem not to be *needed* outisde of general topology itself. They may well be a topic that other mathematicians *want* to know, but the combination of lack of applicability together with my own lack of insight into these topics has not impelled me to include coverage of them here.

When it comes to coverage of uniform spaces, I am full of ambivalence. The identificaton of a structure intermediate between metric structure and topological structure suitable for discussions of uniform continuity, completeness and so forth sunds like an absolutely fundamental discovery. Uniform structures are a key part of the topological vocabulary of several generations of French mathematics thanks to the central role they play in Bourbaki's treatment, and at the beginning of my own mathematical career I often read mathematical texts and papers that seemed to require a knowledge of uniform spaces. However, I know of few instances where uniform structures are *crucially used* but many theorems about metric spaces that can be generalized to uniform spaces in a rather routine way but at the cost of adding a layer of technicality to the exposition. It would be nice to include uniform spaces *with a sufficient payoff.*

This brings me to two topics that are *not* standard in general topology textbooks but whose absence here is nevertheless regretted:

- Topological groups
- Categorical concepts in topology

Topological groups form a significant part of the foundational material in topology that is used in many other mathematical fields. It would be lovely to cover Pontrjagin duality, which is a significant application of the quasi-compact open topology. Indeed the fact that topological groups carry canonical uniform structures is a better reason to discuss uniform spaces than any other one that I know. I hope that this material will appear at some point in the future.

Just about every standard general topology textbook I know assumes that the reader knows a lot of set theory and does not assume that the reader knows (or cares to know) any category theory whatsoever. For instance many topological texts seem to want the reader to either know or quickly learn about ordinal numbers, whereas in fact they have quite a minor role to play. However even the basic idea that topological concepts should systematically be explored in a functorial way is absent from standard texts — it is interesting that Bourbaki, with its discussion of initial and final topologies, takes one key step down this road and then stops, whereas the material of filters and uniform spaces are, honestly, quite *over*-developed. (Filters and uniform spaces were both invented in 1937, by Cartan and Weil respectively, just as [**Bo**] was first being written. The main ideas of category theory were invented by Eilenberg and Mac Lane in the early 1940's. Although the Bourbaki books continued to be written and significantly revised for decades, they never revisited their basic proto-categorical approach.) To this day one can see that researchers working in general topology seem to prefer sets over categories...but this is an odd rivalry since there is absolutely no need to choose one over the other. I have some material on categorical concepts in topology that is slowly being written and hope to release in a future version.

Moving on to the features:

• We begin with **real induction**, which is a relative of mathematical induction that applies to intervals on the real line. This material is based on the article [**Cl19**], which itself has a lot of precedents in the literature. Our goal here is to give an introduction to basic topological concepts like connectedness and compactness on the real line in a way that is light and quick but also (I hope) fresh enough to pique interest.

Real induction generalizes to topologies on ordered spaces, for which we give a quick introduction.

We then turn to metric spaces. For the reasons alluded to above (and because we like the material), we give a considerably more substantial treatment than in most other topology texts. Some highlights:

• In §2.6 we give a substantial treatment of product metrics, including Theorem 2.36 that characterizes "good metrics" on a product of metric spaces.

• In §2.12 we explore the relationship between uniform continuity and extension of maps.

• In §2.16 we cover the Contraction Mapping Theorem and related fixed point theorems in certain metric spaces.

(A warning: in a one semester course, in order to get on to general topological spaces, one should omit parts of Chapter 2.)

• In §3.6 we discuss the coproduct (a.k.a. direct sum, a.k.a. disjoint union) topology. This simple and useful construction is missing from most of the older texts (though perfectly understood by those who wrote them). From a modern categorical perspective it is really essential to include this construction.

(• We should also include a discussion of direct and inverse limits of topological spaces. Unfortunately, as of February 2021 this has not been written.)

• In §3.10 we discuss initial and final topologies. This is another absolutely basic construction that is somehow missing from [**Ke**] and [**Mu**], the most standard American texts, and it is another example of building categorical thinking in the subject before categories are explicitly introduced.

CHAPTER 1

Introduction to Topology

1. Introduction to Real Induction

1.1. Real Induction.

Consider for a moment "conventional" mathematical induction. To use it, one thinks in terms of *predicates* – i.e., statements $P(n)$ indexed by the natural numbers – but the cleanest enunciation comes from thinking in terms of *subsets* of $\mathbb{N}$. The same goes for real induction.

Let $a < b$ be real numbers. We define a subset $S \subset [a, b]$ to be **inductive** if:

(RI1) $a \in S$.
(RI2) If $a \leq x < b$, then $x \in S \implies [x, y] \subset S$ for some $y > x$.
(RI3) If $a < x \leq b$ and $[a, x) \subset S$, then $x \in S$.

THEOREM 1.1. *(Real Induction) For $S \subset [a, b]$, the following are equivalent:*
(i) S is inductive.
(ii) $S = [a, b]$.

PROOF. (i) $\implies$ (ii): let $S \subset [a, b]$ be inductive. Seeking a contradiction, suppose $S' = [a, b] \setminus S$ is nonempty, so $\inf S'$ exists and is finite.
Case 1: $\inf S' = a$. Then by (RI1), $a \in S$, so by (RI2), there exists $y > a$ such that $[a, y] \subset S$, and thus y is a greater lower bound for S' then $a = \inf S'$: contradiction.
Case 2: $a < \inf S' \in S$. If $\inf S' = b$, then $S = [a, b]$. Otherwise, by (RI2) there exists $y > \inf S'$ such that $[\inf S', y] \subset S$, contradicting the definition of $\inf S'$.
Case 3: $a < \inf S' \in S'$. Then $[a, \inf S') \subset S$, so by (RI3) $\inf S' \in S$: contradiction!
(ii) $\implies$ (i) is immediate. □

Theorem 1.1 is due to D. Hathaway [**Ha11**] and, independently, to me. But mathematically equivalent ideas have been around in the literature for a long time: see [**Ch19**], [**Kh23**], [**Pe26**], [**Kh49**], [**Du57**], [**Fo57**], [**MR68**], [**Sh72**], [**Be82**], [**Le82**], [**Sa84**], [**Do03**]. Especially, I acknowledge my indebtedness to a work of Kalantari [**Ka07**]. I read this paper early in the morning of Tuesday, September 7, 2010 and found it fascinating. Kalantari's formulation works with subsets $S \subset [a, b)$, replaces (RI2) and (RI3) by the single axiom

(RIK) For $x \in [a, b)$, if $[a, x) \subset S$, then there exists $y > x$ with $[a, y) \subset S$,[1]

and the conclusion is that a subset $S \subset [a, b)$ satisfying (RI1) and (RIK) must be equal to $[a, b)$. Unfortunately I was a bit confused by Kalantari's formulation, and

[1] One also needs the convention $[x, x) = \{x\}$ here.

I wrote to Professor Kalantari suggesting the "fix" of replacing (RIK) with (RI2) and (RI3). He wrote back later that morning to set me straight. I was scheduled to give a general interest talk for graduate students in the early afternoon, and I had planned to speak about binary quadratic forms. But I found real induction to be too intriguing to put down, and my talk at 2 pm that day was on real induction (in the formulation of Theorem 1.1). This was, perhaps, the best received non-research lecture I have ever given, and I was motivated to develop these ideas in more detail.

In 2011 D. Hathaway published a short note "Using Continuity Induction" [**Ha11**] giving an all but identical formulation: instead of (RI2), he takes

(RI2H) If $a \leq x < b$, then $x \in S \implies [x, x+\delta) \subset S$ for some $\delta > 0$.

(RI2) and (RI2H) are equivalent: $[x, x + \frac{\delta}{2}] \subset [x, x+\delta) \subset [x, x+\delta]$. Hathaway and I arrived at our formulations completely independently. Moreover, when first formulating real induction I too used (RI2H), but soon changed it to (RI2) with an eye to a certain more general inductive principle that we will meet later.

2. Real Induction in Calculus

We begin with the "interval theorems" from honors (i.e., theoretical) calculus: these fundamental results all begin the same way: "Let $f : [a,b] \to \mathbb{R}$ be a continuous function." Then they assert four different conclusions. One of these conclusions is truly analytic in character, but the other three are really the source of all topology.

To be sure, let's begin with the definition of a continuous real-valued function $f : I \to \mathbb{R}$ defined on a subinterval of $\mathbb{R}$: let x be a point of I. Then f is **continuous at x** if for all $\epsilon > 0$, there is $\delta > 0$ such that for all $y \in I$, if $|x - y| \leq \delta$ then $|f(x) - f(y)| \leq \epsilon$. f is **continuous** if it is continuous at every point of I.

Let us also record the following definition: $f : I \to \mathbb{R}$ is **uniformly continuous** if for all $\epsilon > 0$, there is $\delta > 0$ such that for all $y \in I$, if $|x - y| \leq \delta$ then $|f(x) - f(y)| \leq \epsilon$. Note that this is stronger than continuity in a rather subtle way: the only difference is that in the definition of continuity, the δ is allowed to depend on ϵ but also on the point x; in uniform continuity, δ is only allowed to depend on ϵ: there must be one δ which works simultaneously for all $x \in I$.

EXERCISE 1.1. *a) Show – from scratch – that each of the following functions is continuous but not uniformly continuous.*
(i) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$.
(ii) $g : (0,1) \to \mathbb{R}$, $f(x) = \frac{1}{x}$.
b) Recall that a subset $S \subset \mathbb{R}$ *is* ***bounded*** *if* $S \subset [-M, M]$ *for some* $M \geq 0$. *Show that if* I *is a bounded interval and* $f : I \to \mathbb{R}$ *is uniformly continuous, then* f *is bounded (i.e.,* $f(I)$ *is bounded). Notice: by part a), continuity is not enough.*

THEOREM 1.2. *(Intermediate Value Theorem (IVT)) Let* $f : [a,b] \to \mathbb{R}$ *be a continuous function, and let* L *be any number in between* $f(a)$ *and* $f(b)$. *Then there exists* $c \in [a,b]$ *such that* $f(c) = L$.

PROOF. It is easy to reduce the theorem to the following special case: let $f : [a,b] \to \mathbb{R}$ be continuous and nowhere zero. If $f(a) > 0$, then $f(b) > 0$.

Let $S = \{x \in [a,b] \mid f(x) > 0\}$. Then $f(b) > 0$ iff $b \in S$. We will show $S = [a,b]$.
(RI1) By hypothesis, $f(a) > 0$, so $a \in S$.
(RI2) Let $x \in S$, $x < b$, so $f(x) > 0$. Since f is continuous at x, there exists $\delta > 0$ such that f is positive on $[x, x+\delta]$, and thus $[x, x+\delta] \subset S$.
(RI3) Let $x \in (a,b]$ be such that $[a,x) \subset S$, i.e., f is positive on $[a,x)$. We claim that $f(x) > 0$. Indeed, since $f(x) \neq 0$, the only other possibility is $f(x) < 0$, but if so, then by continuity there would exist $\delta > 0$ such that f is negative on $[x-\delta, x]$, i.e., f is both positive and negative at each point of $[x-\delta, x]$: contradiction! □

In the first examples of mathematical induction the statement itself is of the form "For all $n \in \mathbb{N}$, $P(n)$ holds", so it is clear what the induction hypothesis should be. However, mathematical induction is much more flexible and powerful than this once one learns to *try to find a statement* $P(n)$ whose truth for all n will give the desired result. She who develops skill at "finding the induction hypothesis" acquires a formidable mathematical weapon: for instance the Arithmetic-Geometric Mean Inequality, the Fundamental Theorem of Arithmetic, and the Law of Quadratic Reciprocity have all been proved in this way; in the last case, the first proof given (by Gauss) was by induction.

Similarly, to get a Real Induction proof properly underway, we need to find a subset $S \subset [a,b]$ for which the conclusion $S = [a,b]$ gives us the result we want, and for which our given hypotheses are suitable for "pushing from left to right". If we can find the right set S then we are, quite often, more than halfway there: the rest may take a little while to write out but is relatively straightforward to produce.

THEOREM 1.3. *(Extreme Value Theorem (EVT))*
Let $f : [a,b] \to \mathbb{R}$ be continuous. Then:
a) f is bounded.
b) f attains a minimum and maximum value.

PROOF. a) Let $S = \{x \in [a,b] \mid f : [a,x] \to \mathbb{R} \text{ is bounded}\}$.
(RI1): Evidently $a \in S$.
(RI2): Suppose $x \in S$, so that f is bounded on $[a,x]$. But then f is continuous at x, so is bounded near x: for instance, there exists $\delta > 0$ such that for all $y \in [x-\delta, x+\delta]$, $|f(y)| \leq |f(x)| + 1$. So f is bounded on $[a,x]$ and also on $[x, x+\delta]$ and thus on $[a, x+\delta]$.
(RI3): Suppose $x \in (a,b]$ and $[a,x) \subset S$. Now **beware**: this *does not say* that f is bounded on $[a,x)$: rather it says that for all $a \leq y < x$, f is bounded on $[a,y]$. These are different statements: for instance, $f(x) = \frac{1}{x-2}$ is bounded on $[0,y]$ for all $y < 2$ but it is not bounded on $[0,2)$. But of course this f is not continuous at 2. So we can proceed almost exactly as we did above: since f is continuous at x, there exists $0 < \delta < x - a$ such that f is bounded on $[x-\delta, x]$. But since $a < x - \delta < x$ we know f is bounded on $[a, x-\delta]$, so f is bounded on $[a,x]$.
b) Let $m = \inf f([a,b])$ and $M = \sup f([a,b])$. By part a) we have

$$-\infty < m \leq M < \infty.$$

We want to show that there exist $x_m, x_M \in [a,b]$ such that $f(x_m) = m$, $f(x_M) = M$, i.e., that the infimum and supremum are actually attained as values of f. Suppose that there does not exist $x \in [a,b]$ with $f(x) = m$: then $f(x) > m$ for all $x \in [a,b]$

and the function $g_m : [a, b] \to \mathbb{R}$ by $g_m(x) = \frac{1}{f(x)-m}$ is defined and continuous. By the result of part a), g_m is bounded, but this is absurd: by definition of the infimum, $f(x) - m$ takes values less than $\frac{1}{n}$ for any $n \in \mathbb{Z}+$ and thus g_m takes values greater than n for any $n \in \mathbb{Z}^+$ and is accordingly unbounded. So indeed there must exist $x_m \in [a, b]$ such that $f(x_m) = m$. Similarly, assuming that $f(x) < M$ for all $x \in [a, b]$ gives rise to an unbounded continuous function $g_M : [a, b] \to \mathbb{R},\ x \mapsto \frac{1}{M-f(x)}$, contradicting part a). So there exists $x_M \in [a, b]$ with $f(x_M) = M$. $\square$

EXERCISE 1.2. *Consider the* ***Hansen Interval Theorem (HIT)****: let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there are real numbers $m \leq M$ such that $f([a, b]) = [m, M]$.*
a) Show that HIT is equivalent to the conjunction of IVT and EVT: that is, prove HIT using IVT and EVT and then show that HIT implies both of them.
b) Can you give a direct proof of HIT?

Let $f : I \to \mathbb{R}$. For $\epsilon, \delta > 0$, let us say that f is (ϵ, δ)-uniformly continuous on I – abbreviated (ϵ, δ)-UC on I – if for all $x_1, x_2 \in I$, $|x_1 - x_2| < \delta$ implies $|f(x_1) - f(x_2)| < \epsilon$. This is a halfway unpacking of the definition of uniform continuity: $f : I \to \mathbb{R}$ is uniformly continuous iff for all $\epsilon > 0$, there is $\delta > 0$ such that f is (ϵ, δ)-UC on I.

LEMMA 1.4. *(Covering Lemma) Let $a < b < c < d$ be real numbers, and let $f : [a, d] \to \mathbb{R}$. Suppose that for real numbers $\epsilon, \delta_1, \delta_2 > 0$,*
- *f is (ϵ, δ_1)-UC on $[a, c]$ and*
- *f is (ϵ, δ_2)-UC on $[b, d]$.*

Then f is $(\epsilon, \min(\delta_1, \delta_2, c - b))$-UC on $[a, b]$.

PROOF. Suppose $x_1 < x_2 \in I$ are such that $|x_1 - x_2| < \delta$. Then it cannot be the case that both $x_1 < b$ and $c < x_2$: if so, $x_2 - x_1 > c - b \geq \delta$. Thus we must have either that $b \leq x_1 < x_2$ or $x_1 < x_2 \leq c$. If $b \leq x_1 < x_2$, then $x_1, x_2 \in [b, d]$ and $|x_1 - x_2| < \delta \leq \delta_2$, so $|f(x_1) - f(x_2)| < \epsilon$. Similarly, if $x_1 < x_2 \leq c$, then $x_1, x_2 \in [a, c]$ and $|x_1 - x_2| < \delta \leq \delta_1$, so $|f(x_1) - f(x_2)| < \epsilon$. $\square$

THEOREM 1.5. *(Uniform Continuity Theorem) Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous on $[a, b]$.*

PROOF. For $\epsilon > 0$, let $S(\epsilon)$ be the set of $x \in [a, b]$ such that there exists $\delta > 0$ such that f is (ϵ, δ)-UC on $[a, x]$. To show that f is uniformly continuous on $[a, b]$, it suffices to show that $S(\epsilon) = [a, b]$ for all $\epsilon > 0$. We will show this by Real Induction.
(RI1): Trivially $a \in S(\epsilon)$: f is (ϵ, δ)-UC on $[a, a]$ for all $\delta > 0$!
(RI2): Suppose $x \in S(\epsilon)$, so there exists $\delta_1 > 0$ such that f is (ϵ, δ_1)-UC on $[a, x]$. Moreover, since f is continuous at x, there exists $\delta_2 > 0$ such that for all $c \in [x, x+\delta_2]$, $|f(c)-f(x)| < \frac{\epsilon}{2}$. Why $\frac{\epsilon}{2}$? Because then for all $c_1, c_2 \in [x-\delta_2, x+\delta_2]$,

$$|f(c_1) - f(c_2)| = |f(c_1) - f(x) + f(x) - f(c_2)| \leq |f(c_1) - f(x)| + |f(c_2) - f(x)| < \epsilon.$$

In other words, f is (ϵ, δ_2)-UC on $[x-\delta_2, x+\delta_2]$. We apply the Covering Lemma to f with $a < x - \delta_2 < x < x + \delta_2$ to conclude that f is $(\epsilon, \min(\delta, \delta_2, x - (x - \delta_2))) = (\epsilon, \min(\delta_1, \delta_2))$-UC on $[a, x + \delta_2]$. It follows that $[x, x + \delta_2] \subset S(\epsilon)$.
(RI3): Suppose $[a, x) \subset S(\epsilon)$. As above, since f is continuous at x, there exists $\delta_1 > 0$ such that f is (ϵ, δ_1)-UC on $[x-\delta_1, x]$. Since $x - \frac{\delta_1}{2} < x$, by hypothesis there exists δ_2 such that f is (ϵ, δ_2)-UC on $[a, x-\frac{\delta_1}{2}]$. We apply the Covering Lemma to f

with $a < x-\delta_1 < x-\frac{\delta_1}{2} < x$ to conclude that f is $(\epsilon, \min(\delta_1, \delta_2, x-\frac{\delta_1}{2}-(x-\delta_1))) = (\epsilon, \min(\frac{\delta_1}{2}, \delta_2))$-UC on $[a,x]$. Thus $x \in S(\epsilon)$. $\square$

THEOREM 1.6. *Let $f : [a,b] \to \mathbb{R}$ be a continuous function. Then f is Riemann integrable.*

PROOF. We will use Darboux's Integrability Criterion: we must show that for all $\epsilon > 0$, there exists a partition $\mathbb{P}$ of $[a,b]$ such that $U(f,\mathbb{P}) - L(f,\mathbb{P}) < \epsilon$. It is convenient to prove instead the following equivalent statement: for every $\epsilon > 0$, there exists a partion $\mathbb{P}$ of $[a,b]$ such that $U(f,\mathbb{P}) - L(f,\mathbb{P}) < (b-a)\epsilon$.

Fix $\epsilon > 0$, and let $S(\epsilon)$ be the set of $x \in [a,b]$ such that there exists a partition $\mathbb{P}_x$ of $[a,b]$ with $U(f,\mathbb{P}_x) - L(f,\mathbb{P}_x) < \epsilon$. We want to show $b \in S(\epsilon)$, so it suffices to show $S(\epsilon) = [a,b]$. In fact it is necessary and sufficient: observe that if $x \in S(\epsilon)$ and $a \leq y \leq x$, then also $y \in S(\epsilon)$. We will show $S(\epsilon) = [a,b]$ by Real Induction.
(RI1) The only partition of $[a,a]$ is $\mathbb{P}_a = \{a\}$, and for this partition we have $U(f,\mathbb{P}_a) = L(f,\mathbb{P}_a) = f(a) \cdot 0 = 0$, so $U(f,\mathbb{P}_a) - L(f,\mathbb{P}_a) = 0 < \epsilon$.
(RI2) Suppose that for $x \in [a,b)$ we have $[a,x] \subset S(\epsilon)$. We must show that there is $\delta > 0$ such that $[a,x+\delta] \subset S(\epsilon)$, and by the above observation it is enough to find $\delta > 0$ such that $x+\delta \in S(\epsilon)$: we must find a partition $\mathbb{P}_{x+\delta}$ of $[a,x+\delta]$ such that $U(f,\mathbb{P}_{x+\delta}) - L(f,\mathbb{P}_{x+\delta}) < (x+\delta-a)\epsilon)$. Since $x \in S(\epsilon)$, there is a partition $\mathbb{P}_x$ of $[a,x]$ with $U(f,\mathbb{P}_x) - L(f,\mathbb{P}_x) < (x-a)\epsilon$. Since f is continuous at x, we can make the difference between the maximum value and the minimum value of f as small as we want by taking a sufficiently small interval around x: i.e., there is $\delta > 0$ such that $\max(f,[x,x+\delta]) - \min(f,[x,x+\delta]) < \epsilon$. Now take the smallest partition of $[x,x+\delta]$, namely $\mathbb{P}' = \{x,x+\delta\}$. Then $U(f,\mathbb{P}') - L(f,\mathbb{P}') = (x+\delta-x)(\max(f,[x,x+\delta]) - \min(f,[x,x+\delta])) < \delta\epsilon$. Thus if we put $\mathbb{P}_{x+\delta} = \mathbb{P}_x + \mathbb{P}'$ and use the fact that upper / lower sums add when split into subintervals, we have

$$U(f,\mathbb{P}_{x+\delta}) - L(f,\mathbb{P}_{x+\delta}) = U(f,\mathbb{P}_x) + U(f,\mathbb{P}') - L(f,\mathbb{P}_x) - L(f,\mathbb{P}')$$
$$= U(f,\mathbb{P}_x) - L(f,\mathbb{P}_x) + U(f,\mathbb{P}') - L(f,\mathbb{P}') < (x-a)\epsilon + \delta\epsilon = (x+\delta-a)\epsilon.$$

(RI3) Suppose that for $x \in (a,b]$ we have $[a,x) \subset S(\epsilon)$. We must show that $x \in S(\epsilon)$. The argument for this is the same as for (RI2) except we use the interval $[x-\delta,x]$ instead of $[x,x+\delta]$. Indeed: since f is continuous at x, there exists $\delta > 0$ such that $\max(f,[x-\delta,x]) - \min(f,[x-\delta,x]) < \epsilon$. Since $x-\delta < x$, $x-\delta \in S(\epsilon)$ and thus there exists a partition $\mathbb{P}_{x-\delta}$ of $[a,x-\delta]$ such that $U(f,\mathbb{P}_{x-\delta}) = L(f,\mathbb{P}_{x-\delta}) = (x-\delta-a)\epsilon$. Let $\mathbb{P}' = \{x-\delta,x\}$ and let $\mathbb{P}_x = \mathbb{P}_{x-\delta} \cup \mathbb{P}'$. Then

$$U(f,\mathbb{P}_x) - L(f,\mathbb{P}_x) = U(f,\mathbb{P}_{x-\delta}) + U(f,\mathbb{P}') - (L(f,\mathbb{P}_{x-\delta}) + L(f,\mathbb{P}'))$$
$$= (U(f,\mathbb{P}_{x-\delta}) - L(f,\mathbb{P}_{x-\delta})) + \delta(\max(f,[x-\delta,x]) - \min(f,[x-\delta,x]))$$
$$< (x-\delta-a)\epsilon + \delta\epsilon = (x-a)\epsilon.$$
$\square$

REMARK 1.7. *The standard proof of Theorem 1.6 is to use Darboux's Integrability Criterion and UCT: this is a short, straightforward argument that we leave to the interested reader. In fact this application of UCT is probably the one place in which the concept of uniform continuity plays a critical role in calculus. (Challenge: does your favorite – or least favorite – freshman calculus book discuss uniform continuity? In many cases the answer is "yes" but the treatment is very well hidden from anyone who is not expressly looking for it.) Uniform continuity is hard to fake – how do you explain it without ϵ's and δ's? – so UCT is probably destined to be the black sheep of the interval theorems. This makes it an appealing challenge to give a* *uniform continuity-free proof of Theorem 1.6. In fact Spivak's text does so* [**S**, pp. 292-293]*: he establishes equality of the upper and lower integrals by differentiation. This sort of proof goes back at least to M.J. Norris* [**No52**].

3. Real Induction in Topology

Our task is now to "find the topology" in the classic results of the last section. In calculus, the standard moral one draws from them is that they are (except for Theorem 1.5) properties that are satisfied by our intuitive notion of continuous function, and the fact that they are theorems is a sign of the success of the ϵ, δ definition of continuity.

I want to argue against that – not for all time, but here at least, because it will be useful to our purposes to do so. I claim that there is something much deeper going on in the previous results than just the formal definition of continuity. To see this, let us suppose that we replace the closed interval $[a, b]$ with the *rational closed interval*

$$[a, b]_{\mathbb{Q}} = \{x \in \mathbb{Q} \mid a \leq x \leq b\}.$$

Nothing stops us from defining continuous and uniformly continuous functions $f : [a, b]_{\mathbb{Q}} \to \mathbb{Q}$ in exactly the same way as before: namely, using the ϵ, δ definition. (Soon we will see that this is a case of the ϵ, δ definition of continuity for functions between *metric spaces*.)

Here's the punchline: by switching from the real numbers to the rationals, none of the interval theorems are true. We will except Theorem 1.6 because it is not quite clear what the definition of integrability of a rational function should be, and it is not our business to try to mess with this here. But as for the others:

Example 1.1. *Let*

$$X = \{x \in [0, 2]_{\mathbb{Q}} \mid 0 \leq x^2 < 2\},\ Y = \{x \in [0, 2]_{\mathbb{Q}} \mid 2 < x^2 \leq 4\}.$$

Define $f : [0, 2]_{\mathbb{Q}} \to \mathbb{Q}$ by $f(x) = -1$ if $x \in X$ and $f(x) = 1$ if $x \in Y$. The first thing to notice is that f is indeed well-defined on $[0, 2]_{\mathbb{Q}}$: initially one worries about the case $x^2 = 2$, but – I hope you've heard! – there are in fact no such rational numbers, so no worries. In fact f is continuous: in fact, suppose $x^2 < 2$. Then for any $\epsilon > 0$ we can choose any $\delta > 0$ such that $(x + \delta)^2 < 2$. But clearly f does not satisfy the Intermediate Value Property: it takes exactly two values!

Notice that our choice of δ has the strange property that it is independent of ϵ! This means that the function f is ***locally constant****: there is a small interval around any point at which the function is constant. However the δ cannot be taken independently of ϵ so f is not uniformly continuous. More precisely, for every $\delta > 0$ there are rational numbers x, y with $x^2 < 2 < y^2$ and $|x - y| < \delta$, and then $|f(x) - f(y)| = 2$.*

Exercise 1.3. *Construct a locally constant (hence continuous!) function $f : [0, 2]_{\mathbb{Q}} \to \mathbb{Q}$ which is unbounded. Deduce the EVT does not hold for continuous functions on $[a, b]_{\mathbb{Q}}$. Deduce that such a function cannot be uniformly continuous.*

The point of these examples is that there must be some good property of $[a, b]$ itself that $[a, b]_{\mathbb{Q}}$ lacks. Looking back at the proof of Real Induction we quickly find it: it is the celebrated **least upper bound** axiom. The least upper bound axiom is in fact the source of all the goodness of $\mathbb{R}$ and $[a, b]$, but because in analysis one doesn't study structures which don't have this property, this can be a bit hard to appreciate. Moreover, there are actually several pleasant topological properties that are all implied by the least upper bound axiom, but become distinct in a more general topological context.

To go further, we now introduce some rudimentary topological concepts for intervals in the real line and show how real induction works nicely with these concepts.

A subset $U \subset \mathbb{R}$ is **open** if for all $x \in U$, there is $\epsilon > 0$ such that $(x - \epsilon, x + \epsilon) \subset U$. That is, a subset is open if whenever it contains a point it contains all points sufficient close to it. In particular the empty set $\varnothing$ and $\mathbb{R}$ itself are open.

Exercise 1.4. *An interval is open in $\mathbb{R}$ iff it is of the form (a, b), $(-\infty, b)$ or (a, ∞).*

We also want to define open subsets *of intervals*, especially of the closed bounded interval $[a, b]$. In this course we will define open sets in several different contexts before arriving at the final (for us!) level of generality of topological spaces, but one easy common property is that when we are trying to define open subsets of a set X, we always want to include X as an open subset of itself. Notice that if we directly extend the above definition of open sets to $[a, b]$ then this doesn't work: $a \in [a, b]$ but there is no $\epsilon > 0$ such that $(a - \epsilon, a + \epsilon) \subset [a, b]$.

For now we fix this in the kludgiest way: let $I \subset \mathbb{R}$ be an interval.[2] A subset $U \subset I$ is open if:
• For every point $x \in U$ which is not an endpoint of I, we have $(x - \epsilon, x + \epsilon) \subset U$ for some $\epsilon > 0$;
• If $x \in U$ is the left endpoint of I, then there is some $\epsilon > 0$ such that $[x, x + \epsilon) \subset I$.
• If $x \in U$ is the right endpoint of I, then there is some $\epsilon > 0$ such that $(x - \epsilon, x] \subset I$.

Exercise 1.5. *Show: a subset U of an interval I is open iff whenever U contains a point, it contains all points* of I *which lie sufficiently close to it.*

Let A be a subset of an interval I. A point $x \in I$ is a **limit point of A in I** if for all $\epsilon > 0$, $(x - \epsilon, x + \epsilon)$ contains a point of A other than x.

Exercise 1.6. *Let I be an interval in I. Show that except in the case in which I consists of a single point, every point of I is a limit point of I.*

Theorem 1.8. *(Bolzano-Weierstrass) Every infinite subset of $[a, b]$ has a limit point in $[a, b]$.*

Proof. Let $\mathcal{A}$ be an infinite subset of $[a, b]$, and let S be the set of x in $[a, b]$ such that *if* $\mathcal{A} \cap [a, x]$ is infinite, it has a limit point. It suffices to show that $S = [a, b]$, which we prove by Real Induction. As usual, (i) is trivial. Since $\mathcal{A} \cap [a, x)$ is finite iff $\mathcal{A} \cap [a, x]$ is finite, (iii) follows. As for (ii), suppose $x \in S$. If $\mathcal{A} \cap [a, x]$ is infinite, then by hypothesis it has a limit point and hence so does $[a, b]$. So we may assume that $\mathcal{A} \cap [a, x]$ is finite. Now either there exists $\delta > 0$ such that $\mathcal{A} \cap [a, x + \delta]$ is finite – okay – or every interval $[x, x + \delta]$ contains infinitely many points of $\mathcal{A}$ in which case x itself is a limit point of $\mathcal{A}$. $\square$

[2] Until further notice, "interval" will always mean interval in $\mathbb{R}$.

A subset $A \subset \mathbb{R}$ is **compact** if given any family $\{U_i\}_{i\in I}$ of open subsets of $\mathbb{R}$, if $A \subset \bigcup_{i\in I} U_i$, then there is a finite subset $J \subset I$ with $A \subset \bigcup_{i\in J} U_i$. We define compact subsets of an interval (in $\mathbb{R}$) similarly.

LEMMA 1.9. *Let $A \subset \mathbb{R}$ be compact. Then A is bounded and every limit point of A is an element of A.*

PROOF. For $n \in \mathbb{Z}^+$, let $U_n = (-n, n)$. Then $\bigcup_{n=1}^{\infty} U_n = \mathbb{R}$, and every finite union of the U_n's is bounded, so if A is unbounded then $A \subset \bigcup_{n=1}^{\infty} U_n$ and is not contained in $\bigcup_{n\in J} A_n$ for any finite $J \subset \mathbb{Z}^+$. Suppose that a is a limit point of A which does not lie in A. Let $U_n = (-\infty, a-\frac{1}{n}) \cup (a+\frac{1}{n}, \infty)$. Then $\bigcup_{n=1}^{\infty} U_n = \mathbb{R} \setminus A$, so $A \subset \bigcup_{n\in\mathbb{Z}^+} U_n$, but since a is a limit point of A, there is no finite subset $J \subset \mathbb{Z}^+$ with $A \subset \bigcup_{n\in J} U_n$. $\square$

THEOREM 1.10. *(Heine-Borel) The interval $[a, b]$ is compact.*

PROOF. For an open covering $\mathcal{U} = \{U_i\}_{i\in I}$ of $[a, b]$, let

$$S = \{x \in [a, b] \mid \mathcal{U} \cap [a, x] \text{ has a finite subcovering}\}.$$

We prove $S = [a, b]$ by Real Induction. (RI1) is clear. (RI2): If $U_1, \ldots, U_n$ covers $[a, x]$, then some U_i contains $[x, x+\delta]$ for some $\delta > 0$. (RI3): if $[a, x) \subset S$, let $i_x \in I$ be such that $x \in U_{i_x}$, and let $\delta > 0$ be such that $[x-\delta, x] \in U_{i_x}$. Since $x - \delta \in S$, there is a finite $J \subset I$ with $\bigcup_{i\in J} U_i \supset [a, x-\delta]$, so $\{U_i\}_{i\in J} \cup U_{i_x}$ covers $[a, x]$. $\square$

PROPOSITION 1.11. *a) IVT implies the connectedness of $[a, b]$: if A, B are open subsets of $[a, b]$ such that $A \cap B = \varnothing$ and $A \cup B = [a, b]$, then either $A = \varnothing$ or $B = \varnothing$.*
b) The connectedness of $[a, b]$ implies IVT.

PROOF. In both cases we will argue by contraposition.
a) Suppose $[a, b] = A \cup B$, where A and B are nonempty open subsets such that $A \cap B = \varnothing$. Then function $f : [a, b] \to \mathbb{R}$ which sends $x \in A \mapsto 0$ and $x \in B \mapsto 1$ is continuous but does not have the Intermediate Value Property.
b) If IVT fails, there is a continuous function $f : [a, b] \to \mathbb{R}$ and $A < B < C$ such that $A, C \in f([a, b])$ but $B \notin f([a, b])$. Let

$$U = \{x \in [a, b] \mid f(x) < B\},\ V = \{x \in [a, b] \mid B < f(x)\}.$$

Then U and V are open sets – the basic principle here is that if a continuous function satisfies a strict inequality at a point, then it satisfies the same strict inequality in some small interval around the point – which partition $[a, b]$. $\square$

4. The Miracle of Sequences

LEMMA 1.12. *(Rising Sun [**NP88**]) Every infinite sequence in the real line[3] has a monotone subsequence.*

PROOF. Let us say that $m \in \mathbb{Z}^+$ is a **peak** of the sequence $\{a_n\}$ if for all $n > m$ we have $a_n < a_m$. Suppose first that there are infinitely many peaks. Then any sequence of peaks forms a strictly decreasing subsequence, hence we have found a monotone subsequence. So suppose on the contrary that there are only finitely many peaks, and let $N \in \mathbb{N}$ be such that there are no peaks $n \geq N$. Since $n_1 = N$ is not a peak, there exists $n_2 > n_1$ with $a_{n_2} \geq a_{n_1}$. Similarly, since n_2 is not a peak,

[3]Or any ordered set: see §5.

there exists $n_3 > n_2$ with $a_{n_3} \geq a_{n_2}$. Continuing in this way we construct an infinite (not necessarily strictly) increasing subsequence $a_{n_1}, a_{n_2}, \ldots, a_{n_k}, \ldots$. Done! □

Theorem 1.13. *(Sequential Bolzano-Weierstrass) Every sequence in $[a, b]$ admits a convergent subsequence.*

Proof. Let $\{x_n\}$ be a sequence in $[a, b]$. By the Rising Sun Lemma, $\{x_n\}$ admits a monotone subsequence. A bounded increasing (resp. decreasing) sequence converges to its supremum (resp. infimum). □

Exercise 1.7. *(Bolzano-Weierstrass = Sequential Bolzano-Weierstrass)*
a) Suppose that every infinite subset of $[a, b]$ has a limit point in $[a, b]$. Show that every sequence in $[a, b]$ admits a convergent subsequence.
b) Suppose that every sequence in $[a, b]$ admits a convergent subsequence. Show that every infinite subset of $[a, b]$ has a limit point in $[a, b]$.

Theorem 1.14. *Sequential Bolzano-Weierstrass implies EVTa).*

Proof. Seeking a contradiction, let $f : [a, b] \to \mathbb{R}$ be an *unbounded* continuous function. Then for each $n \in \mathbb{Z}^+$ we may choose $x_n \in [a, b]$ such that $|f(x_n)| \geq n$. By Theorem 4.1, after passing to a subsequence (which, as usual, we will suppress from our notation) we may suppose that x_n converges, say to $\alpha \in [a, b]$. Since f is continuous, $f(x_n) \to f(\alpha)$, so in particular $\{f(x_n)\}$ is bounded...contradiction!
(With regard to the attainment of extrema, we have no improvement to offer on the simple argument using suprema / infima given in the proof of Theorem 1.3. □

Theorem 1.15. *Sequential Bolzano-Weierstrass implies UCT (Theorem 1.5).*

Proof. Seeking a contradiction, let $f : [a, b] \to \mathbb{R}$ be continuous but *not* uniformly continuous. Then there is $\epsilon > 0$ such that for all $n \in \mathbb{Z}^+$, there are $x_n, y_n \in [a, b]$ with $|x_n - y_n| < \frac{1}{n}$ and $|f(x_n) - f(y_n)| \geq \epsilon$. By Theorem 1.13, after passing to a subsequence (notationally suppressed!) x_n converges to some $\alpha \in [a, b]$, and thus also $y_n \to \alpha$. Since f is continuous $f(x_n)$ and $f(y_n)$ both converge to $f(\alpha)$, hence for sufficiently large n, $|f(x_n) - f(y_n)| < \epsilon$: contradiction! □

5. Induction and Completeness in Ordered Sets

5.1. Introduction to Ordered Sets.

Consider the following properties of a binary relation $\leq$ on a set X:
(Reflexivity) For all $x \in X$, $x \leq x$.
(Anti-Symmetry) For all $x, y \in X$, if $x \leq y$ and $y \leq x$, then $x = y$.
(Transitivity) For all $x, y, z \in X$, if $x \leq y$ and $y \leq z$, then $x \leq z$.
(Totality) For all $x, y \in X$, either $x \leq y$ or $y \leq x$.
A relation which satisfies reflexivity and transitivity is called a **quasi-ordering**. A relation which satisfies reflexivity, anti-symmetry and transitivity is called a **partial ordering**. A relation which satisfies all four properties is called an **ordering** (sometimes a **total** or **linear ordering**). An **ordered set** is a pair $(X, \leq)$ where X is a set and $\leq$ is an ordering on X.

Rather unsurprisingly, we write $x < y$ when $x \leq y$ and $x \neq y$. We also write $x \geq y$ when $y \leq x$ and $x > y$ when $x \geq y$ and $x \neq y$.

Ordered sets are a basic kind of mathematical structure which *induces* a topological structure. (It is not yet supposed to be clear exactly what this means.) Moreover they allow an inductive principle which generalizes Real Induction.

A **bottom element** of an ordered set is an element which is strictly less than every other element of the set. Clearly bottom elements are unique if they exist, and clearly they may or may not exist: the natural numbers $\mathbb{N}$ have 0 as a bottom element, and the integers do not have a bottom element. We will denote the bottom element of an ordered set, when it exists, by $\mathbb{B}$.

If an ordered set does not have a bottom element, we can add one. Let X be an ordered set without a bottom element, choose any $\mathbb{B} \notin X$, let $X_{\mathbb{B}} = X \cup \{\mathbb{B}\}$, and extend the ordering to $\mathbb{B}$ by taking $\mathbb{B} < x$ for all $x \in X$.[4]

There is an entirely parallel discussion for **top elements** $\mathbb{T}$.

Example 1.2. *Starting from the empty set – which* is *an ordered set! – and applying the bottom element construction n times, we get a linearly ordered set with n elements. Similarly for applying the top element construction n times.*

We will generally suppress the $\leq$ when speaking about ordered sets and simply refer to "the ordered set X".

Let X and Y be ordered sets. A function $f : X \to Y$ is:
- **increasing** (or **isotone**) if for all $x_1 \leq x_2 \in X$, $f(x_1) \leq f(x_2)$ in Y;
- **strictly increasing** if for all $x_1 < x_2$ in X, $f(x_1) < f(x_2)$;
- **decreasing** (or **antitone**) if for all $x_1 \leq x_2$ in X, $f(x_1) \geq f(x_2)$;
- **strictly decreasing** if for all $x_1 < x_2$ in X, $f(x_1) > f(x_2)$.

This directly generalizes the use of these terms in calculus. But now we take the concept more seriously: we think of orderings on X and Y as giving *structure* and we think of isotone maps as being the maps which *preserve that structure.*

Exercise 1.8. *Let $f : X \to Y$ be an increasing function between ordered sets. Show that f is strictly increasing iff it is injective.*

Exercise 1.9. *Let X, Y and Z be ordered sets.*
a) Show that the identity map $1_X : X \to X$ is isotone.
b) Suppose that $f : X \to Y$ and $g : Y \to Z$ are isotone maps. Show that the composition $g \circ f : X \to Z$ is an isotone map.
c) Suppose that $f : X \to Y$ and $g : Y \to Z$ are antitone maps. Show that the composition $g : X \to Z$ is an isotone (not antitone!) map.

Let X and Y be ordered sets. An **order isomorphism** is an isotone map $f : X \to Y$ for which there exists an isotone inverse function $g : Y \to X$.

Exercise 1.10. *Let X and Y be ordered sets, and let $f : X \to Y$ be an isotone bijection. Show that f is an order isomorphism.*

[4]This procedure works even if X already has a bottom element, except with the minor snag that our suggested notation now has us denoting two different elements by $\mathbb{B}$. We dismiss this as being beneath our pay grade.

EXERCISE 1.11. *a) Which linear functions $f : \mathbb{R} \to \mathbb{R}$ are order isomorphisms?*
b) Let $d \in \mathbb{Z}^+$. Show that there is a degree d polynomial order isomorphism $P : \mathbb{R} \to \mathbb{R}$ iff d is odd.

EXERCISE 1.12. *a) Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous bijection. Show: f is either increasing or decreasing.*
b) Let $f : \mathbb{R} \to \mathbb{R}$ be a bijection which is either increasing or decreasing. Show: f is continuous.

We say that a property of an ordered set is **order-theoretic** if whenever an ordered set possesses that property, every order-isomorphic set has that property.

EXAMPLE 1.3. *The following are all order-theoretic properties: being nonempty, being finite, being infinite, having a given cardinality (indeed these are all properties preserved by bijections of sets), having a bottom element, having a top element, being dense.*

Let $a < b$ be elements in an ordered set X. We define

$$[a, b] = \{x \in X \mid a \leq x \leq b\},$$

$$(a, b] = \{x \in X \mid a < x \leq b\},$$

$$[a, b) = \{x \in X \mid a \leq x < b\},$$

$$(a, b) = \{x \in X \mid a < x < b,$$

$$(-\infty, b] = \{x \in S \mid x \leq b\},$$

$$(-\infty, b) = \{x \in S \mid x < b\},$$

$$[a, \infty) = \{x \in S \mid a \leq x\},$$

$$(a, \infty) = \{x \in S \mid a < x\}.$$

An **interval** in X is any of the above sets together with $\varnothing$ and X itself. We call intervals of the form $\varnothing$, (a, b), $(-\infty, b)$, (a, ∞) and X **open intervals**. We call intervals of the form $[a, b]$, $(-\infty, b]$ and $[a, \infty)$ **closed intervals**.

EXERCISE 1.13. *a) Let $a < b$ and $c < d$ be real numbers. Show that $[a, b]$ and $[c, d]$ are order-isomorphic.*
b) Let $a < b$. Show that (a, b) is order-isomorphic to $\mathbb{R}$.
c) Classify all intervals in $\mathbb{R}$ up to order-isomorphism.

Let $x < y$ be elements in an ordered set. We say that **y covers x** if $(x, y) = \varnothing$: in other words, y is the bottom element of the subset of all elements which are greater than x. We say y is the **successor** of x and write $y = x^+$. Similarly, we say that x is the **predecessor** of y and write $x = y^-$. Clearly an element in an ordered set may or may not have a successor or a predecessor: in $\mathbb{Z}$, every element has both; in $\mathbb{R}$, no element has either one. An element x of an ordered set is **left-discrete** if $x = \mathbb{B}$ or x has a predecessor and **right discrete** $x = \mathbb{T}$ or x has a successor. An ordered set is **discrete** if every element is left-discrete and right-discrete.

At the other extreme, an ordered set X is **dense** if for all $a < b \in X$, there exists c with $a < c < b$.

Exercise 1.14. *Let X be an ordered set with at least two elements. Show that the following are equivalent:*
(i) No element $x \neq \mathbb{B}$ of X is left-discrete.
(ii) No element $x \neq \mathbb{T}$ of X is right-discrete.
(iii) X is dense.

Exercise 1.15. *For a linearly ordered set X, we define the* ***order dual*** *$X^\vee$ to be the ordered set with the same underlying set as X but with the ordering reversed: that is, for $x, y \in X^\vee$, we have $x \leq y \iff y \leq x$ in X.*
a) Show that X has a top element (resp. a bottom element) $\iff$ $X^\vee$ has a bottom element (resp. a top element).
b) Show that X is well-ordered iff $X^\vee$ satisfies the ascending chain condition.
c) Suppose X is finite. Show that $X \cong X^\vee$ (order-isomorphic).
d) Determine which intervals I in $\mathbb{R}$ are isomorphic to their order-duals.

Let $(X_1, \leq_1)$ and $(X_2, \leq_2)$ be ordered spaces. We define the **lexicographic order** $\leq$ on the Cartesian product $X_1 \times X_2$ as follows: $(x_1, x_2) \leq (y_1, y_2)$ iff $x_1 < y_1$ or ($x_1 = y_1$ and $x_2 \leq y_2$).

Exercise 1.16.
a) Show: the lexicographic ordering is indeed an ordering on $X_1 \times X_2$.
b) Show: if X_1 and X_2 are both well-ordered, so is $X_1 \times X_2$.
c) Extend the lexicographic ordering to finite products $X_1 \times X_n$.
(N.B.: It can be extended to infinite products as well...)

5.2. Completeness and Dedekind Completeness.

The characteristic property of the real numbers among ordered fields is the least upper bound axiom: every nonempty subset which is bounded above has a least upper bound. But this axiom says nothing about the algebraic operations $+$ and $\cdot$: it is purely order-theoretic. In fact, by pursuing its analogue in an arbitrary ordered set we will get an interesting and useful generalization of Real Induction.

For a subset S of a linearly ordered set X, a **supremum** $\sup S$ of S is a least element which is greater than or equal to every element of S, and an **infimum** $\inf S$ of S is a greatest element which is less than or equal to every element of S. X is **complete** if every subset has a supremum; equivalently, if every subset has an infimum. If X is complete, it has a least element $\mathbb{B} = \sup \varnothing$ and a greatest element $\mathbb{T} = \inf \varnothing$. X is **Dedekind complete** if every nonempty bounded above subset has a supremum; equivalently, if every nonempty bounded below subset has an infimum. X is complete iff it is Dedekind complete and has $\mathbb{B}$ and $\mathbb{T}$.

Exercise 1.17. *Show that an ordered set X is Dedekind complete iff the set obtained by adjoining top and bottom elements to X is complete.*

Exercise 1.18. *Let X be an ordered set with order-dual $X^\vee$.*
a) Show that X is complete iff $X^\vee$ is complete.
b) Show that X is Dedekind complete iff $X^\vee$ is Dedekind complete.

Exercise 1.19. *In the following problem, X and Y are nonempty ordered sets, and Cartesian products are given the lexicographic ordering.*
a) Show: if X and Y are complete, then $X \times Y$ is complete.

b) Show: $\mathbb{R} \times \mathbb{R}$ *in the lexicographic ordering is* not *Dedekind complete.*
c) Show: if $X \times Y$ *is complete, then* X *and* Y *are complete.*
d) Suppose $X \times Y$ *is Dedekind complete. What can be said about* X *and* Y*?*

5.3. Principle of Ordered Induction.

We give an inductive characterization of Dedekind completeness in linearly ordered sets, and apply it to prove three topological characterizations of completeness which generalize familiar results from elementary analysis.

Let X be an ordered set. A set $S \subset X$ is **inductive** if it satisfies:
(IS1) There exists $a \in X$ such that $(-\infty, a] \subset S$.
(IS2) For all $x \in S$, either $x = \mathbb{T}$ or there exists $y > x$ such that $[x, y] \subset S$.
(IS3) For all $x \in X$, if $(-\infty, x) \subset S$, then $x \in S$.

Exercise 1.20. *Let* X *be an ordered set with a bottom element* $\mathbb{B}$*. Show that (IS3)* $\implies$ *(IS1).*

Theorem 1.16. *(Principle of Ordered Induction) For a linearly ordered set* X*, the following are equivalent:*
(i) X *is Dedekind complete.*
(ii) The only inductive subset of X *is* X *itself.*

Proof. (i) $\implies$ (ii): Let $S \subset X$ be inductive. Seeking a contradiction, we suppose $S' = X \setminus S$ is nonempty. Fix $a \in X$ satisfying (IS1). Then a is a lower bound for S', so by hypothesis S' has an infimum, say y. Any element less than y is strictly less than every element of S', so $(-\infty, y) \subset S$. By (IS3), $y \in S$. If $y = 1$, then $S' = \{1\}$ or $S' = \varnothing$: both are contradictions. So $y < 1$, and then by (IS2) there exists $z > y$ such that $[y, z] \subset S$ and thus $(-\infty, z] \subset S$. Thus z is a lower bound for S' which is strictly larger than y, contradiction.
(ii) $\implies$ (i): Let $T \subset X$ be nonempty and bounded below by a. Let S be the set of lower bounds for T. Then $(-\infty, a] \subset S$, so S satisfies (IS1).
Case 1: Suppose S does not satisfy (IS2): there is $x \in S$ with no $y \in X$ such that $[x, y] \subset S$. Since S is downward closed, x is the top element of S and $x = \inf(T)$.
Case 2: Suppose S does not satisfy (IS3): there is $x \in X$ such that $(-\infty, x) \in S$ but $x \notin S$, i.e., there exists $t \in T$ such that $t < x$. Then also $t \in S$, so t is the least element of T: in particular $t = \inf T$.
Case 3: If S satisfies (IS2) and (IS3), then $S = X$, so $T = \{1\}$ and $\inf T = 1$. □

Exercise 1.21. *Use the fact that an ordered set* X *is Dedekind complete iff its order dual is to state a* ***downward version*** *of Theorem 1.16.*

Exercise 1.22. *a) Show that when* X *is well-ordered, Theorem 1.16 becomes the principle of transfinite induction.*
b) Show that when $X = [a, b]$*, we recover Real Induction.*

6. Dedekind Cuts

Let S be a nonempty ordered set.

A **quasi-cut** in S is an ordered pair (S_1, S_2) of subsets $S_1, S_2 \subset S$ with $S_1 \leq S_2$ and $S = S_1 \cup S_2$. It follows immediately that S_1 is initial, S_2 is final and S_1 and

S_2 intersect at at most one point.

A **cut** $\Lambda = (\Lambda^L, \Lambda^R)$ is a quasi-cut with $\Lambda^L \cap \Lambda^R = \varnothing$; Λ is a **Dedekind cut** if Λ^L and Λ^R are both nonempty. We call Λ^L and Λ^R the **initial part** and **final part** of the cut, respectively. Any initial (resp. final) subset $T \subset S$ is the initial part of a unique cut: the final (resp. initial) part is $S \setminus T$.

For any subset $M \subset S$, we define the **downward closure**

$$D(M) = \{x \in S \mid x \leq m \textit{ for some } m \in M\}$$

and the **upward closure**

$$U(M) = \{x \in S \mid m \leq x \textit{ for some } m \in M\}.$$

Exercise 1.23. *Let $M \subset S$.*
a) Show that $D(M)$ is the intersection of all initial subsets of S containing M and thus the unique minimal initial subset containing M.
b) Show that $U(M)$ is the intersection of all final subsets of S containing M and thus the unique minimal final subset containing M.

Thus any subset of M determines two (not necessarily distinct) cuts: a cut M^+ with initial part $D(M)$ and a cut M^- with final part $U(M)$. For $\alpha \in S$, we write α^+ for $\{\alpha\}^+$ and α^- for $\{\alpha\}^-$.

Exercise 1.24. *Let $\alpha \in S$. Show that α^+ is the unique cut in which α is the maximum of the initial part and that α^- is the unique cut in which α is the minimum of the final part.*

We call cuts of the form α^+ and α^- **principal**. Thus a cut fails to be principal iff its initial part has no maximum and its final part has no minimum.

Example 1.4. *Let S be a nonempty ordered set.*
a) The cut $(S, \varnothing)$ is principal iff S has a top element. The cut $(\varnothing, S)$ is principal iff S has a bottom element.
b) Let $S = \mathbb{R}$. Then the above two cuts are not principal, but let $\Lambda = (\Lambda_L, \Lambda_R)$ be a Dedekind cut. Then Λ_L is bounded above (by any element of Λ_R), so let $\alpha = \sup \Lambda_L$. Then either $\alpha \in \Lambda_L$ and $\Lambda = \alpha^+$ or $\alpha \in \Lambda_R$ and $\Lambda = \alpha^-$. Thus every Dedekind cut in $\mathbb{R}$ is principal.
c) Let $S = \mathbb{Q}$. Then $\{(-\infty, \sqrt{2}) \cap \mathbb{Q}, (\sqrt{2}, \infty)\}$ is a nonprincipal Dedekind cut.

One swiftly draws the following moral.

Theorem 1.17. *Let S be an ordered set. Then:*
a) Every cut in S is principal iff S is complete.
b) Every Dedekind cut in S is principal iff S is Dedekind complete.

Exercise 1.25. *Prove it.*

Now let T be an ordered set, let $S \subset T$ be nonempty, and let $\Lambda = (\Lambda^L, \Lambda^R)$ be a cut in S. We say that $\gamma \in T$ **realizes Λ in T** if $\Lambda^L \leq \gamma \leq \Lambda^R$. Conversely, to each $\gamma \in T$ we associate the cut

$$\Lambda(\gamma) = (\{x \in S \mid x \leq \gamma\}, (x \in S \mid x > \gamma\})$$

in S. This is a sinister definition: if $\gamma \in S$ we get $\Lambda(\gamma) = \gamma^+$. (We could have set things up the other way, but we do need to make a choice one way or the other.) The cuts in S which are realized by some element of S are precisely the principal cuts, and a principal cut is realized by either one or two elements of S (the latter cannot happen if S is order-dense). Conversely, every element $\gamma \in S$ realizes precisely two cuts, γ^+ and γ^-.

EXAMPLE 1.5. *Let $S = \mathbb{Q}$ and $T = \mathbb{R}$. The cut $\{(-\infty, \sqrt{2}) \cap \mathbb{Q}, (\sqrt{2}, \infty)\}$, which is non-principal in S, is realized in T by $\sqrt{2}$.*

If in an ordered set S we have a nonprincipal cut $\Lambda = (\Lambda_L, \Lambda_R)$, up to order-isomorphism there is a unique way to add a point γ to S which realizes Λ: namely we adjoin γ with $\Lambda_L < \gamma < \Lambda_R$.

For an ordered set S, we denote by $\tilde{S}$ the set of all cuts in S. We equip $\tilde{S}$ with the following binary relation: for $\Lambda_1, \ \Lambda_2 \in \tilde{S}$, we put $\Lambda_1 \leq \Lambda_2$ if $\Lambda_1^L \subset \Lambda_2^L$.

PROPOSITION 1.18. *Let S be an ordered set, and let $\tilde{S}$ be the set of cuts of S.*
a) The relation $\leq$ on S is an ordering.
b) Each of the maps

$$\iota_+ : S \to \tilde{S}, \ x \mapsto x^+$$
$$\iota_- : S \to \tilde{S}, \ x \mapsto x^-$$

is an isotone injection.

PROOF. a) The inclusion relation $\subset$ is a partial ordering on the power set 2^S; restricting to initial sets we still get a partial ordering. A cut is determined by its initial set, so $\leq$ is certainly a *partial* ordering on $\tilde{S}$. The matter of it is to show that we have a total ordering: equivalently, given any two initial subsets Λ_1^L and Λ_2^L of an ordered set, one is contained in the other. Well, suppose not: if neither is contained in the other, there is $x_1 \in \Lambda_1^L \setminus \Lambda_2^L$ and $x_2 \in \Lambda_2^L \setminus \Lambda_1^L$. We may assume without loss of generality that $x_1 < x_2$ (otherwise, switch Λ_1^L and Λ_2^L): but since Λ_2^L is initial and contains x_2, it also contains x_1: contradiction.
b) This is a matter of unpacking the definitions, and we leave it to the reader. $\square$

THEOREM 1.19. *Let S be a totally ordered set. The map $\iota^+ : S \hookrightarrow \tilde{S}$ gives an order completion of S. That is:*
a) $\tilde{S}$ is complete: every subset has a supremum and an infimum.
b) If X is a complete ordered set and $f : S \to X$ is an isotone map, there is an isotone map $\tilde{f} : \tilde{S} \to X$ such that $f = \tilde{f} \circ \iota_+$.

PROOF. It will be convenient to identify a cut Λ with its initial set Λ_L.
a) Let $\{\Lambda_i\}_{i \in I} \subset \tilde{S}$. Put $\Lambda_1^L = \bigcup_{i \in I} \Lambda_i^L$ and $\Lambda_2^L = \bigcap_{i \in I} \Lambda_i^L$. Since unions and intersections of initial subsets are initial, Λ_1^L and Λ_2^L are cuts in S, and clearly

$$\Lambda_1^L = \inf\{\Lambda_i\}_{i \in I}, \ \Lambda_2^L = \sup\{\Lambda_i\}_{i \in I}.$$

b) For $\Lambda_L \in \tilde{S}$, we define

$$\tilde{f}(\Lambda_L) = \sup_{x \in \Lambda_L} f(x).$$

It is easy to see that defining $\tilde{f}$ in this way gives an isotone map with $\tilde{f} \circ \iota_+ = f$. $\square$

THEOREM 1.20. *Let F be an ordered field, and let $D(F)$ be the Dedekind completion of F. Then $D(F)$ can be given the structure of a field compatible with its ordering iff the ordering on F is Archimedean.*

CHAPTER 2

Metric Spaces

1. Metric Geometry

A **metric** on a set X is a function $d : X \times X \to [0, \infty)$ satisfying:

(M1) (Definiteness) For all $x,\ y \in X$, we have $d(x, y) = 0 \iff x = y$.
(M2) (Symmetry) For all $x,\ y \in X$, we have $d(x, y) = d(y, x)$.
(M3) (Triangle Inequality) For all $x,\ y,\ z \in X$, we have $d(x, z) \leq d(x, y) + d(y, z)$.

A **metric space** is a pair (X, d) consisting of a set X and a metric d on X. By the usual abuse of notation, when only one metric on X is under discussion we will typically refer to "the metric space X."

EXAMPLE 2.1. *(Discrete Metric) Let X be a set, and for any $x, y \in X$, put*

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

*This is a metric on X which we call the **discrete metric**. We warn the reader that we will later study a property of metric spaces called **discreteness**. A set endowed with the discrete metric is a discrete space, but there are discrete metric spaces which are not endowed with the discrete metric.*

In general showing that a given function $d : X \times X \to \mathbb{R}$ is a metric is nontrivial. More precisely verifying the Triangle Inequality is often nontrivial; (M1) and (M2) are usually very easy to check.

EXAMPLE 2.2.
a) Let $X = \mathbb{R}$ and take $d(x, y) = |x - y|$.
*b) More generally, let $N \geq 1$, let $X = \mathbb{R}^N$, and take $d(x, y) = ||x - y|| = \sqrt{\sum_{i=1}^N (x_i - y_i)^2}$. It is very well known but not very obvious that d satisfies the triangle inequality. This is a special case of **Minkowski's Inequality**, which will be studied later.*
c) More generally let $p \in [1, \infty)$, let $N \geq 1$, let $X = \mathbb{R}^N$ and take

$$d_p(x, y) = ||x - y||_p = \left(\sum_{i=1}^N (x_i - y_i)^p \right)^{\frac{1}{p}}.$$

*The assertion that d_p satisfies the triangle inequality is **Minkowski's Inequality**.*

EXAMPLE 2.3. *Let (X, d) be a metric space, and let $Y \subset X$ be any subset. Then the restricted function*

$$d|_{Y \times Y} : Y \times Y \to \mathbb{R}$$

is a metric function on Y. Indeed, because the three properties (M1), (M2) and (M3) are all universally quantified statements, since they hold for all $(x_1, x_2) \in X \times X$ or all $(x_1, x_2, x_3) \in X \times X \times X$, they necessarily hold for all $(y_1, y_2) \in Y \times Y$ or all $(y_1, y_2, y_3) \in Y \times Y \times Y$.

*The set Y endowed with its restricted metric d_Y is called a **subspace** of the metric space X. We also say that the metric d_Y is **induced** from the metric d on Y.*

EXAMPLE 2.4. *Let $a \leq b \in \mathbb{R}$. Let $\mathcal{C}[a, b]$ be the set of all continuous functions $f : [a, b] \to \mathbb{R}$. For $f \in \mathcal{C}[a, b]$, let*

$$||f|| = \sup_{x \in [a,b]} |f(x)|.$$

Then $d(f, g) = ||f - g||$ is a metric function on $\mathcal{C}[a, b]$.

PROPOSITION 2.1. *(Reverse Triangle Inequality) Let (X, d) be a metric space, and let $x, y, z \in X$. Then we have*

$$|d(x, y) - d(x, z)| \leq d(y, z). \tag{1}$$

PROOF. The triangle inequality gives

$$d(x, y) \leq d(x, z) + d(z, y)$$

and thus

$$d(x, y) - d(x, z) \leq d(y, z).$$

Similarly, we have

$$d(x, z) \leq d(x, y) + d(y, z)$$

and thus

$$d(x, z) - d(x, y) \leq d(y, z).$$ □

1.1. Exercises.

EXERCISE 2.1.
a) Let (X, d_x) and (Y, d_y) be metric spaces. Show that the function

$$d_{X \times Y} : (X \times Y) \times (X \times Y) \to \mathbb{R},\ ((x_1, y_1), (x_2, y_2)) \mapsto \max(d_X(x_1, x_2), d_Y(y_1, y_2))$$

is a metric on $X \times Y$.
b) Extend the result of part a) to finitely many metric spaces $(X_1, d_{X_1}), \ldots, (X_n, d_{X_n})$.
c) Let $N \geq 1$, let $X = \mathbb{R}^N$ and take $d_\infty(x, y) = \max_{1 \leq i \leq N} |x_i - y_i|$. Show that d_∞ is a metric.
d) For each fixed $x, y \in \mathbb{R}^N$, show

$$d_\infty(x, y) = \lim_{p \to \infty} d_p(x, y).$$

Use this to give a second (more complicated) proof of part c).

Let (X, d_x) and (Y, d_Y) be metric spaces. A function $f : X \to Y$ is an **isometric embedding** if for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)$. That is, the distance between any two points in X is the same as the distance between their images under f. An **isometry** is a surjective isometric embedding.

EXERCISE 2.2. *a) Show that every isometric embedding is injective.*
b) Show that every isometry is bijective and thus admits an inverse function.
c) Show that if $f : (X, d_X) \to (Y, d_Y)$ is an isometry, so is $f^{-1} : (Y, d_Y) \to (X, d_X)$.

For metric spaces X and Y, let $\mathrm{Iso}(X,Y)$ denote the set of all isometries from X to Y. Put $\mathrm{Iso}(X) = \mathrm{Iso}(X,X)$, the isometries from X to itself. According to more general mathematical usage we ought to call elements of $\mathrm{Iso}(X)$ "autometries" of X...but almost no one does.

Exercise 2.3. *a) Let $f : X \to Y$ and $g : Y \to Z$ be isometric embeddings. Show that $g \circ f : X \to Z$ is an isometric embedding.*
b) Show that $\mathrm{Iso}\, X$ forms a group under composition.
c) Let X be a set endowed with the discrete metric. Show that $\mathrm{Iso}\, X = \mathrm{Sym}\, X$ is the group of all bijections $f : X \to X$.
d) Can you identify the isometry group of $\mathbb{R}$? Of Euclidean N-space?

Exercise 2.4. *a) Let X be a set with $N \geq 1$ elements endowed with the discrete metric. Find an isometric embedding $X \hookrightarrow \mathbb{R}^{N-1}$.*
b) Show that there is no isometric embedding $X \hookrightarrow \mathbb{R}^{N-2}$.*
c) Deduce that an infinite set endowed with the discrete metric is not isometric to any subset of a Euclidean space.

Exercise 2.5. *a) Let G be a finite group. Show that there is a finite metric space X such that $\mathrm{Iso}\, X \cong G$ (isomorphism of groups).*
b) Prove or disprove: for every group G, there is a metric space X with $\mathrm{Iso}\, X \cong G$?

Let A be a nonempty subset of a metric space X. The **diameter** of A is

$$\mathrm{diam}(A) = \sup\{d(x,y) \mid x, y \in A\}.$$

Exercise 2.6. *a) Show that $\mathrm{diam}\, A = 0$ iff A consists of a single point.*
b) Show that A is bounded iff $\mathrm{diam}\, A < \infty$.
c) Show that For any $x \in X$ and $\epsilon > 0$, $\mathrm{diam}\, B(x,\epsilon) \leq 2\epsilon$.

Exercise 2.7. *Recall: for sets X,Y we have the **symmetric difference***

$$X \Delta Y = (X \setminus Y) \coprod (Y \setminus X),$$

the set of elements belonging to exactly one of X and Y ("exclusive or"). Let S be a finite set, and let 2^S be the set of all subsets of S. Show that

$$d : 2^S \times 2^S \to \mathbb{N},\ d(X,Y) = \#(X \Delta Y)$$

*is a metric function on 2^S, called the **Hamming metric**.*

Exercise 2.8. *Let X be a metric space.*
a) Suppose $\#X \leq 2$. Show that there is an isometric embedding $X \hookrightarrow \mathbb{R}$.
b) Let d be a metric function on the set $X = \{a,b,c\}$. Show that up to relabelling the points we may assume

$$d_1 = d(a,b) \leq d_2 = d(b,c) \leq d_3 = d(a,c).$$

Find necessary and sufficient conditions on d_1, d_2, d_3 such that there is an isometric embedding $X \hookrightarrow \mathbb{R}$. Show that there is always an isometric embedding $X \hookrightarrow \mathbb{R}^2$.
c) Let $X = \{\bullet, a, b, c\}$ be a set with four elements. Show that

$$d(\bullet,a) = d(\bullet,b) = d(\bullet,c) = 1,\ d(a,b) = d(a,c) = d(b,c) = 2$$

gives a metric function on X. Show that there is no isometric embedding of X into any Euclidean space.

Exercise 2.9. *Let $G = (V,E)$ be a connected graph. Define $d : V \times V \to \mathbb{R}$ by taking $d(P,Q)$ to be the length of the shortest path connecting P to Q.*

a) *Show that d is a metric function on V.*
b) *Show that the metric of Exercise 2.8c) arises in this way.*
c) *Find necessary and/or sufficient conditions for the metric induced by a finite connected graph to be isometric to a subspace of some Euclidean space.*

EXERCISE 2.10. *Let $d_1, d_2 : X \times X \to \mathbb{R}$ be metric functions.*

a) *Show that $d_1 + d_2 : X \times X \to \mathbb{R}$ is a metric function.*
b) *Show that $\max(d_1, d_2) : X \times X \to \mathbb{R}$ is a metric function.*

EXERCISE 2.11.
a) Show that for any $x, y \in \mathbb{R}$ there is $f \in \operatorname{Iso}\mathbb{R}$ such that $f(x) = y$.
b) Show that for any $x \in \mathbb{R}$, there are exactly two isometries f of $\mathbb{R}$ such that $f(x) = x$.
c) Show that every isometric embedding $f : \mathbb{R} \to \mathbb{R}$ is an isometry.
d) Find a metric space X and an isometric embedding $f : X \to X$ which is not surjective.

EXERCISE 2.12.
Consider the following property of a function $d : X \times X \to [0, \infty)$:
(M1′) For all $x \in X$, $d(x, x) = 0$.
*A **pseudometric function** is a function $d : X \times X \to [0, \infty)$ satisfying (M1′), (M2) and (M3), and a **pseudometric space** is a pair (X, d) consisting of a set X and a pseudometric function d on X.*
a) Show that every set X admits a pseudometric function.
b) Let (X, d) be a pseudometric space. Define a relation $\sim$ on X by $x \sim y$ iff $d(x, y) = 0$. Show that $\sim$ is an equivalence relation.
c) Show that the pseudometric function is well-defined on the set $X/\sim$ of $\sim$-equivalence classes: that is, if $x \sim x'$ and $y \sim y'$ then $d(x, y) = d(x', y')$. Show that d is a metric function on $X/\sim$.

1.2. Constructing Metrics.

In this section as well as later on we will make use of basic results on convex functions $f : I \to \mathbb{R}$, where I is an interval on the real line. A reference for this is [**HC**, §7.3].

EXERCISE 2.13. a) *Let (X, d) be a metric space, let Y be a subset of X, and let $d_Y : Y \times Y \to \mathbb{R}$ be the induced metric, as above. Show that inclusion of Y into X gives an isometric embedding $(Y, d_Y) \hookrightarrow (X, d_X)$.*
b) *Conversely, let (X', d') be a metric space and $\iota : (X', d') \to (X, d)$ be an isometric embedding. Show that ι induces an isometry $(X', d') \to (\iota(X'), d)$.*

LEMMA 2.2. *Let (X, d) be a metric space, and let $f : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ be an increasing, concave function – i.e., $-f$ is convex – with $f(0) = 0$. Then $d_f = f \circ d$ is a metric on X.*

PROOF. The only nontrivial verification is the triangle inequality. Let $x, y, z \in X$. Since d is a metric, we have

$$d(x, z) \leq d(x, y) + d(y, z).$$

Since f is increasing, we have

$$d_f(x,z) = f(d(x,z)) \leq f(d(x,y) + d(y,z)). \tag{2}$$

Since $-f$ is convex and $f(0) = 0$, by the Generalized Two Secant Inequality [**HC**, Cor. 7.14], we have for all $a \geq 0$ and all $t > 0$ that

$$\frac{f(a+t) - f(a)}{(a+t) - a} \leq \frac{f(t)}{t - 0}$$

and thus

$$f(a+t) \leq f(a) + f(t). \tag{3}$$

Taking $a = d(x,y)$ and $t = d(y,z)$ and combining (2) and (3), we get

$$d_f(x,z) \leq f(d(x,y) + d(y,z)) \leq d_f(x,y) + d_f(y,z). \qquad \square$$

COROLLARY 2.3. *Let (X, d) be a metric space, and let $\alpha > 0$. Let $d_\alpha : X \times X \to \mathbb{R}$ be given by*

$$d_\alpha(x,y) = \frac{\alpha d(x,y)}{d(x,y) + 1}.$$

Then d_α is a metric on X and $\operatorname{diam}(X, d_\alpha) \leq \alpha$.

EXERCISE 2.14. *Prove Corollary 2.3.*

EXERCISE 2.15. *Prove the **Quadrilateral Inequality**: for points x_1, x_2, y_1, y_2 in a metric space (X, d) we have*

$$|d(x_1, y_1) - d(x_2, y_2)| \leq d(x_1, x_2) + d(y_1, y_2).$$

2. Metric Topology

Let (X, d) be a metric space.

For $x \in X$ and $\epsilon \geq 0$ we define the **open ball**

$$B^\circ(x, \epsilon) = \{y \in X \mid d(x,y) < \epsilon\}.$$

and the **closed ball**

$$B^\bullet(x, \epsilon) = \{y \in X \mid d(x,y) \leq \epsilon\}.$$

Notice that

$$B^\circ(x, 0) = \varnothing,$$
$$B^\bullet(x, 0) = \{x\}.$$

A subset Y of a metric space X is **open** if for all $y \in Y$, there is $\epsilon > 0$ such that

$$B^\circ(y, \epsilon) \subset Y.$$

A subset Y of a metric space X is **closed** if its complement

$$X \setminus Y = \{x \in X \mid x \notin Y\}$$

is open.

EXERCISE 2.16. *Find a subsets X_1, X_2, X_3, X_4 of $\mathbb{R}$ such that:*

(i) *X_1 is both open and closed.*
(ii) *X_2 is open and not closed.*
(iii) *X_3 is closed and not open.*
(iv) *X_4 is neither open nor closed.*

Proposition 2.4. *Let X be a metric space, and let $\{Y_i\}_{i \in I}$ be subsets of X.*

a) *The union $Y = \bigcup_{i \in I} Y_i$ is an open subset of X.*
b) *If I is nonempty and finite, then the intersection $Z = \bigcap_{i \in I} Y_i$ is an open subset of X.*

Proof. a) If $y \in Y$, then $y \in Y_i$ for at least one i. So there is $\epsilon > 0$ such that

$$B^\circ(y, \epsilon) \subset Y_i \subset Y.$$

b) We may assume that $I = \{1, \ldots, n\}$ for some $n \in \mathbb{Z}^+$. Let $y \in Z$. Then for $1 \leq i \leq n$, there is $\epsilon_i > 0$ such that $B^\circ(y, \epsilon_i) \subset Y_i$. Then $\epsilon = \min_{1 \leq i \leq n} \epsilon_i > 0$ and $B^\circ(y, \epsilon) \subset B^\circ(y, \epsilon_i) \subset Y_i$ for all $1 \leq i \leq n$, so $B^\circ(y, \epsilon) \subset \bigcap_{i=1}^n Y_i = Z$. $\square$

Let X be a set and $\tau \subset 2^X$ be family of subets of X. We say τ is a **topology** if:
(T1) $\varnothing, X \in \tau$;
(T2) For any set I, if $Y_i \in \tau$ for all $i \in I$ then $\bigcup_{i \in I} Y_i \in \tau$;
(T3) For any nonempty finite set I, if $Y_i \in \tau$ for all $i \in I$, then $\bigcup_{i \in I} Y_i \in \tau$.

The axioms (T2) and (T3) are usually referred to as “arbitrary unions of open sets are open” and “finite intersections of open sets are open”, respectively.

In this language, Proposition 2.4 may be rephrased as follows.

Proposition 2.5. *The open sets of a metric space (X, d) form a topology.*

We say that two metrics d_1 and d_2 on the same set X are **topologically equivalent** if they determine the same topology: that is, every set which is open with respect to d_1 is open with respect to d_2.

Example 2.5. *In $\mathbb{R}$, for $n \in \mathbb{Z}^+$, let $Y_n = (\frac{-1}{n}, \frac{1}{n})$. Then each Y_n is open but $\bigcap_{n=1}^\infty Y_n = \{0\}$ is not. This shows that infinite intersections of open subsets need not be open.*

Exercise 2.17. *In any metric space:*

a) *Show that finite unions of closed sets are closed.*
b) *Show that arbitrary intersections of closed sets are closed.*
c) *Exhibit an infinite union of closed subsets that is not closed.*

Exercise 2.18. *A metric space X is **discrete** if every subset $Y \subset X$ is open.*

a) *Show that any set endowed with the discrete metric is a discrete metric space.*
b) *A metric space X is **uniformly discrete** if there is $\epsilon > 0$ such that for all $x \neq y \in X$, $d(x, y) \geq \epsilon$. Show: every uniformly discrete metric space is discrete.*
c) *Let $X = \{\frac{1}{n}\}_{n=1}^\infty$ as a subspace of $\mathbb{R}$. Show that X is discrete but not uniformly discrete.*

Proposition 2.6. *Let X be a metric space.*

a) *Open balls in X are open sets.*
b) *A subset Y of X is open iff Y is a union of open balls.*

Proof. a) Let $x \in X$, let $\epsilon > 0$, and let $y \in B^\circ(x, \epsilon)$. We claim that $B^\circ(y, \epsilon - d(x, y)) \subset B^\circ(x, \epsilon)$. Indeed, if $z \in B^\circ(y, \epsilon - d(x, y))$, then $d(y, z) < \epsilon - d(x, y)$, so

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + (\epsilon - d(x, y)) = \epsilon.$$

b) If Y is open, then for all $y \in Y$, there is $\epsilon_y > 0$ such that $B^\circ(\epsilon, y) \subset Y$. It follows that $Y = \bigcup_{y \in Y} B^\circ(y, \epsilon_y)$. The fact that a union of open balls is open follows from part a) and the previous result. $\square$

LEMMA 2.7. *Let Y be a subset of a metric space X. Then the map $U \mapsto U \cap Y$ is a surjective map from the open subsets of X to the open subsets of Y.*

EXERCISE 2.19. *Prove Lemma 2.7.*
(Hint: for any $y \in Y$ and $\epsilon > 0$, let $B_X^\circ(y, \epsilon) = \{x \in X \mid d(x, y) < \epsilon\}$ and let $B_Y^\circ(y, \epsilon) = \{x \in Y \mid d(x, y) < \epsilon\}$. Then $B_X^\circ(y, \epsilon) = B_Y^\circ(y, \epsilon)$.)

Let X be a metric space, and let $Y \subset X$. We define the **interior** of Y as

$$Y^\circ = \{y \in X \mid \exists\ \epsilon > 0 \text{ such that } B^\circ(y, \epsilon) \subset Y\}.$$

In words, the interior of a set is the collection of points that not only belong to the set, but for which some open ball around the point is entirely contained in the set.

LEMMA 2.8. *Let Y, Z be subsets of a metric space X.*

a) *All of the following hold:*
  (i) *$Y^\circ \subset Y$.*
  (ii) *If $Y \subset Z$, then $Y^\circ \subset Z^\circ$.*
  (iii) *$(Y^\circ)^\circ = Y^\circ$.*

b) *The interior Y° is the largest open subset of Y: that is, Y° is an open subset of Y and if $U \subset Y$ is open, then $U \subset Y^\circ$.*

c) *Y is open iff $Y = Y^\circ$.*

EXERCISE 2.20. *Prove Lemma 2.8.*

We say that a subset Y is a **neighborhood** of $x \in X$ if $x \in Y^\circ$. In particular, a subset is open precisely when it is a neighborhood of each of its points. (This terminology introduces nothing essentially new. Nevertheless the situation it encapsulates it ubiquitous in this subject, so we will find the term quite useful.)

Let X be a metric space, and let $Y \subset X$. A point $x \in X$ is an **adherent point of Y** if every neighborhood $\mathcal{N}$ of x intersects Y: i.e., $\mathcal{N} \cap Y \neq \varnothing$. Equivalently, for all $\epsilon > 0$, we have $B(x, \epsilon) \cap Y \neq \varnothing$.

We follow up this definition with another, rather subtly different one, that we will fully explore later, but it seems helpful to point out the distinction now. For $Y \subset X$, a point $x \in X$ is an **limit point** of Y if every neighborhood $\mathcal{N}$ of x contains a point of $Y \setminus \{x\}$. Equivalently, for all $\epsilon > 0$, we have

$$(B^\circ(x, \epsilon) \setminus \{x\}) \cap Y \neq \varnothing.$$

EXERCISE 2.21. *Let X be a metric space, let Y be a subset of X, and let x be a point of X. Show: x is a limit point of Y iff every neighborhood of x contains infinitely many points of Y.*

Every $y \in Y$ is an adherent point of Y but not necessarily a limit point. For instance, if Y is finite then it has no limit points.

The following is the most basic and important result of the entire section.

PROPOSITION 2.9.
For a subset Y of a metric space X, the following are equivalent:
(i) *Y is closed: i.e., $X \setminus Y$ is open.*
(ii) *Y contains all of its adherent points.*
(iii) *Y contains all of its limit points.*

PROOF. (i) $\implies$ (ii): Suppose that $X \setminus Y$ is open, and let $x \in X \setminus Y$. Then there is $\epsilon > 0$ such that $B^\circ(x, \epsilon) \subset X \setminus Y$, and thus $B^\circ(x, \epsilon)$ does not intersect Y, i.e., x is not an adherent point of Y.
(ii) $\implies$ (iii): Since every limit point is an adherent point, this is immediate.
(iii) $\implies$ (i): Suppose Y contains all its limit points, and let $x \in X \setminus Y$. Then x is not a limit point of Y, so there is $\epsilon > 0$ such that $(B^\circ(x, \epsilon) \setminus \{x\}) \cap Y = \varnothing$. Since $x \notin Y$ this implies $B^\circ(x, \epsilon) \cap Y = \varnothing$ and thus $B^\circ(x, \epsilon) \subset X \setminus Y$. Thus $X \setminus Y$ contains an open ball around each of its points, so is open, so Y is closed. $\square$

For a subset Y of a metric space X, we define its **closure** of Y as

$$\overline{Y} = Y \cup \{\text{all adherent points of } Y\} = Y \cup \{\text{all limit points of } Y\}.$$

LEMMA 2.10. *Let Y, Z be subsets of a metric space X.*
a) All of the following hold:
(KC1) $Y \subset \overline{Y}$.
(KC2) If $Y \subset Z$, then $\overline{Y} \subset \overline{Z}$.
(KC3) $\overline{\overline{Y}} = \overline{Y}$.
b) The closure $\overline{Y}$ is the smallest closed set containing Y: that is, $\overline{Y}$ is closed, contains Y, and if $Y \subset Z$ is closed, then $\overline{Y} \subset Z$.

EXERCISE 2.22. *Prove Lemma 2.10.*

LEMMA 2.11. *Let Y, Z be subsets of a metric space X. Then:*
a) $\overline{Y \cup Z} = \overline{Y} \cup \overline{Z}$.
b) $(Y \cap Z)^\circ = Y^\circ \cap Z^\circ$.

PROOF. a) Since $\overline{Y} \cup \overline{Z}$ is a finite union of closed sets, it is closed. Since $\overline{Y} \supset Y$ and $\overline{Z} \supset Z$ we have

$$\overline{Y} \cup \overline{Z} \supset Y \cup Z.$$

Now let A be a closed subset of X that contains $Y \cup Z$. Then we have

$$Y \subset A \implies \overline{Y} \subset \overline{A} = A, \; Z \subset A \implies \overline{Z} \subset \overline{A} = A,$$

so $\overline{Y} \cup \overline{Z} \subset A$. Thus $\overline{Y} \cup \overline{Z}$ is the minimal closed subset containing $Y \cup Z$, so

$$\overline{Y} \cup \overline{Z} = \overline{Y \cup Z}.$$

b) The set $Y^\circ \cap Z^\circ$ is a finite intersection of open sets, hence it is an open subset of $Y \cap Z$. Now let U be an open subset of X that is contained in $Y \cap Z$. Then

$$U \subset Y \implies U = U^\circ \subset Y^\circ, \; U \subset Z \implies U = U^\circ \subset Z^\circ,$$

so $U \subset Y^\circ \cap Z^\circ$. Thus $Y^\circ \cap Z^\circ$ is the maximal open subset contained in $Y \cap Z$, so

$$Y^\circ \cap Z^\circ = (Y \cap Z)^\circ.$$

$\square$

EXERCISE 2.23. *Let Y and Z be subsets of a metric space X.*
a) *Show: $\overline{Y \cap Z} \subset \overline{Y} \cap \overline{Z}$.*

b) *Let* $X = \mathbb{R}$ *with the standard Euclidean medtric. Find subsets* Y *and* Z *such that* $\overline{Y \cap Z} = \varnothing$ *and* $\overline{Y} \cap \overline{Z} = \mathbb{R}$.

EXERCISE 2.24. *Formulate and prove an analogue of Exercise 2.23 for unions of interiors of subsets* Y, Z *of a metric space* X.

The similarity between the proofs of parts a) and b) of the preceding result is meant to drive home the point that just as open and closed are "dual notions" – one gets from one to the other via taking complements – so are interiors and closures.

PROPOSITION 2.12. *Let* Y *be a subset of a metric space* Z. *Then we have*

$$Y^\circ = X \setminus \overline{X \setminus Y} \tag{4}$$

and

$$\overline{Y} = X \setminus (X \setminus Y)^\circ. \tag{5}$$

PROOF. We will prove (4) and leave (5) as an exercise. To prove (4) our strategy is to show that $X \setminus \overline{X \setminus Y}$ is the largest open subset of Y and apply Lemma 2.8b). Since $X \setminus \overline{X \setminus Y}$ is the complement of a closed set, it is open. Moreover, if $x \in X \setminus \overline{X \setminus Y}$, then $x \notin \overline{X \setminus Y}$ and $\overline{X \setminus Y} \supset X \setminus Y$, so $x \in Y$. Now let $U \subset Y$ be open. Then $X \setminus U$ is closed and contains $X \setminus Y$, so $X \setminus U$ contains $\overline{X \setminus Y}$. Taking complements again we get $U \subset (X \setminus \overline{X \setminus Y})$. □

EXERCISE 2.25. *Prove (5).*

PROPOSITION 2.13. *For a subset* Y *of a metric space* X, *consider the following:*

(i) $B_1(Y) = \overline{Y} \setminus Y^\circ$.
(ii) $B_2(Y) = \overline{Y} \cap \overline{X \setminus Y}$.
(iii) $B_3(Y) = \{x \in X \mid$ *every neighborhood* $\mathcal{N}$ *of* x *intersects both* Y *and* $X \setminus Y\}$.

Then $B_1(Y) = B_2(Y) = B_3(Y)$ *is a closed subset of* X, *called the* ***boundary of Y*** *and denoted* ∂Y.

EXERCISE 2.26. *Prove Proposition 2.13.*

EXERCISE 2.27. *Let* Y *be a subset of a metric space* X.

a) *Show* $\overline{X} = X^\circ \coprod \partial X$ *(disjoint union).*
b) *Show* $(\partial X)^\circ = \varnothing$.
c) *Show that* $\partial(\partial Y) = \partial Y$.

EXERCISE 2.28. *Show: for all closed subsets* B *of* $\mathbb{R}^N$, *there is a subset* Y *of* $\mathbb{R}^N$ *with* $B = \partial Y$.

EXAMPLE 2.6. *Let* $X = \mathbb{R}$, $A = (-\infty, 0)$ *and* $B = [0, \infty)$. *Then* $\partial A = \partial B = \{0\}$, *and*

$$\partial(A \cup B) = \partial\mathbb{R} = \emptyset \neq \{0\} = (\partial A) \cup (\partial B);$$
$$\partial(A \cap B) = \partial\emptyset = \emptyset \neq \{0\} = (\partial A) \cap (\partial B).$$

Thus the boundary is not as well-behaved as either the closure or interior.

EXERCISE 2.29. *For a subset* Y *of a metric space* Y, *we define the* ***exterior of Y*** *as*

$$\mathrm{ext}(Y) := X \setminus \overline{Y}.$$

a) *Show that we have* $X = Y^\circ \coprod \partial Y \coprod \mathrm{ext}(Y)$.
b) *Show:* $\mathrm{ext}(Y) = (X \setminus Y)^\circ$.

A subset Y of a metric space X is **dense** if $\overline{Y} = X$.

Exercise 2.30. *Show: a subset Y of a metric space X is dense iff Y intersects every nonempty open subset of X.*

Example 2.7. *Let X be a discrete metric space. The only dense subset of X is X itself.*

Example 2.8. *The subset $\mathbb{Q}^N = (x_1, \ldots, x_N)$ is dense in $\mathbb{R}^N$.*

Exercise 2.31. *Let X be a metric space, and let $Z \subset Y \subset X$. Suppose that Z is dense in Y (we give Y the induced metric) and that Y is dense in X. Show: Z is dense in X.*

The **weight** of a metric space is the least cardinality of a dense subspace.

Exercise 2.32. *Let X be a metric space.*

a) *Show: if X is discrete, then the weight of X is $\#X$. Deduce that every cardinal number occurs as the weight of a metric space.*
b) *Show: if X is finite, then the weight of X is $\#X$.*
c) *Show: the weight of Euclidean N-space $\mathbb{R}^N$ is $\aleph_0$.*
(In other words, no finite subset of $\mathbb{R}^N$ is dense, and there is a countably infinite dense subset.)

Explicit use of cardinal arithmetic is popular in some circles but not in others. Much more commonly used is the following special case: a metric space is **separable** if it admits a countable dense subspace. Thus the previous example shows that Euclidean N-space is separable, and a discrete space is separable iff it is countable.

2.1. Further Exercises.

Exercise 2.33. *Let Y be a subset of a metric space X. Show:*

$$(\overline{Y^\circ})^\circ = Y^\circ$$

and

$$\overline{\overline{Y}^\circ} = \overline{Y}.$$

Exercise 2.34. *A subset Y of a metric space X is **regularly closed** if $Y = \overline{Y^\circ}$ and **regularly open** if $Y = (\overline{Y})^\circ$.*

a) *Show that every regularly closed set is closed, every regularly open set is open, and a set is regularly closed iff its complement is regularly open.*
b) *Show that a subset of $\mathbb{R}$ is regularly closed iff it is a disjoint union of closed intervals.*
c) *Show that for any subset Y of a metric space X, $\overline{Y^\circ}$ is regularly closed and $\overline{Y}^\circ$ is regularly open.*

Exercise 2.35. *A metric space is a **door space** if every subset is either open or closed (or both). In a topologically discrete space, every subset is both open and closed, so such spaces are door spaces, however of a rather uninteresting type. Show that there is a subset of $\mathbb{R}$ which, with the induced metric, is a door space which is not topologically discrete.*

Exercise 2.36. *A subset Y of a metric space (X, d) is a G_δ-set if there is a sequence of open subsets $\{U_n\}_{n=1}^\infty$ such that $Y = \bigcap_{n=1}^\infty U_n$.*

a) *Show: every open subset of a metric space is a G_δ-set.*
b) *Show: every closed subset of a metric space is a G_δ-set.*
c) *Show: $\mathbb{R} \setminus \mathbb{Q}$ is a G_δ-set in $\mathbb{R}$ that is neither open nor closed.*
d) *Show: $\mathbb{Q}$ is not a G_δ-set in $\mathbb{R}$.*

3. Convergence

In any set X, a sequence in X is just a mapping a mapping $\mathbf{x} : \mathbb{Z}^+ \to X$, $n \mapsto \mathbf{x}_n$. If X is endowed with a metric d, a sequence $\mathbf{x}$ in X is said to **converge** to an element x of X if for all $\epsilon > 0$, there exists an $N = N(\epsilon)$ such that for all $n \geq N$, $d(x, x_n) < \epsilon$. We denote this by $\mathbf{x} \to x$ or $\mathbf{x}_n \to x$.

Exercise 2.37. *Let $\mathbf{x}$ be a sequence in the metric space X, and let $L \in X$. Show that the following are equivalent.*
a) The $\mathbf{x} \to L$.
b) Every neighbhorhood $\mathcal{N}$ of $\mathbf{x}$ contains all but finitely many terms of the sequence. More formally, there is $N \in \mathbb{Z}^+$ such that for all $n \geq N$, $x_n \in \mathcal{N}$.

Proposition 2.14. *In any metric space, the limit of a convergent sequence is unique: if $L, M \in X$ are such that $\mathbf{x} \to L$ and $\mathbf{x} \to M$, then $L = M$.*

Proof. Seeking a contradiction, we suppose $L \neq M$ and put $d = d(L, M) > 0$. Let $B_1 = B^\circ(L, \frac{d}{2})$ and $B_2 = B^\circ(M, \frac{d}{2})$, so B_1 and B_2 are disjoint. Let N_1 be such that if $n \geq N_1$, $x_n \in B_1$, let N_2 be such that if $n \geq N_2$, $x_n \in B_2$, and let $N = \max(N_1, N_2)$. Then for all $n \geq N$, $x_n \in B_1 \cap B_2 = \varnothing$: contradiction! $\square$

A **subsequence** of $\mathbf{x}$ is obtained by choosing an infinite subset of $\mathbb{Z}^+$, writing the elements in increasing order as $n_1, n_2, \ldots$ and then restricting the sequence to this subset, getting a new sequence $\mathbf{y}$, $k \mapsto \mathbf{y}_k = \mathbf{x}_{n_k}$.

Exercise 2.38. *Let $n : \mathbb{Z}^+ \to \mathbb{Z}^+$ be strictly increasing: for all $k_1 < k_2$, $n_{k_1} < n_{k_2}$. Let $\mathbf{x} : \mathbb{Z}^+ \to X$ be a sequence in a set X. Interpret the composite sequence $\mathbf{x} \circ n : \mathbb{Z}^+ \to X$ as a subsequence of $\mathbf{x}$. Show that every subsequence arises in this way, i.e., by precomposing the given sequence with a unique strictly increasing function $n : \mathbb{Z}^+ \to \mathbb{Z}^+$.*

Exercise 2.39. *Let $\mathbf{x}$ be a sequence in a metric space.*
a) Show that if $\mathbf{x}$ is convergent, so is every subsequence, and to the same limit.
b) Show that conversely, if every subsequence converges, then $\mathbf{x}$ converges. (Hint: in fact this is not a very interesting statement. Why?)
c) A more interesting converse would be: suppose that there is $L \in X$ such that: every subsequence of $\mathbf{x}$ which is convergent *converges to L. Then $\mathbf{x} \to L$. Show that this* fails *in $\mathbb{R}$. Show however that it holds in $[a, b] \subset \mathbb{R}$.*

Let $\mathbf{x}$ be a sequence in a metric space X. A point $L \in X$ is a **partial limit** of $\mathbf{x}$ if every neighborhood $\mathcal{N}$ of L contains infinitely many terms of the sequence: more formally, for all $N \in \mathbb{Z}^+$, there is $n \geq N$ such that $\mathbf{x}_n \in \mathcal{N}$.

Lemma 2.15. *For a sequence $\mathbf{x}$ in a metric space X and $L \in X$, the following are equivalent:*
(i) L is a partial limit of $\mathbf{x}$.
(ii) There is a subsequence $\mathbf{x}_{n_k}$ converging to L.

PROOF. (i) Suppose L is a partial limit. Choose $n_1 \in \mathbb{Z}^+$ such that $d(x_{n_1}, L) < 1$. Having chosen $n_k \in \mathbb{Z}^+$, choose $n_{k+1} > n_k$ such that $d(x_{n_{k+1}}, L) < \frac{1}{k+1}$. Then $\mathbf{x}_{n_k} \to L$.
(ii) Let $\mathcal{N}$ be any neighborhood of L, so there is $\epsilon > 0$ such that $L \subset B^\circ(L, \epsilon) \subset \mathcal{N}$. If $\mathbf{x}_{n_k} \to L$, then for every $\epsilon > 0$ and all sufficiently large k, we have $d(x_{n_k}, L) < \epsilon$, so infinitely many terms of the sequence lie in $\mathcal{N}$. $\square$

The following basic result shows that closures in a metric space can be understood in terms of convergent sequences.

PROPOSITION 2.16. *Let Y be a subset of (X, d). For $x \in X$, the following are equivalent:*
(i) $x \in \overline{Y}$.
(ii) There exists a sequence $\mathbf{x} : \mathbb{Z}^+ \to Y$ such that $\mathbf{x}_n \to x$.

PROOF. (i) $\implies$ (ii): Suppose $y \in \overline{Y}$, and let $n \in \mathbb{Z}^+$. There is $x_n \in Y$ such that $d(y, x_n) < \epsilon$. Then $x_n \to y$.
$\neg$ (i) $\implies$ $\neg$ (ii): Suppose $y \notin \overline{Y}$: then there is $\epsilon > 0$ such that $B^\circ(y, \epsilon) \cap Y = \varnothing$. Then no sequence in Y can converge to y. $\square$

COROLLARY 2.17. *Let X be a set, and let $d_1, d_2 : X \times X \to X$ be two metrics. Suppose that for every sequence $\mathbf{x} \in X$ and every point $x \in X$, we have $\mathbf{x} \xrightarrow{d_1} x \iff \mathbf{x} \xrightarrow{d_2} x$: that is, the sequence $\mathbf{x}$ converges to the point x with respect to the metric d_1 it converges to the point x with respect to the metric d_2. Then d_1 and d_2 are topologically equivalent: they have the same open sets.*

PROOF. Since the closed sets are precisely the complements of the open sets, it suffices to show that the closed sets with respect to d_1 are the same as the closed sets with respect to d_2. So let $Y \subset X$, and suppose that Y is closed with respect to d_1. Then, still with respect to d_1, Y is its own closure, so by Proposition 2.16 for $x \in X$ we have that x lies in Y iff there is a sequence $\mathbf{y}$ in Y such that $\mathbf{y} \to x$ with respect to d_1. But by assumption this latter characterization is also valid with respect to d_2, so Y is closed with respect to d_2. And conversely, of course. $\square$

4. Continuity

Let $f : X \to Y$ be a map between metric spaces, and let $x \in x$. We say **f is continuous at x** if for all $\epsilon > 0$, there is $\delta > 0$ such that for all $x' \in X$, if $d(x, x') < \delta$ then $d(f(x), f(x')) < \epsilon$. We say **f is continuous** if it is continuous at every $x \in X$.

A map $f : X \to Y$ of metric spaces is **uniformly continuous** if for all $\epsilon > 0$, there is $\delta > 0$ such that for all $x_1, x_2 \in X$, if $d(x_1, x_2) < \delta$ then $d(f(x_1), f(x_2)) < \epsilon$.

EXERCISE 2.40. *Let $f : X \to Y$ be a map between metric spaces. We define a function*

$$\Delta_f : X \times (0, \infty) \to [0, \infty]$$

as follows: for $x \in X$, let

$$\Delta_f(x, \epsilon) := \sup\{\delta \geq 0 \mid d(x_1, x_2) \leq \delta \implies d(f(x_1), f(x_2)) \leq \epsilon\}.$$

a) *Show: we have $\infty \in \Delta_f(X \times (0, \infty))$ iff f is bounded.*
b) *Show: f is continuous at x iff $\Delta_f(x, \epsilon) > 0$ for all $\epsilon > 0$.*

c) *Show: f is uniformly continuous iff* $\inf\{\Delta_f(x,\epsilon) \mid x \in X\} > 0$ *for all* $\epsilon > 0$.

Let $f : X \to Y$ be a map between metric spaces. A real number $C \geq 0$ is a **Lipschitz constant for f** if for all $x, y \in X$, $d(f(x), f(y)) \leq Cd(x,y)$. A map f is **Lipschitz** if some $C \geq 0$ is a Lipschitz constant for f.

A map $f : X \to Y$ between metric spaces is a **contraction** if it is Lipschitz with a Lipschitz constant $C < 1$, is **weakly contractive** if for all $x_1 \neq x_2 \in X$ we have $d(f(x_1), f(x_2)) < d(x_1, x_2)$, and is a **short map** if it is Lipschitz with a Lipschitz constant $C \leq 1$. (Thus contractive $\implies$ weakly contractive $\implies$ short.)

EXERCISE 2.41. *Exhibit a map of metric spaces $f : X \to Y$ that is short but is neither a contraction nor an isometric embedding.*

EXERCISE 2.42. *Let I be an interval in $\mathbb{R}$, and let $f : I \to I$.*

a) *Show: if f' exists and is bounded, then f is Lipschitz.*
b) *Deduce: if $I = [a,b]$ and f has a continuous derivative, then f is Lipschitz.*

EXERCISE 2.43. a) *Show: a Lipschitz function is continuous.*
b) *Show that if f is Lipschitz, the infimum of all Lipschitz constants for f is a Lipschitz constant for f.*
c) *Show that an isometry is Lipschitz.*

LEMMA 2.18. *For a map $f : X \to Y$ of metric spaces, the following are equivalent:*

(i) *f is continuous.*
(ii) *For every open subset $V \subset Y$, $f^{-1}(V)$ is open in X.*

PROOF. (i) $\implies$ (ii): Let $x \in f^{-1}(V)$, and choose $\epsilon > 0$ such that $B^\circ(f(x), \epsilon) \subset V$. Since f is continuous at x, there is $\delta > 0$ such that for all $x' \in B^\circ(x, \delta)$, $f(x') \in B^\circ(f(x), \epsilon) \subset V$: that is, $B^\circ(x, \delta) \subset f^{-1}(V)$.
(ii) $\implies$ (i): Let $x \in X$, let $\epsilon > 0$, and let $V = B^\circ(f(x), \epsilon)$. Then $f^{-1}(V)$ is open and contains x, so there is $\delta > 0$ such that

$$B^\circ(x, \delta) \subset f^{-1}(V).$$

That is: for all x' with $d(x, x') < \delta$, $d(f(x), f(x')) < \epsilon$. □

A map $f : X \to Y$ between metric spaces is **open** if for all open subsets $U \subset X$, $f(U)$ is open in Y. A map $f : X \to Y$ is a **homeomorphism** if it is continuous, is bijective, and the inverse function $f^{-1} : Y \to X$ is continuous. A map $f : X \to Y$ is a **topological embedding** if it is continuous, injective and open.

EXERCISE 2.44. *For a metric space X, let X_D be the same underlying set endowed with the discrete metric.*

a) *Show that the identity map $1 : X_D \to X$ is continuous.*
b) *Show that the identity map $1 : X \to X_D$ is continuous iff X is discrete (in the topological sense: every point of x is an isolated point).*

EXAMPLE 2.9. a) *Let X be a metric space that is not discrete. Show: the the identity map $1 : X_D \to X$ is bijective and continuous but not open. The identity map $1 : X \to X_D$ is bijective and open but not continuous.*

b) *Show: the map* $f : \mathbb{R} \to \mathbb{R}$ *by* $x \mapsto |x|$ *is continuous – indeed, Lipschitz with* $C = 1$ *– but not open:* $f(\mathbb{R}) = [0, \infty)$.

EXERCISE 2.45. *Let* $f : \mathbb{R} \to \mathbb{R}$.
a) Show that at least one of the following holds:
(i) f *is increasing: for all* $x_1 \leq x_2$, $f(x_1) \leq f(x_2)$.
(ii) f *is decreasing: for all* $x_1 \leq x_2$, $f(x_1) \geq f(x_2)$.
(iii) f *is of "*Λ*-type": there are* $a < b < c$ *such that* $f(a) < f(b) > f(c)$.
(iv) f *is of "*V*-type": there are* $a < b < c$ *such that* $f(a) > f(b) < f(c)$.
b) Suppose f *is a continuous injection. Show that* f *is strictly increasing or strictly decreasing.*
c) Let $f : \mathbb{R} \to \mathbb{R}$ *be increasing. Show that for all* $x \in \mathbb{R}$

$$\sup_{y<x} f(y) \leq f(x) \leq \inf_{y>x} f(y).$$

Show that

$$\sup_{y<x} f(y) = f(x) = \inf_{y>x} f(y)$$

iff f *is continuous at* x.
d) Suppose f *is bijective and strictly increasing. Show that* f^{-1} *is strictly increasing.*
e) Show that if f *is strictly increasing and surjective, it is a homeomorphism. Deduce that every continuous bijection* $f : \mathbb{R} \to \mathbb{R}$ *is a homeomorphism.*

LEMMA 2.19. *For a map* $f : X \to Y$ *between metric spaces, the following are equivalent:*

(i) *The map* f *is a homeomorphism.*
(ii) *The map* f *is continuous, bijective and open.*

EXERCISE 2.46. *Prove it.*

PROPOSITION 2.20. *Let* X, Y, Z *be metric spaces and* $f : X \to Y$, $g : Y \to Z$ *be continuous maps. Then* $g \circ f : X \to Z$ *is continuous.*

PROOF. Let W be open in Z. Since g is continuous, $g^{-1}(W)$ is open in Y. Since f is continuous, $f^{-1}(g^{-1}(W)) = (g \circ f)^{-1}(W)$ is open in X. □

PROPOSITION 2.21. *For a map* $f : X \to Y$ *of metric spaces, the following are equivalent:*

(i) *The map* f *is continuous.*
(ii) *If* $\mathbf{x}_n \to x$ *in* X, *then* $f(\mathbf{x}_n) \to f(x)$ *in* Y.

PROOF. (i) $\implies$ (ii) Let $\epsilon > 0$. Since f is continuous, by Lemma 2.18 there is $\delta > 0$ such that if $x' \in B^\circ(x, \delta)$, $f(x') \in B^\circ(x, \epsilon)$. Since $\mathbf{x}_n \to x$, there is $N \in \mathbb{Z}^+$ such that for all $n \geq N$, $\mathbf{x}_n \in B^\circ(x, \delta)$, and thus for all $n \geq N$, $f(\mathbf{x}_n) \in B^\circ(x, \epsilon)$.
$\neg$ (i) $\implies$ $\neg$ (ii): Suppose that f is not continuous: then there is $x \in X$ and $\epsilon > 0$ such that for all $n \in \mathbb{Z}^+$, there is $x_n \in X$ with $d(x_n, x) < \frac{1}{n}$ and $d(f(x_n), f(x)) \geq \epsilon$. Then $x_n \to x$ and $f(x_n)$ does not converge to $f(x)$. □

In other words, continuous functions between metric spaces are precisely the functions that preserve limits of convergent sequences.

EXERCISE 2.47. a) *Let* $f : X \to Y$, $g : Y \to Z$ *be maps of topological spaces. Let* $x \in X$. *Use* ϵ*'s and* δ*'s to show that if* f *is continuous at* x *and* g *is continuous at* $f(x)$ *then* $g \circ f$ *is continuous at* x. *Deduce another proof of Proposition 2.20 using the* (ϵ, δ)*-definition of continuity.*

b) *Give (yet) another proof of Proposition 2.20 using Proposition 2.21.*

In higher mathematics, one often meets the phenomenon of rival definitions which are equivalent in a given context (but may not be in other contexts of interest). Often a key part of learning a new subject is learning which versions of definitions give rise to the shortest, most transparent proofs of basic facts. When one definition makes a certain proposition harder to prove than another definition, it may be a sign that in some other context these definitions are not equivalent and the proposition is true using one but not the other definition. We will see this kind of phenomenon often in the transition from metric spaces to topological spaces. However, in the present context, all definintions in sight lead to immediate, straightforward proofs of "compositions of continuous functions are continuous". And indeed, though the concept of a continuous function can be made in many different general contexts (we will meet some, but not all, of these later), to the best of my knowledge it is always clear that compositions of continuous functions are continuous.

4.1. Further Exercises.

EXERCISE 2.48. *Let X be a metric space, and let $f, g : X \to \mathbb{R}$ be continuous functions. Show that $\{x \in X \mid f(x) < g(x)\}$ is open and $\{x \in X \mid f(x) \leq g(x)\}$ is closed.*

EXERCISE 2.49. a) *Let X be a metric space, and let $Y \subset X$. Let $\mathbf{1}_Y : X \to \mathbb{R}$ be the* ***characteristic function*** *of Y: for $x \in X$, $\mathbf{1}_Y(x) = 1$ if $x \in Y$ and 0 otherwise. Show that $\mathbf{1}_Y$ is* not *continuous at $x \in X$ iff $x \in \partial Y$.*

b) *Let $Y \subset \mathbb{R}^N$ be a bounded subset. Deduce that $\mathbf{1}_Y$ is Riemann integrable iff ∂Y has measure zero. (Such sets are called* ***Jordan measurable****.)*

EXERCISE 2.50. *Show that for a metric space X, the following are equivalent:*
(i) Every function $f : X \to X$ is continuous.
(ii) X is topologically discrete.

Recall that a G_δ-subset in a metric space (X, d) is a subset that can be expressed as the intersection of a countable family of open sets.

EXERCISE 2.51. *Let $f : X \to Y$ be a function between metric spaces, and let $Z := \{x \in X \mid f \text{ is continuous at } x\}$. Show: Z is a G_δ-subset of X.*

EXERCISE 2.52. *This exercise gives a* ***sequential characterization of uniform continuity****. For a function $f : X \to Y$ between metric spaces, show that the following are equivalent:*

(i) *The function $f : X \to Y$ is uniformly continuous.*
(ii) *If $x_\bullet$ and $y_\bullet$ are sequences in X such that $\lim_{n\to\infty} d(x_n, y_n) = 0$, then $\lim_{n\to\infty} d(f(x_n), f(y_n)) = 0$.*

5. Equivalent Metrics

It often happens in geometry and analysis that there is more than one natural metric on a set X and one wants to compare properties of these different metrics. Thus we are led to study equivalence relations on the class of metrics on a given set...but in fact it is part of the natural richness of the subject that there is more than one natural equivalence relation. We have already met the coarsest one we will consider here: two metrics d_1 and d_2 on X are **topologically equivalent** if they determine the same topology; equivalently, in view of X.X, for all sequences $\mathbf{x}$ in X and points x of X, we have $\mathbf{x} \stackrel{d_1}{\to} x \iff \mathbf{x} \stackrel{d_2}{\to} x$. Since continuity is characterized in terms of open sets, equivalent metrics on X give rise to the same class of continuous functions on X (with values in any metric space Y).

LEMMA 2.22. *Two metrics d_1 and d_2 on a set X are topologically equivalent iff the identity function $1_X : (X, d_1) \to (X, d_2)$ is a homeomorphism.*

PROOF. To say that 1_X is a homeomorphism is to say that 1_X is continuous from (X, d_1) to (X, d_2) and that its inverse – which also happens to be 1_X! – is continuous from (X, d_2) to (X, d_1). This means that every d_2-open set is d_1-open and every d_1-open set is d_2-open. $\square$

The above simple reformulation of topological equivalence suggests other, more stringent notions of equivalence of metrics d_1 and d_2, in terms of requiring $1_X : (X, d_1) \to (X, d_2)$ to have stronger continuity properties. Namely, we say that two metrics d_1 and d_2 are **uniformly equivalent** (resp. **Lipschitz equivalent**) if 1_X is uniformly continuous with a uniformly continuous inverse (resp. Lipschitz and with a Lipschitz inverse).

LEMMA 2.23. *Let d_1 and d_2 be metrics on a set X.*
a) The metrics d_1 and d_2 ae uniformly equivalent iff for all $\epsilon > 0$ there are $\delta_1, \delta_2 > 0$ such that for all $x_1, x_2 \in X$ we have

$$d_1(x_1, x_2) \leq \delta_1 \implies d_2(x_1, x_2) \leq \epsilon \text{ and } d_2(x_1, x_2) \leq \delta_2 \implies d_1(x_1, x_2) \leq \epsilon.$$

b) The metrics d_1 and d_2 are Lipschitz equivalent iff there are constants $C_1, C_2 \in (0, \infty)$ such that for all x_1, x_2 in X we have

$$C_1 d_2(x_1, x_2) \leq d_1(x_1, x_2) \leq C_2 d_2(x_1, x_2).$$

EXERCISE 2.53. *Prove it.*

REMARK 2.24. *The typical textbook treatment of metric topology is not so careful on this point: one must read carefully to see which of these equivalence relations is meant by "equivalent metrics".*

EXERCISE 2.54. *a) Explain how the existence of a homeomorphism of metric spaces $f : X \to Y$ which is not uniformly continuous can be used to construct two topologically equivalent metrics on X which are not uniformly equivalent. Then construct such an example, e.g. with $X = \mathbb{R}$ and $Y = (0, 1)$.*
b) Explain how the existence of a uniformeomorphism of metric spaces $f : X \to Y$ which is not a Lipschitzeomorphism can be used two construct two uniformly equivalent metrics on X which are not Lipschitz equivalent.
c) Exhibit a uniformeomorphism $f : \mathbb{R} \to \mathbb{R}$ which is not a Lipschitzeomorphism.
d) Show that $\sqrt{x} : [0, 1] \to [0, 1]$ is a uniformeomorphism and not a Lipschitzeomorphism.[1]

PROPOSITION 2.25. *Let (X, d) be a metric space. Let $f : [0, \infty) \to [0, \infty)$ be a continuous strictly increasing function with $f(0) = 0$, and suppose that $f \circ d : X \times X \to \mathbb{R}$ is a metric function. Then the metrics d and $f \circ d$ are uniformly equivalent.*

[1] In particular, compactness does *not* force continuous maps to be Lipschitz!

Proof. Let $A = f(1)$. The function $f : [0,1] \to [0,A]$ is continuous and strictly increasing, hence it has a continuous and strictly increasing inverse function $f^{-1} : [0,A] \to [0,1]$. Since $[0,1]$ and $[0,A]$ are compact metric spaces, f and f^{-1} are in fact uniformly continuous. The result follows easily from this, as we leave to the reader to check. $\square$

In particular that for any metric d on a set X and any $\alpha > 0$, the metric $d_\alpha(x,y) = \frac{\alpha d(x,y)}{d(x,y)+1}$ of Corollary 2.3 is uniformly equivalent to d. In particular, every metric is uniformly equivalent to a metric with diameter at most α. The following exercise gives a second, convexity-free approach to this.

Exercise 2.55. *Let (X,d) be a metric space, and let $d_b : X \times X \to \mathbb{R}$ be given by $d_b(x,y) = \min(d(x,y), 1)$.*

a) *Show that d_b is a bounded metric on X that is uniformly equivalent to d.*
b) *Show that d_b is Lipschitz equivalent to d iff (X,d) is bounded.*

We will call the metric d_b of Exercise 2.55 the **standard bounded metric** associated to the metric d. Having it at our disposal makes for a good time to mildly generalize our notion of a metric. One sees many places in topology and analysis where working with extended real numbers rather than just real numbers conveys technical advantages. What about allowing a metric to take $+\infty$ as a value? For a set X, we say that a function $d : X \times X \to [0,\infty]$ is an **extended metric** (or **emetric**) if it satisfies the properties (M1), (M2) and (M3) as a metic: notice that addition is well-defined on the non-negative extended real numbers: for all $x \in [0,\infty]$ we have $x + \infty = \infty + x = \infty$. We can then define open and closed balls for an emetric in exactly the same way. Also as for a standard metric, we define a subset of an emetric space to be open if it is a union of open balls, and we call the family of open sets the **emetric topology**.

The following exercise shows that emetrics can be reduced to metrics to the same extent that metrics can be reduced to bounded metrics.

Exercise 2.56. *Let (X,d) be an emetric space. Define $d_b : X \times X \to \mathbb{R}$ by $d_b(x,y) = \min(d(x,y), 1)$.*

a) *Show: (X, d_b) is a bounded metric space.*
b) *Show: the emetric topology on (X,d) is the metric topology on (X, d_b).*
c) *Say what it means for two emetric spaces to be uniformly equivalent, and show that (X,d) and (X,d_b) are uniformly equivalent.*

6. Product Metrics

6.1. Minkowski's Inequality.

Theorem 2.26. *(Jensen's Inequality) Let $f : I \to \mathbb{R}$ be continuous and convex. For any $x_1, \ldots, x_n \in I$ and any $\lambda_1, \ldots, \lambda_n \in [0,1]$ with $\lambda_1 + \ldots + \lambda_n = 1$, we have*

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n).$$

Proof. We go by induction on n, the base case $n = 1$ being trivial. So suppose Jensen's Inequality holds for some $n \in \mathbb{Z}^+$, and consider $x_1, \ldots, x_{n+1} \in I$ and $\lambda_1, \ldots, \lambda_{n+1} \in [0,1]$ with $\lambda_1 + \ldots + \lambda_{n+1} = 1$. If $\lambda_{n+1} = 0$ we are reduced to the case of n variables which holds by induction. Similarly if $\lambda_{n+1} = 1$ then

$\lambda_1 = \ldots = \lambda_n = 0$ and we have, trivially, equality. So we may assume $\lambda_{n+1} \in (0,1)$ and thus also that $1 - \lambda_{n+1} \in (0,1)$. Now for the big trick: we write

$$\lambda_1 x_1 + \ldots + \lambda_{n+1} x_{n+1} = (1-\lambda_{n+1}) \left(\frac{\lambda_1}{1-\lambda_{n+1}} x_1 + \ldots + \frac{\lambda_n}{1-\lambda_{n+1}} x_n \right) + \lambda_{n+1} x_{n+1},$$

so that

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) = f((1-\lambda_{n+1})(\frac{\lambda_1}{1-\lambda_{n+1}} x_1 + \ldots + \frac{\lambda_n}{1-\lambda_{n+1}} x_n) + \lambda_{n+1} x_{n+1})$$

$$\leq (1-\lambda_{n+1}) f \left(\frac{\lambda_1}{1-\lambda_{n+1}} x_1 + \ldots + \frac{\lambda_n}{1-\lambda_{n+1}} x_n \right) + \lambda_{n+1} f(x_{n+1}).$$

Since $\frac{\lambda_1}{1-\lambda_{n+1}}, \ldots, \frac{\lambda_n}{1-\lambda_{n+1}}$ are non-negative numbers that sum to 1, by induction the n variable case of Jensen's Inequality can be applied to give that the above expression is less than or equal to

$$(1-\lambda_{n+1}) \left(\frac{\lambda_1}{1-\lambda_{n+1}} f(x_1) + \ldots + \frac{\lambda_n}{1-\lambda_{n+1}} f(x_n) \right) + \lambda_{n+1} f(x_{n+1})$$

$$= \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n) + \lambda_{n+1} f(x_{n+1}). \qquad \square$$

THEOREM 2.27. *(Weighted Arithmetic Geometric Mean Inequality) Let* $x_1, \ldots, x_n \in [0, \infty)$ *and* $\lambda_1, \ldots, \lambda_n \in [0,1]$ *be such that* $\lambda_1 + \ldots + \lambda_n = 1$. *Then:*

$$x_1^{\lambda_1} \cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \ldots + \lambda_n x_n. \tag{6}$$

Taking $\lambda_1 = \ldots = \lambda_n = \frac{1}{n}$, *we get the* ***arithmetic geometric mean inequality****:*

$$(x_1 \cdots x_n)^{\frac{1}{n}} \leq \frac{x_1 + \ldots + x_n}{n}.$$

PROOF. We may assume $x_1, \ldots, x_n > 0$. For $1 \leq i \leq n$, put $y_i = \log x_i$. Then

$$x_1^{\lambda_1} \cdots x_n^{\lambda_n} = e^{\log(x_1^{\lambda_1} \cdots x_n^{\lambda_n})} = e^{\lambda_1 y_1 + \ldots + \lambda_n y_n} \leq \lambda_1 e^{y_1} + \ldots + \lambda_n e^{y_n} = \lambda_1 x_1 + \ldots + \lambda_n x_n.$$

$\square$

THEOREM 2.28. *(Young's Inequality)*
Let $x, y \in [0, \infty)$ *and let* $p, q \in (1, \infty)$ *satisfy* $\frac{1}{p} + \frac{1}{q} = 1$. *Then*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}. \tag{7}$$

PROOF. When either $x = 0$ or $y = 0$ the left hand side is zero and the right hand side is non-negative, so the inequality holds and we may thus assume $x, y > 0$. Now apply the Weighted Arithmetic-Geometric Mean Inequality with $n = 2$, $x_1 = x^p$, $x_2 = y^q$, $\lambda_1 = \frac{1}{p}$, $\lambda_2 = \frac{1}{q}$. We get

$$xy = (x^p)^{\frac{1}{p}} (y^q)^{\frac{1}{q}} = x_1^{\lambda_1} x_2^{\lambda_2} \leq \lambda_1 x_1 + \lambda_2 x_2 = \frac{x^p}{p} + \frac{y^q}{q}. \qquad \square$$

THEOREM 2.29. *(Hölder's Inequality)*
Let $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$ *and let* $p, q \in (1, \infty)$ *satisfy* $\frac{1}{p} + \frac{1}{q} = 1$. *Then*

$$|x_1 y_1| + \ldots + |x_n y_n| \leq (|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} (|y_1|^q + \ldots + |y_n|^q)^{\frac{1}{q}}. \tag{8}$$

Proof. Again the result is clear if $x_1 = \ldots = x_n = 0$ or $y_1 = \ldots = y_n = 0$, so we may assume that neither of these is the case. For $1 \leq i \leq n$, apply Young's Inequality with

$$x = \frac{|x_i|}{(|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}}}, y = \frac{|y_i|}{(|y_1|^q + \ldots + |y_n|^q)^{\frac{1}{q}}},$$

and sum the resulting inequalities from $i = 1$ to n, getting

$$\frac{\sum_{i=1}^n |x_i y_i|}{(|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} (|y_1|^q + \ldots + |y_n|^q)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1. \qquad \square$$

Theorem 2.30. *(Minkowski's Inequality)*
For $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$ and $p \geq 1$, we have

$$(9) \quad (|x_1 + y_1|^p + \ldots + |x_n + y_n|^p)^{\frac{1}{p}} \leq (|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} + (|y_1|^p + \ldots |y_n|^p)^{\frac{1}{p}}$$

Proof. When $p = 1$, the inequality reads

$$|x_1 + y_1| + \ldots + |x_n + y_n| \leq |x_1| + |y_1| + \ldots + |x_n| + |y_n|$$

and this holds just by applying the triangle inequality: for all $1 \leq i \leq n$, $|x_i + y_i| \leq |x_i| + |y_i$. So we may assume $p > 1$. Let q be such that $\frac{1}{p} + \frac{1}{q} = 1$, and note that then $(p-1)q = p$. We have

$$|x_1 + y_1|^p + \ldots + |x_n + y_n|^p$$

$$\leq |x_1||x_1+y_1|^{p-1} + \ldots + |x_n||x_n+y_n|^{p-1} + |y_1||x_1+y_1|^{p-1} + \ldots + |y_n||x_n+y_n|^{p-1} \overset{\text{HI}}{\leq}$$

$$(|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} (|x_1+y_1|^p + \ldots + |x_n+y_n|^p)^{\frac{1}{q}} + (|y_1|^p + \ldots + |y_n|^p)^{\frac{1}{p}} (|x_1+y_1|^p + \ldots + |x_n+y_n|^p)^{\frac{1}{q}}$$

$$= \left((|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} + (|y_1|^p + \ldots |y_n|^p)^{\frac{1}{p}} \right) (|x_1 + y_1|^p + \ldots + |x_n + y_n|^p)^{\frac{1}{q}}.$$

Dividing both sides by $(|x_1 + y_1|^p + \ldots + |x_n + y_n|^p)^{\frac{1}{q}}$ and using $1 - \frac{1}{q} = \frac{1}{p}$, we get the desired result. $\square$

For $p \in [1, \infty)$ and $x \in \mathbb{R}^N$, we put

$$||x||_p = \left(\sum_{i=1}^N |x_i|^p \right)$$

and

$$d_p : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}, \ d_p(x, y) = ||x - y||_p.$$

We also put

$$||x||_\infty = \max_{1 \leq i \leq N} |x_i|$$

and

$$d_\infty : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}, \ d_\infty(x, y) = ||x - y||_\infty.$$

Lemma 2.31. *a) For each fixed nonzero $x \in \mathbb{R}^N$, the function $p \mapsto ||x||_p$ is decreasing and $\lim_{p \to \infty} ||x||_p = ||x||_\infty$.*
b) For all $1 \leq p \leq \infty$ and $x \in \mathbb{R}^N$ we have

$$||x||_\infty \leq ||x||_p \leq ||x||_1 = |x_1| + \ldots + |x_N| \leq N||x||_\infty.$$

Proof. a) Let $1 \leq p \leq p' < \infty$, and let $0 \neq x = (x_1, \ldots, x_N) \in \mathbb{R}^N$. For any $\alpha \geq 0$ we have $||\alpha x||_p = |\alpha| ||x||_p$, so we are allowed to rescale: put $y = (\frac{1}{||x||_{p'}})x$, so $||y||_{p'} \leq 1$. Then $|y_i| \leq 1$ for all i, so $|y_i|^{p'} \leq |y_i|^p$ for all i, so $||y||_p \geq 1$ and thus $||x||_p \geq ||x||_{p'}$.

Similarly, by scaling we reduce to the case in which the maximum of the $|x_i|$'s is equal to 1. Then in $\lim_{p \to \infty} |x_1|^p + \ldots + |x_N|^p$, all of the terms $|x_i|^p$ with $|x_i| < 1$ converge to 0 as $p \to \infty$; the others converge to 1; so the given limit is the number of terms with absolute value 1, which lies between 1 and N: that is, it is always at least one and it is bounded independently of p. Raising this to the $1/p$ power and taking the limit we get 1.
b) The inequalities $||x||_\infty \leq ||x||_p \leq ||x||_1$ follow from part a). For the latter inequality, let $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and suppose that i is such that $|x_i| = \max_{1 \leq i \leq N} |x_i|$. Then

$$|x_1| + \ldots + |x_N| \leq |x_i| + \ldots |x_i| = N||x||_\infty. \qquad \square$$

Theorem 2.32. *For each $p \in [1, \infty)$, d_p is a metric on $\mathbb{R}^N$, and all of these metrics are Lipschitz equivalent.*

Proof. For any $1 \leq p \leq \infty$ and $x, y, z \in \mathbb{R}^N$, Minkowski's Inequality gives

$$d_p(x,z) = ||x-z||_p = ||(x-y)+(y-z)||_p \leq ||x-y||_p + ||y-z||_p = d_p(x,y) + d_p(x,z).$$

Thus d_p satisfies the triangle inequality; that $d_p(x,y) = d_p(y,x)$ and $d_p(x,y) = 0 \iff x = y$ is immediate. So each d_p is a metric on $\mathbb{R}^N$. Lemma 2.31 shows that for all $1 \leq p \leq \infty$, d_p is Lipschitz equivalent to d_∞. Since Lipschitz equivalence is indeed an equivalence relation, this implies that all the metrics d_p are Lipschitz equivalent. $\square$

The metric d_2 on $\mathbb{R}^N$ is called the **Euclidean metric**. The topology that it generates is called the **Euclidean topology**. The point of the above discussion is that all metrics d_p are close enough to the Euclidean metric so as to generate the Euclidean topology.

6.2. Product Metrics.

Let $(X_i, d_i)_{i \in I}$ be an indexed family of metric spaces. Our task is to put a metric on the Cartesian product $X = \prod_{i \in I} X_i$.[2]

Well, but that can't be right: we have already put *some* metric on an arbitrary set, namely the discrete metric. Rather we want to put a metric on the product which usefully incorporates the metrics on the factors, in a way which generalizes the metrics d_p on $\mathbb{R}^N$.

This is still not precise enough. We are lingering over this point a bit to emphasize the fundamental perspective of general topological spaces that we currently lack: eventually we will discuss the **product topology**, which is a canonically defined topology on any Cartesian product of topological spaces. With this perspective, the

[2] Trivial remark: We have $X = \varnothing$ iff $X_i = \varnothing$ for some $i \in I$. When this holds, there is a unique metric on X – evidently this is a trivial case. From now until the end of this section, when we consider an indexed family $\{(X_i, d_i)\}_{i \in I}$ of metric spaces, we will tacitly assume that $X_i \neq \varnothing$ for all $i \in I$ (and also that $I \neq \varnothing$!)

problem can then be gracefully phrased as that of finding a metric on a Cartesian product of metric spaces that induces the product topology. For now we bring out again our most treasured tool: sequences. Namely, convergence in the Euclidean metric on $\mathbb{R}^N$ has the fundamental property that a sequence $\mathbf{x}$ in $\mathbb{R}^N$ converges iff for all $1 \leq i \leq N$, its ith component sequence $\mathbf{x}^{(i)}$ converges in $\mathbb{R}$.

In general, let us say that a metric on $X = \prod_{i \in I} X_i$ is **good** if for any sequence $\mathbf{x}$ in X and point $x \in X$, we have $\mathbf{x} \to x$ in X iff for all $i \in I$, the component sequence $\mathbf{x}^{(i)}$ converges to the ith component $x^{(i)}$ of x.

In the case of finite products, we have already done almost all of the work.

LEMMA 2.33. *For $1 \leq i \leq N$, let $\{x_n^{(i)}\}$ be a sequence of non-negative real numbers, and for $n \in \mathbb{Z}^+$ let $m_n = \max_{1 \leq i \leq N} x_n^{(i)}$. Then $m_n \to 0 \iff x_n^{(i)} \to 0$ for all $1 \leq i \leq N$.*

EXERCISE 2.57. *Prove it.*

THEOREM 2.34. *Let $(X_1, d_1), \ldots, (X_N, d_N)$ be a finite sequence of metric spaces, and put $X = \prod_{i=1}^N X_i$. Fix $p \in [1, \infty]$, and consider the function*

$$d_p : X \times X \to \mathbb{R}, \ d_p((x_1, \ldots, x_N), (y_1, \ldots, y_N)) = \left(\sum_{i=1}^N |d_i(x_i, y_i)|^p \right)^{\frac{1}{p}}.$$

a) The function d_p is a metric function on X.
b) For $p, p' \in [1, \infty]$, the metrics d_p and $d_{p'}$ are Lipschitz equivalent.
*c) The function d_p is a **good metric** on X.*

PROOF. If each X_i is $\mathbb{R}$ with the standard Euclidean metric, then parts a) and b) reduce to Theorem 2.32 and part c) is a familiar (and easy) fact from basic real analysis: a sequence in $\mathbb{R}^N$ converges iff each of its component sequences converge. The proofs of parts a) and b) in the general case are almost identical and are left to the reader as a straightforward but important exercise.

In view of part b), it suffices to establish part c) for any one value of p, and the easiest is probably $p = \infty$, since $d_\infty(x, y) = \max_i d_i(x_i, y_i)$. If $\mathbf{x}$ is a sequence in X and x is a point of X, we are trying to show that

$$d_\infty(\mathbf{x}_n, x) = \max d_i(\mathbf{x}_n^{(i)}, x^{(i)}) \to 0 \iff \forall 1 \leq i \leq N, \ d_i(\mathbf{x}_n^{(i)}, x^{(i)}) \to 0.$$

This follows from Lemma 2.33. □

Here is one simple but useful application.

PROPOSITION 2.35. *Let (X, d) be a metric space. Endowing $X \times X$ with the good metric d_∞, the metric function $d : X \times X \to \mathbb{R}$ is Lipschitz continuous.*

PROOF. Fix $\delta > 0$, and let $P_1 = (x_1, y_1), \ P_2 = (x_2, y_2) \in X \times X$ be such that $d_\infty(P_1, P_2) < \delta$. Then

$$d_\infty(P_1, P)2) = \max\left(d(x_1, x_2), \ d(y_1, y_2)\right) < \delta.$$

So the Quadrilateral Inequality (Exercise 2.15) implies

$$d_{\mathbb{R}}(d(P_1), d(P_2)) = |d(P_1) - d(P_2)| \leq d(x_1, x_2) + d(y_1, y_2) < \delta + \delta = 2\epsilon.$$

This shows that d is Liptschitz with Lipshitz constant 2. □

EXERCISE 2.58. *Let $N \in \mathbb{Z}^+$.*
a) Show that the standard maps $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto x + y$ and $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto xy$ are continuous. (Here we may take any good metric on $\mathbb{R} \times \mathbb{R}$.)
b) Show that the standard maps $+ : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$, $((x_1, \ldots, x_N), (y_1, \ldots, y_N)) \mapsto (x_1 + y_1, \ldots, x_N + y_N)$, $\cdot : \mathbb{R} \times \mathbb{R}^N$, $(\alpha, (x_1, \ldots, x_N)) \mapsto (\alpha x_1, \ldots, \alpha x_N))$ are continuous.

For infinite families of metric spaces, things get more interesting. The following is a variant of [**Du**, Thm. IX.7.2].

THEOREM 2.36. *Let I be an infinite set, let $(X_i, d_i)_{i \in I}$ be an indexed family of metric spaces, let $X = \prod_{i \in I} X_i$, and let*

$$d : X \times X \to [0, \infty],\ d(x, y) = \sup_{i \in I} d_i(x_i, y_i).$$

a) *The following are equivalent:*
  (i) *There is a finite subset $J \subset I$ and $D < \infty$ such that for all $i \in I \setminus J$ we have* $\operatorname{diam} X_i \leq D$.
  (ii) *We have $d(x, y) < \infty$ for all $x, y \in X$.*
  (iii) *The function d is a metric on X.*

b) *The following are equivalent:*
  (i) *d is a good metric.*
  (ii) *For all $\delta > 0$, $\{i \in I \mid \operatorname{diam} X_i \geq \delta\}$ is finite.*

PROOF. a) (i) $\iff$ (ii): If (i) holds, then for all $x, y \in X$, $\sup_i d_i(x_i, y_i)$ is the supremum over the union of a finite set and a bounded set of real numbers, hence it is finite. If (i) fails, then there is an injective function $i_\bullet : \mathbb{Z}^+ \to I$ such that for all $n \in \mathbb{Z}^+$ there are points $x_{i_n}, y_{i_n} \in X_{i_n}$ with $d(x_{i_n}, y_{i_n}) \geq n$. Then if x (resp. y) is any elements of X with i_n coordinate equal to x_{i_n} (resp. y_{i_n}), then $d(x, y) = \infty$.
(ii) $\implies$ (iii): This is quite straightforward. We will show the least trivial (M3): let $x = \{x_i\}$, $y = \{y_i\}$, $z = \{z_i\}$ be three points of X. Then

$$d(x, z) = \sup_i d_i(x_i, z_i) \leq \sup_i d_i(x_i, y_i) + d_i(y_i, z_i)$$

$$\leq \sup_i d_i(x_i, y_i) + \sup_i d_i(y_i, z_i) = d(x, y) + d(y, z).$$

(iii) $\implies$ (ii): In order to be a metric, d must be finite-valued.
b) $\neg$ (ii) $\implies$ $\neg$ (i): If (ii) fails, then there is $\delta > 0$ and an injection $i_\bullet : \mathbb{Z}^+ \hookrightarrow I$ and for all $n \in \mathbb{Z}^+$ points $x_{i_n}, y_{i_n} \in X_{i_n}$ such that $d_{i_n}(x_{i_n}), y_{i_n}) \geq \delta$. For every $i \in J := I \setminus x_\bullet(\mathbb{Z}^+)$, fix a point $z_i \in X_i$. We build a sequence $\{x^{(n)}\}$ in X as follows: for each $j \in J$, we let $(x^{(n)})_j = z_j$ for all $n \in \mathbb{Z}^+$; that is, the jth component sequence is constant. For $m, n \in \mathbb{Z}^+$, we put

$$x_{i_m}^{(n)} = \begin{cases} x_{i_n} & n \leq m \\ y_{i_n} & n > m. \end{cases}$$

That is, the i_m-component sequence has x_{i_m} as its first m values and y_{i_m} for all subsequent values; in particular it converges to y_{i_n}. However, the sequence $\{x^{(n)}\}$ does not converge to the element x with i_n-component y_{i_n} for all $n \in \mathbb{Z}^+$ and j-component z_j for all $j \in J$, since for all $n \in \mathbb{Z}^+$, we have

$$d(x^{(n)}, x) \geq d_{i_n}(x_{i_n}^{(n)}, x_{i_n}) = d_{i_n}(x_{i_n}, y_{i_n}) \geq \delta.$$

(ii) $\implies$ (i): Let $\{x^{(n)}\}$ be a sequence in X such that for all $i \in I$, the ith component sequence $\{x_i^{(n)}\}$ converges to $x_i \in X_i$. Put $x \coloneqq \{x_i\}_{i \in I}$; we will show that $x^{(n)} \to x$. Fix $\epsilon > 0$, and let J be the finite subset of i such that for $j \in J$ we have $\operatorname{diam}(X_j) > \epsilon$. For each $j \in J$, choose $N_j \in \mathbb{Z}^+$ such that for all $n \geq N_j$ we have $d_j(x_j^{(n)}, x_j) \leq \epsilon$. Then for all $n \geq N \coloneqq \max_{j \in J} N_j$ and all $i \in I$, we have $d_i(x_i^{(n)}, x_i) \leq \epsilon$ an thus $d(x^{(n)}, x) = \sup_{i \in I} d_i(x_i^{(n)}, x_i) \leq \epsilon$. □

COROLLARY 2.37. *Let $\{X_n, d_n\}_{n=1}^{\infty}$ be an infinite sequence of metric spaces. Then there is a good metric on the Cartesian product $X = \prod_{i=1}^{\infty} X_n$.*

PROOF. The sequence of metrics need not satisfy the hypotheses of Theorem 2.36, but we can replace each d_n with a topologically equivalent metric so that the hypotheses hold. Indeed, the metric $d_n' = \frac{d_n}{2^n(d_n+1)}$ of Corollary 2.3 is topologically equivalent to d_n and has diameter at most $\frac{1}{2^n}$. The family (X_n, d_n') satisfies the hypotheses of Theorem 2.36b), so $d = \sup_n d_n'$ is a good metric on X. □

Corollary 2.37 shows in particular that $\prod_{i=1}^{\infty} \mathbb{R}$ and $\prod_{i=1}^{\infty}[a, b]$ can be given metrics so that convergence amounts to convergence in each factor. These are highly interesting and important examples in the further study of analysis and topology. The space $\prod_{i=1}^{\infty}[0, 1]$ is often called the **Hilbert cube**.

PROPOSITION 2.38. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of nonempty metric spaces, and let $X = \prod_{n=1}^{\infty} X_n$, endowed with a good metric via Corollary 2.37. For $n \in \mathbb{Z}^+$, let $\pi_n : X \to X_n$ be the projection map $\{x^{(n)}\} \mapsto x_n$.*
a) The map $\pi_n : X \to X_n$ is continuous.
b) Let M be a metric space, and let $f : M \to X$ be a function. The following are equivalent:
(i) The map $f : M \to X$ is continuous.
(ii) For all $n \in \mathbb{Z}^+$, the map $\pi_n \circ f : M \to X_n$ is continuous.

PROOF. The key is Proposition 2.21, which characterizes continuous maps between metric spaces as those that preserve limits of sequences.
a) By definition of a good metric, if $x^{(m)} \to x$, then for all $n \in \mathbb{Z}^+$ we have $\pi_n(x^{(m)}) = x_n^{(m)} \to x_n = \pi_n(x)$, so π_n is continuous.
b) (i) $\implies$ (ii): The composition of continuous functions is continuous.
(ii) $\implies$ (i): Let $\{m_\bullet\}$ be a sequence in M that converges to M. By our assumption, for all $n \in \mathbb{Z}^+$, the sequence $\pi_n(f(m_\bullet)$ converges to $\pi_n(f(m))$. Then, by the definition of a good metric, $f(m_\bullet) \to f(m)$, so f is continuous. □

There is a case left over: what happens when we have a family of metrics indexed by an uncountable set I? In this case the condition that all but finitely many factors have diameter less than any given positive constant turns out to be prohibitively strict.

EXERCISE 2.59. *Let $\{X_i, d_i\}_{i \in I}$ be a family of metric spaces indexed by an uncountable set I. Suppose that* $\operatorname{diam} X_i > 0$ *for uncountably many $i \in I$ – equivalently, uncountably many X_i contains more than one point. Show that there is $\delta > 0$ such that $\{i \in I \mid \operatorname{diam} X_i \geq \delta\}$ is uncountable.*

Thus Theorem 2.37 can never be used to put a good metric on an uncountable product except in the trivial case that all but countably many of the spaces X_i

consist of a single point. (Nothing is gained by taking Cartesian products with one-point sets: this is the multiplicative equivalent of repeatedly adding zero!) At the moment this seems like a weakness of the result. Later we will see that is is essential: the Cartesian product of an uncountable family of metric spaces each consisting of more than a single point cannot in fact be given any good metric. In later terminology, this is an instance of nonmetrizability of large Cartesian products.

Exercise 2.60. *Let X and Y be metric spaces, and let $X \times Y$ be endowed with any good metric. Let $f : X \to Y$ be a function.*
a) Show that if f is continuous, its ***graph*** *$G(f) = \{(x, f(x) \mid x \in X\}$ is a closed subset of $X \times Y$.*
b) Give an example of a function $f : [0, \infty) \to [0, \infty)$ which is discontinuous at 0 but for which $G(f)$ is closed in $[0, \infty) \times [0, \infty)$.

7. Compactness

7.1. Basic Properties of Compactness.

Let X be a metric space, and let $A \subset X$. A family $\{Y_i\}_{i \in I}$ of subsets of X is a **covering** of A if $A \subset \bigcup_{i \in I} Y_i$. A subset $A \subset X$ is **compact** if for every open covering $\{U_i\}_{i \in I}$ of A there is a finite subset $J \subset I$ such that $\{U_i\}_{i \in J}$ covers A.

Exercise 2.61. *Show (directly) that $A = \{0\} \cup \{\frac{1}{n}\}_{n=1}^{\infty} \subset \mathbb{R}$ is compact.*

Exercise 2.62. *Let X be a metric space, and let $A \subset X$ be a finite subset. Show that A is compact.*

Lemma 2.39. *Let X be a metric space, and let $K \subset Y \subset X$. Then K is compact as a subset of Y if and only if K is compact as a subset of X.*

Proof. Suppose K is compact as a subset of Y, and let $\{U_i\}_{i \in I}$ be a family of open subsets of X such that $K \subset \bigcup_{i \in I} U_i$. Then $\{U_i \cap Y\}_{i \in I}$ is a covering of K by open subsets of Y, and since K is compact as a subset of Y, there is a finite subset $J \subset I$ such that $K \subset \bigcup_{i \in J} U_i \cap Y \subset \bigcup_{i \in J} U_i$.

Suppose K is compact as a subset of X, and let $\{V_i\}_{i \in I}$ be a family of open subsets of Y such that $K \subset \bigcup_{i \in I} V_i$. By X.X we may write $V_i = U_i \cap Y$ for some open subset of X. Then $K \subset \bigcup_{i \in I} V_i \subset \bigcup_{i \in I} U_i$, so there is a finite subset $J \subset I$ such that $K \subset \bigcup_{i \in J} U_i$. Intersecting with Y gives

$$K = K \cap Y \subset \left(\bigcup_{i \in J} U_i \right) \cap Y = \bigcup_{i \in J} V_i. \qquad \square$$

A sequence $\{A_n\}_{n=1}^{\infty}$ of subsets of X is **expanding** if $A_n \subset A_{n+1}$ for all $n \geq 1$. We say the sequence is **properly expanding** if $A_n \subsetneq A_{n+1}$ for all $n \geq 1$. An **expanding open cover** is an expanding sequence of open subsets with $X = \bigcup_{n=1}^{\infty} A_i$; we define a properly expanding open covering similarly.

Exercise 2.63. *Let $\{A_n\}_{n=1}^{\infty}$ be a properly expanding open covering of X.*
a) Let $J \subset \mathbb{Z}^+$ be finite, with largest element N. Show that $\bigcup_{i \in J} A_i = A_N$.
b) Suppose that an expanding open covering $\{A_n\}_{n=1}^{\infty}$ admits a finite subcovering. Show that there is $N \in \mathbb{Z}^+$ such that $X = A_N$.
c) Show that a properly expanding open covering has no finite subcovering, and thus if X admits a properly expanding open covering it is not compact.

An open covering $\{U_i\}_{i\in I}$ is **disjoint** if for all $i \neq j$, $U_i \cap U_j = \varnothing$.

Exercise 2.64. *a) Let $\{U_i\}_{i\in I}$ be a disjoint open covering of X. Show that the covering admits no* proper *subcovering.*
b) Show: if X admits an infinite disjoint open covering, it is not compact.
c) Show: a discrete space is compact iff it is finite.

Any property of a metric space formulated in terms of open sets may, by taking complements, also be formulated in terms of closed sets. Doing this for compactness we get the following simple but useful criterion.

Proposition 2.40. *For a metric space X, the following are equivalent:*
(i) *The space X is compact.*
(ii) *The space X satisfies the* ***finite intersection property****: if $\{A_i\}_{i\in I}$ is a family of closed subsets of X such that for all finite subsets $J \subset I$ we have $\bigcap_{i\in J} A_i \neq \varnothing$, then $\bigcap_{i\in I} A_i \neq \varnothing$.*

Exercise 2.65. *Prove it.*

Another easy but crucial observation is that compactness is somehow antithetical to discreteness. More precisely, we have the following result.

Proposition 2.41. *For a metric space X, the following are equivalent:*
(i) X is both compact and topologically discrete.
(ii) X is finite.

Exercise 2.66. *Prove it.*

Lemma 2.42. *Let X be a metric space and $A \subset X$.*
a) If X is compact and A is closed in X, then A is compact.
b) If A is compact, then A is closed in X.
c) If X is compact, then X is bounded.

Proof. a) Let $\{U_i\}_{i\in I}$ be a family of open subsets of X that covers A: i.e., $A \subset \bigcup_{i\in I} U_i$. Then the family $\{U_i\}_{i\in I} \cup \{X \setminus A\}$ is an open covering of X. Since X is compact, there is a finite subset $J \subset I$ such that $X = \bigcup_{i\in J} U_i \cup (X \setminus A)$, and it follows that $A \subset \bigcup_{i\in J} U_i$.
b) Let $U = X \setminus A$. For each $p \in U$ and $q \in A$, let $V_q = B(p, \frac{d(p,q)}{2})$ and $W_q = B(p, \frac{d(p,q)}{2})$, so $V_q \cap W_q = \varnothing$. Moreover, $\{W_q\}_{q\in A}$ is an open covering of the compact set A, so there are finitely many points $q_1, \ldots, q_n \in A$ such that

$$A \subset \bigcup_{i=1}^{n} W_i =: W,$$

say. Put $V = \bigcap_{i=1}^{n} V_i$. Then V is a neighborhood of p which does not intersect W, hence lies in $X \setminus A = U$. This shows that $U = X \setminus A$ is open, so A is closed.
c) Let $x \in X$. Then $\{B^\circ(x,n)\}_{n=1}^\infty$ is an expanding open covering of X; since X is compact, we have a finite subcovering. By Exercise 2, we have $X = B^\circ(x, N)$ for some $N \in \mathbb{Z}^+$, and thus X is bounded. $\square$

Example 2.10. *Let $X = [0, 10] \cap \mathbb{Q}$ be the set of rational points on the unit interval. As a subset of itself, X is closed and bounded. For $n \in \mathbb{Z}^+$, let*

$$U_n = \{x \in X \mid d(x, \sqrt{2}) > \frac{1}{n}\}.$$

Then $\{U_n\}_{n=1}^\infty$ is a properly expanding open covering of X, so X is not compact.

Proposition 2.43. *Let $f : X \to Y$ be a surjective continuous map of topological spaces. If X is compact, so is Y.*

Proof. Let $\{V_i\}_{i\in I}$ be an open cover of Y. For $i \in I$, put $U_i = f^{-1}(V_i)$. Then $\{U_i\}_{i\in I}$ is an open cover of X. Since X is compact, there is a finite $J \subset I$ such that $\bigcup_{i\in J} U_i = X$, and then $Y = f(X) = f(\bigcup_{i\in J} U_i) = \bigcup_{i\in J} f(U_i) = \bigcup_{i\in J} V_i$. □

Theorem 2.44 (Extreme Value Theorem). *Let X be a compact metric space. A continuous function $f : X \to \mathbb{R}$ is bounded and attains its maximum and minimum: there are $x_m, x_M \in X$ such that for all $x \in X$, $f(x_m) \leq f(x) \leq f(x_M)$.*

Proof. Since $f(X) \subset \mathbb{R}$ is compact, it is closed and bounded. Thus $\inf f(X)$ is a finite limit point of $f(X)$, so it is the minimum; similarly $\sup f(X)$ is the maximum. □

7.2. Heine-Borel.

When one meets a new metric space X, it is natural to ask: which subsets A of X are compact? Lemma 2.42 gives the necessary condition that A must be closed and bounded. In an arbitrary metric space this is nowhere near sufficient, and one need look no farther than an infinite set endowed with the discrete metric: every subset is closed and bounded, but the only compact subsets are the finite subsets. In fact, compactness is a topological property whereas we saw in §6 that given any metric space there is a topologically equivalent bounded metric.

Nevertheless in *some* metric spaces it is indeed the case that every closed, bounded set is compact. In this section we give a concrete treatment that Euclidean space $\mathbb{R}^N$ has this property: this is meant to be a reminder of certain ideas from honors calculus / elementary real analysis that we will shortly want to abstract and generalize.

A sequence $\{A_n\}_{n=1}^{\infty}$ of subsets of X is **nested** if $A_{n+1} \supset A_n$ for all $n \geq 1$.

Let $a_1 \leq b_1, a_2 \leq b_2, \ldots, a_n \leq b_n$ be real numbers. We put

$$\prod_{i=1}^{n}[a_i, b_i] = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid \forall 1 \leq i \leq n,\ a_i \leq x_i \leq b_i\}.$$

We will call such sets **closed boxes**.

Exercise 2.67.
a) Show: a subset $A \subset \mathbb{R}^n$ is bounded iff it is contained in some closed box.
b) Show that

$$\operatorname{diam}\left(\prod_{i=1}^{n}[a_i, b_i]\right) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}.$$

Lemma 2.45 (Lion-Hunting Lemma). a) *Let $\{\mathcal{B}_m\}_{m=1}^{\infty}$ be a nested sequence of closed boxes in $\mathbb{R}^n$. Then there is $x \in \bigcap_{m=1}^{\infty} \mathcal{B}_m$.*
b) *If $\lim_{m\to\infty} \operatorname{diam} \mathcal{B}_m = 0$, then $\bigcap_{m=1}^{\infty} \mathcal{B}_m$ consists of a single point.*

Proof. Write $\mathcal{B}_m = \prod_{i=1}^{n}[a_i(m), b_i(m)]$. Since the sequence is nested, we have

$$a_i(m) \leq a_i(m+1) \leq b_i(m+1) \leq b_i(m)$$

for all i and m. Then $x_m = (x_m(1), \ldots, x_m(n)) \in \bigcap_{m=1}^{\infty} \mathcal{B}_m$ iff for all $1 \leq i \leq n$ we have $a_m(i)) \leq x_m \leq b_m(i)$. For $1 \leq i \leq n$, put

$$A_i = \sup_m a_m(i), \ B_i = \inf_m b_m(i).$$

It then follows that

$$\bigcap_{m=1}^{\infty} \mathcal{B}_m = \prod_{i=1}^{n} [A_i, B_i],$$

which is nonempty. □

Exercise 2.68. *In the above proof it is implicit that $A_i \leq B_i$ for all $1 \leq i \leq n$. Convince yourself that you could write down a careful proof of this (e.g. by writing down a careful proof!).*

Exercise 2.69. *Under the hypotheses of the Lion-Hunting Lemma, show that the following are equivalent:*
(i) $\inf\{\operatorname{diam} \mathcal{B}_m\}_{m=1}^{\infty} = 0$.
(ii) $\bigcap_{m=1}^{\infty} \mathcal{B}_m$ *consists of a single point.*

Theorem 2.46 (Heine-Borel). *A closed, bounded subset of $\mathbb{R}^n$ is compact.*

Proof. Because every closed bounded subset is a subset of a closed box and closed subsets of compact sets are compact, it is sufficient to show the compactness of every closed box $\mathcal{B} = \prod_{i=1}^{n} [a_i, b_i]$. Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open covering of $\mathcal{B}$. Seeking a contradiction we suppose $\mathcal{U}$ admits no finite subcovering. We **bisect** $\mathcal{B}$ into 2^n closed subboxes of equal size, so that e.g. the bottom leftmost one is $\prod_{i=1}^{n} [a_i, \frac{a_i+b_i}{2}]$. It must be that at least one of the subboxes cannot be covered by any finite number of sets in $\mathcal{U}$: if all 2^n of them have finite subcoverings, taking the union of 2^n finite subcoverings, we get a finite subcovering of $\mathcal{B}$. Identify one such subbox $\mathcal{B}_1$, and notice that $\operatorname{diam} \mathcal{B}_1 = \frac{1}{2} \operatorname{diam} \mathcal{B}$. Now bisect $\mathcal{B}_1$ and repeat the argument: we get a nested sequence $\{\mathcal{B}_m\}_{m=1}^{\infty}$ of closed boxes with

$$\operatorname{diam} \mathcal{B}_m = \frac{\operatorname{diam} \mathcal{B}}{2^m}.$$

By the Lion-Hunting Lemma there is $x \in \bigcap_{m=1}^{\infty} \mathcal{B}_m$.[3] Choose $U_0 \in \mathcal{U}$ such that $x \in U_0$. Since U_0 is open, for some $\epsilon > 0$ we have

$$x \in B^\circ(x, \epsilon) \subset U_0.$$

For sufficiently large m we have – formally, by the Archimedean property of $\mathbb{R}$ – that $\operatorname{diam} \mathcal{B}_m < \epsilon$. Thus every point in $\mathcal{B}_m$ has distance less than ϵ from x so

$$\mathcal{B}_m \subset B^\circ(x, \epsilon) \subset U_0.$$

This contradicts the heck out of the fact that $\mathcal{B}_m$ admits no finite subcovering. □

Proposition 2.47. *Let X be a compact metric space, and let $A \subset X$ be an infinite subset. Then A has a limit point in X.*

Proof. Seeking a contradiction we suppose that A has no limit point in X. Then also no subset $A' \subset A$ has any limit points in X. Since a set is closed if it contains all of its limit points, every subset of A is closed in X. In particular A is closed in X, hence A is compact. But since for all $x \in A$, $A \setminus \{x\}$ is closed in A, we

[3]Though we don't need it, it follows from Exercise 1.9 that the intersection point x is unique.

have that $\{x\}$ is open in A. (In other words, A is discrete.) Thus $\{\{x\}\}_{x \in A}$ is an infinite cover of A without a finite subcover, so A is not compact: contradiction. □

THEOREM 2.48. *(Bolzano-Weierstrass for Sequences) Every bounded sequence in $\mathbb{R}^N$ admits a convergent subsequence.*

PROOF. Step 1: Let $N = 1$. I leave it to you to carry over the proof of Bolzano-Weierstrass in $\mathbb{R}$ given in § 2.2 to our current sequential situation: replacing the Monotonicity Lemma with the Rising Sun Lemma, the endgame is almost identical. Step 2: Let $N \geq 2$, and let $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence in $\mathbb{R}^N$. Then each coordinate sequence $\{x_n(i)\}_{n=1}^{\infty}$is bounded, so Step 1 applies to each of them.

However, if we just extract subsequences for each component separately, we will have N different subsequences, and it will in general not be possible to get one subsequence out of all of them. So we proceed in order: first we extract a subsequence such that the first coordinates converge. Then we extract a subsequence *of the subsequence* such that the second coordinates converge. This does not disturb what we've already done, since every subsequence of a convergent sequence is convergent (we're applying this in the familiar context of real sequences, but it is equally true in any metric space). Thus we extract a sub-sub-sub...subsequence (N "subs" altogether) which converges in every coordinate and thus converges. But a sub-sub....subsequence is just a subsequence, so we're done. □

A metric space is **sequentially compact** if every sequence admits a convergent subsequence.

A metric space X is **limit point compact** if every infinite subset $A \subset X$ has a limit point in X.

8. Completeness

8.1. Lion Hunting In a Metric Space.

Recall the Lion-Hunting Lemma: any nested sequence of closed boxes in $\mathbb{R}^N$ has a common intersection point; if the diameters approach zero, then there is a unique intersection point. This was the key to the proof of the Heine-Borel Theorem.

Suppose we want to hunt lions in an arbitrary metric space: what should we replace "closed box" with? The following exercise shows that we should at least keep the "closed" part in order to get something interesting.

EXERCISE 2.70. *Find a nested sequence $A_1 \supset A_2 \supset \ldots \supset A_n \ldots$ of nonempty subsets of $[0, 1]$ with $\bigcap_{n=1}^{\infty} A_i = \varnothing$.*

So perhaps we should replace "closed box" with "closed subset"? Well...we could. However, even in $\mathbb{R}$, if we replace "closed box" with "closed set", then lion hunting need not succeed: for $n \in \mathbb{Z}^+$, let $A_n = [n, \infty)$. Then $\{A_n\}_{n=1}^{\infty}$ is a nested sequence of closed subsets with $\bigcap_{n=1}^{\infty} A_n = \varnothing$.

Suppose however that we consider nested covers of nonempty closed subsets with the additional property that diam $A_n \to 0$. In particular, all but finitely many A_n's are bounded, so the previous problem is solved. Indeed, Lion-Hunting works under these hypothesis in $\mathbb{R}^N$ because of Heine-Borel: some A_n is closed and bounded, hence compact, so we revisit the previous case.

A metric space is **complete** if for every nested sequence $\{A_n\}$ of nonempty closed subsets with diameter tending to 0 we have $\bigcap_n A_n \neq \varnothing$.

EXERCISE 2.71. *Let* $\{A_n\}_{n=1}^{\infty}$ *be any sequence of sets in a metric space* X *with* $\operatorname{diam}(A_n) \to 0$. *Show*

$$\# \bigcap_{n=1}^{\infty} A_n \leq 1.$$

The following result shows that completeness, like compactness, is a kind of intrinsic closedness property.

LEMMA 2.49. *Let* Y *be a subset of a metric space* X.

a) *If* X *is complete and* Y *is closed, then* Y *is complete.*
b) *If* Y *is complete, then* Y *is closed.*

PROOF. a) If Y is closed in X, then a nested sequence $\{A_n\}_{n=1}^{\infty}$ of nonempty closed subsets of Y with diameter approaching 0 is also a nested sequence of nonempty closed subsets of X with diameter approaching 0. Since X is complete, there is $x \in \bigcap_n A_n$.
b) If Y is not closed, let $\mathbf{y}$ be a sequence in Y converging to an element $x \in X \setminus Y$. Put $A_n = \{\mathbf{y_k} \mid \mathbf{k} \geq \mathbf{n}\}$. Then $\{A_n\}_{n=1}^{\infty}$ is a nested sequence of nonempty closed subsets of Y of diameter approaching 0 and with empty intersection. $\square$

THEOREM 2.50. *For a metric space* X, *the following are equivalent:*

(i) *The space* X *is complete.*
(ii) *Let* $\{A_i\}_{i\in I}$ *be a family of closed subsets of* X *satisfying the following properties:*
   (a) *(Finite Intersection Condition) For all finite subsets* $J \subset I$ *we have* $\bigcap_{i\in J} A_i \neq \varnothing$.
   (b) *For all* $\epsilon > 0$, *there is* $i \in I$ *such that that diameter of* A_i *is at most* ϵ.

   Then $\bigcap_{i\in I} A_i$ *is nonempty and consists of a single point* $x \in X$.

PROOF. (i) $\implies$ (ii): For $n \in \mathbb{Z}^+$, choose $i_n \in I$ such that the diameter of A_{i_n} is at most $\frac{1}{n}$. Put $\tilde{B}_n \coloneqq A_{i_n}$ and $B_n = \bigcap_{i=1}^{n} \tilde{B}_i$. By the completeness of X and Exercise 2.71, we have $\bigcap_{n=1}^{\infty} B_n = \{x\}$ for some $x \in X$. Thus

$$\bigcap_{i\in I} A_i \subset \bigcap_{n=1}^{\infty} B_n = \{x\}.$$

Seeking a contradiction, suppose that there is some $i \in I$ such that $x \notin A_i$. Then

$$\bigcap_{n=1}^{\infty} (A \cap B_n) = A \cap \bigcap_{n=1}^{\infty} B_n = \varnothing.$$

But $\{A \cap B_n\}_{n=1}^{\infty}$ is also a nested sequence of nonempty closed subsets of X of diameter tending to 0, so this contradicts the completeness of X.
(ii) $\implies$ (i): It is clear that a nested sequence $\{A_n\}_{n=1}^{\infty}$ of nonempty closed subsets with diameter tending to 0 satisfies conditions (a) and (b) of (ii), so $\bigcap_{n=1}^{\infty} A_n = \{x\}$ for some $x \in X$. $\square$

8.2. Cauchy Sequences.

Our Lion Hunting definition of completeness is conceptually pleasant, but it seems like it could be a lot of work to check in practice. It is also – we now admit – not the standard one. We now make the transition to the standard definition.

LEMMA 2.51. *A metric space in which each sequence of closed balls with diameters tending to zero has nonempty intersection is complete.*

PROOF. Let $\{A_n\}_{n=1}^{\infty}$ be a nested sequence of nonempty closed subsets with diameter tending to zero. We may assume without loss of generality that each A_n has finite diameter, and we may choose for all $n \in \mathbb{Z}^+$, $x_n \in A_n$ and a positive real number r_n such that $A_n \subset B^\bullet(x_n, r_n)$ and $r_n \to 0$. By assumption, there is a unique point $x \in \bigcap_n B^\bullet(x_n, r_n)$. Then $x_n \to x$. Fix $n \in \mathbb{Z}^+$. Then x is the limit of the sequence $x_n, x_{n+1}, \ldots$ in A_n, and since A_n is closed, $x \in A_n$. □

Let us nail down which sequences of closed balls we can use for lion hunting.

LEMMA 2.52. *Let $\{B^\bullet(x_n, r_n)\}_{n=1}^{\infty}$ be a nested sequence of closed balls in a metric space X with $r_n \to 0$. Then for all $\epsilon > 0$, there is $N = N(\epsilon)$ such that for all $m, n \geq N$, we have $d(x_m, x_n) \leq \epsilon$.*

PROOF. Fix $\epsilon > 0$, and choose N such that $r_N \leq \frac{\epsilon}{2}$. Then if $m, n \geq N$ we have $x_n, x_m \in B^\bullet(x_N, r_N)$, so $d(x_n, x_m) \leq 2r_N \leq \epsilon$. □

At last, we have motivated the following definition. A sequence $\{x_n\}$ in am metric space X is **Cauchy** if for all $\epsilon > 0$, there is $N = N(\epsilon)$ such that for all $m, n \geq N$, $d(x_m, x_n) < \epsilon$. Thus in a nested sequence of closed balls with diameter tending to zero, the centers of the balls form a Cauchy sequence. Moreover:

LEMMA 2.53. *Let $\{x_n\}$ be a sequence in a metric space X, and for $n \in \mathbb{Z}^+$ put $A_n = \{x_k \mid k \geq n\}$. The following are equivalent:*

(i) *The sequence $\{x_n\}$ is Cauchy.*
(ii) We have $\operatorname{diam} A_n \to 0$.

EXERCISE 2.72. *Prove it.*

EXERCISE 2.73. *Show that every convergent sequence is Cauchy.*

LEMMA 2.54. *Let $x_\bullet$ and $y_\bullet$ be two sequences in the metric space (X, d) such that* $\lim_{n\to\infty} d(x_n, y_n) = 0$.
a) *If $x_\bullet$ is Cauchy, then so is $y_\bullet$.*
b) *If for some $L \in X$ we have that $y_n \to L$, then also $x_n \to L$.*

PROOF. a) Let $\epsilon > 0$, and choose N such that for all $m, n \geq N$ we have $d(x_m, x_n) < \frac{\epsilon}{3}$ and for all $n \geq N$ we have $d(x_n, y_n) < \frac{\epsilon}{3}$. Then for all $m, n \geq N$,

$$d(y_m, y_n) \leq d(y_m, x_m) + d(x_m, x_n) + d(x_n, y_n) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

b) Let $\epsilon > 0$, and choose N such that for all $n \geq N$ we have $d(y_n, L) < \frac{\epsilon}{2}$ and $d(y_n, x_n) < \frac{\epsilon}{2}$. Then for all $n \geq N$ we have

$$d(x_n, L) \leq d(x_n, y_n) + d(y_n, L) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$ □

LEMMA 2.55. *Every partial limit of a Cauchy sequence is a limit.*

PROOF. Let $\{x_n\}$ be a Cauchy sequence, and let $x \in X$ be such that some subsequence $x_{n_k} \to x$. Fix $\epsilon > 0$, and choose N such that for all $m, n \geq N$, $d(x_m, x_n) < \frac{\epsilon}{2}$. Choose K such that $n_K \geq N$ and for all $k \geq K$, $d(x_{n_k}, x) < \frac{\epsilon}{2}$. Then for all $n \geq N$,

$$d(x_n, x) \leq d(x_n, x_{n_K}) + d(x_{n_K}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \qquad \square$$

THEOREM 2.56 (Cantor Intersection Theorem). *For a metric space X, the following are equivalent:*

(i) *The space X is complete: that is, for every nested sequence $\{A_n\}_{n=1}^{\infty}$ of nonempty closed subsets of X with* $\operatorname{diam}(A_n) \to 0$, *we have $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.*
(ii) *Every Cauchy sequence in X is convergent.*

PROOF. $\neg$ (ii) $\implies$ $\neg$ (i): Suppose $\{x_n\}$ is a Cauchy sequence which does not converge. By Lemma 2.55, the sequence $\{x_n\}$ has no partial limit, so $A_n = \{x_k \mid k \geq n\}$ is a nested sequence of closed subsets with diameter tending to 0 and $\bigcap_n A_n = \varnothing$, so X is not complete.
(ii) $\implies$ (i): By Lemma 2.51, it is enough to show that any nested sequence of closed balls with diameters tending to zero has nonempty intersection. By Lemma 2.52, the sequence of centers $\{x_n\}$ is Cauchy, hence converge to $x \in X$ by assumption. For each $n \in \mathbb{Z}^+$, the sequence $x_n, x_{n+1}, \ldots$ lies in $B^\bullet(x_n, r_n)$, hence the limit, x, lies in $B^\bullet(x_n, r_n)$. $\square$

8.3. Very good metrics. Let I be a nonempty set, and for each $i \in I$ let (X_i, d_i) be a nonempty metric space. Let $X := \prod_{i \in I} X_i$ be the Cartesian product. Recall that a metric d on X is called **good** if for every sequence $\mathbf{x}$ in X and point $x \in X$, we have $\mathbf{x} \to x$ in X iff for all $i \in I$, the component sequence $\mathbf{x}^{(i)}$ converges to the ith component $x^{(i)}$ of x. Let us call a metric d on X **Cauchy good** if for every sequence $\mathbf{x}$ in X we have that $\mathbf{x}$ is Cauchy iff $\mathbf{x}^{(i)}$ is Cauchy in X_i for all $i \in I$ and **very good** if it is both good and Cauchy good.

EXERCISE 2.74. *We maintain the above setup: $X = \prod_{i \in I} X_i$, d_i is a metric on X_i for all $i \in I$ and d is a metric on X.*

a) *Exhibit a metric that is good but not Cauchy good. (Suggestion: take $I = \{1\}$ and $X_1 = \mathbb{R}$.)*
b) *Suppose d is Cauchy good. Show: if $\mathbf{x} \to x$ in X, then for all $i \in I$ we have $\mathbf{x}^{(i)} \to x^{(i)}$.*
c) *Exhibit a metric that is Cauchy good but not good. (Suggestion: take $I = \{1\}$ and $X = \{\frac{1}{n} \mid n \in \mathbb{Z}^+\}$.)*

PROPOSITION 2.57. *Let I be a nonempty set, for $i \in I$ let (X_i, d_i) be a nonempty metric space, and let $X := \prod_{i \in I} X_i$.*

a) *Let d be a very good metric on X. Then X is complete iff X_i is complete for all $i \in I$.*
b) *Suppose $I = \mathbb{Z}^+$ and that for all $n \in \mathbb{Z}^+$ we have* $\operatorname{diam}(X_n, d_n) \leq 1$. *Then*

$$d(x, y) := \sup_n \frac{d_n(x_n, y_n)}{2^n}$$

is a very good metric on X.

Proof. a) Suppose X is complete, let $i \in I$, and let $\{x_{n,i}\}_{n=1}^{\infty}$ be a Cauchy sequence in X_i. For all $j \in I \setminus \{i\}$, choose $x_j \in X_j$, and for $n \in \mathbb{Z}^+$ let $\mathbf{x}_n$ be the element of X whose ith coordinate is $x_{n,i}$ and for all $j \neq i$ has jth coordinate x_j. In other words, $\mathbf{x} = \{xx_n\}_{n=1}^{\infty}$ is a sequence in X whose ith component is the Cauchy sequence $\{x_{n,i}\}$ in X_i and has every other component a constant (hence convergent) sequence. Because d is Cauchy good, the sequence $\mathbf{x}$ is Cauchy. Since X is complete, the sequence $\mathbf{x}$ converges. Since d is good, the sequence $\{x_{n,i}\}$ converges, so X_i is complete. The converse is similar but easier: suppose each X_i is complete, and let $\mathbf{x}$ be a Cauchy sequence in X. Then each component is Cauchy, hence converges, hence the original sequence converges.
b) For $n \in \mathbb{Z}^+$, put $d_n' := \frac{d_n}{2^n}$. Then $\operatorname{diam}(X_n, d_n') \leq 2^{-n}$ and $d = \sup_n d_n'$, so Theorem 2.36 applies to show that d is good. For all $i \in \mathbb{Z}^+$ the projection map $\pi_i : X \to X_i$ is 2^i-Lipschitz, hence uniformly continuous, so if $\mathbf{x}$ is a Cauchy sequence in X then $\mathbf{x}^{(i)}$ is a Cauchy sequence in X_i. Conversely, if $\mathbf{x}$ is a sequence in X such that $\mathbf{x}^{(i)}$ is Cauchy for all i, let $\epsilon > 0$. We may choose $M \in \mathbb{Z}^+$ such that $2^{-M} < \epsilon$ and choose $N \in \mathbb{Z}^+$ such that for all $m, n \geq N$ and all $1 \leq i \leq M$ we have $d_i(\mathbf{x}_m^{(i)}, \mathbf{x}_n^{(i)}) < \epsilon$. Then for all $m, n \geq N$ we have

$$d(\mathbf{x}_m, \mathbf{x}_n) = \sup_i \frac{d_i(\mathbf{x}_m^{(i)}, \mathbf{x}_n^{(i)})}{2^i} < \epsilon. \qquad \square$$

8.4. Baire's Theorem.

A subset A of a metric space X is **nowhere dense** if $\overline{A}$ contains no nonempty open subset, or in other (fewer) words, if $\overline{A}^{\circ} = \varnothing$.

Exercise 2.75. *Let x be a point of a metric space X. Show that x is a limit point of X iff $\{x\}$ is nowhere dense.*

Theorem 2.58 (Baire I). *Let X be a complete metric space.*

a) *Let $\{U_n\}_{n=1}^{\infty}$ be a sequence of dense open subsets of X. Then $U = \bigcap_{n=1}^{\infty} U_n$ is also dense in X.*
b) *Let $\{A_n\}_{n=1}^{\infty}$ be a countable collection of nowhere dense subsets of X. Then $A = \bigcup_{n=1}^{\infty} A_n$ has empty interior.*

Proof. a) We must show that for every nonempty open subset W of X we have $W \cap U \neq \varnothing$. Since U_1 is open and dense, $W \cap U_1$ is nonempty and open and thus contains some closed ball $B^{\bullet}(x_1, r_1)$ with $0 < r_1 \leq 1$. For $n \geq 1$, having chosen x_n and $r_n \leq \frac{1}{n}$, since U_{n+1} is open and dense, $B(x_n, r_n) \cap U_{n+1}$ is nonempty and open and thus contains some closed ball $B^{\bullet}(x_{n+1}, r_{n+1})$ with $0 < r_{n+1} \leq \frac{1}{n+1}$. Since X is complete, there is a (unique)

$$x \in \bigcap_{n=2}^{\infty} B^{\bullet}(x_n, r_n) \subset \bigcap_{n=1}^{\infty} B(x_n, r_n) \cap U_n \subset \bigcap_{n=1}^{\infty} U_n = U.$$

Moreover

$$x \in B^{\bullet}(x_1, r_1) \subset W \cap U_1 \subset W,$$

so

$$x \in U \cap W.$$

b) Without loss of generality we may assume that each A_n is closed, because A_n is nowhere dense iff $\overline{A_n}$ is nowhere dense, and a subset of a nowhere dense set is certainly nowhere dense. For $n \in \mathbb{Z}^+$, let $U_n = X \setminus A_n$. Each U_n is open; moreover, since $\overline{A_n}$ contains no nonempty open subset, every nonempty open subset must intersect U_n and thus U_n is dense. By part a), $\bigcap_{n=1}^{\infty} U_n = \bigcap_{n=1}^{\infty} X \setminus A_n = X \setminus A$ is dense. Again, this means that every nonempty open subset of X meets the complement of A so no nonempty open subset of X is contained in A. □

Corollary 2.59. *Let X be a nonempty complete metric space in which every point is a limit point. Then X is uncountable.*

Proof. Let $A = \{a_n \mid n \in \mathbb{Z}^+\}$ be a countably infinite subset of X. Then each $\{a_n\}$ is nowhere dense, so by Theorem 2.58, $A = \bigcup_{n=1}^{\infty} \{a_n\}$ has empty interior. In particular, $A \subsetneq X$. □

Corollary 2.59 applies to $\mathbb{R}$ and gives a purely topological proof of its uncountability!

Corollary 2.60. *A countably infinite complete metric space has infinitely many isolated points.*

Proof. Step 1: Let X be a metric space with finitely isolated points, say $a_1, \ldots, a_n$, and put $Y := X \setminus A$. We claim that Y is a perfect subset of X: that is, it is closed and has no isolated points. Indeed, for each isolated point p in any metric space the singleton set $\{p\}$ is open and thus the set of all isolated points is open in any metric space, which implies that Y is closed. If $y \in Y$, then y is not an isolated point of X, so each neighborhood of y in X contains infinitely many points. However, intersecting with Y involves removing only finitely many points, so each neighobrhood of y in Y is infinite. Thus y is not an isolated point of Y.
Step 2: Let X be a countably infinite complete metric space whose set A of isolated points is finite. By Step 1 and Lemma 2.49a), the subset $X \setminus A$ is a countably infinite complete metric space without isolated points, contradicting Corollary 2.59. □

One interesting consequence of these results is that we can deduce purely topological consequences of the metric condition of completeness.

Example 2.11. *Let $\mathbb{Q}$ be the rational numbers, equipped with the usual Euclidean metric $d(x, y) = |x - y|$. As we well know, $(\mathbb{Q}, d)$ is not complete. But here is a more profound question: is there some topologically equivalent metric d' on $\mathbb{Q}$ which is complete? Now in general a complete metric can be topologically equivalent to an incomplete metric: e.g. this happens on $\mathbb{R}$. But that does not happen here: any topologically equivalent metric is a metric on a countable set in which no point is isolated (the key observation being that the latter depends only on the topology), so by Corollary 2.59 cannot be complete.*

9. Total Boundedness

We saw above that the property of boundedness is not only not preserved by homeomorphisms of metric spaces, it is not even preserved by uniformeomorphisms of metric spaces (and also that it is preserved by Lipschitzeomorphisms). Though this was as simple as replacing any unbounded metric by the standard bounded metric $d_b(x, y) = \min d(x, y), 1$, intuitively it is still a bit strange: e.g. playing around a bit with examples, one soon suspects that for subspaces of Euclidean space $\mathbb{R}^N$, the property of boundedness *is* preserved by uniformeomorphisms.

The answer to this puzzle lies in identifying a property of metric spaces: perhaps the most important property that does not get "compactness level PR".

A metric space X is **totally bounded** if for all $\epsilon > 0$, it admits a finite cover by open ϵ-balls: there is $N \in \mathbb{Z}^+$ and $x_1, \ldots, x_N \in X$ such that $X = \bigcup_{i=1}^N B(x_i, \epsilon)$.

Since any finite union of bounded sets is bounded, certainly total boundedness implies boundedness (thank goodness).

Notice that we could require the balls to be closed without changing the definition: just slightly increase or decrease ϵ. (And indeed, sometimes we will want to use one form of the definition and sometimes the other.) In fact we don't really need balls at all: consider the following reformulation.

Lemma 2.61. *For a metric space X, the following are equivalent:*
(i) For all $\epsilon > 0$, there exists a finite family $S_1, \ldots, S_N$ of subsets of X such that $\operatorname{diam} S_i \leq \epsilon$ *for all i and $X = \bigcup_{i=1}^N S_i$.*
(ii) X is totally bounded.

Proof. (i) $\implies$ (ii): We may assume each S_i is nonempty, and choose $x_i \in S_i$. Since $\operatorname{diam} S_i \leq \epsilon$, $S_i \subset B^\bullet(x_i, \epsilon)$ and thus $X = \bigcup_{i=1}^N B^\bullet(x_i, \epsilon)$.
(ii) $\implies$ (i): For every $\epsilon > 0$, choose $x_1, \ldots, x_N$ such that $\bigcup_{i=1}^N B^\bullet(x_i, \frac{\epsilon}{2}) = X$. We have covered X by finitely many sets each of diameter at most ϵ. □

Corollary 2.62.
a) *Every subset of a totally bounded metric space is totally bounded.*
b) *Let $f : X \to Y$ be a uniformeomorphism of metric spaces. Then X is totally bounded iff Y is totally bounded.*

Proof. a) Suppose that X is totally bounded, and let $Y \subset X$. Since X is totally bounded, for each $\epsilon > 0$ there exist $S_1, \ldots, S_N \subset X$ such that $\operatorname{diam} S_i < \epsilon$ for all i and $X = \bigcup_{i=1}^N S_i$. Then $\operatorname{diam}(S_i \cap Y) < \epsilon$ for all i and $Y = \bigcup_{i=1}^N (S_i \cap Y)$.
b) Suppose X is totally bounded. Let $\epsilon > 0$, and choose $\delta > 0$ such that f is (ϵ, δ)-uniformly continuous. Since X is totally bounded there are finitely many sets $S_1, \ldots, S_N \subset X$ with $\operatorname{diam} S_i \leq \delta$ for all $1 \leq i \leq N$ and $X = \bigcup_{i=1}^N S_i$. For $1 \leq i \leq N$, let $T_i = f(S_i)$. Then $\operatorname{diam} T_i \leq \epsilon$ for all i and $Y = \bigcup_{i=1}^N T_i$. It follows that Y is uniformly bounded. Using the uniformly continuous function $f^{-1} : Y \to X$ gives the converse implication. □

Lemma 2.63. *(Archimedes) A subset of $\mathbb{R}^N$ is bounded iff it is totally bounded.*

Proof. Total boundedness always implies boundedness. Moreover any bounded subset of $\mathbb{R}^N$ lies in some cube $C_n = [-n, n]^N$ for some $n \in \mathbb{Z}^+$, so by Corollary 2.62 it is enough to show that C_n is totally bounded. But C_n can be written as the union of finitely many subcubes with arbitrarily small side length and thus arbitrarily small diameter. Provide more details if you like, but this case is closed. □

Let $\epsilon > 0$. An ϵ-**net** in a metric space X is a subset $N \subset X$ such that for all $x \in X$, there is $n \in N$ with $d(x, n) < \epsilon$. An ϵ-**packing** in a metric space X is a subset $P \subset X$ such that $d(p, p') \geq \epsilon$ for all $p, p' \in P$.

These concepts give rise to a deep duality in discrete geometry between **packing** – namely, placing objects in a space without overlap – and **covering** – namely, placing objects in a space so as to cover the entire space. Notice that already we can cover the plane with closed unit balls or we can pack the plane with closed unit balls but we cannot do both at once. The following is surely the simplest possible **duality principle** along these lines.

PROPOSITION 2.64. *Let X be a metric space, and let $\epsilon > 0$.*

a) *The space X admits either a finite ϵ-net or an infinite ϵ-packing.*
b) *If X admits a finite ϵ-net then it does not admit an infinite (2ϵ)-packing.*
c) *Thus X is totally bounded iff for all $\epsilon > 0$, there is no infinite ϵ-packing.*

PROOF. a) First suppose that we do not have a finite ϵ-net in X. Then X is nonempty, so we may choose $p_1 \in X$. Since $X \neq B(p_1, \epsilon)$, there is $p_2 \in X$ with $d(p_1, p_2) \geq \epsilon$. Inductively, having constructed an n element ϵ-packing $P_n = \{p_1, \ldots, p_n\}$, since P_n is not a finite ϵ-net there is $p_{n+1} \in X$ such that $d(p_i, p_{n+1}) \geq \epsilon$ for all $1 \leq i \leq n$, so $P_{n+1} = P_n \cup \{p_{n+1}\}$ is an $n+1$ element ϵ-packing. Then $P = \bigcup_{n \in \mathbb{Z}^+} P_n$ is an infinite ϵ-packing.
b) Seeking a contradiction, suppose that we have both an infinite (2ϵ)-packing P and a finite ϵ-net N. Since P is infinite, N is finite and $X = \bigcup_{n \in N} B(n, \epsilon)$, there must be distinct points $p \neq p' \in P$ each lying in $B(n, \epsilon)$ for some $n \in N$, and then by the triangle inequality $d(p, p') \leq d(p, n) + d(n, p') < 2\epsilon$.
c) To say that $N \subset X$ is an ϵ-net means precisely that if we place an open ball of radius ϵ centered at each point of N, then the union of these balls covers X. Thus X is totally bounded iff it admits a finite ϵ-net for all $\epsilon > 0$, and then by part b) there is no $\epsilon > 0$ such that X admits an infinite ϵ-packing. Conversely, if for no $\epsilon > 0$ does X admit an infinite ϵ-packing then by part a) X admits a finite ϵ-net for all $\epsilon > 0$ and thus X is totally bounded. $\square$

THEOREM 2.65. *A metric space X is totally bounded iff each sequence $\mathbf{x}$ in X admits a Cauchy subsequence.*

PROOF. If X is not totally bounded, then by Proposition 2.64 there is an infinite ϵ-packing for some $\epsilon > 0$. Passing to a countably infinite subset $P = \{p_n\}_{n=1}^{\infty}$, we get a sequence such that for all $m \neq n$, $d(p_m, p_n) \geq \epsilon$. This sequence has no Cauchy subsequence.

Now suppose that X is totally bounded, and let $\mathbf{x}$ be a sequence in X. By total boundedness, for all $n \in \mathbb{Z}^+$, we can write X as a union of finitely many closed subsets $Y_1, \ldots, Y_N$ each of diameter at most $\frac{1}{n}$ (here N is of course allowed to depend on n). An application of the Pigeonhole Principle gives us a subsequence all of whose terms lie in Y_i for some i, and thus we get a subsequence each of whose terms have distance at most ϵ. Unfortunately this is not quite what we want: we need one subsequence each of whose sufficiently large terms differ by at most $\frac{1}{n}$. We attain this via a **diagonal construction**: namely, let

$$x_{1,1}, x_{1,2}, \ldots, x_{1,n}, \ldots$$

be a subsequence each of whose terms have distance at most 1. Since subspaces of totally bounded spaces are totally bounded, we can apply the argument again inside the smaller metric space Y_i to get a subsubsequence

$$x_{2,1}, x_{2,2}, \ldots, x_{2,n}, \ldots$$

each of whose terms differ by at most $\frac{1}{2}$ and each $x_{2,n}$ is selected from the subsequence $\{x_{1,n}\}$; and so on; for all $m \in \mathbb{Z}^+$ we get a subsub...subsequence

$$x_{m,1}, x_{m,2}, \ldots, x_{m,n}, \ldots$$

each of whose terms differ by at most $\frac{1}{n}$. Now we choose the **diagonal subsequence**: put $y_n = x_{n,n}$ for all $n \in \mathbb{Z}^+$. We allow the reader to check that this is a subsequence of the original sequence $\mathbf{x}$.. This sequence satisfies $d(y_n, y_{n+k}) \leq \frac{1}{n}$ for all $k \geq 0$, so we get a Cauchy subsequence. $\square$

9.1. Further Exercises. The exercises in this section develop a proof of the following results of S.S. Kim [**Ki99**].

LEMMA 2.66. *For a metric space X, the following are equivalent:*

(i) *The space X has no isolated points.*
(ii) *There is a subset $A \subset X$ such that both A and $X \setminus A$ are dense in X.*

THEOREM 2.67. *Let X be a metric space without isolated points. For $Y \subset X$, the following are equivalent:*

(i) *There is a function $f : X \to \mathbb{R}$ such that $Y = \{x \in X \mid f \text{ is continuous at } x\}$.*
(ii) *We have that Y is a G_δ-subset of X, i.e., a countable intersection of open subsets of X.*

Exercise 2.51 shows that (i) $\implies$ (ii) for all metric spaces.

EXERCISE 2.76. *Show that (ii) $\implies$ (i) in Lemma 2.66.*

EXERCISE 2.77. *Let X be a metric space, and let $\epsilon > 0$. Show that X admits a* ***maximal ϵ-packing****, i.e., an ϵ-packing that is not properly contained in any other ϵ-packing. (Suggestion: apply Zorn's Lemma.)*

The following exercise shows that (i) $\implies$ (ii) in Lemma 2.66.

EXERCISE 2.78. *Let X be a nonempty metric space without isolated points.*

a) *For $\epsilon > 0$, let P be an ϵ-packing in X. Show that $X \setminus P$ has no isolated points.*
b) *Inductively construct: an infinite sequence $\{P_n\}_{n=1}^\infty$ of pairwise disjoint subsets of X such that for all $n \in \mathbb{Z}^+$, the set P_n is a maximal $\frac{1}{n}$-packing.*
c) *Put $A := \bigcup_{n=1}^\infty P_{2n}$. Show: A and $X \setminus A$ are both dense in X.*

In the following exercise we work through Kim's construction of a function whose locus of continuity is any G_δ-set in a metric space without isolated points.

EXERCISE 2.79. *Let X be a metric space without isolated points, let Y be a G_δ-set in X, and let $Z := X \setminus Y$. We may write $Z = \bigcup_{n=1}^\infty F_n$ with $F_1 \subset F_2 \subset \ldots F_n \subset \ldots$ open subsets, and put $U_n := X \setminus F_n$, so $Y = \bigcup_{n=1}^\infty U_n$. By Lemma 2.66, there is a subset A of X such that both A and $X \setminus A$ are dense in X. Let*

$$\mathbf{1}_A : X \to \mathbb{R},\ x \mapsto \begin{cases} 1 & x \in A \\ 0 & x \in X \setminus A \end{cases}$$

be the characteristic function of A. We put

$$g : X \to \mathbb{R}, x \mapsto \sum_{n \in \mathbb{Z}^+ \mid x \in F_n} 2^{-n}$$

and

$$f : X \to \mathbb{R}, x \mapsto g(x)(\mathbf{1}_A(x) - 1/2).$$

a) *Show: For $x \in X$, we have $g(x) = 0$ iff $x \in Y$.*
b) *Show: If $x \in U_n$, then $|g(x)| \leq 2^{-n}$. Deduce that g is continuous at x for all $x \in Y$ and then that f is continuous at x for all $x \in Y$.*
c) *Suppose that x lies in the interior of Z. Show: $f(x) \neq 0$ and every open ball about x contains points at which f is positive and also points at which f is negative. Deduce that f is not continuous at x.*
d) *Suppose that x lies in $Z \cap \partial Z = Z \cap \partial Y$. Show: $f(x) \neq 0$ and every open ball about x contains points at which f is 0. Deduce that f is not continuous at x.*

10. Separability

We remind the reader that we are an ardent fan of [**Ka**]. The flattery becomes especially sincere at this point: c.f. [**Ka**, §5.2].

Recall that a metric space is **separable** if it admits a countable dense subset.

Exercise 2.80. *Let $f : X \to Y$ be a continuous surjective map between metric spaces. Show that if X is separable, so is Y.*

We want to compare this property with two others that we have not yet introduced.

A **base** $\mathcal{B} = \{B_i\}$ for the topology of a metric space X is a collection of open subsets of X such that every open subset U of X is a union of elements of $\mathcal{B}$: precisely, there is a subset $J \subset \mathcal{B}$ such that $\bigcup_{i \in J} B_i = U$. (We remark that taking $J = \varnothing$ we get the empty union and thus the empty set.)

The example *par excellence* of a base for the topology of a metric space X is to take $\mathcal{B}$ to be the family of all open balls in X. In this case, the fact that $\mathcal{B}$ is a base for the topology is in fact the very definition of the metric topology: the open sets are precisely the unions of open balls.

A **countable base** is just what it sounds like: a base which, as a set, is countable (either finite or countably infinite).

Proposition 2.68.

a) *Let X be a metric space, let $\mathcal{B} = \{B_i\}$ be a base for the topology of X, and let $Y \subset X$ be a subset. Then $\mathcal{B} \cap Y := \{B_i \cap Y\}$ is a base for the topology of Y.*
b) *If X admits a countable base, then so does all of its subsets.*

Proof. a) This follows from the fact that the open subsets of Y are precisely those of the form $U \cap Y$ for U open in X. We leave the details to the reader.
b) This follows immediately. □

Theorem 2.69. *For a metric space X, the following are equivalent:*

(i) *The space X is separable.*
(ii) *X admits a countable base.*
(iii) *The space X is Lindelöf.*

(iv) *The space X admits no uncountable discrete subset.*
(v) *The space X admits no pairwise disjoint uncountable family of open balls.*
(vi) *The space X admits no pairwise disjoint uncountable family of nonempty open subsets.*

PROOF. (i) $\implies$ (ii): Let Z be a countable dense subset. The family of open balls with center at some point of Z and radius $\frac{1}{n}$ is then also countable (because a product of two countable sets is countable). So there is a sequence $\{U_n\}_{n=1}^{\infty}$ in which every such ball appears at least once. I claim that every open set of X is a union of such balls. Indeed, let U be a nonempty subset (we are allowed to take the empty union to get the empty set!), let $p \in U$, and let $\epsilon > 0$ be such that $B(p, \epsilon) \subset U$. Choose n sufficiently large such that $\frac{1}{n} < \frac{\epsilon}{2}$ and choose $z \in Z$ such that $d(z, p) < \frac{1}{2n}$. Then $p \in B(z, \frac{1}{n}) \subset B(p, \epsilon) \subset U$. It follows that U is a union of balls as claimed.
(ii) $\implies$ (iii): Let $\mathcal{B} = \{B_n\}_{n=1}^{\infty}$ be a countable base for X, and let $\{U_i\}_{i \in I}$ be an open covering of X. For each $p \in X$, we have $p \in U_i$ for some i. Since U_i is a union of elements of $\mathcal{B}$ and $p \in U_i$, we must have $p \in B_{n(p)} \subset U_i$ for some $n(p)$ depending on p. Thus we have all the essential content for a countable subcovering, and we formalize this as follows: let J be the set of all positive integers n such that B_n lies in U_i for some i: notice that J is countable! For each $n \in J$, choose $i_n \in I$ such that $B_n \subset U_{i_n}$. It then follows that $X = \bigcup_{n \in J} U_{i_n}$.
(iii) $\implies$ (i): For each $n \in \mathbb{Z}^+$, the collection $\{B(p, \frac{1}{n})\}_{p \in X}$ certainly covers X. Since X is Lindelöf, there is a countable subcover. Let Z_n be the set of centers of the elements of this countable subcover, so Z_n is a countable $\frac{1}{n}$-net. Put $Z = \bigcup_{n \in \mathbb{Z}^+} Z_n$. Then Z is a countable dense subset.
(ii) $\implies$ (iv): If X has a countable base, then so does every subspace. But an uncountable discrete space admits no countable base.
(iv) $\implies$ (v): By contrapositive: if X does admit a pairwise disjoint uncountable family of open balls, then taking the center of each ball yields an uncountable discrete subset.
(v) $\implies$ (vi) is immediate.
(vi) $\implies$ (i): Let $n \in \mathbb{Z}^+$. By Exercise 2.77 there is a maximal $\frac{1}{n}$-packing $A_n \subset X$. Then the family of open $\frac{1}{2n}$-balls centered at the points of A_n are pairwise disjoint, so by assumption A_n is countable. Therefore $A := \bigcup_{n=1}^{\infty} A_n$ is countable. Moreover A is dense: if not, then $X \setminus A$ would contain an open $\frac{1}{n}$-ball for some $n \in \mathbb{Z}^+$ and then the center of that ball could be added to A_n to obtain a larger $\frac{1}{n}$-packing. $\square$

We now get to play the good properties of separability, existence of countable bases, and Lindelöfness off against one another. For instance, we get:

COROLLARY 2.70.

a) *Every subset of a separable metric space is separable.*
b) *Every subset of a Lindelöf metric space is Lindelöf.*
c) *If $f : X \to Y$ is a continuous surjective map of metric spaces and X has a countable base, so does Y.*

We suggest that the reader pause and try to give a proof of Corollary 2.70 directly from the definition.

COROLLARY 2.71. *A compact metric space is separable.*

Proof. Since X is compact, it is Lindelöf, so by Theorem 2.69 X is separable. (Alternately, we can rerun the proof of (iii) $\implies$ (i) in Theorem 2.69 in this context: for each $n \in \mathbb{Z}^+$, X has a finite covering by open balls of radius $\frac{1}{n}$; taking the union of the centers of these balls over all $n \in \mathbb{Z}^+$ gives a countable dense subset.) $\square$

Exercise 2.81. *Let X be a separable metric space, and let $E \subset X$ be a discrete subset: every point of E is an isolated point. Show that E is countable.*

Recall that point p in a metric space is **isolated** if $\{p\}$ is an open set. If we like, we can rephrase this by saying that p admits a neighborhood of cardinality 1. Otherwise p is a **limit point**: every neighborhood of p contains points other than p. Because finite metric spaces are discrete, we can rephrase this by saying that every neighborhood of p is infinite. This little discussion perhaps prepares us for the following more technical definition.

A point p of a metric space X is an ω-**limit point** if every neighborhood of p in X is uncountable.

Theorem 2.72. *A separable metric space has at most continuum cardinality.*

Exercise 2.82. *Prove it. (Hint: think about limits of sequences.)*

Theorem 2.73. *Let X be an uncountable separable metric space. Then all but countably many points of X are ω-limit points.*

Proof. Step 1: We show that X at at least one ω-limit point. Seeking a contradiction we suppose this is not the case. Then, for every $x \in X$, let U_x be a countable neighborhood of X. By Theorem 2.69 X is Lindelöf, so the open covering $\{U_x\}_{x \in X}$ has a countable subcovering. Thus X is countable, a contradiction.
Step 2: Let Z be the set of all ω-limit points of X. Seeking a contradiction, we suppose that $X \setminus Z$ is uncountable. Then by Corollary 2.70 and Step 1, there is $x \in X \setminus Z$ that is an ω-limit point. But then *a fortiori* x is an ω-limit point of X, so $x \in Z$, a contradiction. $\square$

Theorem 2.74. *Let X be an uncountable, complete separable metric space. Then X has continuum cardinality.*

Proof. By Theorem 2.72, X has at most continuum cardinality, so it will suffice to exhibit continuum-many points of X.
Step 1: We claim that for all $\delta > 0$, there is $0 \leq \epsilon \leq \delta$ and $x, y \in X$ such that the closed ϵ-balls $B^\bullet(x, \epsilon)$ and $B^\bullet(y, \epsilon)$ are disjoint and each contain uncountably many points. Indeed, by Theorem 2.73, X has uncountably many ω-limit points. Choose two of them $x \neq y$ and take any $\epsilon < d(x, y)$.
Step 2: Applying the above construction with $\delta = 1$ we get uncountable disjoint closed subsets A_0 and A_1 each of diameter at most 1. Each of A_0 and A_1 is itself uncountable, complete and separable, so we can run the construction in A_0 and in A_1 to get uncountable disjoint closed subsets $A_{0,0}, A_{0,1}$ in A_1 and $A_{1,0}, A_{1,1}$ in A_2, each of diameter at most $\frac{1}{2}$. Continuing in this way we get for each $n \in \mathbb{Z}^+$ a pairwise disjoint family of 2^n uncountable closed subsets $A_{i_1,\ldots,i_n}$ (with $i_1, \ldots, i_n \in \{0, 1\}$) each of diameter at most 2^{-n}. Any infinite binary sequence $\epsilon \in \{0, 1\}^{\mathbb{Z}^+}$ yields a nested sequence of nonempty closed subsets of diameter approaching zero, so by completeness each such sequence has a unique intersection point p_ϵ. If $\epsilon \neq \epsilon'$

are distinct binary sequences, then for some n, $\epsilon_n \neq \epsilon_n'$, so p_ϵ and p_ϵ' are contained in disjoint subsets and are thus distinct. This gives $2^{\#\mathbb{Z}^+} = \#\mathbb{R}$ points of X. $\square$

Theorem 2.74 applies in particular to show the uncountability of $\mathbb{R}$.

10.1. Further Exercises.

EXERCISE 2.83.

a) *(S. Ivanov) Let X be a complete metric space without isolated points. Show that X has at least continuum cardinality.*
(Suggestion: the lack of isolated points implies that every closed ball of positive radius is infinite. Now run the argument of Step 2 of the proof of Theorem 2.74.)
b) *Explain why the assertion that every uncountable complete metric space has at least continuum cardinality is equivalent to the Continuum Hypothesis: i.e., that every uncountable set has at least continuum cardinality.*[4]

EXERCISE 2.84. [**MO**] *Show: a metric space X is separable if and only if every open set in X is a countable union of open balls.*

11. Compactness Revisited

11.1. Characterization of compactness in metric spaces.

The following is perhaps the single most important theorem in metric topology.

THEOREM 2.75. *Let X be a metric space. The following are equivalent:*

(i) *X is compact: every open covering of X has a finite subcovering.*
(ii) *X is sequentially compact: every sequence in X has a convergent subsequence.*
(iii) *X is limit point compact: every infinite subset of X has a limit point.*
(iv) *X is complete and totally bounded.*

PROOF. We will show (i) $\implies$ (iii) $\iff$ (ii) $\iff$ (iv) $\implies$ (i).
(i) $\implies$ (iii): Suppose X is compact, and let $A \subset X$ have no limit point in X. We must show that A is finite. Recall that $\overline{A}$ is obtained by adjoining the set A' of limit points of A, so in our case we have $\overline{A} = A \cup A' = A \cup \varnothing = A$, i.e., A is closed in the compact space X, so A is itself compact. On the other hand, no point *of* A is a limit point, so A is discrete. Thus $\{\{a\}_{a\in A}\}$ is an open covering of A, which certainly has no proper subcovering: we need all the points of A to cover A! So the given covering must itself be finite: i.e., A is finite.
(iii) $\implies$ (ii): Let $\mathbf{x}$ be a sequence in X; we must find a convergent subsequence. If some element occurs infinitely many times in the sequence, we have a constant subsequence, which is convergent. Otherwise $A = \{\mathbf{x}_n \mid n \in \mathbb{Z}^+\}$ is infinite, so it has a limit point $x \in X$ and thus we get a subsequence of $\mathbf{x}$ converging to x.
(ii) $\implies$ (iii): Let $A \subset X$ be infinite; we must show that A has a limit point in X. The infinite set A contains a countably infinite subset; enumerating these elements gives us a sequence $\{a_n\}_{n=1}^\infty$. By assumption, we have a subsequence converging to some $x \in X$, and this x is a limit point of A.

[4]Exercise 8 in §5.2 of [**Ka**] reads "Prove that every uncountable complete metric space has at least the cardinal number $\mathfrak{c}$". So it asks for a proof of the Continuum Hypothesis! But presumably Kaplansky meant to ask part a) and the absence of "without isolated points" is a typo.

(ii) $\implies$ (iv): Let $\mathbf{x}$ be a Cauchy sequence in X. By assumption $\mathbf{x}$ has a convergent subsequence, which by Lemma 2.55 implies that $\mathbf{x}$ converges: X is complete. Let $\mathbf{x}$ be a sequence in X. Then $\mathbf{x}$ has a convergent, hence Cauchy, subsequence. By Theorem 2.65, the space X is totally bounded.
(iv) $\implies$ (ii): Let $\mathbf{x}$ be a sequence in X. By total boundedness $\mathbf{x}$ admits a Cauchy subsequence, which by completeness is convergent. So X is sequentially compact.
(iv) $\implies$ (i): Seeking a contradiction, we suppose that there is an open covering $\{U_i\}_{i\in I}$ of X without a finite subcovering. Since X is totally bounded, it admits a finite covering by closed balls of radius 1. It must be the case that for at least one of these balls, say A_1, the open covering $\{U_i \cap A_1\}_{i\in I}$ of A_1 does not have a finite subcovering – for if each had a finite subcovering, by taking the finite union of these finite subcoverings we would get a finite subcovering of $\{U_i\}_{i\in I}$. Since A_1 is a closed subset of a complete, totally bounded space, it is itself complete and totally bounded. So we can cover A_1 by finitely many closed balls of radius $\frac{1}{2}$ and run the same argument, getting at least one such ball, say $A_2 \subset A_1$, for which the open covering $\{U_i \cap A_2\}_{i\in I}$ has no finite subcovering. Continuing in this way we build a nested sequence of closed balls $\{A_n\}_{n=1}^{\infty}$ of radii tending to 0, and thus also $\operatorname{diam} A_n \to 0$. By completeness there is a point $p \in \bigcap_{n=1}^{\infty} A_n$. Since $\bigcup_{i\in I} U_i = X$, certainly we have $p \in U_i$ for at least one $i \in I$. Since U_i is open, there is some $\epsilon > 0$ such that $B(p,\epsilon) \subset U_i$. Choose $N \in \mathbb{Z}^+$ such that $\operatorname{diam} A_N < \epsilon$. Then since $p \in A_N$, we have $A_N \subset B(p,\epsilon) \subset U_i$. But this means that $A_N = U_i \cap A_N$ is a one element subcovering of A_N: contradiction. □

Exercise 2.85. *Let X be a metric space. Show the following are equivalent:*

(i) *Every closed, bounded subset of X is compact.*
(ii) *The space X is complete, and bounded subsets of X are totally bounded.*

Exercise 2.86. *A metric space is **countably compact** if every countable open cover admits a finite subcover.*

a) *Show that for a metric space X, the following are equivalent:*
   (i) *X is countably compact.*
   (ii) *For any sequence $\{A_n\}_{n=1}^{\infty}$ of closed subsets, if for all finite nonempty subsets $J \subset \mathbb{Z}^+$ we have $\bigcap_{n\in J} A_n \neq \varnothing$, then $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.*
   (iii) *For any nested sequence $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ of nonempty closed subsets of X, we have $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.*
b) *Show that a metric space is compact iff it is countably compact. (Suggestion: use the assumption that X is not limit-point compact to build a countable open covering without a finite subcovering.)*

Exercise 2.87. *Let X be a metric space.*

a) *Show: every finite subset of X is compact. In particular, if X is finite, then every subset is compact.*
b) *Suppose X is topologically discrete. Show: every compact subset of X is finite.*
c) *Suppose X is infinite and not topologically discrete. Show: X has infinitely many compact subsets.*
d) *Show: a subset Y is closed iff its intersection with every compact subset of X is closed.*

11.2. Partial Limits.

Let $\mathbf{x}$ be a sequence in a metric space X. Recall that a $p \in X$ is a **partial limit** of $\mathbf{x}$ if some subsequence of $\mathbf{x}$ converges to p.

Though this concept has come up before, we have not given it much attention. This section is devoted to a more detailed analysis.

EXERCISE 2.88. *Show that the partial limits of $\{(-1)^n\}_{n=1}^{\infty}$ are precisely -1 and 1.*

EXERCISE 2.89. *Let $\{x_n\}$ be a real sequence which diverges to ∞ or to $-\infty$. Show that there are no partial limits.*

EXERCISE 2.90. *In $\mathbb{R}^2$, let $x_n = (n \cos n, n \sin n)$. Show that there are no partial limits.*

EXERCISE 2.91. *In $\mathbb{R}$, consider the sequence*

$$0, 1, \frac{1}{2}, 0, \frac{-1}{2}, -1, \frac{-3}{2}, -2, \frac{-5}{3}, \frac{-4}{3}, \dots, 3, \frac{11}{4} \dots$$

Show that every real number is a partial limit.

EXERCISE 2.92. *a) Let $\{x_n\}$ be a sequence in a metric space such that every bounded subset of the space contains only finitely many terms of the sequence. Then there are no partial limits.*
b) Show that a metric space admits a sequence as in part a) if and only if it is unbounded.

PROPOSITION 2.76. *In any compact metric space, every sequence has at least one partial limit.*

PROOF. This is a rephrasing of "compact metric spaces are sequentially compact." □

EXERCISE 2.93. *Show that a convergent sequence in a metric space has a unique partial limit: namely, the limit of the sequence.*

In general, the converse is not true: e.g. the sequence $\frac{1}{2}, 2, \frac{1}{3}, 3, \frac{1}{4}, 4, \dots, \frac{1}{n}, n, \dots$ has 0 as its only partial limit, but it does not converge.

PROPOSITION 2.77. *In a compact metric space X, a sequence with exactly one partial limit converges.*

PROOF. Let L be a partial limit of a sequence $\{x_n\}$, and suppose that the sequence does not converge to L. Then there is some $\epsilon > 0$ such that $B^\circ(L, \epsilon)$ misses infinitely many terms of the sequence. Therefore some subsequence lies in $Y = X \setminus B^\circ(L, \epsilon)$. This is a closed subset of a compact space, so it is compact, and therefore this subsequence has a partial limit $L' \in Y$, which is then a partial limit of the original sequence. Since $L \notin Y$, $L' \neq L$. □

PROPOSITION 2.78. *Let $\{x_n\}$ be a sequence in a metric space X. Then the set $\mathcal{L}$ of partial limits of $\{x_n\}$ is a closed subset.*

PROOF. We will show that the complement of $\mathcal{L}$ is open: let $y \in X \setminus \mathcal{L}$. Then there is $\epsilon > 0$ such that $B^\circ(y, \epsilon)$ contains only finitely many terms of the sequence. Now for any $z \in B^\circ(y, \epsilon)$,

$$B^\circ(z, \epsilon - d(y, z)) \subset B^\circ(y, \epsilon)$$

so $B^\circ(z, \epsilon - d(y, z))$ also contains only finitely many points of the sequence and thus z is not a partial limit of the sequence. It follows that $B^\circ(y, \epsilon) \subset X \setminus \mathcal{L}$. $\square$

Now let $\{x_n\}$ be a bounded sequence in $\mathbb{R}$: say $x_n \in [a, b]$ for all n. Since $[a, b]$ is closed, the set $\mathcal{L}$ of partial limits is contained in $[a, b]$, so it is bounded. By the previous result, $\mathcal{L}$ is closed. So $\mathcal{L}$ has a minimum and maximum element, say $\underline{L}$ and $\overline{L}$. The sequence converges iff $\underline{L} = \overline{L}$.

We claim that $\overline{L}$ can be characterized as follows: for any $\epsilon > 0$, only finitely many terms of the sequence lie in $[\overline{L}+\epsilon, b]$; and for any $\epsilon > 0$, infinitely many terms of the sequence lie in $[\overline{L} - \epsilon, b]$. Indeed, if infinitely many terms of the sequence lay in $[\overline{L} + \epsilon, b]$, then by Bolzano-Weierstrass there would be a partial limit in this interval, contradicting the definition of $\overline{L}$. The second implication is even easier: since $\overline{L}$ is a partial limit, then for all $\epsilon > 0$, the interval $[\overline{L} - \epsilon, \overline{L} + \epsilon]$ contains infinitely many terms of the sequence.

We can now relate $\overline{L}$ to the **limit supremum**. Namely, put

$$X_n = \{x_k \mid k \geq n\}$$

and put

$$\limsup s_n = \lim_{n\to\infty} \sup X_n.$$

Let us first observe that *this* limit exists: indeed, each X_n is a subset of $[a, b]$, hence bounded, hence $\sup X_n \in [a, b]$. Since $X_{n+1} \subset X_n$, $\sup X_{n+1} \leq \sup X_n$, so $\{\sup X_n\}$ forms a bounded decreasing sequence and thus converges to its least upper bound, which we call the **limit superior** of the sequence x_n.

We claim that $\limsup x_n = \overline{L}$. We will show this by showing that $\limsup x_n$ has the characteristic property of $\overline{L}$. Let $\epsilon > 0$. Then since $(\limsup x_n) + \epsilon > \limsup x_n$, then for some (and indeed all sufficiently large) N we have

$$x_n \leq supX_N < (\limsup x_n) + \epsilon,$$

showing the first part of the property: there are only finitely many terms of the sequence to the right of $(\limsup x_n) + \epsilon)$. For the second part, fix $N \in \mathbb{Z}^+$; then

$$(\limsup x_n) - \epsilon < (\limsup x_n) \leq \sup X_N,$$

so that $(\limsup x_n - \epsilon)$ is *not* an upper bound for X_N: there is some $n \geq N$ with $(\limsup x_n - \epsilon) < x_n$. Since N is arbitrary, this shows that there are infinitely many terms to the right of $(\limsup x_n - \epsilon)$.

We deduce that $\overline{L} = \limsup x_n$.

THEOREM 2.79. *Let X be a metric space. For a nonempty subset $Y \subset X$, the following are equivalent:*
(i) There is a sequence $\{x_n\}$ in X whose set of partial limits is precisely Y.
(ii) There is a countable subset $Z \subset Y$ such that $Y = \overline{Z}$.

PROOF. STEP 1: First suppose $Y = \overline{Z}$ for a countable, nonempty subset Z. If Z is finite then it is closed and $Y = Z$. In this case suppose the elements of Z are $z_1, \ldots, z_N$, and take the sequence

$$z_1, \ldots, z_N, z_1, \ldots, z_N, \ldots$$

On the other hand, if Z is countably infinite then we may enumerate its elements $\{z_n\}_{n=1}^{\infty}$. We take the sequence

$$z_1, z_1, z_2, z_1, z_2, z_3, \ldots, z_1, \ldots, z_N, \ldots.$$

In either case: since each element $z \in Z$ appears infinitely many times as a term of the sequence, there is a *constant* subsequence converging to $z \in Z$. Since the set $\mathcal{L}$ of partial limits is closed and contains Z, we must have

$$\mathcal{L} \supset \overline{Z} = Y.$$

Finally, every term of the sequence lies in the closed set Y, hence so does every term of every subsequence, and so the limit of any convergent subsequence must also lie in Y. Thus $\mathcal{L} = Y$.
STEP 2: Now let $\{x_n\}$ be any sequence in X and consider the set $\mathcal{L}$ of partial limits of the sequence. We may assume that $\mathcal{L} \neq \varnothing$. We know that $\mathcal{L}$ is closed, so it remains to show that there is a countable subset $Z \subset \mathcal{L}$ such that $\mathcal{L} = \overline{Z}$: in other words, we must show that $\mathcal{L}$ is a separable metric space. Let $W = \{x_n \mid n \in \mathbb{Z}^+\}$ be the set of terms of the sequence. Then W is countable, and arguing as above we find $\mathcal{L} \subset \overline{W}$. Therefore $\mathcal{L}$ is a subset of a separable metric space, so by Corollary 2.70, $\mathcal{L}$ is itself separable. □

Though Theorem 2.79 must have been well known for many years, I have not been able to find it in print (in either texts or articles). In fact two recent articles address the collection of partial limits of a sequence in a metric space: [**Si08**] and [**HM09**]. The results that they prove are along the lines of Theorem 2.79 but not quite as general: the main result of the latter article is that in a separable metric space every nonempty closed subset is the set of partial limits of a sequence. Moreover the proof that they give is significantly more complicated.

11.3. Lebesgue Numbers. Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of a metric space X. A real number $\delta > 0$ is a **Lebesgue number** if for all subsets A of X, if the diameter of A less then δ, then $A \subset U_i$ for at least one $i \in I$.

It is immediate from the definition that if $\delta > 0$ is a Lebesgue number for a covering $\mathcal{U}$ then so is any δ' with $0 < \delta' < \delta$.

EXAMPLE 2.12.
a) *Let $X = [0,1]$, and let $\mathcal{U}$ be the open covering $\mathcal{U} = \{U_1 = [0, \frac{2}{3}),\ U_2 = (\frac{1}{3}, 1]\}$. The subset $A = \{\frac{1}{3}, \frac{2}{3}\}$ lies in neither U_1 nor U_2 and has diameter $\frac{1}{3}$, so no $\delta > \frac{1}{3}$ can be a Lebesgue number for $\mathcal{U}$. Now let $A \subset [0,1]$ be a nonempty subset of diameter less than $\frac{1}{3}$. If $\sup A < \frac{2}{3}$ then $A \subset U_1$, while if $\sup A \geq \frac{2}{3}$ then $\inf A > \frac{1}{3}$ so $A \subset U_2$. This shows that $\delta = \frac{1}{3}$ is the largest Lebesgue number for $\mathcal{U}$.*
b) *Let $X = [0,1)$, and let $\mathcal{U}$ be the open covering $\mathcal{U} = \{U_n = [0, 1 - \frac{1}{n})\}_{n=1}^{\infty}$. This covering has no Lebesgue number: for all $\delta > 0$, the set $[1 - \frac{\delta}{2}, 1)$ has diameter $\frac{\delta}{2} < \delta$ and lies in no U_n.*

c) *Let* $X = [0, \frac{1}{2}) \cup (\frac{1}{2}, 1]$, *and let* $\mathcal{U}$ *be the open covering* $\mathcal{U} = \{U_1 = [0, \frac{1}{2}),\ U_2 = (\frac{1}{2}, 1]\}$. *For* $\delta > 0$, *the subset* $A := \{\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\}$ *has diameter* δ *and lies in neither* U_1 *nor* U_2, *so* $\mathcal{U}$ *is a finite open covering with no Lebesgue number.*

Lemma 2.80. *Every open cover* $\mathcal{U} = \{U_i\}_{i \in I}$ *of a sequentially compact metric space* X *admits a Lebesgue number.*

Proof. Seeking a contradiction, suppose that $\mathcal{U}$ admits no Lebesgue number. Then for all $n \in \mathbb{Z}^+$, $\frac{1}{n}$ is not a Lebesgue number for the covering, so there is a subset $A_n \subset X$ with $\operatorname{diam} A_n < \frac{1}{n}$ and such that for no $i \in I$ do we have $A_n \subset U_i$. Certainly such an A_n is nonempty: choose $a_n \in A_n$. Since X is sequentially compact, there is a subsequence a_{n_k} converging to some $L \in X$. Choose $i \in I$ such that $L \in U_i$, and let $\epsilon > 0$ be such that $B^\circ(L, \epsilon) \subset U_i$. Choose $K \in \mathbb{Z}^+$ such that $d(L, a_{n_K}) < \frac{\epsilon}{2}$ and $\operatorname{diam}(A_{n_K}) < \frac{\epsilon}{2}$. Then we have

$$A_{n_K} \subset B^\circ(L, \epsilon) \subset U_i,$$

a contradiction. □

Proposition 2.81. *Let* $f : X \to Y$ *be a continuous map between metric spaces. For* $\epsilon > 0$, *suppose that the open cover* $\mathcal{U}_\epsilon = \{f^{-1}(B(y, \frac{\epsilon}{2})\}_{y \in Y}$ *of* X *admits a Lebesgue number* δ. *Then* f *is* (ϵ, δ)*-UC.*

Proof. Let $x, x' \in X$ be such that $d(x, x') < \delta$. Since

$$\operatorname{diam}\{x, x'\} = d(x, x') < \delta$$

and δ is a Lebesgue number for the covering $\mathcal{U}_\epsilon$, there is $y \in Y$ such that

$$f(\{x, x'\}) \subset B^\circ(y, \frac{\epsilon}{2})$$

and thus

$$d(f(x), f(x')) < d(f(x), y)) + d(y, f(x')) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$ □

Theorem 2.82 (Uniform Continuity Theorem). *Let* $f : X \to Y$ *be a continuous map betwen metric spaces. If* X *is compact, then* f *is uniformly continuous.*

Proof. Let $\epsilon > 0$. By Lemma 2.80, the covering $\{f^{-1}(B(y, \frac{\epsilon}{2}))\}_{y \in Y}$ of X has a Lebesgue number $\delta > 0$, and then by Proposition 2.81, the function f is (ϵ, δ)-UC. Thus f is uniformly continuous. □

11.4. Further Exercises.

Exercise 2.94. *Use the sequential characterization of uniform continuity (Exercise 2.52) to give another proof of the Uniform Continuity Theorem (Theorem 2.82). (Suggestion: suppose that* $d(x_n, y_n) \to 0$ *and* $d(f(x_n), f(y_n))$ *does not approach* 0. *Then after passing to a subsequence there is* $\epsilon > 0$ *such that* $d(f(x_n), f(y_n)) \geq \epsilon$ *fo all* n. *Extract two further subsequences to get a contradiction.)*

Exercise 2.95. *For a metric space* X, *let*

$$\mathfrak{d}(X) := \inf_{x, x' \in X,\ x \neq x'} d(x, x').$$

a) *Show: if* $\mathfrak{d}(X) > 0$, *then* $\mathfrak{d}(X)$ *is a Lebesgue number for* every *open covering of* X.

b) *Show: if* $\mathfrak{d}(X) > 0$*, then for all* $\delta > 0$*, there is an open covering* $\mathcal{U}_\delta$ *of* X *for which* δ *is* not *a Lebesgue number.*

As this point we are so used to analyzing the property of compactness in metric spaces that it is natural to wonder whether any important consequence of compactness in metric spaces is in fact equivalent to it. (This has turned out to be the case for sequential compactness, limit point compactness and countable compactness.) The preceding exercise shows that this is *not* the case for the last two consequences of compactness we have studied: for instance, consider the integers endowed with the usual Euclidean metric. In the terminology of Exercise 2.95 we have $\mathfrak{d}(\mathbb{Z}) = 1$, so 1 is a Lebesgue number for every open covering of the integers. (This is not a deep fact: in $\mathbb{Z}$, sets of diameter less than 1 consist of at most a single point.) Moreover, not only is every continuous function $f : \mathbb{Z} \to Y$ uniformly continuous, but *every* function $f : \mathbb{Z} \to Y$ is uniformly contiuous. Indeed, since $\mathfrak{d}(\mathbb{Z}) > 0$, the space $\mathbb{Z}$ is **uniformly discrete**.

12. Extension Theorems

Let X and Y be metric spaces, let $A \subset X$ be a subset, and let

$$f : A \to Y$$

be a continuous function. We say f **extends to** X if there is a continuous map

$$F : X \to Y$$

such that

$$\forall x \in A, F(x) = f(x).$$

We also say that **F extends f** and write $F|_A = f$. (That F must be continuous is suppressed from the terminology: this is supposed to be understood.) We are interested in both the **uniqueness** and the **existence** of the extension.

Proposition 2.83. *Let* X *and* Y *be metric spaces, let* $A \subset X$ *and let* $f : A \subset Y$ *be a continuous function. If* A *is dense in* X*, then there is at most one continuous function* $F : X \to Y$ *such that* $F|_A = f$*.*

Proof. Suppose $F_1, F_2 : X \to Y$ both extend $f : A \to Y$, and let $x \in X$. Since A is dense, there is a sequence $\mathbf{a}$ in A which converges to x. Then

$$F_1(x) = F_1(\lim_{n\to\infty} \mathbf{a}_n) = \lim_{n\to\infty} F_1(\mathbf{a}_n) = \lim_{n\to\infty} f(\mathbf{a}_n)$$

$$= \lim_{n\to\infty} F_2(\mathbf{a}_n) = F_2(\lim_{n\to\infty} \mathbf{a}_n) = F_2(x). \qquad \square$$

Exercise 2.96. *Let* $A \subset X$*, and let* $f : A \to Y$ *be a continuous map. Show that* f *has at most one continuous extension to* $F : \overline{A} \to Y$*.*

Proposition 2.84. *Let* $f : X \to Y$ *be a uniformly continuous map of metric spaces. Let* $\mathbf{x}$ *be a Cauchy sequence in* X*. Then* $f(\mathbf{x})$ *is a Cauchy sequence in* Y*.*

Proof. Let $\epsilon > 0$. By uniform continuity, there is $\delta > 0$ such that for all $y, z \in X$, if $d(y, z) \le \delta$ then $d(f(y), f(z)) \le \epsilon$. Since $\mathbf{x}$ is Cauchy, there is $N \in \mathbb{Z}^+$ such that if $m, n \ge N$ then $d(\mathbf{x}_m, \mathbf{x}_n) \le \delta$. For all $m, n \ge N$, $d(f(\mathbf{x}_m), f(\mathbf{x}_n)) \le \epsilon$. $\square$

THEOREM 2.85. *Let X be a metric space, Y a complete metric space, $A \subset X$ a dense subset, and let $f : A \to Y$ be uniformly continuous.*
a) There is a unique continuous map $F : X \to Y$ extending f (i.e., such that $F|_A = f$).
b) The map $F : X \to Y$ is uniformly continuous.
c) If f is an isometric embedding, then so is F.

PROOF. a) Exercise 2.96 shows that if $F : \overline{A} \to Y$ is any continuous extension, then $F(x)$ must be $\lim_{n\to\infty} f(\mathbf{a}_n)$ for any sequence $\mathbf{a} \to x$. It remains to show that this limit actually exists and does not depend upon the choice of sequence $\mathbf{a}$ which converges to x. But we are well prepared for this: since $\mathbf{a} \to x$ in X, as a sequence in A, $\mathbf{a}$ is Cauchy. Since f is *uniformly* continuous, the sequence $f(\mathbf{a})$ is Cauchy. Since Y is complete, the sequence $f(\mathbf{a})$ converges. If $\mathbf{b}$ is another sequence in A converging to x, then $d(\mathbf{a}_n, \mathbf{b}_n) \to 0$, so uniform continuity gives $d(f(\mathbf{a}_n), f(\mathbf{b}_n)) \to 0$.
b) Fix $\epsilon > 0$, and choose $\delta > 0$ such that f is $(\frac{\epsilon}{2}, \delta)$-uniformly continuous. We claim that F is (ϵ, δ)-uniformly continuous. Let $x, y \in X$ with $d(x, y) < \delta$. Choose sequences $\mathbf{a}$ and $\mathbf{b}$ in A converging to x and y respectively. Then

$$d(x, y) = \lim_{n\to\infty} d(\mathbf{a}_n, \mathbf{b}_n),$$

so by our choice of δ for all sufficiently large n we have $d(\mathbf{a}_n, \mathbf{b}_n) < \delta$. For such n we have $d(f(\mathbf{a}_n), f(\mathbf{b}_n)) < \frac{\epsilon}{2}$, so

$$d(f(x), f(y)) = \lim_{n\to\infty} d(f(\mathbf{a}_n), f(\mathbf{b}_n)) \leq \frac{\epsilon}{2} < \epsilon.$$

c) Suppose f is an isometric embedding, let $x, y \in X$ and choose sequences $\mathbf{a}$, $\mathbf{b}$ in A converging to x and y respectively. Then

$$\begin{aligned} d(f(x), f(y)) &= d(f(\lim_{n\to\infty} \mathbf{a}_n), f(\lim_{n\to\infty} \mathbf{b}_n)) = \lim_{n\to\infty} d(f(\mathbf{a}_n), f(\mathbf{b}_n)) \\ &= \lim_{n\to\infty} d(\mathbf{a}_n, \mathbf{b}_n) = d(x, y). \end{aligned}$$

□

EXERCISE 2.97. *The proof of Theorem 2.85 does not quite show the simpler-looking statement that if $f : A \to Y$ is (ϵ, δ)-uniformly continuous then so is the extended function $F : X \to Y$. Show that this is in fact true.*

EXERCISE 2.98. *Maintain the setting of Theorem 2.85.*
a) Show: if $f : A \to Y$ is contractive, so is F.
b) Show: if $f : A \to Y$ is Lipschitz, so is F. Show in fact that the optimal Lipschitz constants are equal: $L(F) = L(f)$.

EXERCISE 2.99. *Let $P : \mathbb{R} \to \mathbb{R}$ be a polynomial function, i.e., there are $a_0, \ldots, a_d \in \mathbb{R}$ such that $P(x) = a_d x^d + \ldots + a_1 x + a_0$.*
a) Show that P is uniformly continuous iff its degree d is at most 1.
b) Taking $A = \mathbb{Q}$, $X = Y = \mathbb{R}$, use part a) to show that uniform continuity is not a necessary condition for the existence of a continuous extension.

EXERCISE 2.100. *Say that a function $f : X \to Y$ between metric spaces is* ***Cauchy continuous*** *if for every Cauchy sequence* $\mathbf{x}$ *in X, $f(\mathbf{x})$ is Cauchy in Y.*

a) *Show: uniform continuity implies Cauchy continuity implies continuity.*
b) *Show: Theorem 2.85 holds if "uniform continuity" is replaced everywhere by "Cauchy continuity".*

c) *Let X be totally bounded. Show: Cauchy continuity implies uniform continuity.*

THEOREM 2.86. *(Tietze Extension Theorem) Let X be a metric space, $Y \subset X$ a closed subset, and let $f : Y \to \mathbb{R}$ be a continuous function. Then there is a continuous function $F : X \to \mathbb{R}$ with $F|_Y = f$. If $f(Y) \subset [a, b]$, we may choose F so as to have $F(X) \subset [a, b]$.*

PROOF. We will give a proof of a more general version of this result later on in these notes: Theorem 7.4. □

COROLLARY 2.87. *For a metric space X, the following are equivalent:*
(i) X is compact.
(ii) Every continuous function $f : X \to \mathbb{R}$ is bounded.

PROOF. (i) $\implies$ (ii): this is the Extreme Value Theorem.
(ii) $\implies$ (i): By contraposition and using Theorem 2.75 it suffices to assume that X is not limit point compact – thus admits a countably infinite, discrete closed subset Y – and from this build an unbounded continuous real-valued function. Namely, write $Y = \{x_n\}_{n=1}^{\infty}$ and define f on A by $f(n) = x_n$. By the Tietze Extension Theorem, there is a continuous function $F : X \to \mathbb{R}$ with $F|_Y = f$. Since F takes on all positive integer values, it is unbounded. □

LEMMA 2.88. *(Transport of Structure) Let (X, d_X) and (Y, d_Y) be metric spaces, and let $\Phi : X \to Y$ be a homeomorphism. Then*

$$d' : (x_1, x_2) \mapsto d_Y(\Phi(x_1), \Phi(x_2))$$

is a metric on X that is topologically equivalent to d. Moreover $\Phi : (X, d') \to (Y, d_Y)$ is an isometry.

EXERCISE 2.101. *Prove Lemma 2.88.*

COROLLARY 2.89. *For a metric space (X, d), the following are equivalent:*
(i) X is compact.
(ii) Every metric d' on X that is topologically equivalent to d is totally bounded.
(iii) Every metric d' on X that is topologically equivalent to d is bounded.

PROOF. (i) $\implies$ (ii): Compactness is a topological property and compact metric spaces are totally bounded.
(ii) $\implies$ (iii) is immediate.
$\neg$ (i) $\implies$ $\neg$ (iii): Suppose X is not compact. Then by Corollary 2.87 there is an unbounded continuous map $f : X \to \mathbb{R}$. We define a function

$$\Phi : X \to X \times \mathbb{R},\ x \mapsto (x, f(x)).$$

Endow $X \times \mathbb{R}$ with the maximum metric

$$\tilde{d}((x_1, y_1), (x_2, y_2)) = \max(d(x_1, x_2), |y_1 - y_2|)$$

and let $Y = \Phi(X)$. Since f is unbounded, so is $(Y, \tilde{d})$. Moreover, $\Phi : X \to Y$ is a homeomorphism: we leave this as an exercise. Apply Lemma 2.88: we get that

$$d' = \Phi^{-1} \circ \tilde{d} \circ \Phi$$

is a metric on X which is topologically equivalent to d. Moreover (X, d') is isometric to $(Y, \tilde{d})$, hence unbounded. □

EXERCISE 2.102. *Show that the map Φ appearing in the proof of Corollary 2.89 is a homeomorphism.*

13. The function space $C_b(X,Y)$

Let X be a nonempty set, and let Y be a nonempty metric space. Put

$$\operatorname{Map}(X,Y) := \{f : X \to Y\},$$

i.e., the set of all functions from X to Y. A function $f : X \to Y$ is **bounded** if $f(X)$ is a bounded subset of Y. We denote by $\operatorname{Map}_b(X,Y) \subset \operatorname{Map}(X,Y)$ the subset of all bounded functions $f : X \to Y$.

We can endow $\operatorname{Map}_b(X,Y)$ with a natural metric, namely,

$$d : f, g \in \operatorname{Map}_b(X,Y) \mapsto \sup_{x \in X} d(f(x), g(x)).$$

Here the boundedness of f and g ensures that the supremum is finite; notice that e.g. we could not do this with $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x$ and $g : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$.

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence in $\operatorname{Map}(X,Y)$, and let $f \in \operatorname{Map}(X,Y)$. We say that f_n **converges uniformly to** f **on** X and write $f_n \xrightarrow{u} f$ if for all $\epsilon > 0$, there is $N \in \mathbb{Z}^+$ such that for all $\epsilon > 0$, we have $\sup_{x \in X} d(f_n(x), f(x)) < \epsilon$. Now we make some observations:

LEMMA 2.90. *Suppose $f_n \xrightarrow{u} f$ and that each f_n is bounded. Then f is bounded and $f_n \to f$ in $\operatorname{Map}_b(X,Y)$.*

PROOF. Choose $N \in \mathbb{Z}^+$ such that for all $x \in X$, we have $d(f_N(x), f(x)) \leq 1$. Let $D := \operatorname{diam}(f_N(X))$, and fix $\star \in f_N(X)$. Then for all $x \in X$, we have

$$d(f(x), \star) \leq d(f(x), f_N(x)) + d(f_N(x), \star) \leq D + 1,$$

so $f(X) \subset B^{\bullet}(\star)(D+1)$ and thus f is bounded. The fact that f_n converges to f with respect to the given metric on $\operatorname{Map}_b(X,Y)$ is immediate. □

Now we suppose that X is also a metric space. Let

$$C(X,Y) := \{\text{continuous functions } f : X \to Y\},$$

$$C_b(X,Y) := C_(X,Y) \cap \operatorname{Map}_b(X,Y) = \{\text{bounded continuous functions } f : X \to Y\}.$$

In particular, $C_b(X,Y)$ is a subset of the metric space $\operatorname{Map}_b(X,Y)$ hence a metric space in its own right.

LEMMA 2.91. *Let X and Y be metric spaces.*

a) *Let $\{f_n : X \to Y\}$ be a sequence of continuous functions converging uniformly on X to a function $f : X \to Y$. Then f is continuous.*
b) *The subset $C_b(X,Y)$ is closed in $\operatorname{Map}_b(X,Y)$.*

PROOF. a) Fix $\epsilon > 0$. Choose $N \in \mathbb{Z}^+$ such that for all $x \in X$, we have $d(f(x), f_N(x)) \leq \frac{\epsilon}{3}$. Choose $\delta > 0$ such that if $d(x,x') \leq \delta$ then $d(f_N(x), f_N(x')) \leq \frac{\epsilon}{3}$. Then, if $d(x,x') \leq \delta$ we have

$$d(f(x), f(x')) \leq d(f(x), f_N(x)) + d(f_N(x), f_N(x')) + d(f_N(x'), f(x')) \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

b) The assertion is equivalent to the fact that a uniform limit of bounded continuous functions from X to Y is a bounded continuous function from X to Y, which is immediate from part a) and Lemma 2.90. $\square$

THEOREM 2.92. *2 Let X and Y be metric spaces. The following are equivalent:*

(i) *The space Y is complete.*
(ii) *The space $\mathrm{Map}_b(X, Y)$ is complete.*
(iii) *The space $C_b(X, Y)$ is complete.*

PROOF. (i) $\implies$ (ii): Let $\{f_n : X \to Y\}_{n=1}^{\infty}$ be a Cauchy sequence in $\mathrm{Map}_b(X, Y)$. Then for all $x \in X$ and all $m, n \in \mathbb{Z}^+$ we have $d(f_m(x), f_n(y)) \leq d(f_m, f_n)$, so the sequence $\{f_n(x)\}$ is Cauchy in the complete metric space Y and thus is convergent; call the limit $f(x)$. This of course defines a function $f : X \to Y$. By Lemma 2.90, it is sufficient to show that $f_n \stackrel{u}{\to} f$.

To see this, fix $\epsilon > 0$, and choose $N \in \mathbb{Z}^+$ such that for all $m, n \geq N$ we have $d(f_m, f_n) < \frac{\epsilon}{2}$. Let $x \in X$, and choose $m(x) \in \mathbb{Z}^+$ such that $d(f_{m(x)}(x), f(x)) < \frac{\epsilon}{2}$. Then, for all $n \geq N$ we have

$$d(f_n(x), f(x)) \leq d(f_n(x), d(f_{m(x)}(x)) + d(f_{m(x)}(x), f(x)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

(ii) $\implies$ (iii): If $\mathrm{Map}_b(X, Y)$ is complete, then by Lemma 2.91b), we have that $C_b(X, Y)$ is a closed subset of a complete metric space, hence itself complete.
$\neg$ (i) $\implies$ $\neg$ (iii): Suppose Y is not complete, and let $\{y_n\}$ be a sequence in Y that is Cauchy but not convergent. For $n \in \mathbb{Z}^+$, let f_n map every $x \in X$ to y_n; clearly f_n is bounded and continuous. Then for all $m, n \in \mathbb{Z}^+$ we have $d(f_m, f_n) = d(y_m, y_n)$, so $\{f_n\}$ is Cauchy in $C_b(X, Y)$. Suppose there is a function $f : X \to Y$ such that $f_n \stackrel{u}{\to} f$, and fix $\star \in X$. Then $y_n = f_n(\star) \to f(\star)$, a contradiction. $\square$

EXERCISE 2.103. *Let X be a set, and let Y be a bounded metric space. Then we have*

$$\mathrm{Map}_b(X, Y) = \mathrm{Map}(X, Y) = Y^X.$$

Suppose X is infinite and $\#Y \geq 2$. Show that our metric d on $\mathrm{Map}_b(X, Y)$, viewed as a metric on the Cartesian product Y^X, is not *good. More precisely, show that there is a sequence $\{f_n\}$ in Y^X such that $f_n(x)$ converges for all $x \in X$ but f_n is not convergent in $\mathrm{Map}_b(X, Y)$. In other words, the sequence converges pointwise but not uniformly.*[5]

14. Completion

Completeness is such a desirable property that given a metric space which is not complete we would like to fix it by adding in the missing limits of Cauchy sequences. Of course, the above description is purely intuitive: although we may visualize $\mathbb{R}$ as being constructed from $\mathbb{Q}$ by "filling in the irrational holes", it is much less clear that something like this can be done for an arbitrary metric space.

The matter of the problem is this: given a metric space X, we want to find a complete metric space Y and an isometric embedding

$$\iota : X \hookrightarrow Y.$$

[5]One can use Theorem 2.36 here, but that is not needed: one can also argue directly.

However this can clearly be done in many ways: e.g. we can isometrically embed $\mathbb{Q}$ in $\mathbb{R}$ but also in $\mathbb{R}^N$ for any N (in many ways, but e.g. as $r \mapsto (r, 0, \ldots, 0)$). Intuitively, the embedding $\mathbb{Q} \hookrightarrow \mathbb{R}$ feels natural while (e.g.) the embedding $\mathbb{Q} \hookrightarrow \mathbb{R}^{17}$ feels wasteful. If we reflect on this for a bit, we see that we can essentially recover the good case from the bad case by passing from Y to the closure of $\iota(X)$ in Y. We then get $\mathbb{R} \times 0^{16}$, which is evidently isometric to $\mathbb{R}$ (and even compatibly with the embedding of $\mathbb{Q}$: more on this shortly).

In general: if $\iota : X \to Y$ is an isometric embedding into a complete metric space, then (because closed subsets of complete metric spaces are complete), $\iota : X \to \overline{\iota(X)}$ is an isometric embedding into a complete metric space with dense image, or for short a **dense isometric embedding**. Remarkably, adding the density condition gives us a uniqueness result.

Lemma 2.93. *Let X be a metric space, and for $i = 1, 2$, let $\iota_i : X \to Y_i$ be dense isometric embeddings into a complete metric space. Then there is a unique isometry $\Phi : Y_1 \to Y_2$ such that $\iota_2 = \Phi \circ \iota_1$.*

Proof. We simplify the notation by identifying X with its isometric image $\iota_1(X)$ in Y_1. Then a continuous map $\Phi : Y_1 \to Y_2$ such that $\iota_2 = \Phi \circ \iota_1$ means a continuous extension of ι_2 from the dense subspace X of Y_1 to all of Y_1, and thus the existence and uniqueness of Φ follows from Theorem 2.85, which moreover shows that since ι_2 is an isometric embedding, so is Φ.

It remains to show that Φ is surjective. We give two proofs of this: the first is rather concrete, and the second is an argument that can be made in many situations for objects defined via "universal mapping properties."

First Proof: Let $y \in Y_2$. Since ι_2 is a dense embedding, there is a sequence $x_\bullet$ in X such that $\iota_2(x_n) \to y_2$. Thus the sequence $\iota_2(x_\bullet)$ is Cauchy, and since ι_2 is an isometric embedding the sequence $\{x_n\}$ is Cauchy in $X \subset Y_1$. Since Y_1 is complete, there is $y_1 \in Y$ such that $x_n \to y_1$. Then we have

$$\Phi(x_n) \to \Phi(y_1)$$

but also, since $x_\bullet$ is a sequence in X,

$$\Phi(x_n) = \iota_2(x_n) \to y_2.$$

Since the limit of a sequence in a metric space is unique, we conclude $\Phi(y_1) = y_2$. Thus Φ is surjective.

Second Proof: Applying what we have done so far with the roles of Y_1 and Y_2 reversed, we get an isometric embedding $\Phi' : Y_2 \to Y_1$ such that $\iota_1 = \Phi' \circ \iota_2$. The compositions $\Phi' \circ \Phi$ is a continuous map that restricts to the identity on the dense subspace $\iota_1(X)$ of Y_1, so by Proposition 2.83 we get $\Phi' \circ \Phi = 1_{Y_1}$. In exactly the same way we get that $\Phi \circ \Phi' = 1_{Y_2}$. Thus Φ and Φ' are mutually inverse bijections, so Φ (and also Φ') is an isometry. $\square$

This motivates the following key definition: a **completion** of a metric space X is a complete metric space $\hat{X}$ and a dense isometric embedding $\iota : X \hookrightarrow \hat{X}$. It follows from Lemma 2.93 that *if* a metric space admits a completion then any two completions are isometric (and even more: the embedding into the completion is essentially unique). Thus for any metric space X we have associated a new metric space $\hat{X}$. Well, not quite: there is the small matter of proving the *existence* of $\hat{X}$!

To know "everything but existence" perhaps seems bizarre (even Anselmian?). In fact it is quite common in modern mathematics to define an object by a characteristic property and then be left with the task of "constructing" the object, which can generally be done in several different ways. In this particular instance there are two standard constructions of "the" completion $\hat{X}$ of a metric space X.

LEMMA 2.94. *Let Y be a dense subset of a metric space X. If every Cauchy sequence in Y converges to an element of X, then X is complete.*

PROOF. Let $x_\bullet$ be a Cauchy sequence in X. Since Y is dense in X, for all $n \in \mathbb{Z}^+$ there is $y_n \in Y$ such that $d(x_n, y_n) < \frac{1}{n}$. Then $\lim_{n\to\infty} d(x_n, y_n) = 0$, so by Lemma 2.54a) the sequence $y_\bullet$ is also Cauchy. By hypothesis, this means that $y_n \to L \in X$. Applying Lemma 2.54b), we get that $x_n \to L$. So X is complete. $\square$

FIRST CONSTRUCTION OF THE COMPLETION:

We will give a detailed sketch of the proof, leaving some "claims" for the reader to verify as exercises. Let (X, d) be a metric space. Put $X^\infty = \prod_{i=1}^\infty X$, the set of all sequences in X. Inside X^∞, we define $\mathcal{X}$ to be the set of all Cauchy sequences. For $x_\bullet, y_\bullet \in \mathcal{X}$, we define

$$d(x_\bullet, y_\bullet) := \lim_{n\to\infty} d(x_n, y_n).$$

We need to check that this limit exists. Here is one slick argument for it: the sequence $x_\bullet \times y_\bullet$ is Cauchy in $X \times X$ and the metric function $d : X \times X \to \mathbb{R}$ is uniformly continuous, so the sequence $n \mapsto d(x_n, y_n)$ is Cauchy in $\mathbb{R}$, hence convergent since $\mathbb{R}$ is complete. It is also possible (indeed, straightforward) to check this directly: let $\epsilon > 0$. For $n \in \mathbb{Z}^+$, let

$$S_{x,N} := \{x_n \mid n \geq N\},\ S_{y,N} := \{y_n \mid n \geq N\}.$$

Since $x_\bullet$ and $y_\bullet$ are Cauchy, there is $N \in \mathbb{Z}^+$ such that

$$\operatorname{diam} S_{x,N}, \operatorname{diam} S_{y,N} < \frac{\epsilon}{2}.$$

But given bounded subsets S and T of a metric space, for any $x_1, x_2 \in S$ and $y_1, y_2 \in S$ we have

$$|d(x_1, y_1) - d(x_2, y_2)| \leq \operatorname{diam} S + \operatorname{diam} T.$$

It follows that for all $m, n \geq N$ we have

$$|d(x_m, y_m) - d(x_n, y_n)| < \epsilon,$$

so the sequence $n \mapsto d(x_n, y_n)$ is Cauchy in $\mathbb{R}$ and thus convergent.

We would like $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ to be a metric function. Unforunately it isn't for a rather shallow reason: there will in general be distinct Cauchy sequences that have distance zero from each other.[6] We fix this as follows: we introduce an equivalence relation on $\mathcal{X}$ by $x_\bullet \sim y_\bullet$ if $\rho(x_n, y_n) \to 0$. Put $\hat{X} = \mathcal{X}/\sim$. For a Cauchy sequence $x_\bullet$ in X, we denote its class in $\hat{X}$ by $[x_\bullet]$.
FIRST CLAIM: d descends to a well-defined function

$$d : \hat{X} \times \hat{X} \to \mathbb{R};$$

[6]In fact this *always* happens when X consists of more than one point!

in other words, it makes sense to define

$$d([x_\bullet],[y_\bullet]) \coloneqq \lim_{n\to\infty} d(x_n,y_n).$$

It is now straightforward to check the SECOND CLAIM: $d : \hat{X}\times\hat{X} \to \mathbb{R}$ is a metric function.

Now we define a map $\iota : X \to \mathcal{X} \to \hat{X}$ by

$$x \mapsto [(x,x,\ldots)],$$

i.e., the equivalence class of the constant sequence at x. Then $\iota : X \to \hat{X}$ is an isometric embedding. Moreover, $\iota(X)$ is dense in $\hat{X}$: given $[x_\bullet] \in \hat{X}$.
THIRD CLAIM: the sequence $\iota(x_\bullet)$ (i.e., each term is the equivalence class of the constant Cauchy sequence x_n) converges to $[x_\bullet]$ in $\hat{X}$.

FOURTH CLAIM: If $x_\bullet$ is a Cauchy sequence in X, then the sequence $\iota(x_n)$ (in which the nth term is the class of the constant sequence x_n) converges in $\hat{X}$ to $[x_\bullet]$. Assuming this, Lemma 2.94 now shows that $\hat{X}$ is complete and thus that $\iota : X \hookrightarrow \hat{X}$ is a completion of X.

EXERCISE 2.104. *Supply proofs of the four claims made above.*

EXERCISE 2.105. *Show that $\hat{X}$ is complete without using Lemma 2.94 but rather by a direct diagonalization-type argument.*

SECOND CONSTRUCTION OF THE COMPLETION:
By Theorem 2.92, the set $C_b(X,\mathbb{R})$ of bounded continuous functions $f : X \to \mathbb{R}$ is a complete metric space under $d(f,g) = \sup_{x\in X} d(f(x),g(x))$. Fix a point $\star \in X$. For $x \in X$, let $D_x : X \to \mathbb{R}$ be given by

$$D_x(y) \coloneqq d(\star,y) - d(x,y).$$

By the Reverse Triangle Inequality (Proposition 1) we have

$$D_x(y) \le |d(\star,y) - d(x,y)| \le d(\star,x),$$

so D_x is bounded. Moreover D_x is continuous: e.g. one may apply Proposition 2.35 and Exercise 2.58. Thus $D_x \in C_b(X,\mathbb{R})$, and we get a map

$$\mathcal{D} : X \to C_b(X,\mathbb{R}),\ x \mapsto D_x.$$

Moreover, for $x,x' \in X$, we have one the one hand that

$$d(D_x,D_{x'}) = \sup_{y\in X} |D_x(y) - D_{x'}(y)| = \sup_{y\in X} |d(x,y) - d(x',y)| \le d(x,x')$$

and on the other that

$$d(D_x,D_{x'}) = \sup_{y\in X} |d(x,y) - d(x',y)| \ge |d(x,x) - d(x',x)| = d(x,x').$$

Thus we have

$$d(D_x,D_{x'}) = d(x,x')$$

i.e., $\mathcal{D} : X \hookrightarrow C_b(X,\mathbb{R})$ is an isometric embedding of X into a complete metric space. Therefore the map $X \hookrightarrow \overline{\mathcal{D}(X)}$ is a completion of X.

COROLLARY 2.95. *(Functoriality of completion)*
a) Let $f : X \to Y$ be a uniformly continuous map between metric spaces. Then there exists a unique map $F : \hat{X} \to \hat{Y}$ making the following diagram commute:

$$X \xrightarrow{f} Y$$

$$\hat{X} \xrightarrow{F} \hat{Y}.$$

b) If f is an isometric embedding, so is F.
c) If f is an isometry, so is F.

Proof. a) The map $f' : X \to Y \hookrightarrow \hat{Y}$, being a composition of uniformly continuous maps, is uniformly continuous. Applying the universal property of completion to f' gives a unique extension $\hat{X} \to \hat{Y}$.
Part b) follows from Theorem 2.85c). As for part c), if f is an isometry, so is its inverse f^{-1}. The extension of f^{-1} to a mapping from $\hat{Y}$ to $\hat{X}$ is easily seen to be the inverse function of F. □

14.1. Total Boundedness Revisited.

Lemma 2.96. *Let Y be a dense subspace of a metric space X. Then X is totally bounded iff Y is totally bounded.*

Proof. If X is totally bounded, then every subspace of X is totally bounded, so we do not need the density of Y for this direction. Conversely, suppose Y is totally bounded, and let $\epsilon > 0$. Then there is a finite ϵ-net N in Y. I claim that for any $\epsilon' > \epsilon$, we have that N is a finite ϵ'-net in X. Indeed, let $x \in X$. Since Y is dense in X, there is $y \in Y$ with $d(x, y) < \epsilon' - \epsilon$, and there is $n \in N$ with $d(y, n) < \epsilon$, so $d(x, n) < \epsilon'$. It follows that X is totally bounded. □

Theorem 2.97. *For a metric space X, the following are equivalent:*
(i) X is totally bounded.
(ii) The completion of X is compact.

Proof. Let $\iota : X \hookrightarrow \widehat{X}$ be "the" isometric embedding of X into its completion.
(i) $\implies$ (ii): By Lemma 2.96, since X is totally bounded and dense in $\widehat{X}$, also $\widehat{X}$ is totally bounded. Of course $\widehat{X}$ is complete, so by Theorem 2.75 $\widehat{X}$ is compact.
(ii) $\implies$ (i): If $\widehat{X}$ is compact, then $\widehat{X}$ is totally bounded by Theorem 2.75, hence so is its subspace X. □

We deduce the following interesting characterization of total boundedness.

Corollary 2.98. *A metric space X can be isometrically embedded in a compact metric space iff it is totally bounded.*

Exercise 2.106. *a) Prove it.*
b) Let X be a metric space. Suppose there is a compact metric space C and a uniform embedding $f : X \to C$ – i.e., the map $f : X \to f(X)$ is a uniformeomorphism. Show that X is totally bounded.

The previous exercise shows that "isometric embedding" can be weakened to "uniform embedding" without changing the result. What about topological embeddings? This time the answer must be different, as e.g. $\mathbb{R}$ can be topologically embedded in a compact space: e.g. the arctangent function is a homeomorphism from $\mathbb{R}$ to $(\frac{-\pi}{2}, \frac{\pi}{2})$ and thus a topological embedding from $\mathbb{R}$ to $[\frac{-\pi}{2}, \frac{\pi}{2}]$. Here is something in the other direction.

Lemma 2.99. *A metric space that can be topologically embedded in a compact metric space is separable.*

Proof. Indeed, let $f : X \hookrightarrow C$ be a topological embedding into a compact metric space C. In particular C is separable. Moreover X is homeomorphic to $f(X)$, which is a subspace of C, hence also separable by Corollary 2.70. $\square$

Much more interestingly, the converse of Lemma 2.99 holds: every separable metric space can be topologically embedded in a compact metric space. This is quite a striking result. In particular implies that separability is precisely the topologically invariant part of the metrically stronger property of total boundedness, in the sense that for a metric space (X, d), there is a topologically equivalent totally bounded metric d' on X iff X is separable.

Unfortunately this result lies beyond our present means. Well, in truth it is not really so unfortunate: we take it as a motivation to develop more purely topological tools. In fact we will later quickly deduce this result from one of the most important theorems in all of general topology: Theorem 7.16.

15. Cantor Space

We begin with the most classical definition of the "middle thirds Cantor set." We will define C as the intersection of a nested family $\{C_n\}_{n=0}^{\infty}$ of closed subsets of the unit interval $[0, 1]$. We define

$$C_0 \coloneqq [0, 1]$$

and

$$C_1 \coloneqq C_0 \setminus (\frac{1}{3}, \frac{2}{3}) = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

Observe that C_1 is obtained from the line segment C_0 by removing the "open middle third." Since C_1 is now a disjoint union of two closed line segments, it makes sense to iterate this process by removing the middle third of each one:

$$C_2 = C_1 \setminus \left((\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9}) \right).$$

And we may continue in this manner: having defined C_n as a disjoint union of 2^n closed subintervals of $[0, 1]$, each of length $\frac{1}{3^n}$, we define C_{n+1} by removing the open middle third of each of these line segments, so that C_{n+1} is a disjoint union of 2^{n+1} closed subintervals of $[0, 1]$, each of length $\frac{1}{3^{n+1}}$. Finally, we define the **Cantor set**

$$C \coloneqq \bigcap_{n=0}^{\infty} C_n.$$

Let us make some observations about the Cantor set C:
(i) C is a closed subset of $[0, 1]$ – indeed, it is the intersection of a family of closed subsets – hence a compact metric space.
(ii) C is nonempty. It has some obvious points: e.g. $0 \in C$ and $1 \in C$. Indeed, because we remove only elements of the interior of each subinterval at each stage, all of the elements that are endpoints of any of the subintervals C_n remain in C: this is a countably infinite set of points.
(iiii) C has continuum cardinality. In fact, let $s : \mathbb{Z}^+ \to \{0, 2\}$ be any function. Then we may use s to define a nested sequence of nonempty closed subintervals inside C: namely, C_1 consists of two closed subintervals; if $s(1) = 0$, we choose the left subinterval, wheras if $s(1) = 2$ we choose the right subinterval. Either way, the intersection of the chosen closed subinterval with C_2 is a union of two subintervals;

if $s(2) = 0$ we choose the left one, and if $s(2) = 2$ we choose the right one. And so forth. By compactness, this sequence of subintervals has a (unique, since the diameters approach 0) common intersection point, which gives rise to a point of C. It is easy to see that this process of assigning to each element of C an element of $\{0,2\}^{\mathbb{Z}^+}$ is a bijection. This shows, in particular, that C has continuum cardinality (though it has more profound consequences as well).

EXERCISE 2.107. *a) Show that the sequence $s : \mathbb{Z}^+ \to \{0,2\}$ associated to an element of the Cantor set C is a* ***trinary*** *(i.e., base 3) expansion of C as an element of $\mathbb{R}$. This explains why we used 2 and not 1.*
b) In this way we can define C as the elements of $[0,1]$ admtiting a trinary expansion in which 1 does not appear. Note though that there may also be a trinary expansion in which 1 appears, e.g.

$$.1222\cdots = .2000\cdots = \frac{2}{3}.$$

Characterize all points of C admitting a trinary expansion in which 1 does *appear.*

(iv) C is a perfect subset of $\mathbb{R}$: i.e., it is closed and every point is a limit point. For this, observe that in the canonical sequence representation of C given above, if $x, y \in C$ are such that the first n terms of the sequence agree, then x and y lie in a common closed subinterval of length $\frac{1}{2^n}$ so have distance at most 2^n. From this it follows easily that every element $x \in C$ is the limit of a sequence in $C \setminus \{x\}$: e.g. choose x_n so as to have the first n coordinates agree with x and to have the $n+1$st coordinate disagree with x.

(v) C is not connected. In fact, it has the following property, which lies at the other extreme: given any $x \neq y \in C$, then for some $n \in \mathbb{Z}^+$, x and y do not lie in the same closed subinterval of $[0,1]$ (equivalently, it is not the case that they agree in their first n coordinates). Let I_n be such a closed subinterval containing x but not y, and put $U = I_n \cap C$, $V = C \setminus U$. Then U and V are disjoint open subets of C such that $x \in U$ and $v \in V$.

LEMMA 2.100. *Endow $\{0,1\}$ with the discrete metric and $\{0,1\}^\infty = \prod_{n=1}^\infty \{0,1\}$ with a good metric as in Corollary 2.37. Then the map $T : C \to \{0,1\}^\infty$ that maps the $\sum_{n=1}^\infty \frac{a_n}{3^n} \in C$ to $\{\frac{a_n}{2}\}_{n=1}^\infty$ is a homeomorphism.*

PROOF. In (iii) above we observed that T is a bijection. For $n \in \mathbb{Z}^+$, let

$$T_n : C \to \{0,1\}, \ \sum_{n=1}^{\infty} \frac{a_n}{3^n} \mapsto \frac{a_n}{2}.$$

The map T_n is locally constant, hence continuous. By Proposition 2.38, T is continuous. To show that T^{-1} is continuous is equivalent to showing that for every open subset U of C, its image $T(C) = (T^{-1})^{-1}(C)$ is open in $\{0,1\}^\infty$. Since U is open, $C \setminus U$ is closed in the compact space C, hence compact, hence $T(C \setminus U)$ is compact in the metric space $\{0,1\}^\infty$, hence $T(C \setminus U) = \{0,1\}^\infty \setminus T(U)$ is closed in $\{0,1\}^\infty$, hence $T(U)$ is open in $\{0,1\}^\infty$. □

16. Contractions and Attractions

Let X, Y be metric spaces. A map $f : X \to Y$ is a **contraction** if there is $\alpha \in (0,1)$ such that for all $x, x' \in X$ we have

$$d(f(x), f(x')) \leq \alpha d(x, x').$$

We say (somewhat clumsily) that α is a "contractive constant" for f. We say that $f : X \to Y$ is a **weak contraction** (or **weakly contractive**) if for all $x, x' \in X$, we have

$$d(f(x), f(x')) < d(x, x').$$

Let X be a set, and let $f : X \to X$ be a map. A **fixed point** of f is a point $x \in X$ such that $f(x) = x$. For $n \in \mathbb{Z}^+$, let $f^{\circ n}$ denote $f \circ f \circ \cdots \circ f$ ($n-1$ $\circ$'s in all). We put $f^{\circ 0} = 1_X$. We say that $\star \in X$ is **attracting** if for all $x \in X$, the sequence of iterates $f^{\circ n}(x)$ converges to $\star$. Clearly there is at most one attracting point.

Lemma 2.101. *Let X be a metric space, let $f : X \to X$ be a continuous function, and let $x \in X$. If the sequence of iterates $f^{\circ n}(x)$ converges to $L \in X$, then L is a fixed point of f.*

Proof. Since $f^{\circ n}(x) \to L$ and f is continuous, we have

$$f^{\circ n+1}(x) = f(f^{\circ n}(x)) \to f(L).$$

Since a sequence in a metric space has at most one limit, we conclude $f(L) = L$. $\square$

Exercise 2.108. *Let X be a metric space, and let $f : X \to X$ be map. Let $\star \in X$ be an attracting point.*
a) Suppose f is continuous. Show: $\star$ is the unique fixed point of f.
b) Exhibit a discontinuous map $f : X \to X$ with an attracting point that is not a fixed point.

Let $f : X \to X$ be a map of metric spaces. A fixed point x for f is **locally attracting** if there is $\delta > 0$ such that for all $y \in B(x, \delta)$, the sequence of iterates $f^{\circ n}(y)$ converges to x.

Exercise 2.109. *Let $f : [0, 1] \to [0, 1]$ by $f(x) = x^2$. Show: 0 is a locally attracting fixed point of f, and 1 is a fixed point of f that is not locally attracting.*

Lemma 2.102. *Let X be a metric space, and let $f : X \to X$ be weakly attractive. Then f has at most one fixed point.*

Proof. Seeking a contradiction, suppose that $x \neq x'$ are two fixed points of f. Then we have

$$d(x, x') = d(f(x), f(x')) < d(x, x'),$$

a contradiction. $\square$

16.1. Banach's Fixed Point Theorem.

Exercise 2.110. *Let X be a metric space, let $f : X \to X$ be a contraction, and let $n \in \mathbb{Z}^+$. Show: $f^{\circ n} : X \to X$ is also a contraction. Moreover, if C is a contractive constant for f, show: C^n is a contractive constant for $f^{\circ n}$.*

Theorem 2.103. *(Banach Fixed Point Theorem* [**Ba22**]*) Let (X, d) be a complete metric space, and let $f : X \to X$ be a contraction mapping with contractive constant $C \in (0, 1)$. Then:*
a) The point $\star$ is an attracting point for f.
b) Let $x \in X$. Then for all $n \in \mathbb{Z}^+$, we have:

$$d(f^{\circ n}(x), \star) \leq \frac{C^n}{1 - C} d(f(x), x), \tag{10}$$

$$d(f^{\circ n}(x), \star) \leq C d(f^{\circ (n-1)}(x), \star), \tag{11}$$

$$d(f^{\circ n}(x), \star) \leq \frac{C}{1-C} d(f^{\circ n}(x), f^{\circ (n-1)}(x)). \tag{12}$$

Proof. a) Let $x \in X$. We abbreviate $x_0 := x$, $x_n := f^{\circ n}(x)$ for $n \in \mathbb{Z}^+$. For integers $n \geq N \geq 1$ and $k \geq 0$, we have

$$d(x_{n+k}, x_n) \leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \ldots + d(x_{n+1}, x_n)$$
$$\leq C^{n+k-1} d(x_1, x_0) + C^{n+k-2} d(x_1, x_0) + \ldots + C^n d(x_1, x_0)$$
$$= d(x_1, x_0) C^n \left(1 + C + \ldots + C^{k-1}\right) = d(x_1, x_0) C^n \left(\frac{1 - C^k}{1 - C}\right) < \left(\frac{d(x_1, x_0)}{1 - C}\right) C^n.$$

Since $|C| < 1$, $C^n \to 0$, and it follows that the sequence of iterates $\{x_n\}$ is Cauchy. Since X is complete, this sequence converges, say to $\star$. By Lemma 2.101, $\star$ is a fixed point of f, and by Exercise 2.108, it is the *unique* fixed point. So every sequence of iterates converges to the same point $\star$, and thus $\star$ is an attracting point for f.
b) Let $x \in X$. Above we showed that

$$d(f^{\circ n}(x), f^{\circ (n+k)}(x)) < \frac{C^n}{1-C} d(f(x), x).$$

Taking the limit as $k \to \infty$ gives (10). Moreover we have

$$d(f^{\circ n}(x), \star) = d(f(f^{\circ (n-1)}(x)), f(\star)) \leq C d(f^{\circ (n-1)}(x), \star),$$

which is (11). Using (11) and the triangle inequality, we get

$$d(f^{\circ n}, \star) \leq C d(f^{\circ (n-1)}(x), f^{\circ n}(x)) + C d(f^{\circ n}(x), \star),$$

which gives (12). □

16.2. Refinements and Variations on Banach's Fixed Point Theorem.

Corollary 2.104. *Let X be a complete metric space, and let $f : X \to X$ be a map such that $f^{\circ N} : X \to X$ is a contraction for some $N \in \mathbb{Z}^+$. Then f has an attracting fixed point.*

Proof. Let $\star$ be the unique fixed point of $f^{\circ N}$.
Step 1: Since

$$f(\star) = f(f^{\circ N}(\star)) = f^{\circ N}(f(\star)),$$

we have that $f(\star)$ is a fixed point of $f^{\circ N}$. By Lemma 2.102 we get $f(\star) = \star$.
Step 2: Let $x \in X$. Fix $0 \leq r \leq N-1$ and consider the sequence of points $\{f^{\circ Nk+r}(x\}_{k=1}^{\infty}$. We have

$$f^{\circ (Nk+r)}(x) = f^{\circ NK}(f^{\circ r}(x)) = (f^{\circ N})^{\circ k}(f^{\circ r}(x).$$

Since $f^{\circ N}$ is a contraction, so is $(f^{\circ N})^{\circ k}$, so it has a unique fixed point which is moreover attracting, and clearly this unique fixed point is $\star$, so

$$f^{\circ (Nk+r)}(x) \to \star.$$

We have thus partitioned the sequence of iterates $f^{\circ n}(x)$ into finitely many subsequences, each of which converges to $\star$. So we have $f^{\circ n}(x) \to \star$ and thus $\star$ is an attracting fixed point for f. □

In Corollary 2.104 we say "attracting fixed point" rather than "attracting point" because f need not be continuous, so *a priori* an attracting point of f need not be a fixed point.

EXAMPLE 2.13. *Let X be a metric space.*

• *If X is topologically discrete, then every function $f : X \to X$ is continuous, and an attracting point for f is necessarily a fixed point.*
• *If X is not topologically discrete, it has a nonisolated point x and a sequence of distinct points $x_n \neq x$ converging to x. Define $f : X \to X$ as follows: for each $n \in \mathbb{Z}^+$, let $f(x_n) = x_{n+1}$; for all other points $y \in X$, we put $f(y) = x_1$. Then x is an attracting point for f that is not a fixed point. It follows that f is not continuous.*

EXAMPLE 2.14. *Let X be a metric space, and let $x \neq y$ be distinct points of X. Define a function $f : X \to X$ as follows: $f(y) = f(x) = x$, and for all $z \notin \{x, y\}$, $f(z) = y$.*

Then x is an attracting point for f: indeed, for all $z \in X$, we have $f^{\circ n}(z) = x$ for all $n \geq 2$. Moreover $f^{\circ 2}$ is constant (hence a contraction!).

The function f is continuous iff x and y are isolated points of X.

EXERCISE 2.111. *a) Let $f : (0, \infty) \to (0, \infty)$ by $f(x) = \frac{x}{2}$. Show that f is contractive, but has no fixed point. Since $(0, \infty)$ is not complete, this does not contradict Theorem 2.103.*
b) But it is hard not to notice that f extends continuously to $[0, \infty)$ and 0 is a fixed point of the extension. Generalize this as follows: let X be a metric space, and let $f : X \to X$ be a Lipschitz map, with Lipschitz constant C. Let $\tilde{X}$ be "the" completion of X. Show that there is a unique continuous extension of f to $\tilde{f} : \tilde{X} \to \tilde{X}$, and moreover C is a Lipschitz contant for $\tilde{f}$. Deduce that if f is a contraction, then after extending to the completion $\tilde{X}$ there is a unique fixed point.

Theorem 2.103 can fail for weakly contractive maps:

EXERCISE 2.112. *(Conrad* [**CdC**]*) Let $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \log(1 + e^x)$.*
a) Show: f is weakly contractive.
b) Show: f has no fixed point.

However, a weakly contractive map on a *compact* space must have a fixed point.

THEOREM 2.105. *(Edelstein* [**Ed62**]*) Let X be a compact metric space, and let $f : X \to X$ be a weakly contractive mapping. Then f has an attacting point.*

PROOF. We follow [**CdC**].
Step 1: We claim f has a fixed point. To see this, let

$$g : X \to \mathbb{R}, x \mapsto d(x, f(x)).$$

Since f is continuous, so is g. By Corollary 2.44, g attains a minimum value: there is $\star \in X$ such that for all $x \in X$, we have $d(\star, f(\star)) \leq d(x, f(x))$. But if $f(\star) \neq \star$, then

$$g(f(\star)) = d(f(\star), f(f(\star))) < d(\star, f(\star)) = g(\star),$$

a contradiction. So $\star$ is a fixed point for f.
Step 2: Let $x \in X$. If for some $N \in \mathbb{Z}^+$ we have $f^{\circ N}(x) = \star$, then for all $n \geq N$ we have $f^{\circ n}(x) = \star$, and certainly we have $f^{\circ n}(x) \to \star$. So we may assume that $f^{\circ n}(x) \neq \star$ for all $n \in \mathbb{Z}^+$. Put $d_n := d(f^{\circ n}(x), \star)$. Since f is weakly contractive and $f(\star) = \star$, we have that $\{d_n\}_{n=1}^{\infty}$ is a strictly decreasing sequence of positive numbers, hence convergent to its infimum $d \geq 0$. We have $d = 0$ iff $f^{\circ n}(x) \to \star$, so seeking a contradiction we assume that d is positive.

Since compact metric spaces are sequentially compact, there is a strictly increasing sequence $\{n_k\}$ of positive integers such that $f^{\circ n_k}(x)$ converges to some $y \in X$. The continuity of the metric function gives $d(x_{n_k}, \star) \to d(y, \star)$. On the one hand, continuity of f and d gives $d(f(x_{n_k}), \star) \to d(f(y), \star)$; while on the other hand, $d(f(x_{n_k}), \star) = d(x_{n_k} + 1, \star) = d_{n_k+1} \to d$, and thus

$$d(f(y), f(\star)) = d(f(y), \star) = d = d(y, \star).$$

Thus $y = \star$ and $d = d(y, \star) = 0$, a contradiction. $\square$

As Conrad writes in [**CdC**], "It is natural to wonder if the compactness of X might force f in Theorem 2.105 to be a contraction after all, so [Theorem 2.103] would apply. For instance, the ratios $\frac{d(f(x), f(x'))}{d(x,x')}$ for $x \neq x'$ are always less than 1, so they should be less than or equal to some definitive constant $C < 1$ by compactness. But this reasoning is bogus, because $\frac{d(f(x), f(x'))}{d(x,x')}$ is not defined on [the diagonal $\Delta = \{(x,x) \mid x \in X\}$, and $X \times X \setminus \Delta$ is open in the compact metric space $X \times X$ and hence *not* compact. There is no way to show [that the] f in [Theorem 2.105] has to be a contraction, since there are examples where it isn't."

EXAMPLE 2.15. *(Conrad) Let $f : [0,1] \to [0,1]$ by $f(x) = \frac{1}{1+x}$.*
a) Show: f is weakly contractive but not a contraction.
b) Show: $\frac{-1+\sqrt{5}}{2}$ is the unique fixed point of f.
c) Show that $f([\frac{1}{2}, 1]) \subset [\frac{1}{2}, 1]$ and that $f : [\frac{1}{2}, 1] \to [\frac{1}{2}, 1]$ is a contraction. Deduce that f has a fixed point.
d) Show that $f^{\circ 2}$ is a contraction. Deduce that f has a fixed point.

We now wish to pursue fixed point / attraction theorems for continuous functions $f : I \to I$, where I is a subinterval of the real line. However, notice that $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x + 1$ is continuous (indeed, an isometry) and has no fixed points. Moreover the map $f :\to \mathbb{R}$ by $f(x) = \log(1 + e^x)$ of Example 2.112 is weakly contractive and has no fixed points. So we must expand our horizons a bit. In these examples, the sequences of iterates still exhibit a simple limiting behavior: for all $x \in \mathbb{R}$, we have $f^{\circ n}(x) < f^{\circ n+1}(x)$ and $f^{\circ n}(x) \to \infty$.

To ease the statement of the result, we introduce the following notation: for a sequence of real numbers $\{x_n\}$ and $L \in \mathbb{R} \cup \{\pm\infty\}$, we write $x_n \uparrow L$ if $x_n < x_{n+1}$ for all n and $x_n \to L$. For a sequence $\{x_n\}$ in $\mathbb{R}$ and $L \in \mathbb{R} \cup \{-\infty\}$, we write $x_n \downarrow L$ if $x_n > x_{n+1}$ for all n and $x_n \to L$ in the extended real numbers.

THEOREM 2.106. *Let $I \subset \mathbb{R}$ be an interval, and let $f : I \to I$ be continuous.*
a) Exactly one the following holds:
(i) The function f has a fixed point in I.
(ii) We have that $\sup I \notin I$ and for all $x \in I$, $f^{\circ n}(x) \uparrow \sup I$.
(iiii) We have that $\inf I \notin I$ and for all $x \in I$, $f^{\circ n}(x) \downarrow \inf I$.
b) If $I = [a,b]$, then f has a fixed point in I.

PROOF. a) Define $g : I \to \mathbb{R}$ by $x \mapsto f(x) - x$. Then $\star \in I$ is a fixed point of f iff it is a root of g, so we may assume g has no root in I and show that either (ii) or (iii) holds. Since g has no roots, by the Intermediate Value Theorem, since g has no roots we either have (I) $f(x) > x$ for all $x \in I$ or (II) $f(x) < x$ for all $x \in I$. If (I) holds, then if $\sup I \in I$ we would have $f(\sup I) > \sup I$, a contradiction. For $x \in I$, the sequence $\{f^{\circ n}(x)\}$ is strictly increasing so converges to its supremum

S. If $S < \sup I$, then $S \in I$ and thus S is a fixed point of f by Lemma 2.101, contradiction. If (II) holds, the argument is very similar and is left to the reader.
b) If $I = [a, b]$ then $\inf I, \sup I \in I$, so the result follows from part a). $\square$

In case (ii) of Theorem 2.106, it is reasonable to call $\sup I$ an attracting point of f. If $\sup I < \infty$ then $I \cup \{\sup I\}$ is still a metric space, and it is not hard to show that when $f(x) > x$ for all $x \in I$, the function f has a unique continuous extension to $I \cup \{\sup I\}$. If $\sup I = \infty$ then we can still give $I \cup \{\sup I\}$ the order topology, and as soon as we discuss continuous functions on arbitrary topological spaces, the reader can check that again $f(x) > x$ implies that f has a unique continuous extension to $I \cup \{\sup I\}$. Of course the analogous discussion holds for case (iii).

We employ this terminology in the following result.

THEOREM 2.107. *Let $I \subset \mathbb{R}$ be an interval, and let $f : I \to I$ be weakly contractive. Then f has an attracting point in* $[\inf I, \sup I]$.

PROOF. In partciular f is continuous. We may assume that neither $\inf I$ nor $\sup I$ is an attracting point for f, so by Theorem 2.106 there is a fixed point $\star \in I$ for f. For $x \in I$, put $d = d(x, \star)$. Then the sequence of iterates $f^{\circ n}(x)$ lies in the closed bounded interval $[\star - d, \star + d]$. By Bolzano-Weierstrass, there is a subsequence x_{n_k} converging to $y \in [\star - d, \star + d]$. Since each x_{n_k} lies in I, y lies in $I \cup \{\inf I, \sup I\}$. If $y = \sup I$ and $\sup I \notin I$ then there must be k such that $\star < x_{n_k} < x_{n_{k+1}}$, contradicting weak contractivity, so $\sup I \in I$; similarly, if $y = \inf I$ then $\inf I \in I$. We can thus argue exactly as in the proof of Theorem 2.105: let $d = \lim_{n\to\infty} d(x_n, \star)$, and suppose $d > 0$. Then

$$|x_{n_k} - \star| \to |y - \star|,$$

$$|f(x_{n_k}) - \star| \to |f(y) - \star|,$$

$$|f(x_{n_k}) - \star| = |x_{n_k+1} - \star| \to d,$$

so

$$|f(y) - f(\star)| = |f(y) - \star| = d = |y - \star|.$$

Thus $y = \star$ and $d = d(y, \star) = 0$, a contradiction. $\square$

When $I = \mathbb{R}$, Theorem 2.107 is due to A. Beardon [**Be06**]. When I is closed and bounded, Theorem 2.107 is a special case of Theorem 2.105. As we have seen, the ideas of Conrad's proof of Theorem 2.105 also work to prove Theorem 2.107.

EXERCISE 2.113. *Let X be a nonempty metric space in which all closed, bounded subsets are compact. Let $f : X \to X$ be weakly contractive. Show: if $\star$ is a fixed point of f, then it is an attracting point for f.*

16.3. Picard's Theorem. In this section we will give one of the most classical and important applications of Banach's Fixed Point Theorem, namely a local existence and uniqueness result for differential equations.

By a "rectangle" in a Euclidean space $\mathbb{R}^N$ we will mean a Cartesian product $\prod_{i=1}^N [a_i, b_i]$ of closed, bounded intervals. We denote by $|| \cdot ||_\infty : \mathbb{R}^N \to \mathbb{R}$ the norm $(x_1, \ldots, x_N) \mapsto \max_{i=1}^N |x_i|$. Recall from §2.6 that the function $d_\infty : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ by $d_\infty(y_1, y_2)) := ||y_1 - y_2||_\infty$ makes $\mathbb{R}^N$ into a complete metric space whose underlying topology is the same as the Euclidean topology.

THEOREM 2.108 (Picard's Local Existence and Uniqueness Theorem).
Let $D = D_t \times D_y \subset \mathbb{R} \times \mathbb{R}^N$ *be a closed rectangle, and let* $(t_0, y_0) \in D^\circ$. *Let* $f : D \to \mathbb{R}^N$ *be a function satisfying:*
(C1) f *is continuous, and*
(C2) There is $C > 0$ *such that:*

$$\forall (t, y_1), (t, y_2) \in D, \ ||f(t, y_1) - f(t, y_2)||_\infty \leq M||y_1 - y_2||_\infty.$$

Then: for some $\delta > 0$ *there is a unique function* $y : [t_0 - \delta, t_0 + \delta] \to \mathbb{R}^N$ *such that* $\frac{dy}{dt} = f(y)$ *and* $y(t_0) = y_0$.

PROOF. In this proof we will make use of the Riemann integral for functions $g : [a, b] \to \mathbb{R}^N$. This vector-valued Riemann integral is a smaller variation on the usual Riemann integral for functions $g : [a, b] \to \mathbb{R}$ than the Riemann integral of multivariate functions $g : \prod_{i=1}^N [a_i, b_i] \to \mathbb{R}$. In fact the theory of Riemann integration extends with very few modifications to functions taking values in any Banach space, but in the case of $\mathbb{R}^N$ it is really nothing else than the integral of each of the component functions: if $g = (g_1, \ldots, g_N)$, then $\int_a^b g := (\int_a^b g_1, \ldots, \int_a^b g_N)$.

The first key observation is this: a function $y : [t_0 - \delta, t_0 + \delta] \to \mathbb{R}^N$ satisfies the Initial Value Problem in the theorem's statement if and only if it satisfies the **integral equation**

$$y(t) = y_0 + \int_{t_0}^t f(x, y(x))dt. \tag{13}$$

Since D is compact and f is continuous, it is bounded: let $M > 0$ be such that

$$\forall (t, y) \in D, \ ||f(t, y)||_\infty \leq M.$$

Also by continuity of f we may choose $\delta > 0$ such that $M\delta < 1$ and $(t, y) \in D$ if $|t - t_0| \leq \delta$ and $||y - y_0||_\infty \leq M\delta$.

Let $Y := B^\bullet(y_0, M\delta)$, so Y is a closed subset of D_y and thus also a closed subset of $(\mathbb{R}^N, d_\infty)$. Let

$$\mathcal{C} := C([t_0 - \delta, t_0 + \delta], Y),$$

be the set of all continuous functions from $[t_0 - \delta, t_0 + \delta]$ to Y. Since $[t_0 - \delta, t_0 + \delta]$ is compact, we have that $\mathcal{C} = C_b([t_0 - \delta, t_0 + \delta], Y)$ is a complete metric space under the supremum metric of §2.13. We claim that

$$I : \mathcal{C} \to \mathcal{C}, \varphi \mapsto y_0 + \int_{t_0}^t f(x, \varphi(x))dx$$

is a contraction mapping on the complete metric space $\mathcal{C}$. If so, by the Banach Fixed Point Theorem (Theorem 2.103), the mapping I has a unique fixed point, which means that our Initial Value Problem has a unique solution.

The verification of the claim has two parts. First we must show that if $\varphi \in \mathcal{C}$ then also $I(\varphi) \in \mathcal{C}$. For this: if $|t - t_0| \leq \delta$, then

$$||I(\varphi)(t) - y_0||_\infty = ||\int_{t_0}^t f(x), \varphi(x))dx||_\infty.$$

For all t, every component of $f(t, \varphi(t))$ has absolute value at most M, so the integral of that component has absolute value at most $M|t - t_0|$. This shows that $I(\varphi)(t)$ is uniformly continuous. Moreover, since $|t - t_0| \leq \delta$, we have $||I(\varphi(t) - y_0||_\infty \leq M\delta$, so $I(\varphi(t)) \in Y$, so $I(\varphi) \in \mathcal{C}$.

Second we show that the map I is a contraction with constant $C\delta < 1$: let $\varphi_1, \varphi_2 \in \mathcal{C}$. Then:

$$d(I(\varphi_1), I(\varphi_2)) = \sup_{t \in [t_0-\delta, t_0+\delta]} ||I(\varphi_1(t) - I(\varphi_2(t))||_\infty$$

$$= \sup_{t \in [t_0-\delta, t_0+\delta]} \left|\left| \int_t^{t_0} (f(x, \varphi_1(x)) - f(x, \varphi_2(x)) \, dx \right|\right|$$

$$\leq \sup_{t \in [t_0-\delta, t_0+\delta]} \int_{t_0}^{t} \left|\left| f(x, \varphi_1(x)) - f(x, \varphi_2(x) \right|\right| dx$$

$$\leq C \sup_{t \in [t_0-\delta, t_0+\delta]} \int_{t_0}^{t} ||\varphi_1(x) - \varphi_2(x)||_\infty dx$$

$$\leq C|t - t_0| \sup_{t \in [t_0-\delta, t_0+\delta]} ||\varphi_1(t) - \varphi_2(t)||_\infty \leq C\delta d(\varphi_1, \varphi_2).$$

Thus, I is indeed a contraction mapping on $\mathcal{C}$, which completes the proof. □

Let us make some remarks on the statement and proof. First, it is amusing to see how little analysis is actually used in the proof of this fundamental result about systems of ODE's. We don't even need careful definitions of the derivative and integral; only their most basic properties are being used.

Second, the proof gives more than the statement of the theorem. It (together with the proof of Banach's Fixed Point Theorem) shows that starting with the constant function $\varphi_{y_0} : t \mapsto y_0$, the desired function y is the limit $\lim_{n\to\infty} I^{\circ n}(\varphi_{y_0})$, with a rate of convergence that is guaranteed to be exponentially fast by (10). The functions $I^{\circ n}(\varphi_{y_0})$ are called **Picard iterates**.

Third, Theorem 2.108 has been variously attributed to Picard, Lindelöf, Lipschitz and Cauchy. The first proofs of this result did not use Banach's Fixed Point Theorem. However, the paper in which Banach introduces his fixed point theorem [**Ba22**] is mainly concerned with integral equations, of which (13) is one. This application of topological fixed point theorems to analysis was however enormously influential; since many many other analytic results have been proven in that way.

Now let us examine the condition (C2) on the function f that appears as a hypothesis in the proof. It is a kind of Lipschitz condition, but in the second variable only. It is natural to wonder first how strong this assumption is and second, whether it is really needed to obtain existence and uniqueness of solutions.

As for the first point:

Proposition 2.109. *Maintain the notation of Theorem 2.108. Then $f : D \to \mathbb{R}^N$ satisifies condition (C2) if each of its partial derivatives $\frac{\partial f}{\partial y_j}$ for $1 \leq j \leq N$ (i.e., its partial derivatives with respect to every variable but the t variable) exist and are continuous on D.*

We leave the proof as an exercise for the interested reader. In general, one should use a multivariate version of the Mean Value Theorem (e.g. [**Rud**, Thm. 9.40], generalized from the two to N variables in the evident way). In the $N = 1$ case one can use the ordinary Mean Value Theorem from calculus, with an argument that quickly reduces to the observation that a continuously differentiable function

$f : [a, b] \to \mathbb{R}$ is Lipschitz, with Lipschitz constant $\max_{t\in[a,b]} |f'(t)|$.

Proposition 2.109 shows that the condition (C2) in Theorem 2.108 is rather mild. But is it really necessary: what happens if we just assume that f is continuous?

EXAMPLE 2.16. *Suppose* $f : [-1, 1] \times [0, 1] \to \mathbb{R}$ *by* $f(t, y) = \sqrt{|y|}$*. An evident solution to the Initial Value Problem* $\frac{dy}{dt} = \sqrt{|y|}$ *is the identically* 0 *function. But there is also another solution:*

$$y(t) := \begin{cases} \frac{t^2}{4} & t \geq 0 \\ \frac{-t^2}{4} & t < 0 \end{cases}.$$

It is easy to see that f *does not satisfy the Lipschitz condition (C2).*

Thus in the absence of (C2) the solution to the Initial Value Problem $\frac{dy}{dt} = f(x, y)$, $y(t_0) = y_0$ need not be *unique*. As for existence, one has the following result:

THEOREM 2.110 (Peano's Local Existence Theorem). *Maintain the notation of Theorem 2.108, but suppose only that* $f : D \to \mathbb{R}^N$ *satisfies (C1): i.e., is continuous. Then for some* $\delta > 0$ *there is* at least one function $y : [t_0 - \delta, t_0 + \delta] \to \mathbb{R}^N$ *such that* $\frac{dy}{dt} = f(x, y)$ *and* $y(t_0) = y_0$.

The proof of Peano's Theorem uses results on the topology of function spaces that are unfortunately not yet covered in these notes. We hope to eventually include them and then give a proof of Peano's Theorem.

17. Distance Between Subsets

Let (X, d) be a metric space, and let A, B be nonempty subsets of X. It is natural to consider the "distance from A to B." There is indeed a useful such notion, although it behaves in sometimes tricky and unexpected ways. We define

$$d(A, B) := \inf\{d(a, b) \mid a \in A, \ b \in B\}.$$

Thus $d(A, B) \in [0, \infty)$. Since the infimum of a nonempty set of non-negative real numbers is an adherent point of that set, in general there are sequences $\{a_n\}$ in A and $\{b_n\}$ in B such that

$$\lim_{n\to\infty} d(a_n, b_n) = d(A, B).$$

Certainly we have $d(A, B) = 0$ if $A \cap B \neq \varnothing$, but the converse does not hold: for instance, in the real numbers with the usual metric, let $A = \{0\}$ and $B = (0, 1]$. Then for all $n \in \mathbb{Z}^+$ we have $0 \in A$ and $\frac{1}{n} \in B$, so $d(A, B) \leq d(0, \frac{1}{n}) = \frac{1}{n}$, and thus $d(A, B) = 0$.

LEMMA 2.111. *For nonempty subsets* A *and* B *of a metric space* X*, we have*

$$d(A, B) = d(\overline{A}, \overline{B}).$$

PROOF. Whenever we have subsets

$$\varnothing \subsetneq A_1 \subset A_2, \ \varnothing \subsetneq B_1 \subset B_2$$

of X, we have

$$d(A_2, B_2) \leq d(A_1, B_1)$$

since upon passage from a subset of real numbers to a larger set, the infimum does not increase. This shows that

$$d(\overline{A}, \overline{B}) \leq d(A, B).$$

Conversely, choose sequences $\{a_n\}$ in $\overline{A}$ and $\{b_n\}$ in $\overline{B}$ such that $d(a_n, b_n) \to d(\overline{A}, \overline{B})$. We may choose sequences $\{y_n\}$ in A and $\{z_n\}$ in B such that $d(a_,y_n) \to 0$ and $d(b_n, z_n) \to 0$. It follows from the Quadrilateral Inequality (Exercise 2.15) that

$$|d(y_n, z_n) - d(a_n, b_n)| \to 0.$$

By Lemma 2.54b) we have $d(a_n, b_n) \to d(\overline{A}, \overline{B})$. $\square$

Henceforth we consider the distance between closed subsets. There are two natural questions to address here: first, under what circumstances do we have $d(A, B) = d(a, b)$ for some $a \in A$, $b \in B$? When this occurs we will say that A and B are **distance-attained**. Second, under what circumstances do we have $d(A, B) > 0$? When this occurs we will say that A and B are **distance-separated**.

In general, we say that two subsets A and B of a metric space (later, of a topological space) are separated if $\overline{A} \cap B = A \cap \overline{B} = \varnothing$. Lemma 2.111 implies that distance-separated subsets must be separated, and our first example of a pair of disjoint subsets $A = \{0\}$, $B = (0, 1]$ was not separated. However, two sets can be separated but not distance-separated.

EXAMPLE 2.17. a) *In $\mathbb{R}^2$, let $A = \{(x, \frac{1}{x} \mid x > 0\}$ and let $B = \mathbb{R} \times \{0\}$. Then A and B are closed and pairwise disjoint hence separated. But for all $n \in \mathbb{Z}^+$ we have $a_n = (n, \frac{1}{n}) \in A$ and $b_n = (n, 0) \in B$ and $d(a_n, b_n) = \frac{1}{n}$. So $d(A, B) = 0$.*

b) *In $\mathbb{R}$, let $A = \mathbb{Z}^+$ and let $B = \{n + 2^{-n} \mid n \in \mathbb{Z}^+\}$. Then A and B are closed subsets: indeed, both are sequences that tend to ∞ so have no limit points in R. But for all $n \in \mathbb{Z}^+$ we have $a_n = n \in A$ and $b_n = n + 2^{-n} \in B$ and $d(a_n, b_n) = 2^{-n}$. So $d(A, B) = 0$.*

In the above example one cannot help but notice that in both cases, each of A and B were closed but unbounded, so in particular neither A nor B was compact. Indeed the compactness of both A and B solves our problems:

THEOREM 2.112. *Let A and B be nonempty compact subsets of a metric space X. Then A and B are distance-attained.*

PROOF. Let $\{a_n\}$ be a sequence in A and $\{b_n\}$ a sequence in B such that $d(a_n, b_n) \to d(A, B)$. Since compact metric spaces are sequentially compact, we may extract a convergent subsequence of $\{a_n\}$ and then a convergent subsubsequence of $\{b_{n_k}\}$. After reindexing, this gives a sequence $\{\alpha_n\}$ in A and a sequence $\{\beta_n\}$ in B such that $\alpha_n \to a \in A$ and $\beta_n \to b \in B$ such that $d(\alpha_n, \beta_n) \to d(A, B)$. For any good metric on $X \times X$ the metric function $d : X \times X \to \mathbb{R}$ is (even uniformly) continuous, so it follows that

$$d(A, B) = \lim_{n \to \infty} d(\alpha_n, \beta_n) = d(a, b). \qquad \square$$

What if only one of the closed subsets A and B are compact?

EXAMPLE 2.18. *Let $X = \{0\} \cup (1, 2]$, viewed as a subspace of $\mathbb{R}$, let $A = \{0\}$ and $B = (1, 2]$. Then A is compact and B is closed (in X, though not in $\mathbb{R}$). We have $d(A, B) = 1$ but for all $y \in B$, we have $d(0, y) = y > 1$, so A and B are not distance-attained. However they are distance-separated.*

THEOREM 2.113. *Let A and B be nonempty closed subsets of a metric space X, at least one of which is compact. If A and B are disjoint, then they are distance-separated.*

PROOF. We may assume that A is compact, B is closed, $d(A, B) = 0$ and show that $A \cap B \neq \varnothing$. Then there are sequences $\{a_n\}$ in A and $\{b_n\}$ in B such that

$$d(a_n, b_n) \to d(A, B) = 0.$$

As in the proof of Theorem 2.17, there is a subsequence a_{n_k} converging to some $a \in A$. It follows that

$$\lim_{k\to\infty} d(a_{n_k}, b_{n_k}) = 0,$$

so by Lemma 2.54b) we have $b_{n_k} \to a$. It follows that a is an adherent point of B. Since B is closed, we conclude that $a \in A \cap B$. □

Looking back at Example 2.18, we notice that the subset B is closed and bounded in X but not compact.

THEOREM 2.114. *Let X be a metric space in which every closed, bounded subset is compact (e.g. $\mathbb{R}^N$). Let A and B be nonempty closed subsets, with A compact. Then A and B are distance-attained.*

PROOF. Suppose A is compact, let $a \in A$, let $b \in B$, and put $D := d(a, b)$. Let $\delta := \operatorname{diam}(A) < \infty$. If $d(A, B) = D$, we're done, so suppose not: then there are $a' \in A$ and $b' \in B$ such that $d(a', b') < D$. Then

$$d(a, b') \leq d(a, a') + d(a', b') \leq \delta + D.$$

It follows that

$$d(A, B) = d(A, B \cap B^\bullet(a, \delta + D)) :$$

in other words, to compute the distance between A and B we need only consider elements of B that lie in the closed ball centered at a with radius $\delta + D$: any other point b' of B is so far away from a that it is farther away from every point of A than b is from a. The set $B' := B \cap B^\bullet(a, \delta + D)$ is closed and bounded, hence compact by our hypothesis on X. By Theorem 2.112 there are points $a' \in A$ and $b' \in B' \subset B$ such that

$$d(a', b') = d(A, B') = d(A, B).$$ □

Now let A be a subset of a metric space X and let x be a point of X. We write $d(x, A)$ for $d(\{x\}, A)$.

EXERCISE 2.114. *For a subset A of a metric space X and $x \in X$, show: $d(x, A) = 0 \iff x \in \overline{A}$.*

LEMMA 2.115. *Let A be a subset of X. Then the function $d(\cdot, A) : X \to \mathbb{R}$ by $x \mapsto d(x, A)$ is a short map: for all $x, y \in X$ we have*

$$|d(x, A) - d(y, A)| \leq d(x, y).$$

PROOF. Fix $\epsilon > 0$. We may choose $a_x, a_y \in A$ such that

$$d(x, a_x) < d(x, A) + \epsilon, \ d(y, a_y) < d(y, A) + \epsilon.$$

Then

$$d(y, A) - d(x, A) \leq d(y, a_x) - d(x, A) \leq d(y, x) + d(x, a_x) - d(x, A) < d(x, y) + \epsilon.$$

Since this holds for all $\epsilon > 0$, we have

$$d(y, A) - d(x, A) \leq d(x, y).$$

Making the same argument with x and y reversed, we get

$$d(x, A) - d(y, A) \leq d(x, y). \qquad \square$$

EXERCISE 2.115. *Show that for a metric space X, the following properties are equvialent:*

(i) *Every open cover of X admits a Lebesgue number.*
(ii) *For every metric space Y, if $f : X \to Y$ is continuous, then it is uniformly continuous.*
(iii) *Every pair Y_1, Y_2 of disjoint closed subsets of X are distance-separated: $d(Y_1, Y_2) > 0$.*

*We call such a metric space an **Atsuji space**.*[7]

EXERCISE 2.116. a) *Show: an Atsuji space is complete.*
b) *Show: a metric space is compact iff it is Atsuji and totally bounded.*

EXERCISE 2.117. *Let $f : X \to Y$ be a map between metric spaces.*

a) *Let $x \in X$. Show that the following are equivalent:*
  (i) *For all subsets $A \subset X$, if $d(x, A) = 0$, then $d(f(x), f(A)) = 0$.*
  (ii) *The function f is continuous at x.*
b *(Cleveland* [**Cl73**]*) Show that the following are equivalent:*
  (i) *If $\varnothing \subsetneq A, B \subset X$, then $d(A, B) = 0 \implies d(f(A), f(B)) = 0$.*
  (ii) *The function f is uniformly continuous.*

18. Slow sequences

Let (X, d) be a metric space. A sequence $\mathbf{x}$ in X is **slow** if $\lim_{n\to\infty} d(\mathbf{x}_{n+1}, \mathbf{x}_n) \to 0$.

Every Cauchy sequence is slow, but the converse is not true. For instance, if $\{a_n\}$ is any real sequence such that $a_n \to 0$ and $\sum_n a_n$ diverges, then the sequence of partial sums $S_n = \sum_{i=1}^n a_i$ is slow but not convergent (thus not Cauchy, since $\mathbb{R}$ is complete).

We are interested in the set of partial limits of a slow sequence.

PROPOSITION 2.116. *Let $\mathbf{x}$ be a slow sequence in $\mathbb{R}$. Then the set of partial limits of $\mathbf{x}$ in $\mathbb{R}$ is connected.*

[7] In the literature, such spaces are also called **UC-spaces** and **Lebesgue spaces**. The former is not really a name, and the latter seems already taken by the normed spaces $L^p(\mu)$ associated to a measure μ in real analysis.

Proof. In view of our characterization of connected subsets of $\mathbb{R}$ as intervals, we must show: if $x < z$ are partial limits of $\mathbf{x}$, and $y \in (x, z)$, then y is also a partial limit. Let $\epsilon > 0$. A slow sequence with only finitely many terms in $(y - \epsilon, y + \epsilon)$ must eventually have all its terms at least $y + \epsilon$ or at most $y - \epsilon$, since otherwise $\{n \in \mathbb{Z}^+ \mid x_n \leq y - \epsilon \text{ and } x_{n+1} \geq y + \epsilon\}$ is infinite, but for sufficiently large n this is impossible. Thus $\mathbf{x}$ has infinitely many terms in $(y - \epsilon, y + \epsilon)$. $\square$

Exercise 2.118. a) *Let $Y \subset \mathbb{R}$, and let $\mathbf{x}$ be a slow sequence in Y. Show that the set of partial limits of $\mathbf{x}$ in Y is convex.*

b) *Investigate generalizations of part a) to suitable ordered topological spaces: e.g. ordered fields, ordered commutative groups...*

Proposition 2.117. *Let X be a compact metric space, and let $\mathbf{x}$ be a slow sequence in X. Then the set of partial limits of X (in X) is connected.*

Proof. As we know, the set $\mathcal{L}$ of partial limits of $\mathbf{x}$ is closed in X and thus compact. Seeking a contradiction, let $\mathcal{L} = U \coprod V$ be a separation. Then the set distance $d := d(U, V)$ is strictly positive. Let $\mathcal{C}_U$ be the set of $x \in X$ that have distance at most $\frac{d}{3}$ from U, and let $\mathcal{C}_V$ be the set of $x \in X$ that have distance at most $\frac{d}{3}$ from V. Then $\mathcal{C}_U$, $\mathcal{C}_V$ are disjoint open sets and $d(\mathcal{C}_U, \mathcal{C}_V) \geq \frac{d}{3}$. There are infinitely many terms of the sequence that lie in $\mathcal{C}_U$ and infinitely many that lie in $\mathcal{C}_V$, and thus there must be infinitely many $n \in \mathbb{Z}^+$ such that $\mathbf{x}_n \in \mathcal{C}_U$ and $\mathbf{x}_{n+1} \in \mathcal{C}_V$, contradicting the slowness of $\mathbf{x}$. $\square$

Let $\mathbf{x}$ be a sequence in $\mathbb{R}^d$, and let $\mathcal{L}$ be the set of partial limits of $\mathbf{x}$. If $\mathbf{x}$ is bounded, then $\mathcal{L}$ is connected by Proposition 2.117. If $d = 1$, then even if $\mathbf{x}$ is unbounded, $\mathcal{L}$ is connected by Proposition 2.116. However this need not hold when $d \geq 2$.

Example 2.19. *For $n \in \mathbb{Z}^+$, let R_n be the rectangle centered at the origin with sides parallel to the coordinate axes, of width 2 and height $2n$. Consider the finite sequence $a_1, \ldots, a_{N_1}$ starting at $(1, 0)$ and proceeding counterclockwise around R_1 in steps of length 1. Now consider the finite sequence $a_{N_1+1}, \ldots, a_{N_2}$ starting at $(1, 0)$ and proceeding counterclockwise around R_2 in steps of length $\frac{1}{2}$. We continue on in this manner for all $n \in \mathbb{Z}^+$. The resulting sequence has as its set of partial limits the union of the two lines $x = -1$ and $x = 1$.*

Theorem 2.118. *(Azouzi* [**Az18**]*)*
For a closed subset $\mathcal{L} \subset \mathbb{R}^d$, the following are equivalent:

(i) *$\mathcal{L}$ is the set of all partial limits of a slow sequence $\mathbf{x}$ in $\mathbb{R}^d$.*
(ii) *At least one of the following holds:*
  (a) *$\mathcal{L}$ is connected.*
  (b) *$d \geq 2$ and every connected component of $\mathcal{L}$ is unbounded.*

The proof of Theorem 2.118 requires some preliminaries.

Exercise 2.119. *Let (X, d) be a metric space. For $\epsilon > 0$, an ϵ-**chain** from a point $x \in X$ to a point $y \in X$ is a finite sequence $x_0 = x, x_1, \ldots, x_n = y$ such that $d(x_i, x_{i+1}) < \epsilon$ for all $0 \leq i \leq n - 1$. We define $x \sim_\epsilon y$ if there is an ϵ-chain from x to y.*

a) *Show: $\sim_\epsilon y$ is an equivalence relation on X whose equivalence classes are clopen subsets of X.*

b) *For* $x, y \in X$*, we put* $x \sim y$ *if* $x \sim_\epsilon y$ *for all* $\epsilon > 0$*. Show that* $\sim$ *is an equivalence relation on* X*. For each* $x \in X$*, denote the* $\sim$ *equivalence class of* x *by* $C_c(x)$ *and call it the* ***chain equivalence class of x****. Show that the chain equivalence classes are always closed, and show by example that they need not be open.*
c) *For* $x \in X$*, denote by* $C(x)$ *the connected component of* x*. We say that* X *is* ***well-chained*** *if it has a single chain equivalence class. Show that if* X *is connected then it is well-chained, and deduce that in general, for all* $x \in X$*, we have* $C(x) \subseteq C_c(x)$*.*
d) *Exhibit a totally disconnected metric space that is well-chained.*
e) *Show: if* X *is well-chained and compact, then* X *is connected. (Hint: use Theorem 2.113.)*

EXERCISE 2.120. *Let* $X \subseteq \mathbb{R}^N$ *be a nonempty closed, well-chained subset.*

a) *Show: if* $N = 1$*, show that* X *is connected.*
b) *For each* $N \geq 2$*, exhibit a nonempty closed, well-chained subset of* $\mathbb{R}^N$ *that is* not *connected.*

CHAPTER 3

Introducing Topological Spaces

1. In Which We Meet the Object of Our Affections

Part of the rigorization of analysis in the 19th century was the realization that notions like continuity of functions and convergence of sequences (e.g. $f : \mathbb{R}^n \to \mathbb{R}^m$) were most naturally formulated by paying close attention to the mapping properties between subsets U of the domain and codomain with the property that when $x \in U$, there exists $\epsilon > 0$ such that $||y - x|| < \epsilon$ implies $y \in U$. Such sets are called **open**. In the early twentieth century it was realized that many of the constructions formerly regarded as "analytic" in nature could be carried out in a very general context of sets and maps between them, provided only that the sets are endowed with a distinguished family of subsets, decreed to be open, and satisfying some very mild axioms. This led to the notion of an abstract topological space, as follows.

Let X be a set. A **topology** on X is a family $\tau = \{U_i\}_{i \in I}$ is a of subsets of X satisfying the following axioms:

(T1) $\emptyset,\ X \in \tau$.
(T2) $U_1,\ U_2 \in \tau \implies U_1 \cap U_2 \in \tau$.
(T3) For any subset $J \subset I$, $\bigcup_{i \in J} U_i \in \tau$.

It is pleasant to also be able to refer to axioms by a descriptive name. So instead of "Axiom (T2)" one generally speaks of a family $\tau \subset 2^X$ being closed under binary intersections. Similarly, instead of "Axiom (T3)", one says that the family τ is closed under arbitrary unions.

Remark 3.1. *Consider the following variant of (T2):*

(T2′) For any finite subset $J \subset I$, we have $\bigcap_{i \in J} U_i \in \tau$.

Evidently (T2′) $\implies$ (T2), and at first glance the converse seems to hold. This is almost, but not quite, true: (T2′) also allows the empty intersection, which is – by convention – defined as $\bigcap_{Y \in \emptyset} Y = X$. Since we also have that $\bigcup_{Y \in \emptyset} Y = \emptyset$, it follows that (T2′) + (T3) $\implies$ (T1). None of this is of any particular importance, but the reader should be aware of it because this alternative ("more efficient") axiomatization appears in some texts, e.g. [**Bo**].

A **topological space** (X, τ) consists of a set X and a topology τ on X. The elements of τ are called **open sets**.

If (X, τ_X) and (Y, τ_Y) are topological spaces, a map $f : X \to Y$ is *continuous* if for all $V \in \tau_Y$, $f^{-1}(V) \in \tau_X$. A function $f : X \to Y$ between topological spaces is a **homeomorphism** if it is bijective, continuous, and has a continuous inverse. A function f is **open** if for all $U \in \tau_X$, $f(U) \in \tau_Y$.

EXERCISE 3.1. *For a function $f : X \to Y$ between topological spaces (X, τ_X) and (Y, τ_Y), show that the following are equivalent:*

(i) *The map f is a homeomorphism.*
(ii) *The map f is bijective and for all $V \subset Y$, $V \in \tau_Y \iff f^{-1}(V) \in \tau_X$.*
(iii) *The map f is bijective and for all $U \subset X$, $U \in \tau_X \iff f(U) \in \tau_Y$.*
(iv) *The map f is bijective, continuous and open.*

Tournant dangereux: A continuous bijection need not be a homeomorphism!

EXERCISE 3.2. *Let (X, τ_X), (Y, τ_Y), $(Z, \tau_Z$ be topological spaces, and $f : X \to Y$, $g : Y \to Z$ be continuous functions. Show: $g \circ f : X \to Z$ is continuous.*

Those who are familiar with the basic notions of **category theory** will recognize that we have verified that we get a category **Top** with objects the topological spaces and morphisms the continuous functions between them. Our definition of homeomorphism is chosen so as to coincide with the notion of isomorphism in the categorical sense.

We hasten to add that we by no means expect readers to have prior familiarity with this terminology. On the contrary, some of the material presented in these notes will provide readers with much of the experience and examples necessary to facilitate a later learning of this material.

The previous chapter was devoted to the following example.

EXAMPLE 3.1. *Let (X, d) be a metric space. We define τ to be the set of unions of open balls in X. Then τ is a topology on X, called the **metric topology** on X. We also say that τ is **induced from** the metric d.*

The following definition is all-important in the interface between metric spaces and topological spaces: a topological space (X, τ) is **metrizable** if there is some metric d on X such that τ is induced from d.

EXERCISE 3.3. *Let (X, τ) be a metrizable topological space. Show: if $\#X \geq 2$, then the set of metrics d on X which induce τ is uncountably infinite.*

The above exercise makes clear that passing from a metric space to its associated metric topology involves a great loss of information: in all nontrivial cases there will be many, many metrics inducing the topology. From this perspective a metric looks "better" than a topology. However, it turns out when studying continuous functions – which one naturally does in many branches of mathematics – the topology is sufficient and the extra information of the metric can be awkward or distracting. A good example of this comes up in the discussion of products. We previously explored this in the case of metric spaces and found the phenomenon of **embarrassment of riches**: there is simply not one preferred product metric but a whole class of them. Built into our discussion of "good product metrics" was that they should satisfy a simple property of convergent sequences which uniquely characterizes the resulting topology. We will revisit the discussion of products of topological spaces and see that it is decidedly *simpler*: on *any* product of topological spaces there is a canonically defined product topology, which in the case of finite or countably infinite products of metric spaces is metrizable via any one of the "good product metrics" we constructed before, but if all we want to see is that this product topology is metrizable then we can just concentrate on the $p = \infty$ case and most of the difficulty evaporates. Moreover the product topology is defined also on uncountable products, for which we did not succeed in constructing a good metric. In fact we will show that an uncountable product of metrizable spaces (each with at least two points) is not metrizable. Thus such products provide an example of a construction that can be performed in the class of topological spaces and not in the class of metric spaces. Of course one can ask why we want to consider uncountable products of spaces. This has a good answer but a remarkably deep one: it involves **Tychonoff's Theorem** and the **Stone-Cech compactification**, which are probably the most important results in the entire subject.

Another key construction in the class of topological spaces which we have not met yet because it has absolutely no analogue in metric spaces is the **identification** or **quotient** construction. In geometric applications – especially, in the study of manifolds – this construction is all-important.

On the other hand, there are times when having a metric is more convenient than just a topology: it cannot be denied that a metrizable space is in many respects much more tractable than an arbitrary topological space, and certain purely topological constructions are considerably streamlined by making use of a metric – any metric! – that induces the given topology. For this and other reasons it is of interest to have sufficient (or ideally, necessary and sufficient) conditions for the metrizability of a topological space. This is in fact one of the main problems in general topology and will be addressed later, though, we warn, not in as much detail as most classical texts: we do not discuss the *general* metrization theorems of Nagata-Smirnov or Bing – each of which gives necessary and sufficient conditions for an arbitrary topological space to be metrizable – but only the easier Urysohn Theorem which gives conditions for a space to be metrizable and separable.

The following is the most important example of a property which is possessed by all metrizable spaces but not by all topological spaces.

A topological space X is **Hausdorff** if given distinct points x, y in X, there exist open sets $U \ni x$, $V \ni y$ such that $U \cap Y = \emptyset$.

EXERCISE 3.4. *Show that metrizable topologies are Hausdorff.*

The task of giving an example of a non-Hausdorff topologies brings us to the more general problem of amassing a repertoire of topological spaces sufficiently rich so as to be able to use to see that any number of plausible-sounding implications among properties of topological spaces do not hold. It turns out that the concept of a topological space is – even by the standards of abstract mathematical structure – remarkably inclusive. There are some strange topological spaces out there, and it will be useful to our later study to amass a repertoire of them. This turns out to be a cottage industry in its own right, for which the canonical text is [**SS**]. But let us meet some of the more interesting specimens.

2. A Topological Bestiary

EXAMPLE 3.2. *(Indiscrete Topology) For a set X, $\tau = \{\emptyset, X\}$ is a topology on X, called the **indiscrete topology** (and also the **trivial topology**). If X has more than one element, this topology is not Hausdorff.*

EXAMPLE 3.3. *(Discrete Topology) For a set X, $\tau = 2^X$, the collection of* all *subsets of X, forms a topology, called the **discrete** topology.*

The discrete and indiscrete topologies coincide iff X has at most one element. Otherwise they are distinct and indeed give rise to non-homeomorphic spaces.

EXERCISE 3.5. *Let X be a topological space.*

a) *Show: X is discrete iff for all $x \in X$, $\{x\}$ is open.*
b) *Show: if X is discrete, then X is metrizable.*

EXERCISE 3.6. *Suppose X is a finite topological space: by this we mean that the underlying set X is finite. It then follows that 2^X is finite hence $\tau \subset 2^X$ is finite. (On the other hand there are topological spaces (X, τ) with X infinite and τ is still finite: e.g. indiscrete topologies.) Show: if X Hausdorff, then it is discrete. In particular, finite metrizable spaces are discrete.*

On the other hand, as soon as $n \geq 2$, an n-point set carries non-Hausdorff topologies: e.g. the indiscrete topology. In fact it carries other topologies as well. Here is the first example.

EXAMPLE 3.4 (Sierpinski Space). *Consider the two element set $X = \{\circ, \bullet\}$. We take $\tau = \{\emptyset, \{\circ\}, X\}$. This gives a topology on X in which the point $\circ$ is open but the point $\bullet$ is not, so X is finite and nondiscrete, hence nonmetrizable.*

EXERCISE 3.7. *Let X a set.*

a) *Show that, up to homeomorphism, there are precisely three topologies on a two-element set.*
b) *For $n \in \mathbb{Z}^+$, let $T(n)$ denote the number of homeomorphism classes of topologies on $\{1, \ldots, n\}$. Show that $\lim_{n \to \infty} T(n) = \infty$. (Only one of these topologies, the discrete topology, is metrizable.)*
c) *Can you describe the asymptotics of $T(n)$, or even give reasonable lower and/or upper bounds?*[1]

That the the number $T(n)$ of homeomorphism classes of n-point topological spaces approaches infinity with n has surely been known for some time. The realization that non-Hausdorff finite topological spaces are in fact natural and important and not just a curiosity permitted by a very general definition is more recent. Later we will study a bit about such spaces as an important subclass of **Alexandroff spaces** (these are spaces in which arbitrary intersections of closed sets qualify; this is a very strong and unusual property for an infinite topological space to have, but of course it holds automatically on all finite topological spaces).

EXAMPLE 3.5. *(Particular Point Topology) Let X be a set with more than one element, and let $x \in X$. We take τ to be empty set together with all subsets Y of X containing x.*

[1] This question has received a lot of attention but is, to the best of my knowledge, open in general.

EXAMPLE 3.6. *(Cofinite Topology) Let X be an infinite set, and let τ consist of $\emptyset$ together with subsets whose* complement *is finite (or, for short, "cofinite subsets"). This is easily seen to form a topology, in which any two nonempty open sets intersect*[2]*, hence a non-Hausdorff topology.*

EXAMPLE 3.7. *(Cocountable Topology) Let X be an uncountable set. The family of subsets $U \subset X$ with countable complement together with the empty set forms a topology on X, the* ***cocountable topology****. This is a non-discrete topology (since X is uncountable). In fact it is not even Hausdorff, if N_x and N_y are any two neighborhoods of points x and y, then $X \setminus N_x$ and $X \setminus N_y$ are countable, so $X \setminus (N_x \cap N_y) = (X \setminus N_x) \cup (X \setminus N_y)$ is uncountable and $N_x \cap N_y$ is nonempty.*

EXERCISE 3.8. *(Sorgenfrey Line)*
On $\mathbb{R}$, show that intervals of the form $[a, b)$ form a base for a topology τ_S which is strictly finer than the standard (metric) topology on $\mathbb{R}$. The space $(\mathbb{R}, \tau_S)$ is called the ***Sorgenfrey line*** *after Robert Sorgenfrey.*[3]

EXAMPLE 3.8. *(Moore Plane) Let X be the subset of $\mathbb{R}^2$ consisting of pairs (x, y) with $y \geq 0$, endowed with the following "exotic" topology: a subset U of X is open if: whenever it contains a point $P = (x, y)$ with $y > 0$ it contains some open Euclidean disk $B(P, \epsilon)$; and whenever it contains a point $P = (x, 0)$ it contains $P \cup B((x, \epsilon), \epsilon)$ for some $\epsilon > 0$, i.e., an open disk in the upper-half plane tangent to the x-axis at P. The Moore plane satisfies several properties shared by all metrizable spaces – it is* ***first countable*** *and* ***Tychonoff*** *– but not the property of* ***normality****. More on these properties later, of course.*

EXAMPLE 3.9. *(Arens-Fort Space) Let $X = \mathbb{N} \times \mathbb{N}$. We define a topology τ on X by declaring a subset $U \subset X$ to be open if:*

(i) $(0, 0) \notin U$*, or*
(ii) $(0, 0) \in U$ *and* $\exists M \in \mathbb{N}$ *such that* $\forall m \geq M$, $\{n \in \mathbb{N} \mid (m, n) \notin U\}$ *is finite.*

In other words, a set not containing the origin is open precisely when it contains all but finitely many elements of all but finitely many column of the array $\mathbb{N} \times \mathbb{N}$.

EXERCISE 3.9. *Show that the Arens-Fort space is a Hausdorff topological space. (Don't forget to check that τ is actually a topology: this is not completely obvious.)*

EXERCISE 3.10. *(Zariski Topology): Let R be a commutative ring, and let $\operatorname{Spec} R$ be the set of prime ideals of R. For any subset S of R (including $\emptyset$, let $C(S)$ be the set of prime ideals containing S.*

a) *Show that $C(S_1) \cup C(S_2) = C(S_1 \cap S_2)$.*
b) *Show that, for any collection $\{S_i\}_{i \in I}$ of subsets of R, $\bigcap_i C(S_i) = C(\bigcup_i S_i)$.*
c) *Show: $C(\emptyset) = \operatorname{Spec} R$, $C(R) = \emptyset$. Thus the $C(S)$'s form the closed sets for a topology, called the Zariski topology on $\operatorname{Spec} R$.*
d) *If $\varphi : R \to R'$ is a homomorphism of commutative rings, show that $\varphi^* : \operatorname{Spec} R' \to \operatorname{Spec} R$, $P \mapsto \varphi^{-1}(P)$ is a continuous map.*
e) *Let $\operatorname{rad}(R)$ be the radical of R. Show that the natural map $\operatorname{Spec}(R/\operatorname{rad}(R)) \to \operatorname{Spec}(R)$ is a homeomorphism.*

[2]When we say that two subsets intersect, we mean of course that their intersection is nonempty.

[3]The merit of this "weird" topology is that it is often a source of counterexamples.

f) *Let R be a discrete valuation ring. Show that* $\operatorname{Spec} R$ *is the topological space of Example 3.4 above.*
g) *Let k be an algebraically closed field and $R = k[t]$. Show that* $\operatorname{Spec}(R)$ *can, as a topological space, be identified with k itself with the cofinite topology.*

3. Alternative Characterizations of Topological Spaces

3.1. Closed sets.

In a topological space (X, τ), define a **closed** subset to be a subset whose complement is open. Evidently if we know the open sets we also know the closed sets and conversely: just take complements. The closed subsets of a topological space satsify the following properties:

(CTS1) $\emptyset$, X are closed.
(CTS2) Finite unions of closed sets are closed.
(CTS3) Arbitrary intersections of closed sets are closed.

Conversely, given such a family of subsets of X, then taking the open sets as the complements of each element in this family, we get a topology.

3.2. Closure.

If S is a subset of a topological space, we define its **closure** $\overline{S}$ to be the intersection of all closed subsets containing S. Since X itself is closed containing S, this intersection is nonempty, and a moment's thought reveals it to be the minimal closed subset containing S.

Viewing closure as a mapping c from 2^X to itself, it satisfies the following properties, the **Kuratowski closure axioms**:

(KC1) $c(\emptyset) = \emptyset$.
(KC2) For $A \in 2^X$, $A \subset c(A)$.
(KC3) For $A \in 2^X$, $c(c(A)) = c(A)$.
(KC4) For $A, B \in 2^X$, $c(A \cup B) = c(A) \cup c(B)$.

The axiom (KC4) implies the following axiom:

(KC5) If $B \subset A$, then $c(B) \subset c(A)$.

Indeed, $c(A) = c((A \setminus B) \cup B) = c(A \setminus B) \cup c(B)$.

A function $c : 2^X \to 2^X$ satisfying (KC1)-(KC4) is called an "abstract closure operator." Kuratowski noted that any such operator is indeed the closure operator for a topology on X:

Theorem 3.2 (Kuratowski). *Let X be a set, and let $c : 2^X \to 2^X$ be an operator satisfying axioms (KC1), (KC2) and (KC4).*

a) *The subsets $A \in 2^X$ satisfying $A = c(A)$ obey they axioms (CTS1)-(CTS3) and hence are the closed subsets for a unique topology τ_c on X.*

b) *If c also satisfies (KC3), then closure in τ_c corresponds to closure with respect to c: for all $A \subset X$ we have $\overline{A} = c(A)$.*

PROOF. a) Call a set c-closed if $A = c(A)$. By (KC1) the empty set is c-closed; by (KC2) X is c-closed. By (KC2) finite unions of c-closed sets are closed. Now let $\{A_\alpha\}_{\alpha \in I}$ be a family of $c-$closed sets, and put $A = \cap A_\alpha$. Then for all α, $A \subset A_\alpha$, so by (KC5), $c(A) \subset c(A_\alpha)$ for all α, so

$$c(A) \subset \cap\, c(A_\alpha) = \cap A_\alpha = A.$$

Thus the c-closed sets satisfy (CTS1)-(CTS3), so that the family τ_c of complements of c-closed sets form a topology on X.

Now assume (KC3); we wish to show that for all $A \subset X$, $c(A) = \overline{A}$. We have $\overline{A} = \cap_{C = c(C) \supset A} C$, the intersection extending over all closed subsets containing A. By (KC3), $c(A) = c(c(A))$ is a closed subset containing A we have $\overline{A} \subset c(A)$. Conversely, since $A \subset \cap_C C$, $c(A) \subset \cap_C c(C) = \cap_C C = \overline{A}$. So $c(A) = \overline{A}$. $\square$

Later we will see an interesting example of an operator which satisfies (KC1), (KC2), (KC4) but not necessarily (KC3): the **sequential closure**.

The following result characterizes continuous functions in terms of closure.

THEOREM 3.3. *Let $f : X \to Y$ be a map of topological spaces. The following are equivalent:*

(a) *The map f is continuous.*
(b) *For every subset S of X, $f(\overline{S}) \subset \overline{f(S)}$.*

PROOF. Suppose f is continuous, S is a subset of X and $\overline{A} = A \supset f(S)$. If $x \in X$ is such that $f(x) \in Y \setminus A$, then, since f is continuous and $Y \setminus A$ is open in Y, $f^{-1}(Y \setminus A)$ is an open subset of X containing x and disjoint from S. Therefore x is not in the closure of S.

Conversely, if f is not continuous, then there exists some open $V \subset Y$ such that $U := f^{-1}(V)$ is not open in X. Thus, there exists a point $x \in U$ such that every open set containing x meets $S := X \setminus U$. Thus $x \in \overline{S}$ but $f(x)$ is in V hence not in $Y \setminus V$, which is a closed set containing $f(\overline{S})$. $\square$

3.3. Interior operator.

The dual notion to closure is the **interior** of a subset A in a topological space: A° is equal to the union of all open subsets of A. In particular a subset is open iff it is equal to its interior. We have

$$A^\circ = X \setminus \overline{X \setminus A},$$

and applying this formula we can mimic the discussion of the previous subsection in terms of axioms for an "abstract interior operator" $A \mapsto i(A)$, which one could take to be the basic notion for a topological space. But this is so similar to the characterization using the closure operator as to be essentially redundant.

3.4. Boundary operator.

For a subset A of a topological space, one defines the **boundary**[4]

$$\partial A = \overline{A} \setminus A^\circ = \overline{A} \cap \overline{X \setminus A}.$$

Evidently ∂A is a closed subset of A, and, since $\overline{A} = A \cup \partial A$, A is closed iff $A \supset \partial A$. A set has empty boundary iff it is both open and closed, a notion which is important in connectedness and in dimension theory.

EXAMPLE 3.10. *Let X be the real line, $A = (-\infty, 0)$ and $B = [0, \infty)$. Then $\partial A = \partial B = \{0\}$, and*

$$\partial(A \cup B) = \partial\mathbb{R} = \emptyset \neq \{0\} = (\partial A) \cup (\partial B);$$

$$\partial(A \cap B) = \partial\emptyset = \emptyset \neq \{0\} = (\partial A) \cap (\partial B).$$

Thus the boundary operator is not as well-behaved as either the closure or interior operators. We quote from [**Wi**, p. 28]*: "It is possible, but unrewarding, to characterize a topology completely by its frontier [boundary] operation."*

3.5. Neighborhoods.

Let x be a point of a topological space, and let N be a subset of X. We say that N is a **neighborhood** of x if $x \in N^\circ$. Open sets are characterized as being neighborhoods of each point they contain.

Let $\mathcal{N}_x$ be the set of all neighborhoods of x. It enjoys the following properties:

(NS1) $N \in \mathcal{N}_x \implies x \in N$.
(NS2) $N,\ N' \in \mathcal{N}_x \implies N \cap N' \in \mathcal{N}_x$.
(NS3) $N \in \mathcal{N}_x,\ N' \supset N \implies N' \in \mathcal{N}_x$.
(NS4) For $N \in \mathcal{N}_x$, there exists $U \in \mathcal{N}_x$, $U \subset N$, such that $y \in V \implies V \in \mathcal{N}_y$.

Suppose we are given a set X and a function which assigns to each $x \in X$ a family $\mathcal{N}(x)$ of subsets of X satisfying (NS1)-(NS3). Then the collection of subsets U such that $x \in U \implies U \in \mathcal{N}(x)$ form a topology on X. If we moreover impose (NS4), then $\mathcal{N}(x) = \mathcal{N}_x$ for all x.

4. The Set of All Topologies on X

Let X be a set, and let $\mathbf{Top}(X) \subset 2^{2^X}$ be the collection of all topologies on X.

EXERCISE 3.11. *Suppose X is infinite. Show that $\#\mathbf{Top}(X) = 2^{2^X}$.*

As a subset of 2^{2^X}, $\mathbf{Top}(X)$ inherits a partial ordering: we define $\tau_1 \leq \tau_2$ if $\tau_1 \subset \tau_2$, i.e., if every τ_1-open set is also τ_2-open.

If $\tau_1 \leq \tau_2$ we say that τ_1 is **coarser** than τ_2 and also that τ_2 is **finer** than τ_1.[5] We say that two topologies on X are **comparable** if one of them is coarser than the other. Comparability is an equivalence relation.

---

[4]Alternate terminology: **frontier**.

[5]One sometimes also says, especially in functional analysis, that τ_1 is **weaker** than τ_2 and that τ_2 is **stronger** than τ_1. Unfortunately some of the older literature uses the terms "weaker" and "stronger" in exactly the opposite way! So the coarser/finer terminology is preferred.

EXERCISE 3.12. *Let $\mathcal{T} \subset 2^{2^X}$ be any family of topologies on X. Then $\bigcap_{\tau \in \mathcal{T}} \tau$ is a topology on X. (By convention, $\bigcap_{\emptyset} = 2^X$ is the discrete topology.)*

Let $\mathcal{F} \in 2^{2^X}$ be any family of subsets of X. Then among all topologies τ on X containing $\mathcal{F}$ there is a coarsest topology $\tau(\mathcal{F})$, namely the intersection of all topologies containing $\mathcal{F}$. (**Tournant dangereux**: here $\tau(\emptyset) = \{\emptyset, X\}$ is the indiscrete topology.) We call $\tau(\mathcal{F})$ the **topology generated** by $\mathcal{F}$.

In fact $(\mathbf{Top}(X), \leq)$ is a **complete lattice**. We recall what this means:

(i) There is a "top element" in $\mathbf{Top}(X)$, i.e., a topology which is finer than any other topology on X: namely the discrete topology.
(ii) There is a "bottom element" in $\top(X)$, i.e., a topology which is coarser than any other topology on X: namely the indiscrete topology.
(iii$\wedge$) If $\mathcal{T} \subset \mathbf{Top}(X)$ is any family of topologies on X, then the **meet** $\wedge\mathcal{T}$ (or **infiumum**) exists in $\mathbf{Top}(X)$: there is a unique topology $\tau_{\wedge\mathcal{T}}$ on X such that for any $\tau \in \mathbf{Top}(X)$, $\tau \leq \tau_{\wedge\mathcal{T}}$ iff $\tau \leq T$ for all $T \in \mathcal{T}$: namely we just take the intersection $\cap_{T\in\mathcal{T}}T$, as in Exercise 3.12 above.
(iii$\vee$) If $\mathcal{T} \subset \mathbf{Top}(X)$ is any family of topologies on X, then the **join** $\vee\mathcal{T}$ (or **supremum**) exists in $\mathbf{Top}(X)$: there is a unique topology $\tau_{\vee\mathcal{T}}$ such that for any $\tau \in \mathbf{Top}(X)$, $\tau \geq \tau_{\vee\mathcal{T}}$ iff $\tau \geq T$ for all $T \in \mathcal{T}$: we first take $\mathcal{F}(\mathcal{T}) = \bigcup_{T\in\mathcal{T}} T$ and then $\vee\mathcal{T} = \tau(\mathcal{F})$ is the intersection of all topologies containing $\mathcal{F}$.

Let us now look a bit more carefully at the structure of the topology $\tau(\mathcal{F})$ generated by an arbitrary family $\mathcal{F}$ of subsets of X. The above description is a "top down" or an "extrinsic" construction. Such situations occur frequently in mathematics, and it is also useful (maybe more useful) to have a complementary "bottom up" or "intrinsic construction".

By way of comparison, if G is a group and S is a subset of G, then there is a notion of the subgroup $H(S)$ generated by S. The "extrinsic" construction is again just $\bigcap_{H\supset S} H$, the intersection over all subgroups containing S. But there is also an "instrinsic construction" of $H(S)$: namely, as the collection of all group elements of the form $x_1^{\epsilon_1} \cdots x_n^{\epsilon_n}$, where $x_i \in S$ and $\epsilon_i \in \pm 1$. In some sense, this "bottom up" construction is a two-step process: starting with the set S, we first replace S by $S \cup S^{-1}$, and second we pass to all words (including the empty word!) in $S \cup S^{-1}$.

In general, we may not be so lucky. If X is a set and $\mathcal{F}$ is a family of subsets of X, in order to form the σ-algebra generated by $\mathcal{F}$, extrinsically we again just take the intersection over all σ-algebras on X containing $\mathcal{F}$ (in particular there is always 2^X, so this intersection is nonempty). Sometimes one needs the intrinsic description, but this is usually avoided in first courses on measure theory because it is very complicated: one alternates the processes of passing to countable unions and adjoining complements, but in general one must do this uncountably many times, necessitating a transfinite induction![6]

Luckily, the case of topological spaces is much more like that of groups than that of σ-algebras. Namely, starting with $\mathcal{F} \subset 2^X$, we first form $\mathcal{F}_1$ which consists of all

[6]To read more about this, the keyword is **Borel hierarchy.**

finite intersections of elements of elements of $\mathcal{F}$ (employing, as usual, the convention that the empty intersection is all of X). We then form $\mathcal{F}_2$, which consists of all arbitrary unions of elements of $\mathcal{F}_1$ (employing, as usual, the fact that the empty union is $\emptyset$). Clearly $\mathcal{F}_2$ contains $\emptyset$ and X and is stable under arbitrary unions. In fact it is also stable under finite intersections, since for any two families $\{Y_i\}_{i\in I}$, $\{Z_j\}_{j\in J}$ of elements of $\mathcal{F}_1$,

$$\bigcup_i Y_i \cap \bigcup_j Z_j = \bigcup_{i,j} Y_i \cap Z_j,$$

and for all i and j we have $Y_i \cap Z_j \in \mathcal{F}_1$ since $\mathcal{F}_1$ is closed under finite intersections. So we are done in two steps: $\mathcal{F}_2 = \tau(\mathcal{F})$ is the topology generated by $\mathcal{F}$.

EXAMPLE 3.11. *Let X be any nonempty set. If $\mathcal{F} = \emptyset$, then $\tau(\mathcal{F})$ is the trivial topology. If $\mathcal{F} = \{\{x\} \mid x \in X\}$, $\tau(\mathcal{F})$ is the discrete topology. More generally, let S be any subset of X and $\mathcal{F}(S) = \{\{x\} \mid x \in S\}$, then $\tau(S) := \tau(\mathcal{F}(S))$ is a topology whose open points are precisely the elements of S, so this is a different topology for each $S \in 2^X$.*

5. Bases, Subbases and Neighborhood Bases

5.1. Bases and Subbases.

We have found our way to an important definition: if τ is a topology on X and $\mathcal{F} \subset 2^X$ is such that $\tau = \tau(\mathcal{F})$, we say $\mathcal{F}$ is a **subbase** (or **subbasis**) for τ.

EXAMPLE 3.12. *Let X be a set of cardinality at least 2.*
- (a) *Again, if we take $\mathcal{F}$ to be the empty family, then $\tau(\mathcal{F})$ is the indiscrete topology.*
- (b) *If Y is a subset of X and we take $\mathcal{F} = \{Y\}$, then the open sets in the induced topology τ_Y are precisely those which contain Y. Note that these 2^X topologies are all distinct. If $Y = X$ this again gives the indiscrete topology, whereas if $Y = \emptyset$ we get the discrete topology. Otherwise we get a non-Hausdorff topology: indeed for $x \in X$, $\{x\}$ is closed iff $x \in X \setminus Y$.*

EXERCISE 3.13. *Let X be a set and Y, Y' be two subsets of X. Show that the following are equivalent:*
- (i) *The space (X, τ_Y) is homeomorphic to $(X, \tau_{Y'})$.*
- (ii) *We have $\#Y = \#Y'$.*

The nomenclature "subbase" suggests the existence of a cognate concept, that of a "base". Based upon our above intrinsic construction of $\tau(\mathcal{F})$, it would be reasonable to guess that $\mathcal{F}_1$ is a base, or more precisely that a basis for a topology should be a collection of open sets, closed under finite intersection, whose unions recover all the open sets. But it turns out that a weaker concept is much more useful.

Consider the following axioms on a family $\mathcal{B}$ of subsets of a set X:

(B1) $\forall\ U_1,\ U_2 \in \mathcal{B}$ and $x \in U_1 \cap U_2$, $\exists\ U_3 \in \mathcal{B}$ such that $x \in U_3 \subset U_1 \cap U_2$.
(B2) For all $x \in X$, there exists $U \in \mathcal{B}$ such that $x \in U$.

The point here is that (B1) is weaker than the property of being closed under finite intersections, but is just as good for constructing the generated topology:

PROPOSITION 3.4. *Let $\mathcal{B} = (\mathcal{U}_i)_{i \in I}$ be a family of subsets of X satisfying (B1) and (B2). Then $\tau(\mathcal{B})$, the topology generated by $\mathcal{B}$, is given by $\{\bigcup_{i \in J} \mathcal{U}_i | J \subset I\}$, or in other words by the collection of arbitrary unions of elements of $\mathcal{B}$.*

PROOF. Let T be the set of arbitrary unions of elements of $\mathcal{B}$; certainly $T \subset \tau(\mathcal{B})$. It is automatic that $\emptyset \in T$ (take the empty union), and (B2) guarantees that $X = \bigcup_{i \in I} \mathcal{U}_i$. Clearly T is closed under all unions, so it suffices to show that the intersection $\mathcal{U}_1 \cap \mathcal{U}_2$ of any two elements of $\mathcal{B}$ of $\mathcal{B}$ can be expressed as a union over some set of elements of $\mathcal{B}$. But the point is that (B1) visibly guarantees this: for each $x \in \mathcal{U}_1 \cap \mathcal{U}_2$, by (B1) we may choose $\mathcal{U}_x \in \mathcal{B}$ such that $x \in \mathcal{U}_x \subset \mathcal{U}_1 \cap \mathcal{U}_\in$. Then

$$\mathcal{U}_1 \cap \mathcal{U}_2 = \bigcup_{x \in \mathcal{U}_1 \cap \mathcal{U}_2} U_x. \qquad \square$$

A family $\mathcal{B}$ of subsets of X satisfying (B1) and (B2) is a **base** (or **basis**) for the topology it generates. Or, to put it another way, a subcollection $\mathcal{B}$ of the open sets of a topological space (X, τ) which satisfies (B1) and (B2) is called a base, and then every open set is obtained as a union of elements of the base. And conversely:

EXERCISE 3.14. *Let (X, τ) be a topological space and $\mathcal{B}$ be a family of open sets. Suppose that every open set in X may be written as a union of elements of $\mathcal{B}$. Show that $\mathcal{B}$ satisfies (B1) and (B2).*

EXAMPLE 3.13. *In a metric space (X, d), the open balls form a base for the topology: especially, the intersection of two open balls need not be an open ball but contains an open ball about each of its points. Indeed, the open balls with radii $\frac{1}{n}$, for $n \in \mathbb{Z}^+$, form a base.*

EXAMPLE 3.14. *In $\mathbb{R}^d$, the d-fold products $\prod_{i=1}^d (a_i, b_i)$ of open intervals with rational endpoints is a base. In particular this shows that $\mathbb{R}^d$ has a **countable base**, which will turn out to be a key property for a topological space.*

PROPOSITION 3.5. *Let $f : X \to Y$ be a map between topological spaces, and let $\mathcal{F}$ be a subbase for the topology of Y. The following are equivalent:*
- (i) *For all $S \in \mathcal{F}$, we have that $f^{-1}(S)$ is open in X.*
- (ii) *The map f is continuous.*

PROOF. (i) $\implies$ (ii): Let V be open in Y. Then V is a finite intersection of unions of elements of $\mathcal{F}$. The empty intersection leads to $V = Y$, for which the preimage under f is X, which is indeed open in X, so we may assume that for some $N \in \mathbb{Z}^+$ we have subfamilies $\mathcal{F}_1, \ldots, \mathcal{F}_N$ of $\mathcal{F}$ such that if

$$\forall 1 \le i \le N, \ V_i := \bigcup_{S \in \mathcal{F}_i} S,$$

then

$$V = \bigcap_{i=1}^N V_i.$$

It follows that

$$f^{-1}(V) = \bigcap_{i=1}^N f^{-1}(V_i) = \bigcap_{i=1}^N \bigcup_{S \in \mathcal{F}_i} f^{-1}(S)$$

is open in X. So f is continuous.
(ii) $\implies$ (i): This is immediate. $\square$

5.2. Neighborhood bases. Let x be a point of a topological space X. A family $\{N_\alpha\}$ of neighborhoods of x is said to be a **neighborhood base at** x (or a **fundamental system of neigborhoods of** x) if every neighborhood N of x contains some N_α. Suppose we are given for each $x \in X$ a neighborhood basis $\mathcal{N}_x$ at x. The following axioms hold:

(NB1) $N \in \mathcal{B}_x \implies x \in N$.
(NB2) $N,\ N' \in \mathcal{B}_x \implies$ there exists N'' in $\mathcal{B}_x$ such that $N'' \subset N \cap N'$.
(NB3) $N \in \mathcal{B}_x \implies$ there exists $V \in \mathcal{B}_x$, $V \subset N$, such that $y \in V \implies V \in \mathcal{B}_y$.

Conversely:

Proposition 3.6. *Suppose given a set X and, for each $x \in X$, a collection $\mathcal{B}_x$ of subsets satisfying (NB1)-(NB3). Then the collections $\mathcal{N}_x = \{Y \mid \exists N \in \mathcal{B}_x | Y \supset N\}$ are the neighborhood systems for a unique topology on X, in which a subset U is open iff $x \in U \implies U \in \mathcal{N}_x$. Each $\mathcal{N}_x$ is a neighborhood basis at x.*

Exercise 3.15. *Prove Proposition 3.6.*

Remark: Consider the condition

(NB3′) $N \in \mathcal{B}_x$, $y \in N \implies y \in N$.

Replacing (NB3) with (NB3′) amounts to restricting attention to open neighborhoods. Since (NB3′) $\implies$ (NB3), we may specify a topology on X by giving, for each x, a family $\mathcal{N}_x$ of sets satisfying (NB1), (NB2), (NB3′). This is a very convenient way to define a topology: e.g. the metric topology is thus defined just by taking $\mathcal{N}_x$ to be the family $\{B(x,\epsilon)\}$ of ϵ balls about x.

Proposition 3.7. *Suppose that $\varphi : X \to X$ is a self-homeomorphism of the topological space x. Let $x \in X$ and $\mathcal{N}_x$ be a neighborhood basis at x. Then $\varphi(\mathcal{N}_x)$ is a neighborhood basis at $y = \varphi(x)$.*

Proof: It suffices to work throughout with open neighborhoods. Let V be an open neighborhood of y. By continuity, there exists an open neighborhood U of x such that $\varphi(U) \subset V$. Since φ^{-1} is continuous, $\varphi(U)$ is open.

As for any category, the automorphisms of a topological space X form a group, $\mathrm{Aut}(X)$. We say X is **homogeneous** if $\mathrm{Aut}(X)$ acts transitively on X, i.e., for any $x, y \in X$ there exists a self-homeomorphism φ such that $\varphi(x) = y$. By the previous proposition, if a space is homogeneous we can recover the entire topology from the neighborhood basis of a single point. In particular this applies to topological groups.

Nothing stops us from defining **neighborhood subbases**. However we have no need of them in what follows, so we leave this task to the reader.

6. The Subspace Topology

6.1. Defining the Subspace Topology.

Let (X, τ) be a topological space, and let Y be a subset of X. We want to put a topology on Y so as to satisfy the following properties:

• Let $f : X \to Z$ be a continuous function. Then $f|_Y : Y \to Z$ is continuous.

• Let $f : Z \to Y$ be a continuous function. Since $Y \subset Z$ we may view f as giving a map $f : Z \to X$. This map is continuous.

We define the subspace topology on Y as follows:

$$\tau_Y = \{U \cap Y \mid U \in \tau_X\}.$$

Let us check that this is indeed a topology on Y. First, $\varnothing = \varnothing \cap Y \in \tau_Y$. Second, $Y = X \cap Y \in \tau_Y$. Second, if $\{V_i\}_{i \in I}$ is a family of sets in τ_Y then for all i we have $V_i = U_i \cap Y$ for some $U_i \in \tau_X$. Thus

$$\bigcup_{i \in I} V_i = \bigcup_{i \in I} U_i \cap Y = (\bigcup_{i \in I} U_i) \cap Y \in \tau_Y.$$

Finally, if $V_1, V_2 \in \tau_Y$ then $V_1 = U_1 \cap Y$ and $V_2 = U_2 \cap Y$ for $U_1, U_2 \in \tau_X$. Thus

$$V_1 \cap V_2 = (U_1 \cap Y) \cap (U_2 \cap Y) = (U_1 \cap U_2) \cap Y \in \tau_Y.$$

Proposition 3.8. *Let $f : X \to Z$ be a continuous function, and let $Y \subset X$ be a subset. Then the restricted function $f : Y \to Z$ is continuous.*

Proof. For clarity let us denote the restriction of f to Y by $g : Y \to Z$. Let $V \subset Z$ be open. Then

$$g^{-1}(V) = \{x \in Y \mid g(x) \in V\} = \{x \in X \mid g(x) \in V\} \cap Y = f^{-1}(V) \cap Y$$

is open in Y. $\square$

Corollary 3.9. *Let Y be a subset of a topological space X. Then the inclusion map $\iota : Y \hookrightarrow X$ is continuous.*

Proof. Apply the previous result with $Z = X$ and $f = 1_X$. $\square$

Proposition 3.10. *Let Z be a topological space, let $Y \subset Z$ be a subset, and let $\iota : Y \to Z$ be the inclusion map. Let X be a topological space, and let $f : X \to Y$ be a function. The following are equivalent:*
(i) The function $f : X \to Y$ is continuous.
(ii) The function $\iota \circ f : X \to Z$ is continuous.

Proof. (i) $\implies$ (ii): This is immediate from the previous result and the fact that compositions of continuous functions are continuous.
(ii) $\implies$ (i): Let $V \subset Y$ be open. Since $f(X) \subset Y$, we have that

$$(\iota \circ f)^{-1}(V) = \{x \in X \mid \iota(f(x)) \in V\} = \{x \in X \mid f(x) \in V\} = f^{-1}(V)$$

is open in X. $\square$

Let (X, d) be a metric space, and let $Y \subset X$ be a subset. We have a potential embarrassment of riches situation: Y gets a topology, say τ_1, that it inherits as a subspace of the metric topology τ_X on X, and also a topology, say τ_2, that it gets from restricting the metric function to $d : Y \times Y \to \mathbb{R}$.

It is not completely obvious that τ_1 and τ_2 coincide. The following exercise explores the underlying issues.

Exercise 3.16. *Let (X,d) be a metric space, and let $Y \subset X$ be a metric space.*
a) Suppose $y \in Y$, let $\epsilon > 0$, let $B_Y(y,\epsilon)$ be the open ϵ-ball about y in Y, and let $B_X(y,\epsilon)$ be the open ϵ-ball about y in X. Show that

$$B_Y(y,\epsilon) = B_X(y,\epsilon) \cap Y. \tag{14}$$

b) Give an example of a subset Y of $\mathbb{R}^2$ (with the Euclidean topology) a point $x \in \mathbb{R}^2$ and $\epsilon > 0$ such that $B_X(x,\epsilon) \cap Y$ is not an open ball in Y.

Using this exercise we can see that $\tau_2 \subset \tau_1$. Indeed, since every set in τ_2 is a union of open ϵ-balls in Y, it is enough to check that for all $y \in Y$ and all $\epsilon > 0$, we have that $B_Y(y,\epsilon)$ lies in τ_1, and (14) shows this.

That $\tau_1 \subset \tau_2$ lies just a bit deeper. Namely, let $V \in \tau_1$, so there is an open subset $U \subset X$ with $V = U \cap Y$. Suppose $y \in V$. Since $v \in U$ there is $\epsilon > 0$ such that $B_X(v,\epsilon) \subset U$ and then

$$B_Y(v,\epsilon) = B_X(v,\epsilon) \cap Y \subset U \cap Y = V.$$

This shows that V is a union of elements of τ_2 (an empty union, if $V = \varnothing$), hence $V \in \tau_2$. We summarize:

Proposition 3.11. *The metric topology on a subset Y of a metric space X coincides with the topology Y inherits as a subspace of the metric topology on X.*

We remark that later we will introduce a topology τ_X on any ordered set $(X,\leq)$. For a subset $Y \subset X$ we will have an analogous embarrassment of riches situation: we can endow Y with the topology τ_1 it receives as a subspace of X and also the topology τ_2 it receives by restricting $\leq$ to an ordering on Y. Again it will be easy to show that $\tau_2 \subset \tau_1$. In this case though it can happen that $\tau_2 \subsetneq \tau_1$. This more complicated behavior of subspaces is probably one of the main reasons that order topologies are not as widely used as metric topologies.

As in the case of metric spaces, a subset Y of a topological space X is **dense** in X if $\overline{Y} = X$.

Exercise 3.17.
a) *Let $Z \subset Y \subset X$ with X a topological space. Suppose that Z is dense in Y and Y is dense in X. Show: Z is dense in X.*
b) *Let A, U be dense subspaces of a topological space X, with U open. Show: $U \cap A$ is open and dense in A.*

6.2. The Pasting Lemma.

Theorem 3.12. *(Pasting Lemma) Let X be a topological space, and let $\{Y_i\}_{i\in I}$ be a family of subsets of X with $\bigcup_i Y_i = X$. Let Z be a topological space. For each $i \in I$ let $f_i : Y_i \to Z$ be a continuous function. Consider the following conditions:*
(i) There is a continuous function $f : X \to Z$ such that $f|_{Y_i} = f_i$ for all $i \in I$.
(ii) For all $i \neq j \in I$, we have $f_i|_{Y_i \cap Y_j} = f_j|_{Y_i \cap Y_j}$.
a) In all cases we have (i) $\implies$ (ii).
b) If each Y_i is open, then (ii) $\implies$ (i).
c) If I is finite and each Y_i is closed, then (ii) $\implies$ (i).

Proof. Given any collection of maps $f_i : Y_i \to Z$, condition (ii) is necessary and sufficient for the existence of a map $f : X \to Z$ with $f|_{Y_i} = f_i$, and in this case the corresponding map is unique. This establishes part a). To show the remaining

parts we need to show that the unique such map f is continuous.
b) Suppose each Y_i is open. Let $x \in X$. It is enough to show that f is continuous at x. Let $V \subset Z$ be an open neighborhood of $f(x)$. Choose i such that $x \in Y_i$. Since f_i is continuous at x, there is an open neighborhood U_i of x in Y_i with $f_i(U_i) \subset V$. Since Y_i is open in X, U_i is open in X. Since $f|_{Y_i} = f_i$ we have $f_i(U_i) \subset V$.
c) Suppose each Y_i is closed. We will show that for all closed subsets $B \subset Z$, we have that $f^{-1}(B)$ is closed in Y. For each i we have that f_i is continuous, so $f_i^{-1}(B)$ is closed in Y_i. Because $X = \bigcup_{i=1}^n Y_i$ and $f|_{Y_i} = f_i$ for all i, we have

$$f^{-1}(B) = \bigcup_{i=1}^{n} f^{-1}(B) \cap Y_i = \bigcup_{i=1}^{n} f_i^{-1}(B).$$

Since I is finite, $f^{-1}(B)$ is a finite union of closed sets and is thus closed. $\square$

EXERCISE 3.18. *a) Give an example to show that the finiteness of I in part c) of Theorem 3.12 is necessary in order for the conclusion to hold.*
b) A family of subsets $\{Y_i\}_{i \in I}$ of a topological space X is ***locally finite*** *if for all $x \in X$ there is a neigbhorhood U of X such that $\{i \in I \mid Y_i \cap U \neq \varnothing\}$ is finite. Show that Theorem 3.12c) holds for a locally finite family $\{Y_i\}_{i \in I}$ of closed subsets.*

7. The Product Topology

CONVENTION: When we speak of the Cartesian product $\prod_{i \in I} X_i$ of an indexed family of sets $\{X_i\}_{i \in I}$, we will assume that $I \neq \varnothing$.[7]

Let $\{X_i\}_{i \in I}$ be a family of topological spaces, let $X = \prod_{i \in I} X_i$ be the Cartesian product, and for $i \in I$ let $\pi_i : X \to X_i$ be the ith **projection map**, $\pi_i(\{x_i\}) = x_i$. We want to put a topology on the Cartesian product X. Well, as above in the case of metric spaces we really want more than this – we could just put the discrete topology on X, but this is not (in general) what we want.

In the case of metric spaces, we focused on the property that a sequence $\mathbf{x}$ in X converges to p iff for all $i \in I$ the projected sequence $\pi_i(\mathbf{x})$ converges to $p_i = \pi_i(p)$ in X_i. In the context of a general topological space we still want this property, but because in a general topological space the topology need not be determined by the convergence of sequences, this is no longer a *characteristic* property.

We can suss out the right property by reflecting carefully on how functions behave on Cartesian products (of sets: no topologies yet). Going back to multivariable calculus, recall the difference between a function $f : \mathbb{R}^2 \to \mathbb{R}$ and a function $g : \mathbb{R} \to \mathbb{R}^2$. Then f is a "function of two variables" and such things are inherently more complicated than functions of a single variable: being "separately continuous" in the two variables is not enough to imply that f is continuous.

[7]The most reasonable convention for the Cartesian product over an empty family is that it should be a one-point set – e.g. we would then have $\#X^{\varnothing} = (\#X)^{\#\varnothing}$ – but *which* point should it be? Of course it doesn't really matter, but it is simpler to just evade the issue: in so doing we lose out on (literally) nothing.

EXERCISE 3.19. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by*

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}.$$

a) Show: $\forall x_0 \in \mathbb{R}$, the function $a : \mathbb{R} \to \mathbb{R}$ given by $a(y) = f(x_0, y)$ is continuous.
b) Show: $\forall y_0 \in \mathbb{R}$, the function $b : \mathbb{R} \to \mathbb{R}$ given by $b(x) = f(x, y_0)$ is continuous.
c) Show: f is not continuous.

On the other hand, the function g is a "vector-valued function of one variable". Indeed, if $\pi_1, \pi_2 : \mathbb{R}^2 \to \mathbb{R}$ are the two projection maps

$$\pi_1(x,y) = x, \ \ \pi_2(x,y) = y,$$

then we have

$$g = (\pi_1(g), \pi_2(g)).$$

So every $g : \mathbb{R} \to \mathbb{R}^2$ is no more and no less than a pair of functions $g_1, g_2 : \mathbb{R} \to \mathbb{R}$.

This is completely general. For any set Z and any Cartesian product $X = \prod_{i \in I} X_i$, for every function $f : Z \to X$, we have "component functions" $f_i := \pi_i \circ f : Z \to X_i$ and we uniquely recover f from these component functions as

$$f(z) = \{f_i(z)\}_{i \in I}.$$

To be a little fancier about it, recall that we write Y^X for the set of all maps $X \to Y$. Then we have a canonical bijection

$$\left(\prod_{i \in I} X_i\right)^Z = \prod_{i \in I} X_i^Z.$$

Okay, so what? The point is that this means that for any topological space Z and any family $\{X_i\}_{i \in I}$ of topological spaces, we know what we want the continuous functions $f : Z \to X = \prod_{i \in I} X_i$ to be: namely, we want $f : Z \to X$ to be continuous iff each of its projections $f_i = \pi_i \circ f : Z \to X_i$ is continuous. In general, given a topological space (X, τ), we can recover the topology τ from the knowledge of which functions $f : Z \to X$ from a topological space Z to X (for all topological spaces Z) are continuous. So *this* is the characteristic property of the product topology, and our task is to construct such a topology, ideally in a more direct, explicit way.

Let us begin with the case of two topological spaces X and Y.

PROPOSITION 3.13. *Let X and V be topological spaces, and let $\mathcal{B}$ be the family of all subsets $U \times V = \{(x,y) \in X \times Y \mid x \in U, y \in V\}$ as U ranges over all open subsets of X and V ranges over all subsets of Y.*

a) *The family $\mathcal{B}$ is a base for a topology τ on $X \times Y$.*
b) *In the topology τ, for any topological space Z, a function $f : Z \to X \times Y$ is continuous iff $f_1 = \pi_1 \circ f : Z \to X$ and $f_2 = \pi_2 \circ f : Z \to Y$ are both continuous. Thus τ is the desired product topology on $X \times Y$.*
c) *The maps $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ are continuous and open.*
d) *A sequence $x_\bullet$ in $X \times Y$ converges to $p \in X \times Y$ iff $\pi_1(x_\bullet) \to \pi_1(p)$ in X and $\pi_2(x_\bullet) \to \pi_2(p)$ in Y.*

Proof. First we dispose of an annoying technicality: the product $X \times Y$ is empty iff either $X = \varnothing$ or $Y = \varnothing$. The only set Z for which there is a function $Z \to \varnothing$ is when $Z = \varnothing$, and we will allow the reader to check that the result is (quite vacuous but) true in this case. Now suppose X and Y are both nonempty.
a) The elements of $\mathcal{B}$ are closed under finite intersections: $(U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2)$. This is (more than) enough for the set of unions of elements of $\mathcal{B}$ to be a topology on $X \times Y$.
b) Let $f : Z \to X \times Y$. Continuity can be checked on the elements of a base, so f is continuous iff for all U open in X and V open in Y, $f^{-1}(U \times V)$ is open in Z. But writing $f = (f_1, f_2)$ we have that

$$f^{-1}(U \times V) = \{z \in Z \mid (f_1(z), f_2(z)) \in U \times V\}$$

$$= \{z \in Z \mid f_1(z) \in U \text{ and } f_2(z) \in V\} = f_1^{-1}(U) \cap f_2^{-1}(V).$$

Thus if f_1 and f_2 are each continuous, then $f^{-1}(U \times V)$ is open so f is continuous. Conversely, if f is continuous, then for every open $U \subset X$,

$$f^{-1}(U \times Y) = \{z \in Z \mid f_1(z) \in U \text{ and } f_2(z) \in Y\} = \{z \in Z \mid f_1(z) \in U\} = f_1^{-1}(U)$$

is open in X, so $f_1 : Z \to X$ is continuous. Applying this argument with the roles of X and Y interchanged shows that $f_2 : Z \to Y$ is continuous.
c) If $V \subset X$ is open, then $\pi_1^{-1}(V) = V \times Y$ is open in $X \times Y$, so π_1 is continuous. Similarly π_2 is continuous. Since $\bigcup_{i \in I} f(U_i) = f(\bigcup_{i \in I} U_i)$, openness can be checked on a base, and certainly if $U_1 \subset X$ and $U_2 \subset Y$ are open, then $\pi_1(U_1 \times U_2) = U_1$ is open in X and $\pi_2(U_1 \times U_2) = U_2$ is open in Y. So π_1 and π_2 are open.
d) Since continuous functions preserve convergent sequences and by part c) the projection maps π_1 and π_2 are continuous, it is clear that $x_\bullet \to p$ implies $\pi_1(x_\bullet) \to \pi_1(p)$ and $\pi_2(x_\bullet) \to \pi_2(p)$. Conversely, suppose $\pi_1(x_\bullet) \to \pi_1(p)$ and $\pi_2(x_\bullet) \to \pi_2(p)$. Let N be a neighborhood of p in $X \times Y$; then $p \in U_1 \times U_2 \subset N$ with U_1 open in X and U_2 open in Y. Let $N \in \mathbb{Z}^+$ be sufficiently large so that for all $n \geq N$ we have $\pi_1(p) \in U_1$ and $\pi_2(p) \in U_2$. Then $p \in U_1 \times U_2$. This shows that $x_\bullet \to p$. $\square$

Now for any finite product $X = \prod_{i=1}^n X_i$ of topological spaces, we can define the product topology either by considering it as an iterated pairwise product – e.g. $X \times Y \times Z = (X \times Y) \times Z$ – or by modifying the definition of the product topology directly: namely, we may take as a base the collection of all subsets $W = \prod_{i=1}^n U_i$ such that U_i is open in X_i for all $1 \leq i \leq n$. No problem.

Things become more interesting for infinite products. Let us not try to disguise that the obvious first guess is simply to take as the base the collection of all Cartesian products of open sets in the various factors, namely $W = \prod_{i \in I} U_i$ with U_i open in X_i for all $i \in I$. It is certainly still true that these sets are closed under finite intersection and thus form a base for *some* topology on the infinite Cartesian product $X = \prod_{i \in I} X_i$. As is traditional, we call this topology the **box topology**. However, *this is not the correct definition of the product topology*, because it does not satisfy the property that a map $f : Z \to X$ is continuous iff each projection $f_i = \pi_i \circ f : Z \to X_i$ is continuous. Actually it is easier to explain the right thing than to explain why the wrong thing is wrong, so let us pass to the correct definition (with proof!) of the product topology and revisit this issue shortly.

I claim that we want to take as a base $\mathcal{B}$ the collection of all families $\{U_i\}_{i \in I}$

such that for all $i \in I$ U_i is open in X_i and that $U_i = X_i$ for all but finitely many $i \in I$. Again this family is closed under finite intersections so is certainly a base for a topology on the Cartesian product. Note also that this topology is coarser than the above box topology. Let us now check that for this topology, a function $f : Z \rightarrow X = \prod_{i \in I} X_i$ is continuous iff each $f_i = \pi_i \circ f : Z \rightarrow X_i$ is continuous. In fact the half of the argument that f continuous implies each f_i is continuous is essentially the same as above: for each fixed $i_\bullet \in I$, we choose a basis element $W = \prod_{i \in I} U_i$ with $U_i = X_i$ for all $i \neq i_\bullet$ and $U_{i_\bullet}$ an arbitrary open subset of X_i, and then if f is continuous then $f^{-1}(W) = f_i^{-1}(U_i)$ is open in Z. The other direction is also just as easy to do in this generality and very enlightening to do so: for $W = \prod_{i \in I} U_i$, we find

$$f^{-1}(W) = \{z \in Z \mid f_i(z) \in U_i \text{ for all } i \in I\} = \bigcap_{i \in I} f_i^{-1}(U_i).$$

Now we are assuming that each f_i is continuous and U_i is open in X_i, so each $f_i^{-1}(U_i)$ is open in Z. However, infinite intersections of open sets are not required to be open! So thank goodness we have required that $U_i = X_i$ for all but finitely many $i \in I$; since $f_i^{-1}(X_i) = Z$, the intersection is the same as we get by intersecting over the finitely many indices i such that U_i is a proper open subset of X_i, and thus is a finite intersection of open subsets of Z so is open in Z.

Exercise 3.20. *Let $X = \prod_{i \in I} X_i$ be a product of nonempty topological spaces.*

a) *Show that each projection map $\pi_i : X \rightarrow X_i$ is continuous and open.*
b) *Show that a sequence $x_\bullet$ in X converges to $p \in X$ iff for all $i \in I$ we have $\pi_i(x_\bullet) \rightarrow \pi_i(p)$.*

Exercise 3.21. *Let $X = \prod_{i \in I} X_i$ be a product of nonempty topological spaces. For each $i \in I$, let Y_i be a closed subset of X_i. Show: $Y := \prod_{i \in I} Y_i$ is closed in X.f*

Exercise 3.22. *Let $X = \prod_{i \in I} X_i$ be a product of nonempty topological spaces, with $\#X_i \geq 2$ for all $i \in I$. Show: X has no isolated points.*

Exercise 3.23. *Let I be an infinite index set, and for each $i \in I$ let X_i be a nontrivial topological space (i.e., the topology on X_i is not the indiscrete topology: in particular, $\#X_i \geq 2$). Show that the box topology on $\prod_{i \in I} X_i$ is strictly finer than the product topology on X.*

The following exercise gives an especially clear contrast between the behavior of the box topology and the product topology.

Exercise 3.24. *Let $X = \prod_{n=1}^{\infty} \{0, 1\}$. Give each $\{0, 1\}$ the discrete topology.*

a) *Give X the box topology. Show that X is discrete. More generally, show that any product of discrete spaces is discrete in the box topology.*
b) *Give X the product topology. Using the fact that a function $f : Z \rightarrow X$ is continuous iff each $f_n : Z \rightarrow X_n = \{0, 1\}$ is continuous, construct a homeomorphism from X to the classical Cantor set. Deduce that X is compact. More generally...?*

Theorem 3.14. *Let $X = \prod_{i \in I} X_i$ be a Cartesian product of nonempty topological spaces, endowed with the product topology. Let $\mathbf{x}$ be a sequence in X. Let $p \in X$. The following are equivalent:*

(i) *The sequence $\mathbf{x}$ converges to p in X.*

(ii) *For all $i \in I$, the sequence $\pi_i(\mathbf{x})$ converges to $\pi_i(p)$ in X_i.*

PROOF. (i) $\implies$ (ii): For each $i \in I$, $\pi_i : X \to X_i$ is continuous, and continuous functions preserve convergence of sequences.
(ii) $\implies$ (i): Suppose that $\pi_i(\mathbf{x}) \to \pi_i(p)$ for all $i \in I$. We need to show that every neighborhood of p in X contains $\mathbf{x}_n$ for all but finitely many $n \in \mathbb{Z}^+$. It is enough to check this on a base (in fact, on a neighborhood base at p...), so we may assume that $U = \prod_{i \in I} U_i$ with $U_i = X_i$ for all $i \in I \setminus J$, where $J \subset I$ is finite. For each $j \in J$, choose $N_j \in \mathbb{Z}^+$ such that we have $\pi_j(\mathbf{x}_n) \in U_j$ for all $n \geq N_j$, and put $N = \max_{j \in J} N_j$. Then $\mathbf{x}_n \in U$ for all $n \geq N$. $\square$

COROLLARY 3.15. *Let $\{(X_n, d_n)\}_{n=1}^{\infty}$ be an infinite sequence of metric spaces, and let $X = \prod_{n=1}^{\infty} X_n$.*

a) *There is a metric d on X which is **good** in the sense that for all sequences $\mathbf{x}$ in X and all $p \in X$, we have $\mathbf{x} \to p \iff \pi_n(\mathbf{x}) \to \pi_n(p)$ for all $n \in \mathbb{Z}^+$.*
b) *Any good metric on X induces the product topology on X.*

PROOF. a) We have already proven this: it is Corollary 2.37. b) We have already seen that any two metrics on a set which have the same convergent sequences and the same limits induce the same topology. By part a) and Theorem 3.14, this common topology is the product topology. $\square$

A topological space is **completely metrizable** if its topology is induced from some complete metric.

COROLLARY 3.16. *For each $n \in \mathbb{Z}^+$, let X_n be a nonempty completely metrizable space, and give $X := \prod_{n=1}^{\infty} X_n$ the product topology. Then X is completely metrizable.*

PROOF. Let d_n be a complete metric inducing the topology on X_n. Then the metric $d_n' := \max(d_n, 1)$ is uniformly equivalent: thus it is still complete and induces the same topology. Applying Proposition 2.57 to (X_n, d_n') shows that X is completely metrizable. $\square$

EXERCISE 3.25. a) *Show that a countable product of discrete spaces is completely metrizable.*
b) *Show: $\prod_{n=1}^{\infty} \mathbb{Z}^+$ is homeomorphic to the irrational numbers $\mathbb{I}$, and thus the space $\mathbb{I}$ is completely metrizable.*

It is worth comparing our current discussion of the product topology to our previous discussion of product metrics. In fact the present discussion is significantly simpler, as we do not have to resolve issues arising from the "embarrassment of riches". As an exercise, the reader might try to ignore our previous discussion of product metrics and simply prove directly that a countable product of metrizable spaces is metrizable. This takes about half a page!

From now on, whenever we meet a new property P of topological spaces, we will be interested to know whether it behaves nicely with respect to products. More precisely, we say that P is **productive** if whenever we have a family $\{X_i\}_{i \in I}$ of nonempty topological spaces each having property P, then $X = \prod_{i \in I} X_i$ (with the product topology!) has property P. Similarly we say that P is **factorable** if whenever we have a family $\{X_i\}_{i \in I}$ of nonempty topological spaces, if $X = \prod_{i \in I} X_i$ has property P then so does each X_i. Finally, we say that P is **faithfully productive** if it is both productive and factorable.

REMARK 3.17. *We really do want to require each X_i to be nonempty: if any X_i is empty, that makes the Cartesian product empty. The empty space has many good properties but not all: for instance, we will later prove that connectedness is productive, and according to our convention the empty space is not connected.*

The other direction is much more serious: if any one X_i is empty then the product is empty, so it would be the height of folly to try to deduce properties of the other factors from properties of $\varnothing$!

LEMMA 3.18. *Let $X = \prod_{i \in I} X_i$ be a product of nonempty topological spaces.*
a) For $i \in I$, the projection map $\pi_i : X \to X_i$ is open: for all open subsets $U \subset X$ we have $\pi_i(U)$ is open in X_i.
b) In general $\pi_i : X \to X_i$ need not *be closed: we may have a closed subset $A \subset X$ such that $\pi_i(A)$ is not closed in X_i.*

PROOF. a) Since $f(\bigcup_i Y_i) = \bigcup_i f(Y_i)$, a map $f : X \to Y$ of topological spaces is open iff $f(U)$ is open in Y for all U in a base $\mathcal{B}$ for the topology of X. Thus we may take $U = \prod_{j \in I} U_j$ with U_j open in X_j for all j and $U_j = X_j$ for all but finitely many j and then $f(U) = U_j$ is open in X_j.
b) Consider the map $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x$. Let $F : \mathbb{R}^2 \to \mathbb{R}$ by $F(x, y) = xy$. Then F is continuous, so

$$A = \{(x, y) \in \mathbb{R}^2 \mid xy = 1\} = F^{-1}(\{1\})$$

is closed in $\mathbb{R}^2$. But $\pi_1(A) = \mathbb{R} \setminus \{0\}$ is not. $\square$

EXERCISE 3.26. *a) Let $\pi_1 : [0, 1] \times [0, 1] \to [0, 1]$ be projection onto the first factor. Show that π_1 is closed.*
b) Comparing part a) with Lemma 3.18b) suggests that closedness of projection maps has something to do with compactness. Explore this. (We will address this connection in detail later on.)

Let $X = \prod_{i \in I} X_i$ be a Cartesian product of nonempty topological spaces. A **slice** in X is a subset of X of the form $X_i \times \prod_{j \neq i} \{p_j\}$; here we have chosen $i \in I$ and for all $j \neq i$, an element $p_j \in X_j$. Thus a slice is obtained precisely by restricting the values of all but one of the indices to be particular values and not restricting the remaining index. A **subslice** is a subset of a slice, which is thus of the form $Y_i \times \prod_{j \neq i} \{p_j\}$ for some subset $Y_i \subset X_i$.

LEMMA 3.19. *(Slice Lemma) Let $S = Y_i \times \prod_{j \neq i} \{p_j\}$ be a subslice in the product $X = \prod_{i \in I} X_i$ of nonempty topological spaces. Let $\pi_i : X \to X_i$ be the projection map. Then $\pi_i|_S : S \to Y_i$ is a homeomorphism. Thus every subspace of X_i is homeomorphic to a subspace of X.*

PROOF. The map π_i is the restriction of a continuous map so is continuous. It is plainly a bijection. It remains to check that it is open, which we may check on the elements of a base. There is a base for the topology of S consisting of sets V of the form $\prod_{i \in I} U_i \cap S$ in which each U_i is open in X_i and $U_i = X_i$ for all but finitely many i. Intersecting such a V with S we get either the empty set (if some $p_j \notin U_j$ for some $j \neq i$) or $(U_i \cap Y_i) \times \prod_{j \neq i} \{p_j\}$. Thus $\pi_i|_S(V)$ is either empty or of the form $U_i \cap Y_i$; either way we get an open subset of Y_i. $\square$

COROLLARY 3.20. *Let P be a topological property. If P is either hereditary or imagent, then P is factorable.*

PROOF. Let $X = \prod_{i \in I} X_i$ be a product of nonempty topological spaces which satisfies a topological property P.

Suppose P is hereditary. By the Slice Lemma, for each $i \in I$, X_i is homeomorphic to a slice S in X. Since P is hereditary, S has property P, and since P is topological, the homeomorphic space X_i has property P.

Suppose P is imagent. Then for each $i \in I$ we have $X_i = \pi_i(X)$, so X_i is a continuous image of X and thus has property P. $\square$

PROPOSITION 3.21. *For a topological space X, the following are equivalent:*

(i) *X is Hausdorff.*
(ii) *For all $x \in X$, the intersection of all closed neighborhoods of x is equal to $\{x\}$.*
(iii) *The image Δ of X under the diagonal map is closed in $X \times X$.*

PROOF. (i) $\implies$ (ii): Let $y \neq x$ in X and choose disjoint open neighborhoods U_x, U_y of x and y. Then $C_y := X \setminus U_y$ is a closed neighborhood of x which does not contain y.
(ii) $\implies$ (i): Let x and y be distinct points of X, and choose a closed neighborhood C_y of x which does not contain y. Then C_y° and $X \setminus C_y$ are disjoint open neighborhoods of x and y.
(i) $\iff$ (iii): Assume (i), and let $(x, y) \in X \times X \setminus \Delta$, i.e., $x \neq y$. Let U_x and U_y be disjoint open neighborhoods of x and y. Then $U_x \times U_y$ is an open neighborhood of (x, y) disjoint from (x, y), so (x, y) does not lie in the closure of Δ. So (i) $\implies$ (iii). The converse is quite similar and left to the reader. $\square$

7.1. Further Exercises.

EXERCISE 3.27 (Drobot-Sawka [**DS84**]). *Let $\{Y_i\}_{i \in I}$ be a nonempty family of nonempty metric spaces, and let Y be the Cartesian product $\prod_{i \in I} Y_i$ endowed with the* ***box topology****. Let X be a metric space, and let $f : X \to Y$ be a function, so f is given by $\{f_i : X \to Y_i\}_{i \in I}$.*

a) *Show that the following are equivalent:*
   (i) *The function f is continuous (we emphasize, for the* box topology *on $\prod_{i \in I} Y_i$!).*
   (ii) *Each $f_i : X \to Y_i$ is continuous and for all $x \in X$ there is $\delta > 0$ such that all but finitely many f_i's are constant on the open ball $B^\circ(x, \delta)$.*
b) *Suppose that X is moreover compact. Show: f is continuous iff each f_i is continuous and only finitely many f_i are not constant.*

8. The Coproduct Topology

Let $\{X_i\}_{i \in I}$ be an indexed family of sets. All of a sudden it is not critical that each $X_i \neq \varnothing$. In this context, allowing empty spaces is harmless albeit completely uninteresting.) We denote by $\coprod_i X_i$ the **disjoint union** of the X_i's. Roughly speaking, this means that we regard the X_i's as being pairwise disjoint and then take the union. Sadly, set theoretic correctness requires a bit more precision. The following works: for each $i \in I$, let $\tilde{X}_i = X_i \times \{i\}$. Then there is a super-obvious bijection $X_i \to \tilde{X}_i$ given by $x_i \in X_i \mapsto (x_i, i)$; and moreover have $\tilde{X}_i \cap \tilde{X}_j = \varnothing$ for all $i \neq j$ in I. So we may take

$$\coprod_i X_i = \bigcup_{i \in I} \tilde{X}_i.$$

For $i \in I$, we denote by ι_i the map $X_i \to \coprod_i X_i$, $x_i \mapsto (x_i, i)$.

EXERCISE 3.28. *Let Y be a set; for $i \in I$ let $f_i : X_i \to Y$ be a map. Show: there is a unique map $f : \coprod_i X_i \to Y$ such that $f \circ \iota_i = f_i : X_i \to Y$ for all $i \in I$.*

Now suppose that each X_i is a topological space. Our task is to put a useful topology on the coproduct $\coprod_i X_i$. We could motivate this via the preceding exercise but it seems to be simpler in this case just to give the construction. Namely, for each $i \in I$ let $\mathcal{B}_i$ be a base for τ_i (e.g. take $\mathcal{B}_i = \tau_i$). For $i \in I$, let

$$\tilde{\mathcal{B}_i} = \{\iota_i(U) \mid U \in \mathcal{B}_i.$$

(In other words, $\tilde{\mathcal{B}_i}$ is the copy of $\mathcal{B}_i$ in the relabelled copy $\tilde{X}_i$ of X_i.) Put

$$\mathcal{B} = \bigcup_{i \in I} \tilde{\mathcal{B}_i}.$$

Then $\mathcal{B}$ satisfies the axioms (B1) and (B2) for a base: since $\tilde{X}_i \in \tilde{\mathcal{B}_i}$ for all i, we have $\coprod_i X_i = \bigcup_i \tilde{X}_i$ is a union of elements of $\mathcal{B}$. Moreover, if $U_1, U_2 \in \mathcal{B}$ $U_1 \in \tilde{\mathcal{B}_i}$ and $U_2 \in \tilde{\mathcal{B}_j}$ for $i, j \in I$. If $i = j$ then every element of $U_1 \cap U_2$ contains some $U_3 \in \tilde{\mathcal{B}_i}$ because $\mathcal{B}_i$ is a base on X_i. If $i \neq j$ then $U_1 \cap U_2 = \varnothing$. Therefore the set τ of unions of elements of $\mathcal{B}$ is a base on $\coprod_{i \in I} X_i$. We call this the **coproduct topology** (also **the direct sum** and **the disjoint union**).

PROPOSITION 3.22. *Let $X = \coprod_{i \in I} X_i$ endowed with the coproduct topology.*

a) *For all $i \in I$, $\tilde{X}_i$ is open in X.*
b) *For all $i \in I$, the map $\iota_i : X_i \to \tilde{X}_i$ is a homeomorphism, and thus $\iota_1 : X_i \to X$ is an embedding.*
c) *For a subset $U \subset X$, the following are equivalent.*
   (i) *U is open.*
   (ii) *For all $i \in I$, $U \cap \tilde{X}_i$ is open in $\tilde{X}_i$.*
   (iii) *For all $i \in I$, $\iota^{-1}(U)$ is open in X_i.*
d) *Let Y be a topological space. For a map $f : X \to Y$, the following are equivalent:*
   (i) *f is continuous.*
   (ii) *For all $i \in I$, $f|_{\tilde{X}_i} : \tilde{X}_i \to Y$ is continuous.*
   (iii) *For all $i \in I$, $f \circ \iota_i : X_i \to Y$ is continuous.*

PROOF. a) Since $\tilde{X}_i$ is a union of elements of $\tilde{\mathcal{B}_i}$, it is open.
b) The map $\iota_i : X_i \to \tilde{X}_i$ is certainly a bijection. If $U_i \subset X_i$ is open, then U_i is a union of elements of $\mathcal{B}_i$, hence $\iota_i(U_i)$ is a union of the corresponding elements of $\tilde{\mathcal{B}_i}$, so is open in $\tilde{X}_i$. Conversely, if $V_i \subset \tilde{X}_i$ is open, then it is a union of elements of $\mathcal{B}$, but since it is contained in $\tilde{X}_i$ it is a union of elements of $\tilde{\mathcal{B}_i}$. We have $V_i = \iota_i(U_i)$ for a unique U_i (nothing more is going on here than converting from (x, i) to x) which is a union of the corresponding elements of $\mathcal{B}_i$, so it is open. It follows that $\iota_i : X_i \to \tilde{X}_i$ is a homeomorphism, so $\iota_1 : X_i \to X$ is an embedding.
c) (i) $\implies$ (ii) is the definition of the subspace topology.

(ii) $\iff$ (iii) follows from part b).
(ii) $\implies$ (i): By part a), $U \cap \tilde{X}_i$ is open in X, so $U = \bigcup_{i \in I} U \cap \tilde{X}_i$ is open in X.
(i) $\implies$ (ii): Restricting a continuous map to a subspace yields a continuous map.
(ii) $\iff$ (iii): This follows from the fact that $\iota_i : X_i \to \tilde{X}_i$ is a homeomorphism.
(ii) $\implies$ (i): This is a special case of the Pasting Lemma. □

EXERCISE 3.29. *Let X be a topological space and let $\{U_i\}_{i \in I}$ be an open covering of X. Show that a subset U of X is open iff $U \cap U_i$ is open in U_i for all $i \in I$.*

EXERCISE 3.30. *For each i in a nonempty set I, let X_i be a topological space, and let $X := \coprod_{i \in I} X_i$ be the coproduct. Show: X is metrizable iff X_i is metrizable for all $i \in I$.*

9. The Quotient Topology

We come now to one of the most geometrically useful constructions in general topology: the quotient space. This construction allows us to "identify" or "glue" points together in a topological space. We will see many examples later, but here are some basic ones to give the flavor.

• Let $X = [0, 1]$. If we identify 0 and 1 then we get (or will get...) a space homeomorphic to the circle S^1 (let us take our "standard model" of the circle to be the subspace $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$).
• Let $X = \mathbb{R}$. If we identify x and $x + 1$ for all $x \in \mathbb{R}$ then we get (or will get...) a space homeomorphic to S^1.
• Let $X = \mathbb{C}$. If we identify x and $x + 1$ for all $x \in \mathbb{C}$ then we get (or will get...) a space homeomorphic to an infinite open cylinder, i.e., homeomorphic to $S^1 \times \mathbb{R}$.
• Let $X = \mathbb{C}$. If we identify x, $x + 1$ and $x + i$ for all $x \in \mathbb{C}$ then we get (or will get...) a space homeomorphic to the torus, i.e., homeomorphic to $S^1 \times S^1$.
• Let $X = [0, 1]^N$ be the unit cube, viewed as a subset of $\mathbb{R}^N$. If we identify all points on the boundary ∂X of X, then we get a space homeomorphic to the N-sphere S^N. (Imagine raking leaves onto a square sheet and then pulling the edges of the sheet together to pick up the leaves.)

Our first task is of course to formalize this intuition. The first step is to understand "identifications" set theoretically in terms of equivalence relations. Let $\sim$ be an equivalence relation on a (say nonempty, to avoid trivialities) set X. Let $X/\sim$ be the set of equivalence classes, and let $q : X \to X/\sim$ be the natural map which sends x to its $\sim$-equivalence class $[x]$: then $q : X \to X/\sim$ is surjective and its fibers are precisely the equivalence classes.

The idea is that we "identify" x and y precisely when $x \sim y$. In the first example, the equivalence relation corresponds to having each $x \in (0, 1)$ equivalent only to itself and to having $0 \sim 1$. In the second case we said to identify x with $x+1$, which may initially suggest that the equivalence classes should be $\{x, x+1\}$. But this is not an equivalence relation: it is not transitive. This is not really a problem if we interpret the identification instructions as *generating* an equivalence relation rather than giving one: the equivalence relation generated by $x \sim x + 1$ is $x \sim y$ iff $x - y \in \mathbb{Z}$. In general, we must identify x and x whether we are told to or not (and why say it? it's obvious), when we are told to identify x and y we must also identify y and x (again, obviously) and if we identify x and y and then also identify y and z then we want to identify x and z even if not explicitly so directed.

Now we CLAIM that if $f : X \to Y$ is a map such that $x_1 \sim x_2 \implies f(x_1) = f(x_2)$, then there is a unique map $F : X/\sim \to Y$ such that

$$f = F \circ q.$$

In other words, there is a bijective correspondence between maps out of X which preserve $\sim$-equivalence classes and maps out of $X/\sim$. We ask the reader who is unfamiliar with this simple fact to stop and prove it on the spot.

Now suppose X is a topological space and $\sim$ is an equivalence relation on X. Our task is to endow $X/\sim$ with a topology so as to make $q : X \to X/\sim$ continuous and also to render true the topological analogue of the above fact, namely: if $f : X \to Y$ is a *continuous* map such that $x_1 \sim x_2 \implies f(x_1) = f(x_2)$, then the unique function $F : X/\sim \to Y$ such that $f = F \circ q$ *is continuous*.

We have to perform a bit of a balancing act: the coarser the topology is on $X/\sim$, the easier it will be for $q : X \to X/\sim$ to be continuous: indeed if we gave it the indiscrete topology then every map from a topological space into it would be continuous. But if $X/\sim$ has the indiscrete topology then it is very unlikely that the induced map $F : X/\sim \to Y$ will be continuous.

A little thought yields the following thought: of all topologies on $X/\sim$ that make $q : X \to X/\sim$ continuous, we want the finest one – that gives all the maps $F : X \to Y$ the best possible chance of being continuous. It is fairly clear from general nonsense that there will be a finest topology that makes q continuous (we will meet such general nonsense considerations in the following section), but in this case we can be more explicit: if $V \subset X/\sim$ is open, we need $q^{-1}(V)$ to be open. Therefore, if

$$\tau = \{V \subset Y \mid q^{-1}(V) \text{ is open}\}$$

is a topology, it must be the finest such topology. In fact τ is a topology: since $q^{-1}(\varnothing) = \varnothing$ is open in X, $\varnothing \in \tau$; since $q^{-1}(X/\sim) = X$ is open in X, $X/\sim \in \tau$; if for all $i \in I$, $V_i \in \tau$ then $q^{-1}(V_i)$ is open in X, hence so is $\bigcup_{i \in I} q^{-1}(V_i) = q^{-1}(\bigcup_{i \in I} V_i)$ and thus $\bigcup_{i \in I} V_i \in \tau$; and finally if $V_1, V_2 \in \tau$ then $q^{-1}(V_1)$ and $q^{-1}(V_2)$ are open in X so $q^{-1}(V_1 \cap V_2) = q^{-1}(V_1) \cap q^{-1}(V_2)$ is open in X, so $V_1 \cap v_2 \in \tau$.

And now the moment of truth: let $F : X/\sim \to Y$ be a map such that $f = F \circ q$ for a continuous map $f : X \to Y$. Does our "best chance topology"[8] τ on $X/\sim$ make F continuous? Happily, this is easily answered. Let $W \subset Y$ be open. Since $f = F \circ q$ for a continuous function $f : X \to Y$, we have $f^{-1}(W) = (F \circ q)^{-1}(W) = q^{-1}(F^{-1}(W))$ is open in X. Thus by the very definition of τ, because $q^{-1}(F^{-1}(W))$ is open in X we have that $F^{-1}(W)$ is open in $X/\sim$. Thus we have found the right topology τ on $X/\sim$: we call it the **identification space topology**.

Having defined the identification space associated to an equivalence relation on a topological space we now wish to define quotient maps. This is a fine distinction

[8]For A. Russell: the Swan topology? The Pinocchio topology??

but an important one, and it can be explained via the examples above. We found an equivalence relation $\sim$ on $[0,1]$ for which the identification space $[0,1]/\sim$ is *homeomorphic* to the circle S^1; if $\varphi : [0,1]/\sim \rightarrow S^1$ is such a homeomorphism, then we are more interested in the map $\varphi \circ q : [0,1] \rightarrow S^1$ than the map q itself. We would like a definition of "quotient map" which applies to $[0,1] \rightarrow S^1$, and similarly we want quotient maps $\mathbb{R} \rightarrow S^1$, $\mathbb{R}^2 \rightarrow S^1 \times \mathbb{R}$, $\mathbb{R}^2 \rightarrow S^1 \times S^1$ and $[0,1]^N \rightarrow S^N$.

As a side remark, the situation here is a close analogue of one that arises in group theory. If G is a group and H is a subgroup then we use H to define an equivalence relation $\sim_H$ on G: $g_1 \sim_H g_2$ iff $g_1 g_2^{-1} \in H$. In this case the equivalence classes are the cosets $\{gH \mid g \in G\}$, and we have a natural map $q : G \rightarrow G/H = G/\sim_H$, $g \mapsto gH$. If moreover H is normal in G then there is a unique group structure on G/H such that q becomes a surjective group homomorphism. This is the analogue of what we've done so far. But in group theory one goes farther: if $f : G \rightarrow G'$ is any surjective homomorphism of groups, then its kernel H is a normal subgroup, the map f is constant on $\sim_H$-equivalence classes and thus factors through $F : G/H \rightarrow G'$. But now the fundamental isomorphism theorem kicks in to say that F is an isomorphism of groups. As a result, we may regard any surjective group homomorphism $f : G \rightarrow G'$ as realizing G' as a quotient of G...meaning that there is a unique group isomorphism $F : G/\operatorname{Ker}(f) \rightarrow G'$ such that $f = F \circ q$.

We return to the topological situation: let $f : X \rightarrow Y$ be a surjective continuous map of topological spaces. Let $\sim_f$ be the equivalence relation on X given by $x_1 \sim x_2 \iff f(x_1) = f(x_2)$. Then f is constant on $\sim_f$-equivalence classes, so by our above discussion, if $q : X \rightarrow X/\sim_f$ is the identification map, we get a unique continuous function $F : X/\sim_f \rightarrow Y$ such that

$$f = F \circ q.$$

The map F is a bijection: this has nothing to do with topology and holds whenever we factor a map of sets through the associated equivalence relation $\sim_f$. We leave the verification of this as a simple but important exercise. We thus find ourself in a position of *nonanalogy* with the group theoretic case: namely the map $F : X/\sim_f \rightarrow Y$ is a continuous bijection of topological spaces...but it does not automatically follow that F is a homeomorphism! Indeed it is necessary and sufficient that F be an open map, i.e., if $V \subset X/\sim_f$ is open then $F(V)$ is open in Y. Now comes the following simple but important result.

PROPOSITION 3.23. *Let $f : X \rightarrow Y$ be a continuous surjective map of topological spaces. Let $\sim_f$ be the above equivalence relation, $q : X \rightarrow X/\sim_f$ the identification map and $F : X/\sim_f \rightarrow Y$ the unique continuous map such that $f = F \circ q$, which as above is a bijection. The following are equivalent:*
(i) F is a homeomorphism.
(ii) For all $V \subset Y$, we have that V is open if and only if $f^{-1}(V)$ is open in X.
When these equivalent conditions hold we say that $f : X \rightarrow Y$ is a ***quotient map****.*

PROOF. (i) $\implies$ (ii): We have already seen that F is a continuous bijection, so it is a homeomorphism iff it is an open map. Suppose F is open: then for $W \subset X/\sim_f$, W is open iff $F(W)$ is open in Y. Now let $V \subset Y$. If V is open, then since f is continuous, $f^{-1}(V)$ is open in X. On the other hand if V is not open, then $F^{-1}(V)$ is not open in $X/\sim_f$, and then by definition of the quotient topology

$$f^{-1}(V) = (q \circ F)^{-1}(V) = F^{-1}(q^{-1}(V))$$

is not open in X.
(ii) $\implies$ (i): Let $W \subset X/\sim_f$ be open. Since q is continuous, $q^{-1}(W)$ is open. Since F is a bijection, we have

$$f^{-1}(F(W)) = (F \circ q)^{-1}(F(W)) = q^{-1}(F^{-1}(F(W))) = q^{-1}(W).$$

Thus $f^{-1}(F(W))$ is open, which by assumption implies $F(W)$ is open. Thus F is an open map, hence as above it is a homeomorphism. $\square$

Exercise 3.31. *Let $f : X \to Y$ be a surjective map of topological spaces. Show that the following are equivalent:*
(i) f is a quotient map.
(ii) For all subsets $Z \subset Y$, Z is closed iff $f^{-1}(Z)$ is closed in X.

In theory the definition of a quotient map is simple and clean. In practice determining whether a continuous surjection is a quotient map can be a nontrivial task. The following result gives two pleasant sufficient conditions for this.

Proposition 3.24. *Let $f : X \to Y$ be a continuous surjection. If f is either open or closed, it is a quotient map.*

Proof. A continuous surjection $f : X \to Y$ is a quotient map iff for all $V \subset Y$, if $f^{-1}(V)$ then V is open iff for all $Z \subset Y$, if $f^{-1}(Z)$ is closed then Z is closed. Since f is surjective, for *all* $B \subset Y$ we have $f(f^{-1}(B)) = B$. Thus if f is open and $V \subset Y$ is such that $f^{-1}(V)$ is open, then $V = f(f^{-1}(V))$ is open. Similarly, if f is closed and $Z \subset Y$ is such that $f^{-1}(Z)$ is closed, then $Z = f(f^{-1}(Z))$ is closed. $\square$

Exercise 3.32. *Let $f : X \to Y$ be a map of sets. We say that a subset $A \subset X$ is **saturated** if $A = f^{-1}(f(A))$. We say that a subset $B \subset Y$ is **saturated** if $B = f(f^{-1}(B))$.*

a) *Show: $A \subset X$ is saturated iff it is a union of fibers $f^{-1}(y)$ for $y \in Y$.*
b) *Show that $B \subset Y$ is saturated iff $B \subset f(X)$. In particular, if f is surjective then every subset is saturated.*
c) *Show that for all $A \subset X$, $f^{-1}(f(A))$ is the smallest saturated subset of X containing A. Show that for all $B \subset Y$, $f(f^{-1}(B))$ is the largest saturated subset of Y contained in B.*
d) *Let $\mathcal{S}(X)$ be the set of saturated subsets of X and let $\mathcal{S}(Y)$ be the set of saturated subsets of Y (with respect to the map f, in both cases). Show:*

$$A \in \mathcal{S}(X) \mapsto f(A), \; B \in \mathcal{S}(Y) \mapsto f^{-1}(B)$$

give mutually inverse bijections between $\mathcal{S}(X)$ and $\mathcal{S}(Y)$.

Exercise 3.33. *Let $f : X \to Y$ be a surjective map of topological spaces.*

a) *Show that the following are equivalent:*[9]
   (i) *The map f is a quotient map.*
   (ii) *The open subsets of Y are precisely the images $f(U)$ of the saturated open subsets U of X under f.*

[9]This is almost a restatement of what we've already done, but it provides a useful way for thinking about quotient maps.

(iii) *The closed subsets of* Y *are precisely the images* $f(A)$ *of the saturated closed subsets* A *of* X *under* f.

b) *Suppose that* f *is moreover continuous. Show that the following are equivalent:*

(i) *The map* f *is a quotient map.*

(ii) *If* $U \subset X$ *is open and saturated, then* $f(U)$ *is open.*

(iii) *If* $A \subset X$ *is closed and saturated, then* $f(A)$ *is closed.*

c) *Show that a bijective quotient map is a homeomorphism.*

Thus being a quotient map is equivalent to subtly weaker conditions than either openness and closedness. And indeed a quotient map need not be open or closed.

Example 3.15. *For any family* $\{X_i\}_{i \in I}$ *of nonempty topological spaces, every projection map* $\pi_i : \prod_{i \in I} X_i \to X_i$ *is continuous, surjective and open and thus a quotient map. In general projections are not closed, e.g. the projection maps* $\mathbb{R}^2 \to \mathbb{R}$ *are not closed.*

Example 3.16. *Let* $f : [0, 2\pi] \to S^1$ *by* $f(\theta) = (\cos\theta, \sin\theta)$. *Then* f *is a continuous surjection. Like any continuous map of compact spaces, it is closed, so is a quotient map. However* $[0, \pi)$ *is open and* $f([0, \pi))$ *is not, so* f *is not open.*

Producing a quotient map which is *neither* open nor closed takes a little more work. The next two exercises accomplish this.

Exercise 3.34. *Let* $f : X \to Y$ *be a continuous map. A* ***section*** *of* f *is a continuous map* $\sigma : Y \to X$ *such that* $f \circ \sigma = 1_Y$. *Show that if* f *admits a section, it is a quotient map.*

Exercise 3.35. *Let* $\pi_1 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be projection onto the first coordinate:* $(x, y) \mapsto x$. *Let* $A = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0 \text{ or } y = 0\}$. *Let* $f = \pi_1|_A : A \to \mathbb{R}$.

a) *Show that* f *is quotient map. (Suggestion: use the previous exercise.)*

b) *Show that* f *is neither open nor closed.*

Exercise 3.36. a) *For a quotient map* $f : X \to Y$, *show: the following are equivalent:*

(i) *For all* $y \in Y$, $\{y\}$ *is closed.*[10]

(ii) *All fibers* $f^{-1}(y)$ *are closed subsets of* X.

b) *The rational numbers* $\mathbb{Q}$ *are a (normal, since* $\mathbb{R}$ *is commutative) subgroup of* $\mathbb{R}$. *We have a quotient map of groups* $q : \mathbb{R} \to \mathbb{R}/\mathbb{Q}$. *In particular this is a quotient by a continuous relation so we may put the identification space topology on* $\mathbb{R}/\mathbb{Q}$. *Show that in the resulting topology, for no* $y \in \mathbb{R}/\mathbb{Q}$ *is* $\{y\}$ *closed.*

Part b) of the above exercise gives in particular a quotient map $f : X \to Y$ in which X is Hausdorff and Y is not. By part a), this occurs because $\mathbb{Q}$ is not closed in $\mathbb{R}$ (rather it is proper and dense). Having the fibers be closed is a nice, checkable condition. Unfortunately this condition checks for something weaker than what we really want, which is that Y be Hausdorff. There is (much) more to say on Hausdorff quotient spaces, but we will content ourselves with the following result.

Proposition 3.25. *Let* $q : X \to Y$ *be an* open *quotient map, and let* $\sim$ *be the corresponding equivalence relation on* X*: i.e.,* $x_1 \sim x_2 \iff q(x_1) = q(x_2)$. *The following are equivalent:*

[10] We say that Y is "separated"; this property will be studied in detail in the next chapter.

(i) *Y is Hausdorff.*
(ii) *The relation $\sim$ is a closed subset of $X \times X$.*

Proof. We may assume that $Y = X/\sim$.
(i) $\implies$ (ii): The map $q \times q : X \times X \to Y \times Y$ is continuous. Since H is Hausdorff, the diagonal $\Delta \subset Y \times Y$ is closed, so

$$\sim = (q \times q)^{-1}(\Delta)$$

is closed in $X \times X$. (Note that this implication did not use that q is open.)
(ii) $\implies$ (i): Since q is open, so is $q \times q$. Let $U = X \times X \setminus \sim$. By assumption U is open, hence so is

$$(Y \times Y) \setminus \Delta = (q \times q)(U).$$

Thus Δ is closed in $Y \times Y$ so Y is Hausdorff. $\square$

Exercise 3.37. *Let $q : X \to Y$ be a quotient map with corresponding equivalence relation $\sim$ viewed as a subset of $X \times X$. Consider the map*

$$q \times q : X \times X \to Y \times Y, (x_1, x_2) \mapsto (q(x_1), q(x_2)).$$

a) *Show: $q \times q$ is continuous and surjective.*
b) *Let τ be the product topology on $Y \times Y$ and let τ_Q be the quotient topology on $Y \times Y$ induced from $q \times q$. Show that $\tau_Q \supset \tau$.*
c) *Show that if $\tau = \tau_Q$ and $\sim$ is closed in $X \times X$ then Y is Hausdorff.*
d) *Show: if q is open then $\tau_Q = \tau$.*
e) *Give an example in which $\tau_Q \supsetneq \tau$.*

The previous results give us reason to want our quotient maps to be open. The following exercise gives a useful instance in which this is the case.

Exercise 3.38. *Let X be a topological space, and let G be a group acting on X such that for all $g \in G$, $g\bullet : X \to X$ is a homeomorphism.*

a) *Show that the relation $\sim$ on X defined by $x_1 \sim x_2$ iff there is $g \in G$ with $gx_1 = x_2$ is an equivalence relation. We write X/G for $X/\sim$ and call it the* ***orbit space****.*
b) *Show that the quotient map $q : X \to X/G$ is open.*

10. Initial and Final Topologies

10.1. The Initial Topology. Let $\{Y_i\}_{i \in I}$ be a family of topological spaces, let X be a set, and let $\{f_i : X \to Y_i\}_{i \in I}$ be a family of functions. We will use this data to define a topology on X, the **initial topology**. Indeed, consider the family of all topologies τ on X with respect to which $f_i : X \to Y_i$ is continuous for all $i \in I$. The discrete topology is such a topology. It exists and is evidently the finest such topology. Moreover we did not need a family of maps $f_i : X \to Y_i$ to put the discrete topology on X, so this is a clue that going the other way will be more interesting. Namely, we consider the **coarsest possible topology** on X which makes each f_i continuous. It is not hard to see abstractly that such a thing exists: let $\mathcal{T}$ be the set of all topologies on X making each f_i continuous. By Exercise 3.12 we have that $\tau = \bigcap_{\sigma \in \mathcal{T}} \sigma$ is a topology on X. For all $i \in I$, if $V_i \subset Y_i$ is open, then $f^{-1}(V_i) \in \sigma$ for all $\sigma \in \mathcal{T}$, so $f^{-1}(V_i) \in \tau$. Thus (X, τ) is the coarsest possible topology that makes each f_i continuous.

Let us describe τ in a slightly different way. In order for each $f_i : X \to Y_i$ to be continuous it is necessary and sufficient for every open $V_i \subset Y_i$ that $f_i^{-1}(V_i)$ is open in X. Thus $\{f_i^{-1}(V_i)\} \subset \tau$, and since τ is the coarsest possible topology with this property, it must be the topology generated by $f_i^{-1}(V_i)$: that is, τ consists of arbitrary unions of finite intersections of the subbasic sets $f_i^{-1}(V_i)$.

It is natural to ask whether the family $\mathcal{F}$ of subsets $\{f_i^{-1}(V_i) \mid i \in I,\ V_i \in \tau_{X_i}\}$ is already a topology. In general it is not: for instance, let

$$I = \{1, 2\},\ X = \mathbb{R}^2,\ f_1 : \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto x,\ f_2 : \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto y.$$

Then the set $\mathcal{F}$ consists of all subsets of the form $U \times \mathbb{R}$ for U open in $\mathbb{R}$ and all subsets of the form $\mathbb{R} \times U$ for U open in $\mathbb{R}$. If U and V are proper nonempty open subsets of $\mathbb{R}$ then

$$U \times V = (U \times \mathbb{R}) \cap (\mathbb{R} \times V)$$

is an intersection of two elements of $\mathcal{F}$ that is not an element of $\mathcal{F}$. However in the case of one map $f : X \to Y$ the family $\mathcal{F}$ is already a topology:

EXERCISE 3.39. *Let Y be a topological space, let X be a set and let $f : X \to Y$ be a map. Show: $\{f^{-1}(V) \mid V$ is open in $Y\}$ is the initial topology on X.*

PROPOSITION 3.26 (Characteristic Property of Initial Topologies). *Let $\{Y_i\}_{i \in I}$ be a family of topological spaces, let X be a set, and let $\{f_i : X \to Y_i\}_{i \in I}$ be a family of maps. We endow X with the initial topology. Then for any topological space Z and any function $g : Z \to X$, the following are equivalent:*

(i) *We have that $g : Z \to X$ is continuous.*
(ii) *We have that $f_i \circ g : Z \to Y_i$ is continuous for all $i \in I$.*

PROOF. (i) $\implies$ (ii): Since g is continuous and the initial topology makes each f_i is continuous, it follows that each $f_i \circ g$ is continuous.
$\neg$ (i) $\implies$ $\neg$ (ii): The sets $\{f_i^{-1}(V_i) \mid i \in I,\ V_i \in \tau_{Y_i}\}$ form a subbase for the initial topology on X, so by Proposition 3.5, if g is not continuous then for some $i \in I$ and open $V_i \subset Y_i$ we have that $(f_i \circ g)^{-1}(V_i) = g^{-1}(f_i^{-1}(V_i))$ is not open in Z. So $f_i \circ g$ is not continuous. $\square$

EXAMPLE 3.17 (Subspace Topology Revisited). *Let Y be a topological space, let X be a subset of Y and let $f : X \hookrightarrow Y$ be the inclusion map. By Exercise 3.39 the initial topology on X with respect to f is*

$$\{f^{-1}(V) \mid V \in \tau_Y\} = \{V \cap X \mid V \in \tau_Y\}.$$

This is nothing else than the subspace topology on X. In this special case, Proposition 3.26 tells us that if Z is a topological space and $g : Z \to Y$ is a map such that $g(Z) \subset X$, then g is continuous iff the map $\underline{g} : Z \to X$ by $z \mapsto g(z)$ (i.e., the same map but with the codomain taken to be X instead of Y) is continuous.

EXAMPLE 3.18 (Product Topology Revisited). *Let $\{Y_i\}_{i \in I}$ be a nonempty family of nonempty topological spaces, let $X := \prod_{i \in I} Y_i$, and let $f_i : X \to Y_i$ be the ith projection map. Then the initial topology on X with respect to $\{f_i : X \to Y_i\}$ is nothing else than the product topology. Indeed, by definition a subbase for the initial topology is given by the subsets $f_i^{-1}(V_i)$ for $i \in I$ and V_i open in Y_i, and as seen above in a special case the set $f^{-1}(V_i)$ is V_i in the ith factor and Y_j for all $j \neq i$. Therefore if we intersect finitely many such sets we choose a finite subset $J \subset I$ and for each $j \in J$ an open subset V_j of X_j, and get $\prod_{j \in J} V_j \times \prod_{i \in I \setminus J} Y_i$. These* *are precisely the elements of the base we gave earlier for the product topology.*

Notice that by taking this route we are not tempted into the taking the box topology instead. Moreover, in this case Proposition 3.26 is an important and familiar property of the product topology, namely that for a topological space Z and a map $g : Z \to X$, we have that g is continuous iff each of its component maps $f_i \circ g : Z \to Y_i$ is continuous.

Exercise 3.40. a) *Let X have the initial topology induced by a family $\{f_i : X \to Y_i\}_{i\in I}$ of maps. For each $i \in I$, let J_i be a set and let $\{g_{ij} : Y_i \to Z_j\}_{j\in J_i}$ be family of continuous maps. Suppose that each Y_i has the initial topology induced by the family $\{g_{ij} : Y_i \to Z_j\}_{j\in J_i}$ of maps. Show: X has the initial topology induced by the family $\{g_{ij} \circ f_i : X \to Z_j\}_{i\in I, j\in J_i}$ of maps.*

b) *Deduce: if $Z \subset X$ with inclusion map $\iota : Z \to X$, then Z has the initial topology induced by the family $\{f_i \circ \iota : Z \to Y_i\}_{i\in I}$ of maps.*

10.2. The Final Topology. We now pursue a "dual notion" to that of initial topologies. Let $\{X_i\}_{i\in I}$ be a family of topological spaces, let Y be a set, and let $\{f_i : X_i \to Y\}$ be a family of functions. We will use this data to define a topology on X, the **final topology**.

Consider the family of all topologies τ on Y with respect to which $f_i : X_i \to Y$ is continuous for all $i \in I$. The indiscrete topology is such a topology and is evidently the coarsest such topology. Again, we did not need the family of maps $f_i : X_i \to Y$ to be the indiscrete topology on Y, so let's go the other way and consider the **finest possible topology** on Y that makes each f_i continuous. This time it is a little less clear that such a topology exists, but the following result shows that it does.

Lemma 3.27. *Let $\{X_i\}_{i\in I}$ be a family of topological spaces, let Y be a set, and for all $i \in I$ let $f_i : X_i \to Y$ be a function. Then*

$$\mathcal{F} \coloneqq \{V \subset Y \mid \forall i \in I, f_i^{-1}(V) \text{ is open in } X_i\}$$

is a topology on Y and thus the finest topology on Y that makes each f_i continuous.

Proof. For all $i \in I$ we have that $f_i^{-1}(\varnothing) = \varnothing$ is open in X_i, so $\varnothing \in \mathcal{F}$. Similarly, for all $i \in I$ we have that $f_i^{-1}(Y) = X_i$ is open in X_i, so $Y \in \mathcal{F}$.

If $Y_1, Y_2 \in \mathcal{F}$, then for all $i \in I$ we have that $f_i^{-1}(Y_1)$ and $f_i^{-1}(Y_2)$ are both open in X_i, so

$$f_i^{-1}(Y_1 \cap Y_2) = f_i^{-1}(Y_1) \cap f_i^{-1}(Y_2)$$

is open in X_i. Thus $Y_1 \cap Y_2 \in \mathcal{F}$.

If $\{Y_j\}_{j\in J} \in \mathcal{F}$, then for all $i \in I$ and $j \in J$ we have that $f_i^{-1}(Y_j)$ is open in X_i, so

$$f_i^{-1}(\bigcup_{j\in J} Y_j) = \bigcup_{j\in J} f_i^{-1}(Y_j)$$

is open in X_i. Thus $\bigcup_{j\in J} Y_j \in \mathcal{F}$.

This shows that $\mathcal{F}$ is a topology on Y. For any topology τ on Y that is not contained in $\mathcal{F}$, there is $Y \in \tau$ and $i \in I$ such that $f_i^{-1}(Y)$ is not open in X_i and thus $f_i : X \to (Y, \tau)$ is not continuous. $\square$

We call the topology constructed in Lemma 3.27 the **final topology** on Y relative to the family of maps $f_i : X_i \to Y$.

PROPOSITION 3.28 (Characteristic Property of Final Topologies). *Let $\{X_i\}_{i\in I}$ be a family of topological spaces, let Y be a set, and let $\{f_i : X_i \to Y\}_{i\in I}$ be a family of maps. We endow Y with the final topology. For any topological space Z and any function $g : Y \to Z$, the following are equivalent:*

(i) *We have that $g : Y \to Z$ is continuous.*
(ii) *We have that $g \circ f_i : X_i \to Z$ is continuous for all $i \in I$.*

PROOF. (i) $\implies$ (ii): Since g is continuous and the final topology makes each f_i continuous, it follows that each $g \circ f_i$ is continuous.
$\neg$ (i) $\implies$ $\neg$ (ii): If g is not continuous, there is an open subset W of Z such that $g^{-1}(W)$ is not open in Y. By definition of the final topologiy, there is $i \in I$ such that $f_i^{-1}(g^{-1}(W)) = (g \circ f_i)^{-1}(W)$ is not open in X, so $g \circ f_i$ is not continuous. $\square$

EXAMPLE 3.19 (Quotient Topology Revisited). *Let X be a topological space, Y a set and $f : X \to Y$ a surjective map. The final topology on Y is such that a subset V of Y is open if and only if $f^{-1}(V)$ is open in X, so it is nothing else than the quotient topology. In this special case, Proposition 3.28 says that for a topological space Z, a map $g : Y \to Z$ is continuous iff its pullback to X, $g \circ f$, is continuous.*

EXAMPLE 3.20 (Coproduct Topology Revisited). *Let $\{X_i\}_{i\in I}$ be a family of topological spaces, let $Y = \coprod_i X_i$, and let $f_i : X_i \hookrightarrow Y$ denote the map $x_i \mapsto (x_i, i)$ as in §3.8. The final topology on Y is such that a subset V of Y is open iff for all $i \in I$ the subset $f_i^{-1}(V)$ is open in X_i. This is nothing else than the topology we put on the coproduct. In this special case, Proposition 3.28 says that for a topological space Z, a map $g : Y \to Z$ is continuous iff for all $i \in I$ the pullback to X_i, $g \circ f_i : X_i \to Z$ is continuous: this is Proposition 3.22d).*

10.3. Embeddings and the Initial Topology.

A family $\{f_i : X \to Y_i\}_{i\in I}$ of functions on a set X **separates points of X** if for all $x \neq y \in X$ we have $f_i(x) \neq f_i(y)$ for some $i \in I$. The family **separates points from closed sets** if for all closed subsets A of X and points $p \in X \setminus A$, there is $i \in I$ such that $f_i(p) \notin \overline{f_i(A)}$.

THEOREM 3.29. *Let $\{f_i : X \to Y_i\}_{i\in I}$ be a family of continuous maps, and let $f = (f_i) : X \to \prod_{i\in I} Y_i$ be the corresponding continuous map. Then:*

a) *The map f is injective iff $\{f_i\}$ separates points of X.*
b) *The map $f : X \to f(X)$ is open if $\{f_i\}$ separates points from closed subsets.*
c) *[Embedding Lemma] If all points in X are closed and $\{f_i\}$ separates points from closed sets, then $f : X \to Y$ is a topological embedding.*

PROOF. a) This is immediate from the definition.
b) Let $p \in X$, and let U be a neighborhood of p. It is enough to show that $f(U)$ contains the intersection of an open neighborhood V of $f(p)$ with $f(X)$. Let $i \in I$ be such that $f_i(p) \notin \overline{f_i(X \setminus U)}$. We may take $V = \pi_i^{-1}(Y_i \setminus \overline{f_i(X \setminus U)})$.
c) Since points are closed and continuous functions separate points from closed subsets, continuous functions separate points. We apply parts a) and b). $\square$

The Embedding Lemma will be used later in the proof of the all-important Tychonoff Embedding Theorem. Notice that its proof followed almost immediately from the definitions involved. We now wish to go a bit deeper, following [**Wi**], by connecting the condition that a family of maps $\{f_i : X \to Y_i\}$ yields a topological embedding $f = (f_i) : X \to \prod_{i \in I} Y_i$ to initial topologies.

THEOREM 3.30. *Let $\{f_i : X \to X_i\}_{i \in I}$ be a family of continuous maps of topological spaces. Let $f : X \to \prod_{i \in I} X_i$ be the map $x \mapsto \{f_i(x)\}_{i \in I}$. The following are equivalent:*
(i) The map $f : X \to \prod_{i \in I} X_i$ is a topological embedding.
(ii) The space X has the initial topology induced by the family $\{f_i\}_{i \in I}$, and the family $\{f_i\}_{i \in I}$ separates points of X.

PROOF. [**Wi**, p. 56]. □

Let X be a topological space, and let $\{X_i\}_{i \in I}$ be an indexed family of topological spaces. For each $i \in I$, let $f_i : X \to X_i$ be a function.

THEOREM 3.31. *Let $f : X \to X_i$ be a family of continuous functions. The following are equivalent:*
(i) The family separates points from closed sets in X.
(ii) The family $\{f_i^{-1}(V_i) \mid i \in I,\ V_i \text{ open in } X_i\}$ is a base for the topology of X.

PROOF. (i) $\implies$ (ii): Let U be an open set of X and $p \in U$. Let $A = X \setminus U$. Then A is closed and does not contain p, so by hypothesis there is some $i \in I$ such that $f_i(p) \notin \overline{f_i(A)}$, which in turn means that there is some open neighborhood V_i of $f_i(p)$ in X_i which is disjoint from $f(A)$. Then $W = f_i^{-1}(V_i)$ is an open neighborhood of p disjoint from A and thus contained in U.
(ii) $\implies$ (i): let A be closed in X and let $p \in X \setminus A$. Then $U = X \setminus A$ is open and contains p. The given hypothesis implies that U contains an open neighborhood of p of the form $f_i^{-1}(V_i)$ for some $i \in I$ and V_i open in X_i. If $y \in V_i \cap f_i(A)$, then there is $a \in A$ with $f_i(a) \in V_i$, so $a \in f_i^{-1}(V_i) \subset U$, contradiction. Thus $V_i \cap f_i(A) = \varnothing$ and $f_i(p) \in V_i$, so $f_i(p) \notin \overline{f_i(A)}$. □

COROLLARY 3.32. *If $\{f_i : X \to X_i\}_{i \in I}$ is a family of continuous functions on the topological space X which separates points from closed sets, then the topology on X is the initial topology induced by the maps $\{f_i\}_{i \in I}$.*

PROOF. By Theorem 3.31, $\{f_i^{-1}(V_i) \mid i \in I,\ V_i \text{ open in } X_i\}$ is a base for the topology of X. But to say that X has the initial topology is to say that this family forms a subbase for the topology of X: okay. □

EXERCISE 3.41. *Let $\pi_1, \pi_2 : \mathbb{R}^2 \to \mathbb{R}$ be the two coordinate projections.*
a) *Show: $\{\pi_1, \pi_2\}$ does not separate points from closed sets in $\mathbb{R}^2$.*
b) *Show: $\mathbb{R}^2$ has the weak topology induced by $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$, $\pi_2 : \mathbb{R}^2 \to \mathbb{R}$.*

11. Compactness

11.1. First Properties.

A topological space is **quasi-compact** if every open cover admits a finite subcover. A topological space is **compact** if it is quasi-compact and Hausdorff.

EXERCISE 3.42. *Show that a topological space X is quasi-compact iff it satisfies the **finite intersection property**: if $\{F_i\}_{i \in I}$ is a family of closed subsets of X such that for all finite subsets $J \subset I$, $\bigcap_{i \in J} F_i \neq \varnothing$, then $\bigcap_{i \in I} F_i = \varnothing$.*

LEMMA 3.33. *Let C be a compact subset of the Hausdorff space X, and let $p \in X \setminus C$. Then there are disjoint open subsets $U, V \subset X$ with $p \in U$ and $C \subset V$.*

PROOF. Since $p \notin C$ and X is Hausdorff, for each $y \in C$ we may choose disjoint open neighborhoods U_y of p and V_y of y. Then $\{V_y\}_{y \in Y}$ is an open cover of the compact space Y, so there is a finite subcover, say $Y \subset \bigcup_{i=1}^N V_{y_i}$. We may take $U = \bigcap_{i=1}^N U_{y_i}$ and $V = \bigcup_{i=1}^N V_{y_i}$. □

PROPOSITION 3.34.

a) *A closed subspace of a quasi-compact space is quasi-compact, and a closed subspace of a compact space is compact.*
b) *If X is Hausdorff and $C \subset X$ is compact, then C is closed.*

PROOF. a) Let X be quasi-compact and let $Y \subset X$ be closed. Let $\{V_i\}_{i \in I}$ be a family of open subsets of Y which cover Y. By definition of the subspace topology, for each $i \in I$ there is an open subset $U_i \subset X$ with $V_i = U_i \cap Y$. Then $\{U_i\}_{i \in I} \cup \{X \setminus Y\}$ is an open covering of the quasi-compact space X, so there is a finite subcovering:

$$X = \bigcup_{i=1}^N U_i \cup (X \setminus Y).$$

Intersecting with Y gives

$$Y = \bigcup_{i=1}^N (U_i \cap Y) \cup (X \setminus Y) \cap Y = \bigcup_{i=1}^N V_i.$$

Since (all) subspaces of Hausdorff spaces are Hausdorff, a closed subspace of a compact space is compact.
b) Let $p \in X \setminus C$. By Lemma 3.33, there are disjoint open sets U containing p and V containing C. In particular $p \in U \subset X \setminus C$, so $p \in (X \setminus C)^{circ}$. Since this holds for all p, $X \setminus C$ is open and thus C is closed. □

EXERCISE 3.43. *a) Show: a finite union of quasi-compact subsets is quasi-compact.*
b) Show: a countably infinite union of compact subsets need not be compact.
*c) (**WARNING!**) Show: The intersection of two quasi-compact sets need not be quasi-compact.*
d) Show: a finite intersection of compact subsets is compact.

EXERCISE 3.44. *a) Show that compactness is* not *a hereditary property.*
*b) A topological space is **hereditarily compact** if every subspace is compact. Show that a topological space is hereditarily compact iff it is finite.*
*c) Show that any indiscrete space is hereditarily quasi-compact. Deduce that there exist hereditarily quasi-compact spaces of all possible cardinalities. (We will later study hereditarily quasi-compact spaces and see that in particular they are precisely those spaces for which the open subsets satisfy the **Ascending Chain Condition**.)*

THEOREM 3.35. *Quasi-compactness is an imagent property: if X is quasi-compact and $f : X \to Y$ is a continuous surjection, then Y is quasi-compact.*

Proof. Let $\mathcal{V} = \{V_i\}_{i \in I}$ be an open covering of Y. For each $i \in I$, let $U_i = f^{-1}(V_i)$. Then each U_i is open in X and

$$X = f^{-1}(Y) = f^{-1}(\bigcup_{i \in I} V_i) = \bigcup_{i \in I} f^{-1}(V_i) = \bigcup_{i \in I} U_i,$$

so $\mathcal{U} = \{U_i\}_{i \in I}$ is an open covering of X. Since X is quasi-compact, there is a finite subset $J \subset I$ such that $X = \bigcup_{i \in J} U_i$, and then

$$Y = f(X) = f(\bigcup_{i \in J} U_i) = \bigcup_{i \in J} f(U_i) = \bigcup_{i \in J} V_i,$$

so $\{V_i\}_{i \in J}$ is a finite subcovering. $\square$

Corollary 3.36. *(Extreme Value Theorem) If X is quasi-compact and $f : X \to \mathbb{R}$ is continuous, f is bounded and attains its maximum and minimum values.*

Proof. By Theorem 3.35, $f(X)$ is a compact subset of the metric space $\mathbb{R}$, hence is closed and bounded. Thus $f(X)$ contains its infimum and supremum. $\square$

A topological space is **pseudocompact** if every continuous real-valued function on that space is bounded. Thus the Extreme Value Theorem states quasi-compact spaces are pseudocompact. At first glance it seems to give a little more – the attainment of the maximum and minimum – but in fact that comes along for free.

Exercise 3.45. *Let X be a pseudocompact space, and let $f : X \to \mathbb{R}$ be a continuous function. Show that f attains its maximum and minimum values.*

Exercise 3.46. *A topological space is* ***irreducible*** *if it is nonempty and is not the union of two proper closed subsets.*
a) Show that a continuous image of an irreducible space is irreducible.
b) Show that a topological space X is irreducible and Hausdorff iff $\#X = 1$.
c) Show that an irreducible space is pseudocompact.

We saw – well, up to a big theorem whose proof still lies ahead of us – that every pseudocompact metric space is compact. As the terminology suggests, this is far from being true for arbitrary topological spaces: there is quite a menagerie of pseudo-compact noncompact spaces.

Exercise 3.47. *Let X be an infinite set endowed with the particular point topology (you pick the point!).*
a) Show: X is irreducible, hence pseudocompact.
b) Show: X is not quasi-compact.

It follows from Theorem 3.35 that quasi-compactness is a factorable property: it passes from a nonempty Cartesian product to each factor space. Of course the next question to ask is whether quasi-compactness is productive, i.e., must all products of quasi-compact spaces be quasi-compact? Since Hausdorffness is faithfully productive, it would then follow that compactness is faithfully productive.

It turns out that the productivity of quasi-compactness is true, rather difficult to prove, and of absolutely ubiquitous use in the subject: it is perhaps the single most important theorem of general topology! It is certainly easily said:

Theorem 3.37. *(Tychonoff) Arbitrary products of quasi-compact spaces are quasi-compact. It follows that arbitrary products of compact spaces are compact.*

We are not going to prove the general case of Tychonoff's Theorem in this section. On the contrary, a clean conceptual proof of Tychonoff's Theorem will be the main application of our general study of convergence in topological spaces, to which we devote an entire chapter. However it is much easier to prove the result for finite products, and though that will turn out to be logically superfluous (i.e., the proof of the general case will not rely on this) nevertheless the proof showcases some important ideas, so we will give it now.

THEOREM 3.38 (Tube Lemma). *Let X be a topological space and let Y be a quasi-compact topological space. Let $x_0 \in X$, and let $\mathcal{N}$ be a neighborhood of $\{x_0\} \times Y$ in the product space $X \times Y$. Then there is a neighborhood U of x_0 in X such that $U \times Y \subset \mathcal{N}$.*

PROOF. For each $y \in Y$, choose a basic open subset $U_y \times V_y$ of $X \times Y$ with $(x_0, y) \subset U_y \times V_y \subset \mathcal{N}$. Then $\{U_y \times V_y\}_{y \in Y}$ is an open cover of the quasi-compact space[11] $\{x_0\} \times Y$, and we may extract a finite subcover, say $\{U_i \times V_i\}_{i=1}^n$. Then $U = \bigcap_{i=1}^n U_i$ is an open neighborhood of x_0 in X. Let $(x, y) \in U \times Y$. For at least one i, we have $(x_0, y) \in U_i \times V_i$, so

$$(x, y) \in U \cap V_i \subset U_i \cap V_i \subset \mathcal{N}.$$

It follows that $U \times Y \subset \mathcal{N}$. $\square$

COROLLARY 3.39. *(Little Tychonoff Theorem) Let $X_1, \ldots, X_N$ be quasi-compact topological spaces. Then $X = \prod_{i=1}^N X_i$ is quasi-compact in the product topology.*

PROOF. Induction reduces us to the case $N = 2$. Let $\mathcal{U}$ be an open cover of $X_1 \times X_2$. For each $x \in X_1$, let $\mathcal{U}_x$ be a finite subset of $\mathcal{U}$ which covers $\{x\} \times X_2$ (Slice Lemma again). Then $\mathcal{N}_x = \bigcup \mathcal{U}_x$ is an open neighborhood of $\{x\} \times X_2$. Since X_2 is quasi-compact, by the Tube Lemma there is an open neighborhood W_x of x_0 in X such that $W_x \times X_2 \subset \mathcal{N}_x$. Since X_1 is quasi-compact, there is a finite subset set $\{x_1, \ldots, x_m\}$ of X_1 such that $\bigcup_{i=1}^m W_{x_i} = X_1$. Then $\bigcup_{i=1}^m \mathcal{U}_{x_i}$ is a finite subcover of $X_1 \times X_2$. $\square$

EXERCISE 3.48. *Suppose a topological space Y satisfies the conclusion of the Tube Lemma. Show that for all topological spaces X, the projection map $\pi_1 : X \times Y \to X$ is a closed map.*

REMARK 3.40. *It turns out to be true that for a topological space, being quasi-compact, satisfying the conclusion of the Tube Lemma and projection maps being closed are all equivalent. This requires tools we have not yet developed and we will return to it later.*

In fact the Tube Lemma can be stated in a stronger form:

THEOREM 3.41 (Generalized Tube Lemma). *Let X and Y be topological spaces, let A be a quasi-compact subset of X and let B be a quasi-compact subset of Y. Let $\mathcal{N}$ be an open neighborhood of $A \times B$ in the product space $X \times Y$. Then there is an open neighborhood U of A in X and an open neighborhood V of B in Y such that*

$$A \times B \subseteq U \times V \subseteq \mathcal{N}.$$

[11] Here we use the Slice Lemma.

PROOF. (Brandsma) For each $(a, b) \in A \times B$, choose an open subset $U(a, b)$ of X and an open subset $V(a, b)$ of Y such that

$$(a, b) \in U(a, b) \times V(a, b) \subseteq \mathcal{N}.$$

For fixed $b \in B$, $\{U(a, b)\}_{a \in A}$ is an open cover of the quasi-compact subset A, so there is a finite subset $F_b \subseteq A$ such that

$$U(b) := \bigcup_{a \in F_b} U(a, b) \supseteq A.$$

If we put

$$V(b) := \bigcap_{a \in F_b} V(a, b),$$

then $V(b)$ is an open neighborhood of b and

$$(a, b) \in U(b) \times V(b) \subseteq \mathcal{N}.$$

Now $\{V(b)\}_{b \in B}$ is an open cover of the quasi-compact subset B, so there is a finite subset $G \subseteq B$ such that

$$V := \bigcup_{b \in G} V(b) \supseteq B.$$

If we put

$$U := \bigcap_{b \in G} U(b),$$

then

$$A \times B \subseteq U \times V \subseteq \mathcal{N}$$

as required. □

11.2. Variations on a theme.

A topological space X is **sequentially compact** if every sequence admits a convergent subsequence.

A topological space X is **countably compact** if every countable open cover $\{U_n\}_{n=1}^{\infty}$ admits a finite subcover. This is equivalent to the finite intersection property for countable families $\{F_n\}_{n=1}^{\infty}$ of closed subsets. By passing from F_n to $\mathcal{F}_n = \bigcap_{i=1}^{n} F_i$, we see that a space is countably compact iff every nested sequence of nonempty closed subsets has nonempty intersection.

A topological space X is **limit point compact** if every infinite subset $Y \subset X$ has a **limit point** in X, i.e., there exists $x \in X$ such that for every open neighborhood U of x, $U \setminus \{x\} \cap Y \neq \varnothing$.

Thus Bolzano-Weierstrass asserts that $[a, b]$ is limit point compact, whereas Theorem 1.13 asserts, in particular, that $[a, b]$ is sequentially compact.

EXERCISE 3.49. *a) Show: a sequentially compact space is pseudocompact.*
b) Show: a closed subspace of a sequentially compact space is sequentially compact.
c) Must a sequentially compact subspace of a Hausdorff space be closed?

Proposition 3.42. *Let X be a topological space.*
a) If X is countably compact, it is limit point compact.
b) In particular a compact space is limit point compact.
c) If X is sequentially compact, it is countably compact.
d) In particular a sequentially compact space is limit point compact.

Proof. a) We establish the contrapositive: suppose there exists an infinite subset of X with no limit point; then there exists a countably infinite subset $A \subset X$ with no limit point. Such a subset A must be closed, since any element of $\overline{A} \setminus A$ is a limit point of A. Moreover A must be discrete: for each $a \in A$, since a is not a limit point of A, there exists an open subset U such that $A \cap U = \{a\}$. Now write $A = \{a_n\}_{n=1}^{\infty}$, and define, for each $N \in \mathbb{Z}^+$, $F_N = \{a_n\}_{n=N}^{\infty}$. Then each F_N is closed, any finite intersection of F_N's is nonempty, but $\bigcap_{N=1}^{\infty} F_N = \varnothing$, so X is not countably compact.
b) Clearly a compact space is countably compact; now apply part a).
c) Let $\{F_n\}_{n=1}^{\infty}$ be a nested sequence of closed subsets of X, and choose for all $n \in \mathbb{Z}^+$ a point $x_n \in F_n$. By sequential compactness, after passing to a subsequence – let us suppose we have already done so and retain the current indexing – we get $x \in X$ such that $x_n \to x$. We claim $x \in \bigcap_{n=1}^{\infty} F_n$. Suppose not: then there is $N \in \mathbb{Z}^+$ such that $x \notin F_N$. But then $U = X \setminus F_N$ is an open neighborhood of x, so for all sufficiently large n, $x_n \in U$ and thus $x_n \notin F_N$. But as soon as $n \geq N$ we have $F_n \subset F_N$ and thus $x_n \notin F_n$, contradiction.
d) Apply part c) and then part a). □

Proposition 3.43.
A first countable limit point compact space in which every point is closed is sequentially compact.

Proof. Let a_n be a sequence in X. If the image of the sequence is finite, we may extract a constant, hence convergent, subsequence. Otherwise the image $A = \{a_n\}_{n=1}^{\infty}$ has a limit point a, and since every point of X is closed, every limit point is an ω-limit point: every neighborhood U of a contains infinitely many points of A. Let $\{N_n\}_{n=1}^{\infty}$ be a nested countable neighborhood base at x. Choose n_1 such that $x_{n_1} \in N_1$. For all $k > 1$, choose $n_k > n_{k-1}$ with $x_{n_k} \in N_k$. Then $x_{n_k} \to x$. □

Proposition 3.44. *Sequential compactness is an imagent (hence also factorable) property.*

Proof. Let $f : X \to Y$ be a surjective continuous map, with X sequentially compact. Let $\mathbf{y}$ be a sequence in Y. Since f is surjective, for all $n \in \mathbb{Z}^+$ we may choose $\mathbf{x}_n \in f^{-1}(\mathbf{y}_n)$ and get a sequence $\mathbf{x}$ in X. By hypothesis, there is a subsequence $\mathbf{x}_{n_k}$ converging to a point $p \in X$. Then by continuity $\mathbf{y}_{n_k} = f(\mathbf{x}_{n_k})$ converges to $f(p)$. □

Example 3.21. *Let $X = \{0,1\}^{[0,1]}$: we give each factor $\{0,1\}$ the discrete topology and X the product topology. By (a case which we have not yet proved of) Tychonoff's Theorem, X is compact. We claim that X is not sequentially compact. This will show two things: that compact spaces need not be sequentially compact and that – unlike quasi-compactness! – sequential compactness is not productive.*

An element of X is a function $f : [0,1] \to \{0,1\}$. Let $\{f_n : [0,1] \to \{0,1\}\}_{n=1}^{\infty}$ be a sequence. By Exercise 3.20 for $f : [0,1] \to \{0,1\}$ we have $f_n \to f$ iff for all $x \in [0,1]$ we have $\pi_x(f_n) \to \pi_x(f)$. But $\pi_x(f_n) = f_n(x)$ and $\pi_x(f) = f(x)$, *so convergence means that* $f_n(x) \to f(x)$ *for all* $x \in [0,1]$. *In turn, since this convergence takes place in the discrete space* $\{0,1\}$, *we must have that* $f_n(x) = f(x)$ *for all sufficiently large* n. *All in all, a sequence* $f_n \in \{0,1\}^{[0,1]}$ *converges iff for all* $x \in [0,1]$ *the binary sequence* $\{0,1\}$ *is eventually constant.*

Now define $f_n : [0,1] \to \{0,1\}$ *by mapping* $\alpha \in [0,1]$ *to the* n*th digit of its binary expansion; to avoid ambiguity, we never use a binary expansion ending in an infinite sequence of* 1*'s. We claim that* f_n *has no convergent subsequence. Indeed, let* $n_\bullet : \mathbb{Z}^+ \to \mathbb{Z}^+$ *be any strictly increasing function. There is* $x \in [0,1]$ *whose* n_{2k-1}*th binary digit is* 0 *and whose* n_{2k}*th binary digit it* 1. *Indeed, even on this subsequence there are infinitely many* 0*'s, so completing the binary expansion in any way along elements of* $\mathbb{Z}^+ \setminus \{n_k \mid k \in \mathbb{Z}^+\}$ *gives a legal binary expansion. Then the sequence* $\{f_n(x)\}$ *is* $0, 1, 0, 1, 0, 1 \ldots$, *which is not eventually constant. Thus* f_n *has no convergent subsequence, so* $\{0,1\}^{[0,1]}$ *is not sequentially compact.*

Example 3.22. *There are sequentially compact topological spaces which are not compact, but the ones I know involve order topologies, which we will discuss a little later on. For now we just record that the* ***least uncountable ordinal*** *and the* ***long line*** *are sequentially compact but not compact.*

Proposition 3.45. *Sequential compactness is countably productive: if* $\{X_n\}_{n=1}^\infty$ *is a sequence of sequentially compact spaces, then* $X = \prod_{n=1}^\infty X_n$ *is sequentially compact in the product topology.*

Exercise 3.50. *Prove Proposition 3.45. (Suggestion: make a diagonalization argument. See for instance the proof of Theorem 2.65 for something similar.)*

Proof. A diagonalization argument. □

12. Connectedness

12.1. Basics.

Let X be a nonempty topological space. A **presep** on X is an ordered pair (U, V) of open subsets of X with $U \cup V = X$, $U \cap V = \varnothing$. X certainly admits two preseps, namely $(X, \varnothing)$ and $(\varnothing, X)$; any other presep of (U, V) of X – i.e., in which U and V are each nonempty – is called a **separation** of X.

A space X is **connected** if it is nonempty and *does not* admit a separation.

Example 3.23. *Let* X *be a nonempty set endowed with the discrete topology. The preseps on* X *correspond to the subsets of* X, *via* $Y \mapsto (Y, X \setminus Y)$. *"Thus"*[12] *a discrete space* X *is connected iff* $\#X = 1$.

Let $f : X \to Y$ be a continuous map, and let (U, V) be a presep on Y. Then $(f^{-1}(U), f^{-1}(V))$ is a presep on X. If (U, V) is a separation and f is surjective, then $(f^{-1}(U), f^{-1}(V))$ is a separation on X. This shows:

Proposition 3.46. *The continuous image of a connected space is connected.*

In particular, let $\{0,1\}$ be a two-point discrete space, with the separation $(\{0\}, \{1\})$. For a topological space X and a continuous function $f : X \to \{0,1\}$, $(f^{-1}(\{0\}), (f^{-1}(\{1\}))$ is a presep on X. Conversely, for a presep (U, V) on X, mapping $x \in U \mapsto 0$ and

[12]This was rather clear in any event!

$x \in V \mapsto 1$ gives a continuous function $f : X \to \{0,1\}$. These constructions are mutually inverse bijections between $C(X, \{0,1\})$ and the set of preseps on X.

Recall that an ordered space is connected in the order topology if it is nonempty, order-dense and Dedekind complete. It follows that a nonempty subset of $\mathbb{R}$ is connected iff it is compact iff it is an interval.

PROPOSITION 3.47. *Let Y be a connected subset of a topological space X. Then $\overline{Y}$ is connected.*

PROOF. We may assume without loss of generality that $\overline{Y} = X$. We show the contrapositive: suppose X is not connected, and let (U, V) be a separation. Since Y is dense in X, $U \cap Y$ and $V \cap Y$ are nonempty so the presep $(U \cap Y, V \cap Y)$ of Y is a separation. $\square$

PROPOSITION 3.48. *Let $\{Y_i\}_{i \in I}$ be a nonempty family of connected sets in a topological space X.*

a) *If $\bigcap_{i \in I} Y_i \neq \varnothing$, then $\bigcup_{i \in I} Y_i$ is connected.*
b) *If I is a linearly ordered set and for all $i \leq j$, we have $Y_i \subset Y_j$, then $\bigcup_{i \in I} Y_i$ is connected.*

PROOF. We may assume without loss of generality that $X = \bigcup_{i \in I} Y_i$.
a) Let $x \in \bigcap_{i \in I} Y_i$. Let (U, V) be a presep on X with $x \in U$. Then for all $i \in I$, $(U_i, V_i) = (U \cap Y_i, V \cap Y_i)$ is a presep on the connected space Y_i. Since $x \in U_i$ we have $V_i = \varnothing$ for all i and thus

$$V = V \cap (\bigcup_{i \in I} Y_i) = \bigcup_{i \in I} (V \cap Y_i) = \bigcup_{i \in I} V_i = \varnothing.$$

Thus X admits no separation.
b) Choose $i_0 \in I$ and $x \in Y_i$. Then $X = \bigcup_{i \in I} Y_i = \bigcup_{i \geq i_0} Y_i$. Apply part a). $\square$

EXERCISE 3.51. a) *Let $Y_1, Y_2 \subset X$ be connected subsets with $Y_1 \cap Y_2 = \varnothing$. Give examples to show that $Y_1 \cup Y_2$ may or may not be connected.*
b) *Show that in $\mathbb{R}$, the intersection of any family of connected subsets is either connected or empty.*
c) *Show that this fails dramatically in $\mathbb{R}^2$.*

EXERCISE 3.52. *Let $\{Y_n\}_{n=1}^{\infty}$ be a sequence of connected subsets of a topological space X. Show that if for all $n \in \mathbb{Z}^+$ we have $Y_n \cap Y_{n+1} \neq \varnothing$, then $\bigcup_{n=1}^{\infty} Y_n$ is connected.*

LEMMA 3.49. *(Caging Connected Sets) Let (U, V) be a separation of a topological space X and let $Y \subset X$ be connected. Then $Y \subset U$ or $Y \subset V$.*

PROOF. Let $f : X \to \{0,1\}$ be the continuous function corrresponding to (U, V): $f(U) = \{0\}$, $f(V) = \{1\}$. Since Y is connected, $f(Y)$ is connected, so $f(Y) \in \{0\}$ or $f(Y) \in \{1\}$. $\square$

THEOREM 3.50. *Connectedness is faithfully productive: if $\{X_i\}_{i \in I}$ is a family of nonempty spaces and $X = \prod_{i \in I} X_i$ endowed with the product topology, then X is connected iff X_i is connected for all $i \in I$.*

Proof. If X is connected, then for all $i \in I$ we have that $X_i = \pi_i(X)$ is the continuous image of a connected space, hence is connected. It remains to show the converse: if each X_i is connected then so is X. We do this in several steps.
Step 1: Suppose $\#I = 2$. In this case let us rename the factor spaces X and Y. Seeking a contradiction we let (U, V) be a separation of $X \times Y$, and let $(x_1, y_1) \in U$, $(x_2, y_2) \in V$. Then the subset

$$C = (\{x_1\} \times Y) \cup (X \times \{y_2\})$$

is a union of two connected subsets which intersect at (x_1, y_2) so C is a connected subset of $X \times Y$ containing (x_1, y_1) and (x_2, y_2), contradicting Lemma 3.49.
Step 2: The case in which I is finite follows by induction.
Step 3: Finally suppose that I is infinite. Choose a well-ordering on I, let $I^+ := I \cup \{\top\}$, where $\top$ is some element greater than every element in I, and let $X_\top = \{\bullet\}$ be any one point space. For each $i \in I^+$, choose $x_i \in X_i$; for $j \in I$, put

$$Z_j := \prod_{i<j} X_i \times \prod_{i \geq j} \{x_i\}.$$

Let J be the set of $j \in I$ such that Z_j is connected. It suffices to show that $J = I^+$, for then $Z_\top = X \times \{\bullet\}$ is connected and homeomorphic to X. By the Principle of Transfinite Induction, to get $J = I^+$ it is enough to show that:
(TI1) If $\perp$ is the bottom element of I, then $Z_\perp$ is connected.
(TI2) For all $i \in I$, if Z_i is connected, then Z_{i^+} is connected.
(TI3) For $i \in I$, if Z_j is connected for all $j < i$, then Z_i is connected.
Here we go:
(TI1) The space $Z_\perp = \prod_{i \in I} \{x_i\}$ is a one point space, hence connected.
(TI2) The space Z_{i^+} is homeomorphic to $\prod_{j<i} X_i \times X_i$. The space $\prod_{j<i} x_i$ is homeomorphic to Z_j, which is assumed to be connected, and the space X_{j^+} is also connected, so by Step 2 the space Z_{i^+} is connected.
(TI3) Let $i \in I$ and suppose that Z_j is connected for all $j < i$. By Proposition 3.48b), the subspace $Y_i := \bigcup_{j<i} Z_j$ is connected. We claim that Y_i is dense in Z_i, hence Z_i is connected by Proposition 3.47. If $i = j^+$ for some j (i.e., if the collection of all elements less than i has a top element), then $Y_i = Z_i$. Otherwise there is $K < i$ such that every nonempty open subset U of Z_i projects onto X_k for all $K \leq k < i$, while Y_i contains $\prod_{j \leq K} X_j \times \prod_{j > K} \{x_j\}$, so U meets Y_i. $\square$

12.2. Path Connectedness.

A topological space X is **path-connected** if it is nonempty and for all $x, y \in X$ there is a continuous map $\gamma : [0,1] \to X$ with $\gamma(0) = x$, $\gamma(1) = y$. We say that γ is a **path in X from x to y**.

Proposition 3.51. *A path-connected topological space is connected.*

Proof. Seeking a contradiction, let (U, V) be a separation of X, choose $x \in U$, $y \in V$, and let γ be a path in X from x to y. Then the presep $(\gamma^{-1}(U), \gamma^{-1}(V))$ is a separation of $[0,1]$ since $0 \in \gamma^{-1}(U)$ and $1 \in \gamma^{-1}(V)$: contradiction. $\square$

Exercise 3.53. *Let $n \geq 2$, and let $Y \subset \mathbb{R}^n$ be a countable subset. Show that $\mathbb{R}^n \setminus Y$ is path-connected.*

Exercise 3.54. *A point x in a topological space x is a* ***cut point*** *if X is connected but $X \setminus \{x\}$ is not.*

a) *Show that homeomorphic spaces have the same number of cut points. Thus the number of cut points is a* ***cardinal invariant*** *of a topological space.*
b) *Show that every point of* $\mathbb{R}$ *is a cut point.*
c) *Show that for* $n \geq 2$, $\mathbb{R}^n$ *has no cut points.*
d) *Deduce that for* $n \geq 2$, $\mathbb{R}^n$ *and* $\mathbb{R}$ *are not homeomorphic.*[13]

EXERCISE 3.55. *Let* $f : X \to Y$ *be a continuous function, and let* $\Gamma(f) = \{(x, f(x)) \in X \times Y\}$ *be the graph of* f*. Show:*
a) *The space* $\Gamma(f)$ *is homeomorphic to* X*.*
b) *In particular, the set* $\Gamma(f)$ *is connected iff* X *is connected.*

12.3. Components.

Let x be a point in the space X, and let $\{Y_i\}$ be the family of all connected subsets of X containing X. By Proposition 3.48a), the set

$$C(x) := \bigcup Y_i$$

is connected. Evidently $C(x)$ is the unique maximal connected set containing x. It is called the **connected component of x**. By Proposition 3.47, we have that $\overline{C(x)}$ is a connected set containing $C(x)$, so by maximality we deduce that $C(x)$ is closed. Let $x, y \in X$. If $C(x) \cap C(y) \neq \varnothing$ then Proposition 3.48a) applies to show that $C(x) \cup C(y)$ is a connected subset containing x and y. By maximality we have $C(x) = C(x) \cup C(y) = C(y)$. It follows that $\{C(x)\}_{x \in X}$ is a partition of X by closed subsets.

If for all $x \in X$ we have $C(x) = \{x\}$, we say that X is **totally disconnected**. Clearly a discrete space is totally disconnected; interestingly, there are totally disconnected spaces which are very far from being discrete.

We say points x, y in a space X **can be separated in X** if there is a separation (U, V) with $x \in U$, $y \in V$. If we had $C(x) = C(y)$, then $(U \cap C(x), V \cap C(y))$ would give a separation of $C(x)$, contradiction. So two points which lie in the same connected component cannot be separated in X.

Let X be a topological space. We consider the relation $\mathcal{R}$ on X defined by $x\mathcal{R}y$ if x and y cannot be separated in X; it is clearly reflexive and symmetric. Suppose $x\mathcal{R}y$ and $y\mathcal{R}z$ but that there is a separation (U, V) of X with $x \in U$ and $y \in V$. Then either z lies in U, in which case y and z can be separated, or z lies in V, in which case x and y can be separated: either way, a contradiction. It follows that $\mathcal{R}$ is an equivalence relation; we denote the $\mathcal{R}$-equivalence class of x by $C_Q(x)$ and call it the **quasi-component** of x in X. By the previous paragraph, we have

$$\forall x \in X, C(x) \subset C_Q(x).$$

EXERCISE 3.56. *Recall that a subset* Y *of a topological space* X *is* ***clopen*** *if it is both open and closed. Thus, a nonempy proper subset* Y *is clopen iff* $(Y, X \setminus Y)$ *is a separation of* X*.*
a) Let $x \in X$*. Show: the quasi-component* $C_Q(x)$ *is the intersection of all clopen*

[13] Of course we would like to show that if $m \neq n$ then $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic. This is true but – sadly enough – the proof lies beyond the scope of this course!

subsets Y of X containing x.
b) Deduce: for all $x \in X$, $C_Q(x)$ is closed. Show by example that $C_Q(x)$ need not be open.

Exercise 3.57. *Let X be the following subspace of $\mathbb{R}^2$:*

$$\{(0,0),(0,1)\} \cup \bigcup_{n \in \mathbb{Z}^+} \{(1/n, y) \mid y \in [0,1]\}.$$

Show that $C((0,0)) = \{(0,0)\}$ and $C_Q((0,0)) = \{(0,0),(0,1)\}$.

13. Local Compactness and Local Connectedness

13.1. On Properties.

By a **property** P of a topological space, we really mean a subclass P of the class Top of all topological spaces, but rather than saying $X \in P$, we say that X has the property P. In practice this is of course only natural: for instance if P is the class of all compact topological spaces, then rather than say "X lies in the class of all compact topological spaces" we will say "X has the property of compactness" (of course for many purposes it would be better still to say "X is compact").

A property of P of topological spaces is a **topological property** if whenever a topological space X has that property, so does every topological space Y which is homeomorphic to X. Really this just formalizes what is good sense: topology is by definition the study of topological properties of topological spaces. Thus for instance for a space Y, the property "X is a subspace of Y" is not a topological property: for instance assuming that by S^1 we mean precisely the unit circle in $\mathbb{R}^2$, then $\mathbb{R}$ is not a subspace of S^1. However $\mathbb{R}$ is homeomorphic to a subspace of S^1: indeed, removing any point from S^1 we get a homeomorphic copy of $\mathbb{R}$.

The above example shows that any property of topological spaces which is not itself topological can be made so simply by replacing "is" with "is homeomorphic to". In the above case, the topologization (?!?) of the property "X is a subspace of Y" (for fixed Y, say) is "X can be embedded in Y". When someone speaks or writes about a property of topological spaces that is not topological (or not manifestly topological), it is likely that they really mean the topologization of that property. For instance, anyone who asks "Which topological spaces are subspaces of compact spaces?" must surely mean "Which topological spaces are homeomorphic to a subspace of a compact space – i.e., can be embedded in a compact space?" The latter is a great question, by the way. We will answer it later on.

Let P be a topological property. We define a new property, **locally P**, as follows: for a point x in a topological space X, we say that X **is locally** P **at the point** x if there is a neighborhood base at P consisting of sets all of which have property P. We say that X is **locally P** if it is locally P at each of its points. For instance:

Exercise 3.58. *Show that a metrizable space is locally infinite iff it has no isolated points.*

Let P be a topological property. We define a new property, **weakly locally P**, as follows: for a point x in a topological space X, we say that X is **weakly locally** P **at the point** x is there is a neighborhood N of x that has property P. We say that X is **weakly locally** P if it is weakly locally P at each of its points.

From this it follows that any space X that is P is also weakly locally P: X is a neighborhood of every point that has property P. On the other hand, a space may be P but not locally P. We will see examples of spaces that are connected but not locally connected and quasi-compact but not locally quasi-compact. (Beware: we will define "locally compact" to be Hausdorff and locally quasi-compact. This contradicts the above convention because a locally compact Hausdorff space is required to be Hausdorff, not just locally Hausdorff. The line with two origins $\bullet\ell\bullet$ admits a neighborhood base of compact neighboorhods at every point but is not Hausdorff so is not locally compact.)

13.2. Local Compactness.

The only problem with compactness is that it can be too much to ask for: even the real numbers are not compact. However, every closed bounded interval in $\mathbb{R}$ is compact. The goal of this section is to formalize and study the desirable property *of the real numbers* corresponding to the compactness of closed, bounded intervals.

Let X be a topological space, and let p be a point of X. In line with our above conventions about localization of topological properties, we say that X is **weakly locally compact at** p if there is a compact neighborhood C of p. Let us spell that out more explicitly: C is a compact subset of X and p lies in the interior of p. A topological space is **weakly locally compact** if it is Hausdorff and weakly locally compact at every point.

A topological space is **locally compact at** p if there is a neighborhood base $\{C_i\}_{i\in I}$ at p with each C_i compact. Again we spell it out: this means that for every neighborhood N of p, there is $i \in I$ such that

$$p \in C_i \subset N.$$

A topological space is **locally compact** if it is Hausdorff and locally compact at each of its points.

Note that at the cost of preserving one terminological convention – that when a topological property is most important and useful in the presence of the Hausdorff axiom, we give the cleaner name to the version of the property that includes the Hausdorff axiom – we are breaking another.

EXERCISE 3.59. *Show that the line with two origins is locally compact at each of its points but is not locally compact.*

EXAMPLE 3.24. *Let $y \in \mathbb{R}$. Then for all $x, z \in \mathbb{R}$ with $x < y < z$, we have that $[x, z]$ is a compact neighborhood of y. Thus $\mathbb{R}$ is weakly locally compact. Moreover, if N is any neighborhood of y then for some $\epsilon > 0$ we have*

$$[y - \frac{\epsilon}{2}, y + \frac{\epsilon}{2}] \subset (y - \epsilon, y + \epsilon) \subset N.$$

This shows that – as promised – $\mathbb{R}$ is locally compact.

PROPOSITION 3.52. a) *Let X be a Hausdorff space, and let $p \in X$. If X is weakly locally compact at p, then X is locally compact at p.*
b) *A Hausdorff space in which each point admits a compact neighborhood is locally compact.*
c) *Compact spaces are locally compact.*

PROOF. a) Let C be a compact neighborhood of p. Given a neighborhood U of p, our task is to produce a compact set K with

$$p \in K^\circ \subset K \subset U.$$

Notice that if we can complete our task with the open neighborhood U° we can certainly do it with U, so we may assume that U is open. Then $A = C \setminus U$ is closed in the compact space C so is compact. By Lemma 3.33 there are disjoint open subsets W_1 containing p and W_2 containing A. Then $V = W_1 \cap C^\circ$ is an open neighborhood of p disjoint from $A = C \setminus U$ and thus contained in U. Since X is Hausdorff, C is closed, and thus $\overline{V} \subset C$ is compact. Because $V \subset W_1$ and $W_1 \cap W_2 = \varnothing$, we have

$$\overline{V} \cap (C \setminus U) = \overline{V} \cap A \subset \overline{V} \cap W_2 = \varnothing$$

and thus

$$p \in V^\circ \subset (\overline{V})^\circ \overline{V} \subset U.$$

So $K = \overline{V}$ does the job.
b) This follows immediately.
c) So does this: if X is compact, then it is Hausdorff and for all $p \in X$, X is a compact neighborhood of p. □

EXERCISE 3.60. *For a subset Y of a topological space X, the following are equivalent:*

(i) *For all $p \in Y$, there is an open neighborhood U_p of p in X such that $U_p \cap Y$ is closed in U_p.*
(ii) *There is an open subset $U \subset X$ and a closed subset $A \subset X$ with $Y = U \cap A$.*
(iii) *Viewed as a subspace of $\overline{Y}$, Y is open.*

*A subset satisfying these equivalent conditions is called **locally closed**. (The term applies most sensibly to the first condition.)*

EXERCISE 3.61. a) *Show: a finite intersection of locally closed sets is locally closed.*
b) *Show: the complement of a locally closed set need not be locally closed.*
c) *Let X be a topological space, and let $\mathcal{A}$ be the algebra of sets generated by the topology τ_X: that is, $\tau_X \subset \mathcal{A}$, $\mathcal{A}$ is closed under finite union, finite intersection and taking complements, and $\mathcal{A}$ is the minimal family of sets satisfying these two properties. We say that the elements of $\mathcal{A}$ are **constructible sets**. Show: a subset $Y \subset X$ is constructible iff it is a finite union of locally closed sets.*

PROPOSITION 3.53. *If Y_1 and Y_2 are locally compact subspaces of a topological space X, then $Y_1 \cap Y_2$ is locally compact.*

Proof. Since Y_1 is Hausdorff and $Y_1 \subset Y_1 \cap Y_2$, it follows that $Y_1 \cap Y_2$ is Hausdorff. Now let $p \in Y_1 \cap Y_2$, let K_1 be a compact neighborhood of p in Y_1 and let K_2 be a compact neighborhood of p in Y_2. Then $K_1 \cap K_2$ is compact. Moreover write $K_1^\circ = U_1 \cap Y_1$ and $K_2^\circ = U_2 \cap Y_2$ with U_1, U_2 open in X. Then

$$p \in K_1^\circ \cap K_2^\circ = (U_1 \cap U_2) \cap (Y_1 \cap Y_2),$$

so $K_1 \cap K_2$ is a neighborhood of p in $Y_1 \cap Y_2$. □

Proposition 3.54. *Let X be a locally compact topological space.*

a) *If $Y \subset X$ is open, then Y is locally compact.*
b) *If $Y \subset X$ is closed, then Y is locally compact.*
c) *For $Y \subset X$, the following are equivalent:*
   (i) *The subset Y is locally closed.*
   (ii) *The space Y is locally compact.*

Proof. a) For any topological property P, an open subspace of a locally P space is locally P. Local compactness is not quite defined this way, so we also need to mention that a subspace of a Hausdorff space is Hausdorff.
b) If Y is closed and C_p is a compact neighborhood of p in X, then $C_p \cap Y$ a neighborhood of p in Y which is closed in C_p hence compact.
c) (i) $\implies$ (ii): Write $Y = U \cap A$ with U open, A closed, and apply a), b) and c).
(ii) $\implies$ (i): (A. Fischer) For $p \in Y$, let V_p be an open neighborhood of p in Y whose closure in Y, say $\mathrm{cl}_Y(V_p)$, is compact. We have

$$\mathrm{cl}_Y(V_p) = \overline{V_p} \cap Y.$$

Then $\overline{V_p} \cap Y$ is compact, hence closed, in X. Moreover, there is an open neighborhood W_p of p in X such that $V_p = W_p \cap Y$. Then

$$\overline{W_p \cap Y} \cap Y = \mathrm{cl}_Y(V_p).$$

Since W_p is open, we have

$$W_p \cap \overline{Y} = W_p \cap Y.$$

Since $W_p \cap Y \subset \overline{W_p \cap Y} \cap Y$, we have

$$p \in W_p \cap \overline{Y} \subset \overline{W_p \cap \overline{Y}} = \overline{W_p \cap Y} \subset \overline{\overline{W_p \cap Y} \cap Y} = \overline{W_p \cap Y} \cap Y \subset Y.$$

Then $U = \bigcup_{y \in Y} W_p$ is open in X and

$$U \cap \overline{Y} = \bigcup_{y \in Y} W_p \cap \overline{Y} = Y.$$

So Y is locally closed. □

13.3. The Alexandroff Extension.

A **compactification** of a topological space is X is an embedding $\iota : X \to C$ into a compact space C with dense image: $\overline{\iota(X)} = C$. Given any embedding $\iota : X \to C$ into a compact space, we get a compactification by replacing C with $\overline{\iota(C)}$.

Exercise 3.62. *Suppose X is compact. Show: a map $\iota : X \to C$ is a compactification iff it is a homeomorphism.*

Compactifications are a high point of general topology: beautiful, rich and useful.

We will not pursue the general theory just yet but rather one simple, important case. However we will introduce one piece of terminology: if $\iota : X \to C$ is compactification, the **remainder** is $C \setminus \iota(X)$. In other words, the remainder is what we have to add to X in order to compactify it. Here we are interested in the case in which the remainder consists of a single point.

Example 3.25. *Let $n \in \mathbb{Z}^+$. By removing the point $\infty = (0, \ldots, 0, 1)$ from the n-sphere $S^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \mid x_1^2 + \ldots + x_{n+1}^2 = 1\}$, we get a space which is homeomorphic to $\mathbb{R}^n$. (There is an especially nice way of doing this, called the* ***stereographic projection****. We leave it to the reader to look into this. If you don't care about having such a nice map, it is much easier to construct a homeomorphism.) Thus we get an embedding $\iota : \mathbb{R}^n \to S^n$ with* ***one-point remainder****: $S^n \setminus \iota(\mathbb{R}^n) = \{\infty\}$.*

The goal of this section is to view the above example in reverse: i.e., to figure out how, rather than removing a point from S^n to get $\mathbb{R}^n$, to start with $\mathbb{R}^n$ and "add the point at ∞" intrinsically in terms of the topology of $\mathbb{R}^n$.

Let $\iota : X \to C$ be a compactification with one-point remainder – or, as one often says, a **one-point compactification** – say $\{\infty\} = C \setminus \iota(X)$. Since C is Hausdorff, $\{\infty\}$ is closed and thus $\iota(X)$ is open. Being open in a compact space, $\iota(X)$ is locally compact; since $\iota : X \to \iota(X)$ is a homeomorphism, it follows that X is locally compact. Moreover, if X were compact then $\iota(X)$ is a proper closed subset of C, so it cannot be dense. We've shown the following result.

Proposition 3.55. *If a topological space X admits a compactification with one-point remainder, it is locally compact but not compact.*

Rather remarkably, Proposition 3.55 has a converse: if X is locally compact and not compact, it has a compactification with one-point remainder. This was shown by Alexandroff. To see how to do it, let's cheat and run it backwards by contemplating a compactification $\iota : X \to C$ with one-point remainder $\{\infty\}$ until we can understand how to build C out of X. Because the image $\iota(X)$ is open in C, the embedding $\iota : X \to C$ is an open map: this means that for every $p \in X$, a neighborhood base at p is given by $\{\iota(N) \mid N \text{ is a neighborhood of } p \text{ in } X\}$.

If we know a neighborhood base at each point then we know the topology, so the remaining task is to find a neighborhood base of the point ∞. If N is an open neighborhood of ∞ in C, then $C \setminus N$ is on the one hand closed in the compact space C hence compact, and on the other hand a subset of X, hence a compact subset of X by the intrinsic nature of compactness. Conversely, if $K \subset X$ is compact, then $\iota(K)$ is compact in C, hence closed, and thus $C \setminus K$ is an open neighborhood of ∞. It follows that the open neighborhoods of ∞ in C are precisely the sets of the form $C \setminus K = (X \setminus K) \cup \{\infty\}$ for K compact in X.

Exercise 3.63. *Let X be a topological space, and let $\iota_1 : X \to C_1$ and $\iota_2 : X \to C_2$ be compactifications with one-point remainders ∞_1 and ∞_2. Show: there is a unique homeomorphism $\Phi : C_1 \to C_2$ such that*

$$\iota_2 = \Phi \circ \iota_1.$$

Our cheating has paid off: we now have enough information to construct a one-point compactification of any locally compact space. In fact the construction is meaningful on any topological space, and though we know it can't yield a compactification unless X is locally compact and not compact, nevertheless it is of some interest, so following Alexandroff we phrase it in that level of generality.

Let X be a topological space, and let ∞ be anything which is *not* an element of X. Let $X^* = X \cup \{\infty\}$ and let $\iota : X \to X^*$ be the inclusion map. We endow X^* with the following **A-topology**:[14] the open subsets consist of all open subsets U of X together with all subsets of the form $X^* \setminus K = (X \setminus K) \cup \{\infty\}$ for K a closed quasi-compact subset of X.

EXERCISE 3.64. a) *Show: the A-topology on X^* is a topology.*
b) *The A-topology on X^* is quasi-compact.*

THEOREM 3.56. *Let X be a topological space, let $X^* = X \cup \{\infty\}$ endowed with the Alexandroff topology, and let $\iota : X \to X^*$ be the inclusion map. Then:*

a) *The map $\iota : X \to X^*$ is an open embedding, called the* ***Alexandroff extension****.*
b) *The following are equivalent:*
   (i) *The set $\{\infty\}$ is open in X.*
   (ii) *The space X is not dense in X^*.*
   (iii) *The space X is quasi-compact.*
c) *The following are equivalent:*
   (i) *The space X^* is Hausdorff.*
   (ii) *The space X^* is compact.*
   (iii) *The space X is locally compact.*

PROOF. a) More generally, if $f : A \hookrightarrow B$ is an injective map of topological spaces such that for all $U \subset A$ we have that U is open iff $f(U)$ is open, then f is an open embedding: $f : A \to f(A)$ is a bijection and is open with open inverse, so is a homeomorphism.
b) (i) $\iff$ (ii): Since the subset X of X^* has the one point complement $\{\infty\}$, its closure is either itself, in which case it is closed, or X^*, in which case it is dense. The former happens iff the complement $\{\infty\}$ is open.
(ii) $\iff$ (iii): By what we have just seen, X is not dense in X^* iff $\{\infty\}$ is open. By definition of the topology on X^* we have that $\{\infty\}$ is open iff its complement X is quasi-compact.
c) (i) $\iff$ (ii) follows from Exercise 3.64, since X^* is always quasi-compact.
(i) $\implies$ (iii): Suppose X^* is Hausdorff. Then its subspace X is also Hausdorff, so by Proposition 3.52, to show that X is locally compact it is sufficient to show that every point admits a compact neighborhood. Let $x \in X$. Since X^* is Hausdorff we have open neighborhoods U of x and V of ∞ such that $U \cap V = \varnothing$. Since U does not contain ∞ it is an open neibhorhood of x in X. Since V is an open neighborhood of ∞ we have that $X^* \setminus V = X \setminus V$ is quasi-compact, hence compact because X^* is Hausdorff, and contains U, so V is a compact neighborhood of x.
(iii) $\implies$ (i): If X is locally compact then X is Hausdorff and open in X^*, so any two points of X^* that both lie in X can be separated by open neighborhoods. It

[14] A is for Alexandroff. But later "Alexandroff topology" will mean something else!

remains to consider $x \in X$ and the point ∞. Since X is locally compact, there is a compact neighborhood K of x, and then K° is an open neighborhood of x, $X^* \setminus K$ is an open neighborhood of ∞ and $K^\circ \cap (X^* \setminus K) = \varnothing$. $\square$

In particular the Alexandroff extension gives a "quasi-compactification" of any space that is not itself quasi-compact. (However, quasi-compactifications are much less well-behaved and ultimately less interesting than compactifications.)

13.4. Further Exercises. A Hasudorff space X is **σ-compact** if it is covered by a sequence of compact subsets. A **compact exhaustion** in a Hausdorff space X is a sequence of compact subsets $\{K_n\}_{n=1}^\infty$ such that $K_n \subset K_{n+1}^\circ$ for all $n \in \mathbb{Z}^+$ and $X = \bigcup_{n=1}^\infty K_n$. A Hausdroff spcae X is **hemicompact** if there is a sequence $\{K_n\}_{n=1}^\infty$ of compact subsets of X such that every compact subset is contained in some K_n.

Exercise 3.65. *Let X be a Hausdorff space.*

a) *Show: if X admits a compact exhaustion, it is σ-compact.*
b) *Show: if X is second countable and locally compact, then it admits a compact exhaustion.*
c) *Show: if X is second countable and σ-compact, then it admits a countable base of sets with compact closure and a compact exhaustion.*

Exercise 3.66. *Let X be a Hausdorff space.*

a) *Show: if X is σ-compact, then X is Lindelöf.*
b) *Show: if X is locally compact and Lindelöf, then X is hemicompact.*
c) *Show: the discrete topology on an uncountable set is locally compact, not Lindelöf and not σ-compact.*

Exercise 3.67. *Let X be a Hausdorff space.*

a) *Show: if X is hemicompact, then it is σ-compact.*
b) *Show: $\mathbb{Q}$ (with its usual Euclidean topology) is σ-compact but not hemicompact.*

Exercise 3.68. *Show that for a Hausdorff space X, the following are equivalent:*

(i) *The space X is hemicompact.*
(ii) *The Alexandroff extension $X^* = X \coprod \{\infty\}$ is first countable at ∞.*

In some circles – in my experience this is most common for mathematics written in French and then translated into English – one uses "countable at infinity" as a synonym for "σ-compact." The previous exercises show that this terminology is reasonable when dealing with locally compact spaces – such a space is σ-compact iff X^* is first countable ∞ – but otherwise seems to be potentially misleading.

13.5. Local Connectedness.

Example 3.26. *(Topologist's Sine Curve) Let Y be the subspace $\{x, \sin\frac{1}{x}) \mid x \in (0,1]\}$ of $\mathbb{R}^2$ and let X be its closure. X consists of Y together with all points $(0,y)$ with $y \in [-1,1]$. Then:*

(i) *Y is path-connected: indeed, it is the graph of the continuous function $\sin\frac{1}{x} : (0,1] \to \mathbb{R}$ so it is homeomorphic to $(0,1]$.*
(ii) *X is compact (Heine-Borel).*

(iii) *X, being the closure of a connected set, is connected.*
(iv) *X is not path-connected.*
(v) *X is not locally connected.*

This example shows first that unlike connectedness, the closure of a path-connected subset need not be path-connected, and second that a connected space need not be locally connected, even if it is compact. In particular, weakly locally connected does not imply locally connected.

EXERCISE 3.69. *Prove assertions (iv) and (v) of the preceding example.*

PROPOSITION 3.57. *For a topological space X, the following are equivalent:*

(i) *The space X is (homeomorphic to) the coproduct of its components.*
(ii) *Every component of X is open.*
(iii) *The space X is weakly locally connected.*

PROOF. Since every space is the disjoint union of its connected components, (i) $\iff$ (ii) is immediate, as is (ii) $\implies$ (iii): if the component $C(p)$ of p is open, it provides a connected neighborhood of p.
(iii) $\implies$ (ii): Let $p \in X$, let $C(p)$ be the component of p. Let $q \in C(p)$; since the components partition X we have $C(p) = C(q)$. By weak local connectedness, q has a connected neighborhood N and thus

$$q \in N \subset C(q) = C(p),$$

so $q \in C(p)^\circ$. It follows that $C(p)$ is open. □

PROPOSITION 3.58. *For a topological space X, the following are equivalent:*

(i) *The space X is locally connected.*
(ii) *For every open subset U of X and every point p of U, the connected component of p in U is open (in U or equivalently in X).*

PROOF. (i) $\implies$ (ii): Let $U \subset X$ be open, and let C be a component of U. If $p \in C$, there is a connected neighborhood V of p which is contained in U. Since C is the maximal connected subset of U which contains p we must have $V \subset C$, hence $p \in C^\circ$. Since p was arbitrary, C is open.
(ii) $\implies$ (i): For $p \in X$, let N be a neighborhood of p. The component C of p in N° is open, so $C \subset N$ and C is a connected neighborhood of p. □

EXERCISE 3.70. *Show that in a locally connected space, every point admits a neighborhood base of connected* open *neighborhoods.*

PROPOSITION 3.59. *Let I be a nonempty set, for each $i \in I$, let X_i be a nonempty topological space, and let $X := \prod_{i \in I} X_i$, endowed with the product topology. Then the following are equivalent:*

(i) *The product space X is locally connected.*
(ii) *Each factor X_i is locally connected, and all but finitely many factors are connected.*

PROOF. (i) $\implies$ (ii): Let $i \in I$, let $x_i \in X_i$, and let U_i be an open neighborhood of U_i. For each $j \neq i$, choose $x_j \in X_j$, and let $x = (x_i) \in X$. Then $N_x := U_i \times \prod_{j \neq i} X_i$ is an open neighborhood of x in X, so there is a connected open subset U of X such that $x \in U \subset N_x$. Then $x_i = \pi_i(x) \in \pi_i(U) \subset \pi_i(N_x) = U_i$ and $\pi_i(U)$ is connected (since U is connected and π_i is continuous) and open (since

U is open and π_i is an open map). This shows that X_i is locally connected. Now let $U \subset X$ be any connected nonempty open set (which certainly exists, since X is locally connected. Then for all but finitely many $i \in I$ we have $\pi_i(U) = X_i$ and thus X_i is connected.
(ii) $\implies$ (i): Since each X_i is locally connected, there is a base $\mathcal{B}_i$ for X_i consisting of connected open sets. Then the family of subsets

$$\left\{ \prod_{i \in I} U_i \mid U_i \in \mathcal{B}_i \text{ for all } i \in I \text{ and } U_i = X_i \text{ for all but finitely many } i \in I \right\}$$

is a base for the topology of X consisting of connected sets (the latter is by Theorem 3.50). So X is locally connected. □

14. Further Exercises

EXERCISE 3.71. *Let ω_1 be the least uncountable ordinal. In particular ω_1 is an ordered set, so give it the order topology. Show that ω_1 is:*

a) *sequentially compact but not compact.*
b) *pseudocompact.*
c) *first countable but not separable.*
d) *countably compact and not Lindelöf.*
e) *not metrizable.*

EXERCISE 3.72. *Let X be an ordered set, and let $f : X \to X$ be continuous for the order topology. Observe that the statement of Theorem 2.106 is meaningful in this context.*

a) *Give an example in which X is (nonempty!) and complete but the result fails: there is no fixed point, there is $x_1 \in X$ with $f(x_1) > x_1$ and $x_2 \in X$ with $f(x_2) < x_2$.*
b) *Suppose X is Dedekind complete and densely ordered. Does the conclusion of Theorem 2.106 continue to hold?*

We say a topological space X is **halveable** if it is homeomorphic to $X \coprod X$. The space $X \coprod X$ may be identified with $X \times \{1, 2\}$, where we take the discrete topology on $\{1, 2\}$ – an element x in the first copy of X gets mapped to $(x, 1)$ and an element in the second copy of X gets mapped to $(x, 2)$. A topological space is **strongly halveable** if for all $p \neq q$ in X there is a homeomorphism $f : X \to X \times \{1, 2\}$ such that $\pi_2(f(p)) = 1$ and $\pi_2(f(q)) = 2$.

A topological space X is **homogeneous** if for all $p \neq q$ in X, there is a homeomorphism $f : X \to X$ such that $f(p) = q$. A topological space X is **doubly homogeneous** if for all $(p_1, p_2), (q_1, q_2) \in X \times X$ with $p_1 \neq p_2$ and $q_1 \neq q_2$, then there is a homeomorphism $f : X \to X$ such that $f(p_1) = q_1$ and $f(p_2) = q_2$.[15]

EXERCISE 3.73. *Show that all of the following topological spaces are doubly homogeneous:*

a) *Any discrete space.*
b) *The space $\mathbb{R}^N$ for any $N \in \mathbb{N}$.*

[15] Let $\mathrm{Aut}(X)$ be the group of homeomorphisms $f : X \to X$. Then X is homogeneous if $\mathrm{Aut}(X)$ acts transitively on X, and similarly X is doubly homogeneous if $\mathrm{Aut}(X)$ acts doubly transitively on X.

c) *The Cantor set* C.
d) *The rational numbers* $\mathbb{Q}$.
e) *The irrational numbers* $\mathbb{R} \setminus \mathbb{Q}$.

EXERCISE 3.74. a) *Show: if a nonempty topological space is halveable, then it has infinitely many connected components.*
b) *Exhibit a topological space with infinitely many connected components that is not halveable.*

EXERCISE 3.75. a) *Show that for a discrete space* X*, the following are equivalent:*
(i) *The space* X *is strongly halveable.*
(ii) *The space* X *is halveable.*
(iii) *The set* X *is infinite.*
b) *Show: the Cantor set is strongly halveable.*
c) *Show:* $\mathbb{Q}$ *and* $\mathbb{R} \setminus \mathbb{Q}$ *(in the subspace topology from* $\mathbb{R}$*) are both strongly halveable.*

EXERCISE 3.76. *Show that for a weakly locally connected topological* X*, the following are equivalent:*
(i) *If a topological space* C *is homeomorphic to one of the connected components of* X*, then it is homeomorphic to infinitely many of the connected components of* X.
(ii) *The space* X *is halveable.*

EXERCISE 3.77. *Suppose* X *is a halveable topological topological space.*
a) *Show: for all* $2 \leq n < \aleph_0$*, the space* X *is homeomorphic to* $\coprod_{i=1}^{n} X$.
b) *Suppose that* X *is weakly locally connected. Show that* X *is homeomorphic* $\coprod_{n=1}^{\infty} X$.
c) *Suppose that* X *is countably infinite and strongly halveable. Show that* X *is homeomorphic to* $\coprod_{n=1}^{\infty} X$.
d) *Is there a halveable space that is not homeomorphic to* $\coprod_{n=1}^{\infty} X$*?*

CHAPTER 4

Baire Spaces

1. Introduction

A subset Y of a topological space is **nowhere dense** if its closure has empty interior: $\overline{Y}^{\circ} = \varnothing$. In §2.8.4 we gave this definition for metric spaces. A set is **somewhere dense** it is is not nowhere dense.

Exercise 4.1. *Let Y be a subset of a topological space X. Show: the following are eqivalent:*

(i) *The subset Y is nowhere dense in X.*
(ii) *The subset $X \setminus \overline{Y}$ is dense in X.*
(iii) *For each nonempty open subset U of X, there is a nonempty open subset $V \subset U$ such that $V \cap Y = \varnothing$.*

Exercise 4.2.

a) *Show: a subset of a nowhere dense subset is nowhere dense.*
b) *Show: if there are nowhere dense subsets $\{Y_i\}_{i \in I}$ that cover a subset A of a topological space, then there are nowhere dense subsets $\{Z_i\}_{i \in I}$ (same index set!) such that $A = \bigcup_{i \in I} Z_i$.*

Exercise 4.3. a) *Show: if Y is an open subset of X and $A \subset X$ is nowhere dense, then $A \cap Y$ is nowhere dense in Y.*
b) *Show: part a) may fail if the word "open" is removed.*

In §2.8.4 we stated and proved a theorem of Baire (Theorem 2.58): in a complete metric space, no nonempty open subset is a countable union of nowhere dense subsets. Many mathematical theorems have the following form: each of a certain class $\mathcal{C}$ of mathematical objects has a certain desirable property $\mathcal{P}$. When the property $\mathcal{P}$ is known to be useful, it often turns out to be useful to think of $\mathcal{P}$ itself as defining a class of objects (indeed there is no distinction to be made between a class in a property – for any property, one can consider the class of objects that satisfy it, while for any class, membership in the class counts as a property). The theorem then says that $\mathcal{C} \subset \mathcal{P}$ and it can be interesting to ask for more information about $\mathcal{P}$: what other classes does it contain, is it contained in, is it disjoint from, and so forth?

To perform the above process in the setting of topological spaces we define a topological space to be a **Baire space** if no nonempty open subset is contained in a countable union of nowhere dense subsets, and Baire's theorem becomes that a completely metrizable topological space is a Baire space.

Exercise 4.4. *Show: for a topological space X, the following are equivalent:*

(i) *X is a Baire space: no nonempty open subset is contained in a countable union of nowhere dense subsets.*

(i)′ *No nonempty open subset of X is a countable union of nowhere dense subsets.*

(ii) *For every sequence $\{U_n\}_{n=1}^{\infty}$ of dense open subsets of X, their intersection $\bigcap_{n=1}^{\infty} U_n$ is also a dense subset of X.*

(iii) *If $Y \subset X$ is a countable union of nowhere dense subsets, then $X \setminus Y$ is dense in X.*

In many applications of Theorem 2.58 it is sufficient to use that a nonempty complete metric space is not a countable union of nowhere dense subsets. This makes it subtly hard to remember what the Baire property actually is: it is stronger than that: in a nonempty Baire space a countable union of nowhere dense subsets is not only proper but has empty interior. The following terminology will help us maintain this distinction and have other uses as well.

A subset of a topological space is **meager** if it is a countable union of nowhere dense subsets and otherwise **nonmeager**. A subset of a topological space is **instrinsically meager** if it is meager as a subset of itself in the subspace topology. (For instance a finite subset of $\mathbb{R}$ is meager but not intrinsically meager.) A subset Y of a topological space X is **comeager** if $X \setminus Y$ is meager.

Thus a Baire space is precisely a space in which no nonempty open subset is meager.

Exercise 4.5. a) *Show: a subset of a meager subset is meager.*
b) *Show: a countable union of meager subsets is meager.*

Exercise 4.5 says that in a nonmeager topological space it makes sense to view the meager subsets as being "small." In particular it follows that if X is nonmeager, then for all subsets Y of X, at least one of Y and $X \setminus Y$ is nonmeager, since otherwise X would be the union of two meager subsets and thus meager. It may however be that both Y and $X \setminus Y$ are nonmeager: for instance, in a Baire space, any subset Y with nonempty interior is nonmeager, and $X \setminus Y$ may also have nonempty interior: in $\mathbb{R}$, take Y to be any proper interval consisting of more than a single point.

It is important to point out that a topological space can be neither meager nor Baire. After we develop the theory a bit more such examples will come easily: cf. Exercise 4.12, but the reader may wish to work out the following case now.

Exercise 4.6. *Show: the space* $([0,1] \cap \mathbb{Q}) \cup [2,3]$ *is neither meager nor Baire.*

2. Baire's Theorem II

The main result we know about Baire spaces is that complete metric spaces are Baire. Completeness is a metric, not a topological property. Any property can be made topological just by adding "homeomorphic to": we say a topological space is **completely metrizable** if it is homeomorphic to a complete metric space. Thus Theorem 2.58 immediately tells us that completely metrizable spaces are Baire. This raises the question of which topological spaces are completely metrizable. We will answer this later on – interestingly, the answer is somehow more satisfactory than the known answer(s) to the question of which topological spaces are metrizable. For now we look to see what other spaces this suggests may be Baire.

Every compact metric space is complete and thus Baire, so in other words any compact metrizable space is Baire. What if we drop metrizability? In fact a very similar argument to that of the proof of Theorem 2.58 shows this and more.

THEOREM 4.1 (Baire II). *Every locally compact space is a Baire space.*

PROOF. By Exercise 4.4, it is enough to show: if $\{U_n\}_{n=1}^\infty$ is a sequence of dense open subsets of X, then $Y := \bigcap_{n=1}^\infty U_n$ remains dense. Thus, for any nonempty open subset V_0 of X, we must find $x \in Y \cap V_0$.

We call a subset of X **nice** it is is compact (hence closed, since X is Hausdorff) with nonempty interior. Since X is locally compact, every nonempty open subset contains a nice subset. In particular, V_0 contains a nice subset K_0. Let us put

$$V_1 := U_1 \cap K_0^\circ.$$

Since U_1 is open and dense in K_0° is nonempty and open, also V_1 is nonempty and open. We let K_1 be a nice open subset contained in V_1. And we continue to proceed in this way: inductively, having defined a nonempty open subset V_n of U_n and a nice subset K_n of V_n, we put

$$V_{n+1} := U_{n+1} \cap K_n^\circ,$$

so V_{n+1} is a nonempty open subset of U_{n+1}, and then we take K_{n+1} to be any nice subset of V_{n+1}. Since for all $n \geq 0$ we have

$$K_{n+1} \subset V_{n+1} \subset K_n,$$

and thus

$$A := \bigcap_{n=0}^\infty V_n = \bigcap_{n=0}^\infty K_n.$$

Since $\{K_n\}_{n=0}^\infty$ is a nested family of nonempty closed subsets of the compact space K_0, the space A is nonempty, and thus there is

$$p \in \bigcap_{n=0}^\infty V_n \subset Y \cap V_0,$$

completing the proof. □

The following exercise shows that quasi-compactness is not enough.

EXERCISE 4.7. *Let X be a countably infinite set endowed with the cofinite topology: the nonempty open sets are those with finite complement. Show: X is not a Baire space.*

It is interesting that the proof of Baireness for *locally* compact spaces is no harder than the proof for compact spaces. This a clue to a later result: in fact Baireness is a local property.

3. Meager Subsets and Baire Subspaces

We turn now to the problems of characterizing the meager subsets of a Baire space and the Baire subspaces of a Baire space. This turns out to be surprisingly rich, even for the real line $\mathbb{R}$. What can we say about the Baire subspaces of $\mathbb{R}$? Since a closed subset of a complete metric space remains complete in the induced metric, by Theorem 2.58 every closed subspace of $\mathbb{R}$ is a Baire space. Are these all the Baire subspaces of $\mathbb{R}$?

Indeed not. Consider for instance $(0, 1)$. This is not closed, so is not complete in its standard metric. However, it is homeomorphic to $\mathbb{R}$ via the map $x \mapsto \tan(\pi x - \frac{\pi}{2})$, so it is still completely metrizable and thus still a Baire space.

Exercise 4.8. *Show that every interval in $\mathbb{R}$ is completely metrizable and thus is a Baire space.*

Once again this raises the question of precisely which subsets of $\mathbb{R}$ are completely metrizable, and once again we defer this to a later discussion. So perhaps the the next order of business is to find a subspace of $\mathbb{R}$ that is *not* a Baire space. Will a finite subset do for this? No, because such subsets are discrete, and in a discrete topological space the only nowhere dense subset is the empty set. So discrete spaces are Baire. So we need a subset of $\mathbb{R}$ that is neither discrete nor closed nor (it seems from the examples, at least: we will nail this down later) open. The first example that comes to mind is $\mathbb{Q}$: writing $\mathbb{Q}$ as the union of $\{x\}$ for $x \in \mathbb{Q}$ shows that $\mathbb{Q}$ is intrinsically meager and thus not Baire. (The same argument shows that $\mathbb{Q}$ is a meager subset of $\mathbb{R}$.)

The following simple result generalizes both the observation that $\mathbb{Q}$ is intrinsically meager and Exercise 4.7.

Exercise 4.9. *Show: if X is a countably infinite, separated topological space without isolated points, then X is meager.*

Theorem 4.2. a) *The property of being a Baire space is open-hereditary: every open subspace of a Baire space is Baire.*

b) *For a topological space X, the following are equivalent:*
   (i) *The space X is Baire.*
   (ii) *Every point of X has a neighborhood base of of open Baire neighborhoods.*
   (iii) *Every point of X has an open Baire neighborhood.*

c) *If X is meager, then every open subset of X is intrinsically meager.*

Proof. a) Let X be a Baire space, and let U be an open subset of X. Since the empty space is Baire, we may assume that $U \neq \varnothing$. For a subspace Y of a topological space X, we write cl_Y and int_Y for the topological closure and interior operators on Y. If U is open in a topological space X, then for all $Z \subset U$ we have

$$\mathrm{int}_U(Z) = \mathrm{int}_X(Z) \cap U$$

and

$$\mathrm{cl}_U(Z) = \mathrm{cl}_X(Z) \cap U.$$

Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of subsets of U such that $A \coloneqq \bigcup_{n=1}^{\infty} A_n$ has nonempty interior: there is a nonempty open subset V of U such that

$$V \subset \bigcup_{n=1}^{\infty} A_n.$$

Then V is also a nonempty open subset of the Baire space X, so there is $n \in \mathbb{Z}^+$ such that $\overline{A_n}$ has nonempty interior: that is, there is a nonempty open subset W of X such that $W \subset \overline{A_n}$. Since $W \cap \overline{A_n} \neq \varnothing$, we have $\varnothing \subsetneq W \cap A_n \subset W \cap U$. Thus $W \cap U$ is a nonempty open subset of $\overline{A_n} \cap U = \mathrm{cl}_U(A_n)$.
b) (i) $\implies$ (ii) is immediate from part a) and (ii) $\implies$ (iii) is just immediate.
(iii) $\implies$ (i): Suppose that every point of X admits an open Baire neighborhood,

and let $\{A_n\}_{n=1}^{\infty}$ be a sequence of open sets such that $\bigcup_{n=1}^{\infty} A_n$ has nonempty interior U. We wish to show that not all of the A_n's are nowhere dense. Let $x \in U$ and let N be an open Baire neighborhood of x. Then we have

$$U \cap N \subset \bigcup_{n=1}^{\infty} (A_n \cap N).$$

Since N is a Baire space and $U \cap N$ is a nonempty open subset of N, there is some $n \in \mathbb{Z}^+$ such that

$$\varnothing \neq \operatorname{int}_N(\operatorname{cl}_N(A_n \cap N)) \subset \operatorname{cl}_N(A_n \cap N)^{\circ} \subset \overline{A_n \cap N}^{\circ} \subset \overline{A_n}^{\circ}.$$

c) Let U be a nonempty open subset of X. If A is a nowhere dense subset of X, then $A \cap U$ is a nowhere dense subset of U:

$$\operatorname{int}_U(\operatorname{cl}_U(A \cap U)) = \operatorname{int}_U(\overline{A \cap U} \cap U) = (\overline{A \cap U} \cap U)^{\circ} \subset \overline{A}^{\circ} = \varnothing.$$

So if $X = \bigcup_{n=1}^{\infty} A_n$ is a countable union of nowhere dense subsets, then $U = \bigcup_{n=1}^{\infty} (A_n \cap U)$ is a countable union of nowhere dense subsets. □

EXERCISE 4.10. *Find nonempty subsets A, U of $\mathbb{R}$ with U open such that A is somewhere dense in $\mathbb{R}$ but $A \cap U$ is nowhere dense in U.*

EXERCISE 4.11 (A. Kruckman). *Show: a topological space is Baire iff for any sequence $\{U_n\}_{n=1}^{\infty}$ of open sets in X, we have*

$$\left(\overline{\bigcap_{n=1}^{\infty} U_n}\right)^{\circ} = (\bigcap_{n=1}^{\infty} \overline{U_n})^{\circ}.$$

EXERCISE 4.12. *Let I be a nonempty set, and for $i \in I$ let X_i be a nonempty topological space. Put $X := \coprod_{i \in I} X_i$.*

a) *Show: X is Baire iff each X_i is Baire.*
b) *Show: X is meager iff each X_i is meager iff each X_i is intrinsically meager.*

The following result shows in particular that $\mathbb{R}$ has many Baire spaces that are not completely metrizable.

PROPOSITION 4.3.
a) *Let X be a topological space, and let Y be a dense subset of X. If Y is a Baire space, then so is X.*
b) *Let X be a Baire space, and let Z be a dense open subset of X. For any Y with $Z \subset Y \subset X$, the space Y is Baire.*

PROOF. a) Let $\{U_n\}_{n=1}^{\infty}$ be a sequence of dense open subsets of X, and let V be a nonempty open subset of X. For all $n \in \mathbb{Z}^+$ we have by Exercise 3.17b) that $Y_n := U_n \cap Y$ is a dense open subset of the Baire space X, so $\bigcap_{n=1}^{\infty} Y_n$ is still dense in Y. Since Y is dense in X, we have by Exercise 3.17a) that $\bigcap_{n=1}^{\infty} Y_n$ is dense in X and hence so too is the larger subset $\bigcap_{n=1}^{\infty} U_n$.
b) Being open in a Baire space, the space Z is Baire by Theorem 4.2. Since Z is dense and Baire in X, it is also dense and Baire in Y,so Y is Baire by part a). □

Now let $C \subset \mathbb{R}$ be the classical middle thirds Cantor set. It is closed and totally disconnected, hence nowhere dense, so by Exercise 4.1 the complement

$$U_C := \mathbb{R} \setminus C$$

is open in $\mathbb{R}$ – hence Baire – and dense, with complement of continuum cardinality $\mathfrak{c}$. It follows from Proposition 4.3b) that every subset Y with $U_C \subset Y \subset \mathbb{R}$ is Baire and that there are $2^{\mathfrak{c}}$ such sets. On the other hand, as we will see later the completely metrizable subsets of $\mathbb{R}$ are precisely the G_δ-subsets. In fact the cardinality of the entire Borel σ-algebra of $\mathbb{R}$ – i.e., the smallest family of sets containing the topology and closed under complementations, countable unions and countable intersections – is $\mathfrak{c}$ – so most Baire subsets of $\mathbb{R}$ are not even Borel sets, let alone completely metrizable. This entire discussion holds verbatim with $\mathbb{R}$ replaced by any nonempty separable completely metrizable space without isolated points that admits a dense open subset with uncountable complement.

EXAMPLE 4.1. *Let $\mathbb{I} := \mathbb{R} \setminus \mathbb{Q}$ be the irrational numbers. We view $\mathbb{R}$ as being embedded in $\mathbb{R}^2$ as a closed subspace in the usual manner: $x \mapsto (x, 0)$. Let*

$$X := \mathbb{R}^2 \setminus (\mathbb{I} \times \{0\}) \subset \mathbb{R}^2.$$

Then X contains the dense open subset $\mathbb{R}^2 \setminus \mathbb{R}$ so is Baire by Propostion 4.3. Now

$$\mathbb{Q} = X \cap \mathbb{R}$$

is a closed subspace of X. This shows that a closed subspace of a Baire space need not be Baire.

PROPOSITION 4.4. *Let X be a nonmeager topological space. Then every dense G_δ-subset Y of X is nonmeager.*

PROOF. Write $Y = \bigcap_{n=1}^\infty U_n$ with each U_n open and dense in X. By Exercise 4.1 each $X \setminus U_n$ is nowhere dense, so $X \setminus Y = \bigcup_{n=1}^\infty (X \setminus U_n)$ is meager. Since X is nonmeager, at least one out of each subset and its complement must be nonmeager, so Y is nonmeager. $\square$

PROPOSITION 4.5. a) *Let X be a Baire space, and let Y be a subset of X. Then Y is comeager iff it contains a dense G_δ-subset.*

b) *A topological space X is Baire iff for every meager subset Y of X, the complement $X \setminus Y$ is a Baire space.*

c) *A dense G_δ-subspace of a Baire space is a Baire space.*

PROOF. a) If $X \setminus Y$ is meager, there is a sequence $\{A_n\}_{n=1}^\infty$ of nowhere dense subsets such that $X \setminus Y = \bigcup_{n=1}^\infty A_n$. By Exercise 4.1, each $X \setminus \overline{A_n}$ is dense and open, and then $\bigcap_{n=1}^\infty (X \setminus \overline{A_n})$ is (since X is Baire) a dense G_δ set contained in Y.

If $\bigcap_{n=1}^\infty U_n$ is a dense G_δ contained in Y, then $X \setminus Y \subset \bigcup_{n=1}^\infty (X \setminus U_n)$. By Exercise 4.1 again, each $X \setminus U_n$ is nowhere dense, so $X \setminus Y$ is meager.

b) If $X = \varnothing$ then every subspace is both Baire and and meager and the equivalence holds, so suppose that $X \setminus \varnothing$.

Let X be a Baire space, and let Y be a meager subset. By Exercise 4.4 we have that $X \setminus Y$ is dense. To show that $X \setminus Y$ is Baire we will also use Exercise 4.4: it is enough to show that if Z is a meager subset of $X \setminus Y$, then $(X \setminus Y) \setminus Z = X \setminus (Y \cup Z)$ is dense in $X \setminus Y$. Let B be a closed nowhere dense subset of $X \setminus Y$, and suppose its closure $\overline{B}$ in X contains a nonempty open subset U. Then $U \cap (X \setminus Y) \subset \int_{X \setminus Y} B$, so $U \cap (X \setminus Y) = \varnothing$ and $U \subset Y$, contradicting the Baireness of X. So if B is nowhere dense in $X \setminus Y$, then $\mathrm{cl}_{X \setminus Y}(B)$ is nowhere dense in X, hence B is nowhere dense in X. And now it follows that since Z is meager in $X \setminus Y$, also Z is meager in X. Since Y and Z are both meager in X, the set $Y \cup Z$ is also meager in the

Baire space X, so – once again using Exercise 4.4 – we have that $X \setminus (Y \setminus Z)$ is dense in X. The converse is trivial: take $Y = \varnothing$.
c) Let X be a Baire space, and let Y be a dense G_δ-set in X. By part a), the complement $X \setminus Y$ is meager, so Y is a Baire space by part b). $\square$

4. Strongly Baire Spaces

LEMMA 4.6. *For a topological space X, the following are equivalent:*

(i) *No nonempty closed subset of X is intrinsically meager.*
(iii) *Every closed subspace of X is a Baire space.*

A space satisfying these equivalent conditions is called ***strongly Baire****.*

PROOF. (i) $\implies$ (ii): Suppose that nonempty closed subset of X is intrinsically meager, let Y be a closed subset of X, and let V be a nonempty relatively open subset of Y. Then $\mathrm{cl}_Y(V)$ is closed and nonempty in X so is nonmeager and V is dense and open in $\mathrm{cl}_Y(V)$, so V is nonmeager by Proposition 4.4.
(ii) $\implies$ (i): If every closed subspace of X is Baire, and Y is a nonempty closed subset of X then Y is Baire and nonempty and hence not intrinsically meager. $\square$

THEOREM 4.7. *Every G_δ-subspace in a strongly Baire space is strongly Baire.*

PROOF. Let X be strongly Baire, let Y be a G_δ-subset of X, and let Z be a nonempty (relatively) closed subset of Y. Then $\overline{Z}$ is closed in the strongly Baire space X so is strongly Baire. Also $Z = \mathrm{cl}_Y(Z) = Y \cap \overline{Z}$, so Z is a dense G_δ-subset of $\overline{Z}$, so Z is a Baire sapce by Proposition 4.5c). Thus Y is strongly Baire. $\square$

COROLLARY 4.8 (Big Baire). *A topological space that is homeomorphic to a G_δ-subspace of a complete metric space or a locally compact space is a strongly Baire space.*

PROOF. This is immediate from Theorems 2.58, 4.1 and 4.7. $\square$

In particular we find that the irrational numbers $\mathbb{I}$ are a strongly Baire space, being a G_δ-subset in $\mathbb{R}$. We also see that the rational numbers $\mathbb{Q}$ do *not* form a G_δ-subset of $\mathbb{R}$, because they are intrinsically meager and thus not Baire.

5. Pointwise Limits and Baire Classes

Recall that a map $f : X \to Y$ between topological spaces is **continuous at a point** $x \in X$ if for every neighborhood V of $f(x)$ in Y there is a neighborhood U of x in X such that $f(U) \subset V$ and that f is continuous iff it is continuous at every point of X.

In real analysis, one studies sequences of functions $\{f_n : I \to \mathbb{R}\}_{n=1}^\infty$, where I is an interval on the real line. A function $f : I \to \mathbb{R}$ is a **pointwise limit** of the sequence $\{f_n\}$ if, indeed, for all $x \in I$ we have $f_n(x) \to f(x)$. One of the first things one learns – cf. [**Cl-HC**, §13.1.1] – is that if each f_n is continuous, then the limit function f need not be continuous. A standard example is to take

$$f_n : [0,1] \to \mathbb{R}, x \mapsto x^n,$$

in which case the limit function is

$$f(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}.$$

This motivates the concept of uniform limit, among whose virtues is the fact that a uniform limit of continuous functions is continuous [**Cl-HC**, Cor. 13.3].

But this is not the whole story! That it wasn't was brought to my attention in the following way: at the end of 2004 I was making preparations for teaching an elementary real analysis course at McGill University that included pointwise and uniform convergence. Just as the introduction of this text suggests, I had learned this material from Rudin's classic text [**Rud**]. For me, Chapters 7 and 8 of his text, in which he introduces sequences and series of functions and then gives many applications, formed a high point of undergraduate mathematics, and I remembered them well...or so I thought. While preparing the course, I remembered that Rudin also gave a much more interesting example than the above one, in which the characteristic function of the rational numbers

$$\mathbf{1}_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases} \tag{15}$$

was expressed as a pointwise limit of continuous functions. I thought it would be educational to reconstruct such an example myself, so I tried...to no avail. Later I cracked open the text that I had already bought ten years before and found this:

$$\lim_{m\to\infty}\lim_{n\to\infty}(\cos m!\pi x)^{2n} = \begin{cases} 0 & (x \text{ irrational}) \\ 1 & (x \text{ rational}) \end{cases}.$$

Hey, that's a double limit! Isn't that cheating?!? Yes, it is. In fact what I was trying to do was impossible: $\mathbf{1}_{\mathbb{Q}}$ is discontinuous at every point, whereas a pointwise limit of continuous functions $f_n : I \to \mathbb{R}$ must be continuous at some and in fact "many" points. We will now prove a more general result along these lines.

Lemma 4.9. *Let $f : X \to (Y, d)$ be a function from a topological space to a metric space, and let*

$$Z \coloneqq \{x \in X \mid f \text{ is continuous at } x\}.$$

Then Z is a G_δ-set in X containing all isolated points of X.

Proof. For a subset Z of X, we define the **oscillation of f on Z** to be

$$\omega(f, Z) \coloneqq \sup\{d(f(z_1), f(z_2)) \mid z_1, z_2 \in Z\} \in [0, \infty].$$

This is a non-negative extended real number that is finite iff f is bounded on Z. Now we have that f is continuous at $x \in X$ iff for all $\epsilon > 0$ there is a neighborhood N of x such that $f(N) \subset B^\circ(f(x), \epsilon)$ iff for all $\epsilon > 0$ there is a neighborhood N of x such that $\omega(f, N) \le 2\epsilon$ iff for all $\epsilon > 0$ there is a neighborhood N of x such that $\omega(f, N) < \epsilon$. So if we define **the oscillation of f at x** to be

$$\omega(f, x) \coloneqq \inf\{\omega(f, N) \mid N \text{ is a neighborhood of } x\},$$

then we have that f is continuous at x iff $\omega(f, x) = 0$. If for $x \in X$ and $\epsilon > 0$ we have that $\omega(f, x) < \epsilon$ then there is some open neighborhood U of x such that $\omega(f, U) < \epsilon$, and thus

$$U_\epsilon \coloneqq \{x \in X \mid \omega(f, x) < \epsilon\}$$

is an open subset of X. It follows that

$$Z = \bigcap_{n=1}^{\infty} U_{\frac{1}{n}}$$

is a G_δ-set. Finally, if x isolated, then $\{x\}$ is open and for every function f we have $\omega(f,\{x\}) = 0$, so $x \in Z$. $\square$

We call the set Z in the above result the **locus of continuity** of f.

EXERCISE 4.13. *Let $f : X \to Y$ be a map from a Baire space to a metric space. Let Z be the locus of continuity of f. Show: Z is comeager iff it is dense in X.*

THEOREM 4.10. *Let X be a nonempty Baire space, let (Y,d) be a metric space, and let $\{f_n : X_n \to Y\}$ be a sequence of continuous functions converging pointwise to a function $f : X \to Y$: that is, for all $x \in X$, we have $f_n(x) \to f(x)$. Let*

$$Z := \{x \in X \mid f \text{ is continuous at } x\}.$$

a) *The set Z is a dense G_δ subset of X and thus comeager.*
b) *If X has no isolated points, then Z is uncountable.*

PROOF. Step 1: Assuming that Z is dense, we show all the other assertions. By Exercise 4.13, if the locus of continuity Z is dense, then it is comeager. Suppose now that X has no isolated points. If Z were countable, then it would be meager, but also $X \setminus Z$ is meager so X would be meager, which is a contradiction since X is a nonempty Baire space.
Step 2: So it remains to show that Z is dense. For $N \in \mathbb{Z}^+$ and $\epsilon > 0$, put

$$A_N(\epsilon) := \{x \in X \mid \forall m,n \geq N \text{ we have } d(f_m(x), f_n(x)) \leq \epsilon\}.$$

For each fixed m,n, the function $x \mapsto d(f_m(x), f_n(x))$ is continuous by Proposition 2.35 so the set of $x \in X$ satisfying $d(f_m(x), f_n(x)) \leq \epsilon$ is closed. Then given set $A_N(\epsilon)$ is the intersection over these subsets for all $m,n \geq N$ so is also closed. For all $\epsilon > 0$ we have $N_1 \leq N_2$ implies $A_{N_1}(\epsilon) \subset A_{N_2}(\epsilon)$. For each $x \in X$ the sequence $\{f_n(x)\}$ is convergent, hence Cauchy, so $x \in A_N(\epsilon)$ for all sufficiently large N. It follows that

$$\bigcup_{N=1}^{\infty} A_N(\epsilon) = X.$$

For $\epsilon > 0$, put

$$U(\epsilon) := \bigcup_{N=1}^{\infty} A_N(\epsilon)^\circ.$$

Evidently each $U(\epsilon)$ is open. We claim that also each $U(\epsilon)$ is dense and that putting

$$C := \bigcap_{n=1}^{\infty} U(1/n),$$

we have

$$Z \supset C.$$

Since X is a Baire space, it will then follow that C is dense, hence also that Z is dense, completing the proof.
Step 3: To show that each $U(\epsilon)$ is dense, let V be a nonempty open subset of X. Then V is a nonempty Baire space (Theorem 4.2a)) and is the union of the sequence $\{V \cap A_N(\epsilon)\}_{n=1}^{\infty}$ of closed subspaces, so for at least one N we must have

$$\varnothing \subsetneq \operatorname{int}_V(V \cap A_N(\epsilon)) = (V \cap A_N(\epsilon))^\circ \cap V$$

$$\subset (V \cap A_N(\epsilon))^\circ = V^\circ \cap A_N(\epsilon)^\circ = V \cap A_N(\epsilon)^\circ \subset V \cap U(\epsilon).$$

Step 4: Let $x_\bullet \in C$, and fix $\epsilon > 0$. We will find a neighborhood W of $x_\bullet$ such that $d(f(x), f(x_\bullet)) < \epsilon$ for all $x \in W$, which will show that f is continuous at $x_\bullet$. Choose $m \in \mathbb{Z}^+$ such that $\frac{1}{m} < \frac{\epsilon}{3}$. Since $x_\bullet \in C$, we have $x_\bullet \in U(\frac{1}{m})$, so there is $N \in \mathbb{Z}^+$ such that $x_\bullet \in A_N(\frac{1}{m})^\circ$. Because f_N is continuous, there is a neighborhood W of $x_\bullet$ such that $W \subset A_N(\frac{1}{m})^\circ$ and

$$\forall x \in W,\ d(f_N(x), f_N(x_\bullet)) < \frac{\epsilon}{3}. \tag{16}$$

Since $W \subset A_N(\frac{1}{m})$ it follows that

$$\forall x \in W,\ \forall n \geq N,\ d(f_n(x), f_N(x)) \leq \frac{1}{m}.$$

Letting n approach infinity we get:

$$\forall x \in W,\ d(f(x), f_N(x)) \leq \frac{1}{m} < \frac{\epsilon}{3}, \tag{17}$$

and applying (16) with $x = x_\bullet \in W$, we get

$$d(f_N(x_\bullet), f(x_\bullet)) = d(f(x_\bullet), f_N(x_\bullet)) < \frac{\epsilon}{3}. \tag{18}$$

Combining (17), (16) and (18) we get

$$\forall x \in W,\ d(f(x), f(x_\bullet)) < \epsilon,$$

as desired. $\square$

It follows from Theorem 4.10 that if I is any nondegenerate (i.e., $\#I > 1$) interval in the real line and $\{f_n : I \to \mathbb{R}\}$ is a sequence of functions on I converging pointwise to $f : I \to \mathbb{R}$, then f certainly cannot be discontinuous at every point. Rather its set of points of continuity must be a dense, uncountable G_δ-set.

EXERCISE 4.14. *Consider* ***Thomae's function*** $f_T : \mathbb{R} \to \mathbb{R}$*, defined as follows:* $f_T(0) = 1$*; for* $x \in \mathbb{Q} \setminus \{0\}$*, write* $x = \frac{p}{q}$ *with* $\gcd(p, q) = 1$ *and put* $f_T(x) = \frac{1}{q}$*; for* $x \in \mathbb{I} := \mathbb{R} \setminus \mathbb{Q}$*, put* $f_T(x) = 0$*.*

a) *Show: the locus of continuity of* f_T *is* $\mathbb{I}$*, the irrational numbers.*
b) *Show: there is a sequence* $\{f_n : \mathbb{R} \to \mathbb{R}\}$ *of continuous functions whose pointwise limit is* f*.*

These results motivate the following definition: for a topological space X, we put

$$\mathcal{B}_0(X) := \{f : X \to \mathbb{R} \mid f \text{ is continuous}\}$$

and

$$\mathcal{B}_1(X) := \{f : X \to \mathbb{R} \mid \text{there is } \{f_n\}_{n=1}^\infty \text{ in } \mathcal{B}_0(X) \text{ such that } f_n \to f \text{ pointwise}\}.$$

Inductively, having defined $\mathcal{H}_\beta(X)$ for all ordinals $\beta < \alpha$, we put

$$\mathcal{B}_\alpha(X) := \{f : X \to \mathbb{R} \mid \text{there is } \{f_n\}_{n=1}^\infty \text{ in} \bigcup_{\beta<\alpha} \mathcal{B}_\beta(X) \text{ such that } f_n \to f \text{ pointwise}\}.$$

We say that a function $f : X \to \mathbb{R}$ is **Baire** if it lies in $\mathcal{B}_\alpha(X)$ for some ordinal α, and we write $\mathcal{B}(X)$ for the set of all Baire functions. For $f \in \mathcal{B}(X)$ we say that f is **of Baire class** α if α is the least ordinal number such that $f \in \mathcal{B}_\alpha(X)$. Thus the Baire class zero functions are the continuous functions and the Baire class one functions are discontinuous functions that are pointwise limits of continuous functions, and so forth.

In this terminology, Theorem 4.10 shows that in any nonempty Baire space X, every Baire class one function has locus of continuity a dense G_δ-set.

EXERCISE 4.15. *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function, and let $f' : \mathbb{R} \to \mathbb{R}$ be its derivative. Recall that although f must certainly be continuous, f' need not be: a standard example is*

$$f(x) = \begin{cases} x^2 \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases},$$

which is differentiable and for which f' is discontinuous at 0. Show: if f is differentiable and f' is not continuous, then f' has Baire class one.

EXERCISE 4.16. *Show: the function $\mathbf{1}_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ of (15) has Baire class two.*

EXERCISE 4.17. *Let $X \subset \mathbb{R}$, and let $\mathbf{1}_X : \mathbb{R} \to \mathbb{R}$ be the characteristic function of X:*

$$\mathbf{1}_X : x \mapsto \begin{cases} 1 & x \in X \\ 0 & x \notin X \end{cases}.$$

Show: $\mathbf{1}_X$ has Baire class one iff X and $\mathbb{R} \setminus X$ are both G_δ-sets.

EXERCISE 4.18. *Let X be a topological space.*

a) *Show: if for an ordinal α we have $\mathcal{B}_\alpha(X) = \mathcal{B}_{\alpha+1}(X)$, then $\mathcal{B}(X) = \mathcal{B}_\alpha(X)$.*
b) *Show: there must be an ordinal α such that $\mathcal{B}_\alpha(X) = \mathcal{B}_{\alpha+1}(X)$.*

In fact, for any $f \in \mathcal{B}(X)$ the Baire class of f is a countable ordinal and thus, if ω_1 is the least uncountable ordinal then $\mathcal{B}(X) = \mathcal{B}_{\omega_1}(X)$: for this, see [**vRS**, §17].

EXERCISE 4.19. *Show: for any topological space X, we have $\mathcal{B}_{\omega_1}(X) = \bigcup_\alpha \mathcal{B}_\alpha(X)$, where the union extends over all countable ordinals.*

EXERCISE 4.20. *Let X be a separable topological space of continuum cardinality $\mathfrak{c}$ (e.g. the continuum $X = \mathbb{R}$!).*

a) *Show that $\#\mathcal{B}_0(X) = \mathfrak{c}$.*
b) *Show: for every countable ordinal α, we have $\#\mathcal{B}_\alpha(X) = \mathfrak{c}$.*
c) *Using the aforementioned fact that $\mathcal{B}(X) = \mathcal{B}_{\omega_1}(X)$ and the previous exercise, show that $\#\mathcal{B}(X) = \mathfrak{c}$.*
d) *Show that $\#\mathbb{R}^X > \mathfrak{c}$. Deduce that "most" functions $f : X \to \mathbb{R}$ are not Baire.*

We will end our study of them here, but Baire classes are important in analysis and descriptive set theory.

6. The Comeagerness of Nowhere Differentiable Functions

In this section we work in the space $\mathcal{C} = C([0,1], \mathbb{R})$ of continuous functions $f : [0,1] \to \mathbb{R}$. Since $[0,1]$ is compact, the set $\mathcal{C}$ endowed with the metric function

$$d(f,g) := ||f-g||_\infty := \max_{x \in [0,1]} |f(x) - g(x)|$$

is complete. For $f \in f$, we also have the notion of differentiability at $x \in [0,1]$, that is: $\lim_{h \to 0} \frac{f(x+h)-f(h)}{h}$ exists in $\mathbb{R}$.[1] Let $\mathcal{N} \subset \mathcal{C}$ be the subset of functions $f : [0,1] \to \mathbb{R}$ that are **nowhere differentiable**: that is, are not differentiable at any point of $[0,1]$.

Real analysis greatly developed in rigor and sophistication during the 19th century. At its start, mathematicians did not have fully rigorous definitions of the concepts of limit and derivative. Of course that limited the enterprise only in certain specific ways: there was already a *vast* edifice of calculus that was legendarily effective in a variety of applications – certainly to scientific applications outside of mathematics itself but also even to large parts of pure mathematics. If you are differentiating and integrating *analytic* functions (given by convergent power series exapnsions in a neighborhood of every point) then the details of your foundational definitions of limit, derivative and integral are hardly pertinent, but if you are probing into the set-theoretic aspects of continuity and differentiability as in e.g. Exercises 4.14, 4.15 and 4.16 then these foundations certainly are. So it is not surprising that *function theory* developed rapidly in the late 19th century after the modern ϵ–δ notion of continuity arose due to the efforts of Cauchy and Weierstrass.

The first example of a continuous nowhere differentiable function $f : [0,1] \to \mathbb{R}$ was given by Weierstrass in 1872. An exposition of such a function – not Weierstrass's original example, but a somewhat simpler one – is given in [**Cl-HC**, §14.6]. This function f has everywhere unbounded difference quotients: for each $x \in [0,1]$, the function $\frac{f(x+h)-f(x)}{h}$ is unbounded on $[0,1] \cap ((-h,0) \cup (0,h))$. Using this function we will prove the following yet more startling result.

Let us develop some simple preliminaries on piecewise linear functions. A function $f : [0,1] \to \mathbb{R}$ is **piecewise linear** if there is a finite sequence $a = x_0 < x_1 < \ldots < x_{n_1} < x_n = b$ such that for all $0 \leq i \leq n-1$, $f|_{[x_i, x_{i+1}]}$ is a linear function (in the sense of calculus: the restriction of $mx + b$). Here we have the topological space X expressed as a finite union of closed subspaces $\{Y_i\}_{i=1}^n$, any two of which are either disjoint or intersect in a single point. By an immediate application of the Pasting Lemma, any function that restricts to a continuous function on each Y_i is continuous on X, so piecewise linear functions are continuous. Let $\mathcal{L} \subset \mathcal{C}$ denote the class of piecewise linear functions. For $f \in \mathcal{L}$, we define its **ruggedness** $\mathfrak{r}(f)$ to be the minimum of the absolute values of the slopes of its linear components. The point of this is that for $f \in \mathcal{L}$ and all $x \in [0,1]$, for all sufficiently small h we have that $|\frac{f(x+h)-f(x)}{h}| \geq \mathfrak{r}(f)$. For $\mathfrak{r} > 0$, let $\mathcal{L}_{\mathfrak{r}}$ be the piecewise linear functions of ruggedness at least $\mathfrak{r}$.

LEMMA 4.11. *For every $\mathfrak{r} > 0$, the set $\mathcal{L}_{\mathfrak{r}}$ of piecewise linear functions of ruggedness at least $\mathfrak{r}$ are dense in $\mathcal{C}$.*

PROOF. Step 1: First we claim that the piecewise linear functions are dense in $\mathcal{C}$. This is almost obvious: if we approximated f by the linear function $\ell : [0,1] \to \mathbb{R}$ such that $\ell(0) = f(0)$ and $\ell(1) = f(1)$, then

$$||f - \ell|| \leq \omega(f, [0,1]) = \max(f) - \min(f),$$

[1] At the endpoint 0 we require h to take positive values only; at the endpoint 1 require h to take negative values only.

because for all $x \in [0,1]$ we have $\ell(x) \in [\ell(0), \ell(1)] = [f(0), f(1)] \subset [\min(f), \max(f)]$. Now by uniform continuity, for all $\epsilon > 0$ there is $\delta > 0$ such that the oscillation of f is less than ϵ on every subinterval of length at most δ. So if we choose n such that $\frac{1}{n} < \delta$ and break $[0,1]$ up into n subintervals of length $\frac{1}{n}$ and on each subinterval $[\frac{i}{n}, \frac{i+1}{n}]$ let p be the unique function such that $p(\frac{i}{n}) = f(\frac{i}{n})$, $p(\frac{i+1}{n}) = f(\frac{i+1}{n})$, then p is piecewise linear and $||f - p|| < \epsilon$.
Step 2: It remains to approximate a piecewise linear function uniformly by a piecewise linear function of ruggedness at least $\mathfrak{r}$. We leave this as an exercise. $\square$

EXERCISE 4.21. *Show that for any $\mathfrak{r} > 0$, piecewise linear functions can be uniformly approximated by piecewise linear functions of ruggedness at least $\mathfrak{r}$.*

THEOREM 4.12. *The subset $\mathcal{N}$ of nowhere differentiable functions is comeager in the Baire space $\mathcal{C} = C([0,1], \mathbb{R})$. In particular $\mathcal{N}$ is dense in $\mathcal{C}$: for every continuous $f : [0,1] \to \mathbb{R}$ and $\epsilon > 0$, there is a nowhere differentiable $g : [0,1] \to \mathbb{R}$ such that $|f(x) - g(x)| < \epsilon$ for all $x \in [0,1]$.*

PROOF. Since $\mathcal{C}$ is a complete metric space and thus a Baire space, it suffices to show that $\mathcal{C} \setminus \mathcal{N}$ is meager, for then Proposition 4.5a) shows that $\mathcal{N}$ is dense. To this end, for $n \in \mathbb{Z}^+$ we put

$$A_n := \{f \in \mathcal{C} \mid \exists x_0 \in [0,1] \text{ such that } |f(x) - f(x_0)| \leq n|x - x_0| \text{ for all } x \in [0,1]\}.$$

Step 1: We claim that each A_n is closed.
Proof: Let $f \in \overline{A_n}$, so there is a sequence $\{f_k\}_{k=1}^{\infty}$ in A_n such that $f_k \to f$. By definition of A_n, for each $k \in \mathbb{Z}^+$ there is $x_k \in [0,1]$ such that

$$|f_k(x) - f_k(x_k) \leq n|x - x_k|$$

for all $x \in [0,1]$. By Bolzano-Weierstrass, after passing to a subsequence (we neglect to reindex) we may assume that $x_k \to x_\bullet \in [0,1]$. Fix $x \in [0,1] \setminus \{x_\bullet\}$; for all sufficiently large k we have $x_k \neq x$ and

$$\left| \frac{f_k(x) - f_k(x_k)}{x - x_k} \right| \leq n.$$

Taking the limit as k approaches infinity gives

$$\left| \frac{f(x) - f(x_\bullet)}{x - x_\bullet} \right| \leq n,$$

so $f \in A_n$.
Step 2: We claim that each A_n is nowhere dense in $\mathcal{C}$.
Proof: Fix $n \in \mathbb{Z}^+$. Since A_n is closed, it is enough to show that it contains no open ϵ-ball. This follows from Lemma 4.11: for any $f \in \mathbb{C}$, there is a piecewise linear function p with ruggedness at least $n+1$ such that $||f - p|| < \epsilon$, and $p \notin A_n$.
Step 3: We claim that $\mathcal{C} \setminus \mathcal{N} \subset \bigcup_{n=1}^{\infty} A_n$. This is clear: if $f \in \mathcal{C}$ is somewhere differentiable, then it must have bounded difference quotients at some point, so it lies in A_n for some n.
This shows that $\mathcal{C} \setminus \mathcal{N}$ is meager and thus $\mathcal{N}$ is comeager. $\square$

To prove that is dense in $\mathcal{C}$ latter one doesn't need to use any Baire-theoretic notions: given the existence of a single nowhere differentiable function $W \in \mathcal{C}$ and the fact (Weierstrass Approximation Theorem) that polynomial functions are dense in $\mathcal{C}$ [**Cl-HC**, Thm. 14.20], the density of $\mathcal{N}$ in $\mathcal{C}$ follows almost immediately: for $f \in \mathcal{C}$ and $\epsilon > 0$, choose a polynomial $p \in \mathcal{C}$ such that $||f - p|| < \frac{\epsilon}{2}$. Then

$$g := p + \frac{\epsilon}{2||W||} W$$

is nowhere differentiable and satisfies $||g - f|| < \epsilon$. But the proof of Theorem 4.12 is shorter than either the construction of $W \in \mathcal{N}$ or the proof of [**Cl-HC**, Th. 14.20].

It is interesting to ask how much stronger it is to know that $\mathcal{N}$ is comeager in $\mathcal{C}$ than to know that it is dense in $\mathcal{C}$. By Proposition 4.5a), once we know that $\mathcal{C}$ is a Baire space, the statement that $\mathcal{N}$ is comeager is equivalent to the fact that it contains a dense G_δ-set. But the significance of that subtle strengthening still seems up for grabs.

Baire's Theorem is interesting because it gives us a topological notion of "smallness" to be compared to the set-theoretic notion of smallness – a subset Y of an infinite set X is very small indeed if $\#X < \#Y$ – and the measure-theoretic notion of smallness – having measure zero. For me personally the measure-theoretic notion of smallness feels the most convincing, but that is perhaps simply because I was taught to think about it that way from a relatively early point in my mathematical education. A classic text of Oxtoby [**Ox**] has the comparison between largeness in the sense of measure and largeness in the sense of comeagerness as its main topic. Oxtoby shows that there is no necessary connection between these two concepts and in particular that the real line is the disjoint union of a meager set and a set of Lebesgue measure zero. However, in many natural circumstances a set that is small in one sense is liable to be small in other senses as well. In the setting of Theorem 4.12 Norbert Weiner constructed a natural measure on $\mathcal{C}$ with respect to which $\mathcal{C} \setminus \mathcal{N}$ has measure zero [**We23**].

7. Further Remarks on Baire Spaces

Exercise 4.22. *Let $X : \prod_{I \in I} X_i$ be the product of nonempty topological spaces X_i. Show: if X is Baire, so is each X_i.*

The converse is not true. In fact:

Theorem 4.13. *There is a metrizable Baire space X such that $X \times X$ is not a Baire space.*

Proof. This is a result of Oxtoby-Krom-Cohen [**Ox61**], [**Kr74**], [**Ch76**]. □

On the other hand, finite products of completely metrizable spaces are completely metrizable and finite products of locally compact spaces are locally compact, which suggests there should be additional conditions under which a finite product of Baire spaces remains Baire. Here is a result along these lines. A **π-base** $\mathcal{B}$ for a topological space X is a family $\mathcal{B}$ of nonempty open subsets of X such that for every nonempty open subset U of X there is $V \in \mathcal{B}$ with $V \subset U$. Thus every base is a π-base.

Theorem 4.14 (Oxtoby [**Ox61**]). *For $i \in I$, let X_i be a nonempty Baire space admitting a countable π-base. Then $X := \prod_{I \in I} X_i$ is a Baire space.*

In particular any Cartesian product of second countable, completely metrizable spaces is a Baire space, even though when there are uncountably many factors containing at least two points the product is not itself metrizable.

Exercise 4.23. *Is every Cartesian product of discrete spaces a Baire space?*

CHAPTER 5

Convergence

1. Introduction: Convergence in Metric Spaces

Recall the notion of convergence of sequences in metric spaces. In any set X, a sequence in X is just a mapping a mapping $\mathbf{x} : \mathbb{Z}^+ \to X$, $n \mapsto \mathbf{x}_n$. If X is endowed with a metric d, a sequence $\mathbf{x}$ in X is said to **converge** to an element x of X if for all $\epsilon > 0$, there exists an $N = N(\epsilon)$ such that for all $n \geq N$, $d(x, x_n) < \epsilon$. We denote this by $\mathbf{x} \to x$ or $\mathbf{x}_n \to x$. Since the ϵ-balls around x form a local base for the metric topology at x, an equivalent statement is the following: for every neighborhood U of x, there exists an $N = N(U)$ such that for all $n \geq N$, $\mathbf{x}_n \in U$.

We have the allied concepts of limit point and subsequence: we say that x is a **limit point** of a sequence $\mathbf{x}_n$ if for every neighborhood U of x, the set of $n \in \mathbb{Z}^+$ such that $\mathbf{x}_n \in U$ is infinite. A **subsequence** of $\mathbf{x}$ is obtained by choosing an infinite subset of $\mathbb{Z}^+$, writing the elements in increasing order as $n_1, n_2, \ldots$ and then restricting the sequence to this subset, getting a new sequence $\mathbf{y}$, $k \mapsto \mathbf{y}_k = \mathbf{x}_{n_k}$.

The study of convergent sequences in the Euclidean spaces $\mathbb{R}^n$ is one of the mainstays of any basic analysis course. Many of these facts generalize immediately to the context of an arbitrary metric space (X, d).[1]

Proposition 5.1. *Each sequence in (X, d) converges to at most one point.*

Proposition 5.2. *Let Y be a subset of (X, d). For $x \in X$, the following are equivalent:*
a) $x \in \overline{Y}$.
b) There exists a sequence $\mathbf{x} : \mathbb{Z}^+ \to Y$ such that $\mathbf{x}_n \to x$.

In other words, the closure of a set can be realized as the set of all limits of convegent sequences contained in that set.

Proposition 5.3. *Let $f : X \to Y$ be a mapping between two metric spaces. The following are equivalent:*
a) f is continuous.
b) If $\mathbf{x}_n \to x$ in X, then $f(\mathbf{x}_n) \to f(x)$ in Y.

In other words, continuous functions between metric spaces are characterized as those which preserve limits of convergent sequences.

Proposition 5.4. *Let $\mathbf{x}$ be a sequence in (X, d). For $x \in X$, the following are equivalent:*
a) The point x is a limit point of the sequence $\mathbf{x}$.
b) There exists a subsequence $\mathbf{y}$ of $\mathbf{x}$ converging to x.

[1]We recommend that the reader who finds any of these facts unfamiliar should attempt to verify them on the spot. On the other hand, more general results are coming shortly.

Moreover, there are several results in elementary real analysis that exploit, in various ways, the compactness of the unit interval $[0,1]$:

THEOREM 5.5. *(Bolzano-Weierstrass) Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

THEOREM 5.6. *(Heine-Borel) A subset of the Euclidean space $\mathbb{R}^n$ is compact iff it is closed and bounded.*

In any metric space there are several important criteria for compactness. Two of the most important ones are given in the following theorem. Recall that in any topological space X, we say that a point x is a **limit point** of a subset A if for every neighborhood N of x we have $N \setminus \{x\} \cap A \neq \emptyset$. (In other words, x lies in the closure of $A \setminus \{x\}$.)

THEOREM 5.7. *Let (X,d) be a metric space. The following are equivalent:*
a) Every sequence has a convergent subsequence.
b) Every infinite subset has a limit point.
c) Every open covering $\{U_i\}$ of X has a finite subcovering (i.e., X is compact).

Theorem 5.7 is of a less elementary character than the preceding results, and we shall give a proof of it later on.

Taken collectively, these results show that, in a metrizable space, all the important topological notions can be captured in terms of convergent sequences (and subsequences). Since every student of mathematics receives careful training on the calculus of convergent sequences, this provides significant help in the topological study of metric spaces.

It is clearly desirable to have an analogous theory of convergence in arbitrary topological spaces. Using the criterion in terms of neighborhoods, one can certainly formulate the notion of a convergent sequence in a topological space X. However, we shall see that there are counterexamples to each of the above results for sequences in an arbitrary topological space.

There are two reasonable responses to this. First, we can search for sufficient, or necessary and sufficient, conditions on a space X for these results to hold. In fact relatively mild sufficient conditions are not so difficult to find: the Hausdorff axiom ensures the uniqueness of limits; for most of the other properties the key result is the existence of a countable base of neighborhoods at each point.

The other response is to find a suitable replacement for sequences which renders correct all of the above results in an arbitrary topological space. Clearly this is of interest in applications: one certainly encounters "in nature" topological spaces which are not Hausdorff (e.g. Zariski topologies in algebraic geometry) or which do not admit a countable neighborhood base at each point (e.g. weak topologies in functional analysis), and one does not want to live in eternal fear of meeting a space for which sequences are not sufficient.[2] However, the failure of the above results to hold should suggest to the student of topology that there is "something else out there" which is the correct way to think about convergence in topological spaces. Knowing the "correct" notion of convergence leads to positive results in the theory

[2]Unfortunately many of the standard texts used for undergraduate courses on general topology (and there are rarely graduate courses on general topology nowadays) seem content to leave their readers in this state of fear.

as well as the avoidance of negative results: for instance, armed with this knowledge one can prove the important Tychonoff theorem in a few lines, whereas other proofs are significantly longer and more complicated (even in a situation when sequences suffice to describe the topology of the space!). In short, there are conceptual advantages to knowing "the truth" about convergence.

Intriguingly, there are two different theories of convergence which both successfully generalize the convergence of sequences in metric spaces: nets and filters. The theory of nets was developed by the early twentieth century American topologists E.H. Moore and H.L. Smith (their key paper appeared in 1922). In 1950 J.L. Kelley published a paper which made some refinements on the theory, cosmetic and otherwise (in particular the name "net" appears for the first time in his paper). The prominent role of nets in his seminal text **General Topology** cemented the centrality of nets among American (and perhaps all anglophone) topologists. Then there is the rival theory of filters, discovered by Henri Cartan in 1937 amidst a Séminaire Bourbaki. Cartan successfully convinced his fellow Bourbakistes of the elegance and utility of the theory of filters, and Bourbaki's similarly influential **Topologie Generale** introduces filters early and often. To this day most continental mathematicians retain a preference for the filter-theoretic language.

For the past fifty years or so, most topology texts have introduced at most one of nets and filters (possibly relegating the other to the exercises). As Gary Laison has pointed out, since both theories appear widely in the literature, this practice is a disservice to the student. The fact that the two theories are demonstrably equivalent – that is, one can pass from nets to filters and vice versa so as to preserve convergence, in a suitable sense – does not mean that one does not need to be conversant with both of them! In fact each theory has its own merits. The theory of nets is a rather straightforward generalization of the theory of sequences, so that if one has a sequential argument in mind, it is usually *a priori* clear how to phrase it in terms of nets. (In particular, one can make a lot of headway in functional analysis simply by doing a search/replace of "sequence" with "net.") Moreover, many complicated looking limiting processes in analysis can be expressed more simply and cleanly as convergence with respect to a net – e.g., the Riemann integral. One may say that the main nontriviality in the theory of nets is the notion of "subnet", which is more complicated than one at first expects (in particular, a subnet may have larger cardinality!). The corresponding theory of filters is a bit less straightforward, but most experts agree that it is eventually more penetrating. One advantage is that the filter-theoretic analogue of subnet is much mor transparent: it is just set-theoretic containment. Filters have applications beyond just generalizing the notion of convergent sequences: in completions and compactifications, in Boolean algebra and in mathematical logic, where ultrafilters are arguably the single most important (and certainly the most elegant) single technical tool.

2. Sequences in Topological Spaces

In this section we develop the theory of convergence of sequences in arbitrary topological spaces, including an analysis of its limitations.

2.1. Arbitrary topological spaces. A sequence $\mathbf{x}$ in a topological space X **converges** to $x \in X$ if for every neighborhood U of x, $\mathbf{x}_n \in U$ for all sufficiently large n. Note that it would obviously be equivalent to say that all but finitely many terms of the sequence lie in any given neighborhood U of x, which shows that whether a sequence converges to x is independent of the ordering of its terms.[3]

Remark 2.1.1: The convergence of a sequence is a **topological notion**: i.e., if X, Y are topological spaces, $f : X \to Y$ is a homeomorphism, $\mathbf{x}_n$ is a sequence in X and x is a point of X, then $\mathbf{x}_n \to x$ iff $f(\mathbf{x}_n) \to f(x)$. In particular the theory of sequential convegence in metric spaces recalled in the preceding section applies verbatim to all metrizable spaces.

Tournant dangereux: Let us not forget that in a metric space we have the notion of a **Cauchy sequence**, a sequence $\mathbf{x}_n$ with the property that for all $\epsilon > 0$, there exists $N = N(\epsilon)$ such that $m, n \geq N \implies d(\mathbf{x}_m, \mathbf{x}_n) < \epsilon$, togtether with the attendant notion of completness (i.e., that every Cauchy sequence be convergent) and completion. Being a Cauchy sequence is *not* a topological notion: let $X = (0, 1)$, $Y = (1, \infty)$, $f : X \to Y$, $x \mapsto \frac{1}{x}$, and $\mathbf{x}_n = \frac{1}{n}$. Then $\mathbf{x}_n$ is a Cauchy sequence, but $f(\mathbf{x}_n) = n$ is not even bounded so cannot be a Cauchy sequence. (Indeed, the fact that boundedness is not a topological property is certainly relevant here.) This means that what is, for analytic applications, arguably the most important aspect of the theory – what is first semester analysis but an ode to the completeness of the real numbers? – cannot be captured in the topological context. However there is a remedy, Weil's notion of **uniform spaces**, which will be discussed later on.[4]

Example 5.1. *Let X be a set with at least two elements endowed with the indiscrete topology. Let $\{x_n\}$ be a sequence in X and $x \in X$. Then $\mathbf{x}_n$ converges to x.*

Example 5.2. *A sequence is **eventually constant** if there exists an $x \in X$ and an N such that $n \geq N \implies \mathbf{x}_n = x$; we say that x is the **eventual value** of the sequence (note that this eventual value is unique). In any topological space, an eventually constant sequence converges to its eventual value. However, such a sequence may have other limits as well, as in the above example.*

Exercise 5.1. *In a discrete topological space X, a sequence $\mathbf{x}_n$ converges to x iff $\mathbf{x}_n$ is eventually constant and x is its eventual value.*

In particular the limit of a convergent sequence in a discrete space is unique. (Since discrete spaces are metrizable, by Remark 2.1.1 we knew this already.) The following gives a generalization:

Proposition 5.8. *A sequence in a Hausdorff space converges to at most one point.*

Proof. If $\mathbf{x}_n \to x$ and $x' \neq x$, there exist disjoint neighborhoods N of x and N' of x'. Then only finitely many terms of the sequence can lie in N', so the sequence cannot converge to x. □

Let $\iota : \mathbb{Z}^+ \to \mathbb{Z}^+$ be a monotone increasing injection. If $\{x_n\}$ is a sequence in a space X, then so too is $\{x_{\iota(n)}\}$, a **subsequence** of $\{x_n\}$. Immediately from the definitions, if a sequence converges to a point x then every subsequence converges to x. On the other hand, a divergent sequence may admit a convergent subsequence.

[3]This aspect of sequential convergence will *not* be preserved in the theory of nets.

[4]Later on in time, I suppose – not yet later on in these notes.

We say that x is a **limit point** of a sequence $\mathbf{x}_n$ if every neighborhood N of x contains infinitely many terms from the sequence.

A space X is **first countable at $\mathbf{x} \in X$** if there is a countable neighborhood base at x. A space is **first countable** if it is first countable at each of its points.

In a metric space, the family $\{B(x, \frac{1}{n})\}_{n \in \mathbb{Z}^+}$ is a countable neighborhood base at x. So metrizable spaces are first countable. Note that this countable base at x is nested: $N_1 \supset N_2 \supset \ldots$. This is not particular to metric spaces: if $\{N_n\}$ is a countable base at x, then $N'_n = \cap_{i=1}^n N_i$ is a nested countable base at x. This simple observation justifies the role that sequences play in the topology of a first countable space.

PROPOSITION 5.9. *Let X be a first countable space and $Y \subset X$. Then $\overline{Y}$ is the set of all limits of sequences from Y.*

PROOF. Suppose y_n is a sequence of elements of Y converging to x. Then every neighborhood N of x contains some $y_n \in Y$, so that $x \in \overline{Y}$. Conversely, suppose $x \in \overline{Y}$. If X is first countable at x, we may choose a nested collection $N_1 \supset N_2 \supset \ldots$ of open neighborhoods of x such that every neighborhood of x contains some N_n. Each N_n meets Y, so choose $y_n \in N_n \cap Y$, and y_n converges to y. $\square$

PROPOSITION 5.10. *Let f be a map of sets between the topological spaces X and Y. Assume that X is first countable. The following are equivalent:*

a) *The map f is continuous.*
b) *If $\mathbf{x}_n \to x$, $f(\mathbf{x}_n) \to f(x)$.*

PROOF. a) $\implies$ b): Let V be any open neighborhood of $f(x)$; by continuity there exists an open neighborhood U of x such that $f(U) \subset V$. Since $\mathbf{x}_n \to x$, there exists N such that $n \geq N$ implies $\mathbf{x}_n \in U$, so that $f(\mathbf{x}_n) \in V$. Therefore $f(\mathbf{x}_n) \to f(x)$.

b) $\implies$ a): Suppose f is not continuous, so that there exists an open subset V of Y with $U = f^{-1}(V)$ not open in X. More precisely, let x be a non-interior point of U, and let $\{N_n\}$ be a nested base of open neighborhoods of x. By non-interiority, for all n, choose $\mathbf{x}_n \in N_n \setminus U$; then $\mathbf{x}_n \to x$. By hypothesis, $f(\mathbf{x}_n) \to f(x)$. But V is open, $f(x) \in V$, and $f(\mathbf{x}_n) \in Y \setminus V$ for all n, a contradiction. $\square$

PROPOSITION 5.11. *A first countable space in which each sequence converges to at most one point is Hausdorff.*

PROOF. Suppose not, so there exist distinct points x and y such that every neighborhood of x meets every neighborhood of Y. Let U_n be a nested neighborhood basis for x and V_n be a nested neighborhood basis for y. By hypothesis, for all n there exists $\mathbf{x}_n \in U_n \cap V_n$. Then $\mathbf{x}_n \to x$, $\mathbf{x}_n \to y$, contradiction. $\square$

PROPOSITION 5.12. *Let $\{x_n\}$ be a sequence in a first countable space. The following are equivalent:*
a) x is a limit point of the sequence.
b) There exists a subsequence converging to x.

PROOF. a) $\implies$ b): Take a nested neighborhood basis N_n of x, and for each $k \in \mathbb{Z}^+$ choose successively a term $n_k > n_{k-1}$ such that $x_{n_k} \in N_k$. Then $x_{n_k} \to x$. The converse is almost immediate and does not require first countability. $\square$

The following example shows that the hypothesis of first countability is necessary for each of the previous three results.

EXAMPLE 5.3. *Consider again the cocountable topology on an uncountable set X (cf. Example 3.7). In this topology a sequence $\mathbf{x}_n$ converges to x iff $\mathbf{x}_n$ is eventually constant with eventual value x. Indeed, let $\mathbf{x}_n$ be a sequence for which the set of n such that $\mathbf{x}_n \neq x$ is infinite. Then $X \setminus \{x_n \neq x\}$ is a neighborhood of x which omits infinitely many terms $\mathbf{x}_n$ of the sequence, so $\mathbf{x}_n$ does not converge to x. This implies that the set of all limits of sequences from a subset Y is just Y itself, whereas for any uncountable Y, we have $\overline{Y} = X$.*

EXERCISE 5.2. *A point x of a topological space is **isolated** if $\{x\}$ is open.*

a) *If x is isolated, and $\mathbf{x}_n \to x$, then $\mathbf{x}_n$ is eventually constant with limit x.*
b) *Note that Example 2.1.3 shows that the converse is false in general. Show however, that if X is first countable and x is not isolated, then there exists a non-eventually constant sequence converging to x.*

2.2. Sequential spaces. The hypothesis of first countability appeared as a sufficient condition in most of our results on the topological adequacy of convergent sequences. It is natural to ask to what extent it is necessary.

To explore this let us define the **sequential closure** $\mathrm{sc}(Y)$ of a subset Y of X to be the set of all limits of convergent sequences from Y. We have just seen that $sc(Y) \subset \overline{Y}$ in any space, $\mathrm{sc}(Y) = \overline{Y}$ in a first countable space, and in general we may have $\mathrm{sc}(Y) \neq \overline{Y}$.

One calls a space **Fréchet** if $\mathrm{sc}(Y) = \overline{Y}$ for all Y. However, a weaker condition is in some ways more interesting. Namely, define a space to be **sequential** if sequentially closed subsets are closed. Here are some easy facts:

(i) Closed subspaces of sequential spaces are sequential.
(ii) A space is Fréchet iff every subspace is sequential.
(iii) A space is sequential iff $\mathrm{sc}(Y) \setminus Y \neq \emptyset$ for every nonclosed subset Y.
(iv) Let $f : X \to Y$ be a map between topological spaces. If X is sequential, then f is continuous iff $\mathbf{x}_n \to x \implies f(\mathbf{x}_n) \to f(x)$.

Next we note that in any space, $A \mapsto \mathrm{sc}(A)$ satisfies the three Kuratowski closure axioms (KC1), (KC2), (KC4), but not in general (KC3). As the proof of [**Topological Spaces**, Thm. 1] shows, the sequentially closed sets therefore satisfy the axioms (CTS1)-(CTS3) for the closed sets of a new, finer topology τ' on X.

Consider next the prospect of iterating the sequential closure. If X is not sequential, there exists some nonclosed subset A whose sequential closure is equal to A itself, and then no amount of iteration will bring the sequential closure to the closure. Conversely, if X is sequential but not Fréchet, then for some nonclosed subset A of X we have A is properly contained in $\mathrm{sc}(A)$ which is properly contained in $\mathrm{sc}(\mathrm{sc}(A))$. For any ordinal number α, we can define the α-iterated sequential closure sc_α, by $\mathrm{sc}_{\alpha+1}(A) = \mathrm{sc}(\mathrm{sc}_\alpha(A))$, and for a limit ordinal β we define

$$\mathrm{sc}_\beta(A) = \bigcup_{\alpha<\beta} \mathrm{sc}_\alpha(A).$$

There is then some ordinal α such that $\mathrm{sc}_\alpha(A) = \overline{A}$ for all subsets A of X. The least such ordinal is called the **sequential order**, and is an example of an **ordinal invariant** of a topological space.

These ideas have been studied in considerable detail, notably by S.P. Franklin (to whom the term **sequential space** is due) in the 1960's. The wikipedia entry is excellent and contains many references to the literature.

One would think that there could arise, in practice, situations in which one was naturally led to consider sequential closure and not closure. (In fact, it seems to me that this is the case in the theory of **equidistribution** of sequences. But not being too sure of myself, I will say nothing further about it here.) However, we shall not pursue the matter further here, but rather turn next to two ways of "repairing" the notion of convergence by working with more general objects than sequences.

3. Nets

3.1. Nets and subnets.

On a set I equipped with a binary relation $\leq$, consider the following axioms:

(PO1) For all $i \in I$, $i \leq i$. (reflexivity).
(PO2) For all $i,\ j, k \in I$, $i \leq j,\ j \leq k$ implies $i \leq k$. (transitivity).
(PO3) If $i \leq j$ and $j \leq i$, then $i = j$ (anti-symmetry).
(D) For $i,\ j \in I$ there exists $k \in I$ such that $i \leq k$ and $j \leq k$.

If $\leq$ satisfies (PO1), (PO2) and (PO3), it is called a **partial ordering**. We trust that this is a familiar concept. If $\leq$ (PO1) and (PO2) it is called a **quasi-ordering**.[5] Finally, a relation which satisfies (PO1), (PO2) and (D) is said to be **directed**, and a *nonempty set* I endowed with $\leq$ is called a **directed set**.

EXAMPLE 5.4. *A nonempty set I endowed with the "maximal" (discrete??) relation $I \times I$ – i.e., $x \leq y$ for all $x, y \in I$ is directed, but not partially ordered if I has more than one element.*

EXAMPLE 5.5. *Any totally ordered set is a directed set; in particular the positive integers with their standard ordering form a directed set.*

A subset J of a directed set I is **cofinal** if for all $i \in I$, there exists $j \in J$ such that $j \geq i$. For instance, a subset of $\mathbb{Z}^+$ is cofinal iff it is infinite. A cofinal subset of a directed set is itself directed.

EXAMPLE 5.6. *The neighborhoods of a point x in a topological space form a directed (and partially ordered) set under reverse inclusion. More explicitly, we define $N_1 \leq N_2$ iff $N_1 \supset N_2$. A cofinal subset is precisely a neighborhood basis.*

If X has a countable basis at x, then we saw that we could take a nested neighborhood basis. In other words, the directed set of neighborhoods has a cofinal subset which is order isomorphic to the positive integers $\mathbb{Z}^+$, and this structure was the key to the efficacy of sequential convergence in first countable spaces. This suggests

[5] Alternate terminology: **preordering**.

modiyfying the definition of convergence by replacing sequences by functions with domain in an aribtrary directed set:

A **net** $\mathbf{x} : I \to X$ in a set X is a mapping from a directed set I to X.

Some further net-theoretic (but not yet topological) terminology: a net $\mathbf{x} : I \to X$ is **eventually in** a subset A of X if there exists $i \in I$ such that for all $j \geq i$, $x_j \in A$. Moreover, $\mathbf{x}$ is **cofinally** in A if the set of all i such that $x_i \in A$ is cofinal in I.

Exercise 5.3. *For a net $\mathbf{x} : I \to X$ and a subset A of X, show that the following are equivalent:*

(i) $\mathbf{x}$ *is cofinal in A.*
(ii) $\mathbf{x}$ *is* not *eventually in $X \setminus A$.*

Now suppose that we have a net $x_\bullet : I \to X$ in a topological space X. We say that $x_\bullet$ **converges** to $x \in X$ – and write $\mathbf{x} \to x$ or $x_i \to x$ – if for every neighborhood U of x, there is an element $i \in I$ such that for all $j \geq i$, $x_j \in U$. In other words, $x_i \to x$ iff $\mathbf{x}$ is eventually in every neighborhood of x. Moreover, we say that x is a **limit point** of $\mathbf{x}$ if $\mathbf{x}$ is cofinally in every neighborhood of x.

Exercise 5.4. *Check that for nets with $I = \mathbb{Z}^+$ this reduces to the definition of limit and limit point for sequences given in the previous section.*

Now the following result almost proves itself:

Proposition 5.13. *In a topological space X, the closure of any subset S is the set of limits of convergent nets of elements of S.*

Proof. First, if x is the limit of a net $\mathbf{x}$ of elements of S, then if x were not in $\overline{S}$ there would exist an open neighborhood U of x disjoint from S, but the definition of a net ensures that the set of $i \in I$ for which $x_i \in U \cap S$ is nonempty, a contradiction. On the other hand, assume that $x \in \overline{S}$, and let I be the set of open neighborhoods of x. For each i, select any $x_i \in i \cap S$. That the net x_i converges to x is a tautology: each open neighborhood U of x correponds to some $i \in I$, and for all $j \geq i$ – i.e., for all open neighborhoods $V = V(j) \subset U = U(i)$ – we do indeed have $x_j \in V$. $\square$

Proposition 5.14. *For a map $f : X \to Y$ between topological spaces, the following are equivalent:*
(i) f is continuous.
(ii) If $\mathbf{x}$ is a net converging to x, then $f(\mathbf{x})$ is a net converging to $f(x)$ in Y.

Proposition 5.15. *A space is Hausdorff iff each net converges to at most one point.*

Exercise 5.5. *Prove Propositions 5.14 and 5.15.*

We would now like to give the "net-theoretic analogue" of Proposition 5.12. Its statement should clearly be the following:

Proposition 5.16. *Let $\mathbf{x}$ be a net in a topological spcae. The following are equivalent:*
a) x is a limit point of $\mathbf{x}$.
b) There exists a subnet converging to x.

Of course, in order to make proper sense of this we need to define "subnet": how to do this? It is tempting to define a subnet of $\mathbf{x} : I \to X$ as the net obtained by restricting x to a cofinal subset of I. (At any rate, this is what I would have guessed.) However, with this definition, a subnet of a sequence is nothing else than a subsequence, and although this may sound appealing initially, it would mean that Proposition 5.12 is true without the assumption of first countability. This is not the case, as the following example shows.

EXAMPLE 5.7. *(Arens) Let $X = \mathbb{Z}^+ \times \mathbb{Z}^+$, topologized as follows: every one-point subset except $(0,0)$ is open, and the neighborhoods of $(0,0)$ are those subsets N containing $(0,0)$ for which there exists an M such that $m \geq M \implies \{n \mid (m,n) \notin N\}$ is finite: that is, N contains all but finitely many of the elements of all but finitely many of the columns $M \times \mathbb{Z}^+$ of X. Then X is a Hausdorff space in which no sequence in $X \setminus \{(0,0)\}$ converges to $(0,0)$. Moreover, there is a sequence $\mathbf{x}_n \in X \setminus \{(0,0)\}$ which has $(0,0)$ as a limit point, but by the above there is no subsequence which converges to $(0,0)$.*

So we define a **subnet** of a net $\mathbf{x} : I \to X$ to be a net $\mathbf{y} : J \to X$ for which there exists an order homomorphism $\iota : J \to I$ (i.e., $j_1 \leq j_2 \implies \iota(j_1) \leq \iota(j_2)$) with $\mathbf{y} = \mathbf{x} \circ \iota$ such that $\iota(J)$ is cofinal in I. This differs from the expected definition in that ι is not required to be an injection. Indeed, J may have larger cardinality than I, and this is an important feature of the definition.

EXERCISE 5.6. *Let J and I be a directed sets. A function $\iota : J \to I$ is said to be* ***cofinal*** *if for all $i \in I$ there exists $j \in J$ such that $j' \geq j \implies \iota(j') \geq i$. Show that the order homomorphism ι required in the definition of subnet is a cofinal function.*

Remark 3.1.9: Indeed, many treatments of the theory (e.g. Kelley's) require only that the function ι be cofinal, which gives rise to a more inclusive definition of a subnet. The two definitions lead to exactly the same results, so the issue of which one to adopt is purely a matter of taste. Our perspective here is that by restricting as we have to "order-preserving subnets", results of the form "There exists a subnet such that..." become (in the formal sense) slightly stronger.[6]

EXERCISE 5.7. *Let $\mathbf{y}$ be a subnet of $\mathbf{x}$ and $\mathbf{z}$ be a subnet of $\mathbf{y}$. Show that $\mathbf{z}$ is a subnet of $\mathbf{x}$.*

To prove Proposition 5.16 we will build a subnet in terms of the given net and the directed set of neighborhoods of the limit point x. Here is the key result.

LEMMA 5.17. *(Kelley's Lemma) Let $x : I \to X$ be a net in the topological space X, and $\mathcal{A}$ a family of subsets of X. We assume:*
(i) For all $A \in \mathcal{A}$, $I_A := \{i \in I \mid x_i \in A\}$ is cofinal in A.
(ii) The intersection of any two elements of $\mathcal{A}$ contains an element of A.
Then there is a subnet y of x which is eventually in A for all $A \in \mathcal{A}$.

PROOF. Property (ii) implies that the family $\mathcal{A}$ is directed by $\supset$. Let J be the set of all pairs (i, A) such that $i \in I$, $A \in \mathcal{A}$ and $x_i \in A$, endowed with the induced ordering from the product $I \times \mathcal{A}$. It is easy to see that J is a directed set. Indeed: For $(i, A), (i', A')$ in J, we may choose first $A'' \subset A' \cap A''$ and then $i'' \in I$ such

[6] Indeed, after gaining inspiration from the theory of filters, we will offer in §6 a definition of subnet which is more inclusive than even Kelley's definition and seems decidedly simpler: it does not require an auxiliary function ι.

that $i'' \geq i$, $i'' \geq i'$ and $x_{i''} \in A''$, and then (i'', A'') is an element of J dominating (i, A) and (i', A'). Moreover, the natural map $\iota : J \to I$ given by $(i, A) \mapsto i$ is an order homomorphism. Since $I_A \times \{A\} \subset J$ and I_A is cofinal for all $A \in \mathcal{A}$, $\iota(J)$ is cofinal in I, so that $y := x \circ \iota$ is a subnet of x. Fix $A \in \mathcal{A}$ and choose $i \in I$ such that $x_i \in A$. If $(i', A') \geq (i, A)$, then $x_{i'} \in A' \subset A$, so that $y_{(i',A')} = x_{i'} \in A$, and y is eventually in A. □

Now we can prove Proposition 5.16. Let $x \in X$ be a limit point of a net $x_\bullet$. Then the previous lemma applies to the family of all neighborhoods of x, giving us a subnet $y_\bullet : J \to X$ of $x_\bullet$ such that $y_J \to x$. Conversely, if x is not a limit point of $x_\bullet$ then there exists a neighborhood N of x such that I_N is not cofinal in I, meaning that I is eventually in $X \setminus N$. It follows that every subnet is eventually in $X \setminus N$ and hence that no subnet converges to x.

Exercise 5.8. *Define an "eventually constant net" and prove the following: for a topological space X and $x \in X$, the following are equivalent:*
(i) x is an isolated point of X; (ii) Every net converging to x is eventually constant.
Conclude: a nondiscrete space carries a convergent, not eventually constant net.

Exercise 5.9. *Let $\mathbf{x}$ be a net on a set X, $\mathbf{y}$ a subnet of X, x a point of X and A a subset of X.*
a) If $\mathbf{x}$ is eventually in A, then $\mathbf{y}$ is eventually in A.
b) If $\mathbf{x} \to x$, then $\mathbf{y} \to x$.
c) If $\mathbf{y}$ is cofinally in A, so is $\mathbf{x}$.
d) If x is a limit point of $\mathbf{y}$, it is also a limit point of $\mathbf{x}$.

3.2. Two examples of nets in analysis.

Example 5.8. *Let $A = \{a_i\}$ be an indexed family of real numbers, i.e., a function from a naked set S to $\mathbb{R}$. Can we make sense of the infinite series $\sum_{i \in S} a_i$? Note that we are assuming no ordering on the terms of the series, which may look worrisome, since in case $S = \mathbb{Z}^+$ it is well-known that the convergence of the series (and its sum)* will *in general depend upon the ordering relation on I we use to form the sequence of partial sums.*

Nevertheless, there is a nice answer. We say that the series $\sum_{i \in S} a_i$ **converges unconditionally** to $a \in \mathbb{R}$ if: for all $\epsilon > 0$, there exists a finite subset $J(\epsilon)$ of S such that for all finite subsets $J(\epsilon) \subset J \subset S$, we have $|a - \sum_{i \in J} a_i| < \epsilon$.

Exercise 5.10. a) *Show that if $sum_{i \in I} a_i$ is unconditionally convergent, then the set of indices $i \in I$ for which $a_i \neq 0$ is at most countable.*

b) *Suppose $I = \mathbb{Z}^+$. Show that a series converges unconditionally iff it converges absolutely, i.e., iff $\sum_{i=1}^{\infty} |a_i| < \infty$.*

c) *Define unconditional and absolute convergence of series in any real Banach space. Show that absolute convergence implies unconditional convergence, and find an example of a Banach space in which there exists an unconditionally convergent series which is not absolutely convergent.*[7]

The point is that this "new" type of limiting operation can be construed as an instance of net convergence. Namely, let $I(S)$ be the set of all finite subsets J of S,

[7]In fact, the celebrated Dvoretzky-Rogers theorem asserts that a Banach spaces admits an unconditionally but nonabsolutely convergent series iff it is infinite- dimensional.

directed under containment. Then given $\mathbf{a} : S \to \mathbb{R}$, we can define a net $\mathbf{x}$ on $I(S)$ by $J \mapsto \sum_{i \in J} a_i$. Then the unconditional convergence of the series is equivalent to the convergence of the net $\mathbf{x}$ in $\mathbb{R}$.

EXERCISE 5.11. *Suppose that we had instead decided to define $\sum_{i \in S} a_i$ converges unconditionally to a as: for all $\epsilon > 0$, there exists $N = N(\epsilon)$ such that for all finite subsets J of S with $\#J \geq N$ we have $|a - \sum_{iinJ} a_i| < \epsilon$.*
a) Show that this is again an instance of net convergence.
b) Is this equivalent to the definition we gave?

EXAMPLE 5.9. *The collection of all tagged partitions $(\mathcal{P}, x_i^*)$ of $[a, b]$ forms a directed set, under the relation of inclusion $\mathcal{P} \subset \mathcal{P}'$ ("refinement"). A function $f : [a, b] \to \mathbb{R}$ defines a net in $\mathbb{R}$, namely*

$$(\mathcal{P}, x_i^*) \mapsto R(f, \mathcal{P}, x_i^*),$$

the latter being the associated Riemann sum.[8] *The function f is Riemann-integrable to L iff the net converges to L.*

Such examples motivated Moore and Smith to develop their generalized convergence theory.

3.3. Universal nets. A net $\mathbf{x} : I \to X$ in a set X is said to be **universal**[9] if for any subset A of X, x is either eventually in A or eventually in $X \setminus A$.

EXERCISE 5.12. *Show that a net is universal iff whenever it is cofinally in a subset A, it is eventually in A.*

EXERCISE 5.13. *Let $\mathbf{x} : I \to X$ be a net, and let $f : X \to Y$ be a function.*
a) Show that if $\mathbf{x}$ is universal, so is the induced net $f(\mathbf{x}) = f \circ \mathbf{x}$.
b) Show that the converse need not hold.

EXERCISE 5.14. *Show that any subnet of a universal net is universal.*

EXAMPLE 5.10. *An eventually constant net is universal.*

Less trivial examples are difficult to come by. Note that a convergent net need not be universal: for instance, take the convergent sequence $\mathbf{x}_n = \frac{1}{n}$ in $[0, 1]$ and $A = \{1, \frac{1}{3}, \frac{1}{5}, \ldots\}$. Then the sequence is cofinal in both A and its complement so is not eventually in either one. Indeed, the same argument shows that a *sequence* which is universal is eventually constant.

Nevertheless, one has the following result:

THEOREM 5.18. *(Kelley) Every net admits a universal subnet.*

PROOF. Let $\mathbf{x}$ be a net in X, and consider all collections $\mathcal{A}$ of subsets of X such that:
(i) $Y_1, Y_2 \in \mathcal{A} \implies Y_1 \cap Y_2 \in \mathcal{A}$.
(ii) $Y_1 \in \mathcal{A}, Y_2 \supset Y_1 \implies Y_2 \in \mathcal{A}$.
(iii) $Y \in \mathcal{A} \implies \mathbf{x}$ is cofinal in Y.

---

[8] Moreover, all of the standard variations on the definitio of Riemann integrability – e.g. upper and lower sums – can be similarly described in terms of convergence of nets.

[9] Alternate terminology: **ultranet**.

The set of all such families is nonempty, since $\mathcal{A} = \{X\}$ is one. The collection of such families is therefore a nonempty poset under the relation $\mathcal{A}_1 \leq \mathcal{A}_2$ if $\mathcal{A}_1 \subset \mathcal{A}_2$. The union of a chain of such families is is immediately checked to be such family, so Zorn's Lemma entitles us to a family $\mathcal{A}$ which is not properly contained in any other such family. We claim that such an $\mathcal{A}$ has the following additional property: for any $A \subset X$, either $A \in \mathcal{A}$ or $X \setminus A \in \mathcal{A}$.

Indeed, suppose first that for every $Y \in \mathcal{A}$, $\mathbf{x}$ is cofinal in $A \cap Y$. Then the fmaily $\mathcal{A}'$ of all sets containing $A \cap Y$ for some $Y \in \mathcal{A}$ satisfies (i), (ii) and (iii) and contains $\mathcal{A}$, so by maximality $\mathcal{A}' = \mathcal{A}$ and hence $A = A \cap X$ is in $\mathcal{A}$ and $\mathbf{x}$ is cofinal in A.

So we may assume that there exists $Y \in \mathcal{A}$ such that $\mathbf{x}$ is not cofinal in $A \cap Y$, i.e., $\mathbf{x}$ is eventually in (so *a fortiori* is cofinal in) $X \setminus (A \cap Y)$. Then by the previous case, $X \setminus (Z \cap Y) \in \mathcal{A}$; by (ii) so too is

$$Y \cap (X \setminus A \cap Y) = Y \setminus (A \cap Y),$$

and then by (ii) we get $X \setminus A \in \mathcal{A}$.

Now we apply Kelley's Lemma (Lemma 5.17) to the net $\mathbf{x} : I \to X$ and the family $\mathcal{A}$: we get a subnet $y_\bullet$ which is eventually in each $A \in \mathcal{A}$. Since $\mathcal{A}$ has the property that for all A, either A or $X \setminus A$ lies in $\mathcal{A}$, this subnet is universal. $\square$

At this point, the reader who is not wondering "What on earth is the point of universal nets?" is either a genius, has seen the material before or is pathologically uncurious. The following results provide a hint:

PROPOSITION 5.19. *For a universal net* $\mathbf{x}$ *in a topological space, and* $x \in X$, *the following are equivalent:*
(i) x *is a limit point of* $\mathbf{x}$.
(ii) $\mathbf{x} \to x$.

PROOF. Of course (ii) $\implies$ (i) for all nets. Conversely, if x is a limit point of $\mathbf{x}$, then $\mathbf{x}$ is eventually in every neighborhood U of x. But then, by Exercise 3.3.1, universality implies that $\mathbf{x}$ is eventually in N. So $\mathbf{x} \to x$. $\square$

PROPOSITION 5.20. *Let* X *be a topological space. The following are equivalent:*
(i) Every net in X *admits a convergent subnet.*
(ii) Every net in X *has a limit point.*
(iii) Every universal net in X *is convergent.*

PROOF. This follows from previous results. Indeed, by Proposition 5.16 (i) $\implies$ (ii); by Proposition 5.19 (ii) $\implies$ (iii); and by Theorem 5.18 (iii) $\implies$ (i). $\square$

Recall that in the special case of metric spaces these conditions hold with "net" replaced by "sequence", and moreover they are equivalent to the **Heine-Borel condition** that every open cover admits a finite subcover (Theorem 5.7, which we have not yet proved). We shall now see that, for any topological space, our net-theoretic analogues of Proposition 5.20 are equivalent to the Heine-Borel condition.

4. Convergence and (Quasi-)Compactness

4.1. Net-theoretic criteria for quasi-compactness.

Definition: A family $\{U_i\}_{i\in I}$ of subsets of a set X is said to **cover** X (or **be a covering** of X) if $X = \bigcup_{i\in I} U_i$. A family $\{F_i\}_{i\in I}$ of subsets of a set X is said to satisfy the **finite intersection property** (FIP) if for every finite subset $J \subset I$, $\cap_{i\in J} F_i \neq \emptyset$.

Theorem 5.21. *For a topological space X, the following are equivalent:*
a) Every net in X admits a convergent subnet.
b) Every net in X has a limit point.
c) Every universal net in X is convergent.
d) X is quasi-compact: every open covering admits a finite subcovering.
e) For every family $\{F_i\}_{i\in I}$ of closed subsets satisfying the finite intersection property, we have $\cap_{i\in I} F_i \neq \emptyset$.

Proof. The equivalence of a), b) and c) has already been shown. The equivalence of d) and e) is "due to de Morgan": property d) becomes property e) upon setting $F_i = X \setminus U_i$, and conversely. Thus it suffices to show b) $\implies$ e) $\implies$ b).

Assume b), and let $\{F_i\}_{i\in I}$ be a family of closed subsets satisfying the finite intersection property. Then the index set I is directed under reverse inclusion. For each $i \in I$, choose any $x_i \in F_i$; the assignment $i \mapsto x_i$ is then a net $\mathbf{x}$ in X. Let x be a limit point of $\mathbf{x}$, and assume for a contradiction that there exists i such that x does not lie in F_i. Then $x \in U_i = X \setminus F_i$, and by definition of limit point there exists some index $j > i$ such that $x_j \in U_i$. But $j > i$ means $F_j \subset F_i$, so that $x_j \in F_j \cap U_i \subset F_i \cap U_i = (X \setminus U_i) \cap U_i = \emptyset$, contradiction! Therefore $x \in \cap_{i\in I} F_i$.

Now assume e) and let $\mathbf{x} : I \to X$ be a net in X. For each $i \in I$, define $F_i = \overline{\{x_j \mid j \geq i\}}$. Since directedness implies that given any finite subset J of I there exists some $i \in I$ such that $i \geq j$ for all $j \in J$, the family $\{F_i\}_{i\in I}$ of closed subsets satisfies the finite intersection condition. Thus by our assumption there exists $x \in \cap_{i\in I} F_i$. Let U be any neighborhood of x and take any $i \in I$. Then $x \in \overline{F_i}$, so that $F_i \cap U$ is nonempty. In other words, there exists $j \geq i$ such that $x_j \in U$, and this means that $\mathbf{x}$ is cofinal in U. Since U was arbitrary, we conclude that x is a limit point of $\mathbf{x}$. □

Theorem 5.22. a) *In a first countable space, limit point compactness implies sequential compactness.*
b) *In a metrizable space, sequential compactness implies quasi-compactness, and hence quasi-compactness, sequential compactness, limit point compactness, and countable compactness are all equivalent properties.*

Proof. Suppose first that X is first countable and limit point compact, and let $\mathbf{x}$ be a sequence in X. If the image of the sequence is finite, we can extract a constant, hence convergent, subsequence. Otherwise the image is an infinite subset of X, which (since quasi-compactness implies limit point compactness) has a limit point x, which is in particular a partial limit of the sequence. Then, as in any first countable space, Proposition 5.12 implies that there exists a subsequence converging to x.

Now suppose that X is sequentially compact. For each positive integer n, let T_n be a subset which is maximal with respect to the property that the distance between any two elements is at least $\frac{1}{n}$. (Such subsets exist by Zorn's Lemma.) It is clear that the set T_n can have no limit points, so (because sequential compactness implies limit point compactness) it must be finite. Since every point of X lies at a distance at most $\frac{1}{n}$ from some element of T_n, the set $\bigcup_n T_n$ is a countable dense subset. By Proposition **??** this implies that every open covering has a countable subcovering. But since sequential compactness implies countable compactness, this countable subcovering in turn has a finite subcovering, so altogether we have shown that X is quasi-compact. □

4.2. Products of quasi-compact spaces. Let $\{X_i\}_{i \in I}$ be a family of topological spaces. Recall that the product topology on the Cartesian product $X = \prod_i X_i$ is the topology whose subbase is the collection of all sets of the form $\pi_i^{-1}(U_i)$, where $\pi_i : X \to X_i$ is projection onto the ith factor and U_i is an open set in X_i.

An easy and important fact:

THEOREM 5.23. *Let $\mathbf{x} : J \to \prod_i X_i$ be a net in the product space $X = \prod_i X_i$. The following are equivalent:*

a) *The net $\mathbf{x}$ converges to $x = (x_i)$ in X.*
b) *For all i, the image net $\pi_i(\mathbf{x})$ converges to x_i in X_i.*

PROOF. Continuous functions preserve net convergence, so a) $\implies$ b). Conversely, suppose that $\mathbf{x}$ does not converge to x. Then there exists a finite subset $\{i_1, \ldots, i_n\}$ of I and open subsets U_{i_k} of x_{i_k} in X_{i_k} such that $\mathbf{x}$ is not eventually in $\cap_{k=1}^n \pi_{i_k}^{-1}(U_{i_k})$, which in fact means that for some k $\mathbf{x}$ is not eventually in $\pi_{i_k}^{-1}(U_{i_k})$. But then $\pi_{i_k}(\mathbf{x})$ is not eventually in U_{i_k} and hence does not converge to x_{i_k}. □

We can now prove one of the truly fundamental theorems in general topology.

THEOREM 5.24. *(Tychonoff Theorem) For a family $\{X_i\}_{i \in I}$ of nonempty topological spaces, the following are equivalent:*
a) Each factor space X_i is quasi-compact.
b) The Cartesian product $X = \prod_{i \in I} X_i$ is quasi-compact in the product topology.

PROOF. That b) implies a) follows from Exercise 4.1.1, since X_i is the image of X under the projection map X_i. Conversely, assume that each factor space X_i is quasi-compact. To show that X is quasi-compact, we shall use the notion of universal nets: by Theorem 5.21 it suffices to show that every universal net $\mathbf{x}$ on X is convergent. But since $\mathbf{x}$ is universal, by Exercise 3.3.2 each projected net $\pi_i(\mathbf{x})$ is universal on X_i. Since X_i is quasi-compact, Theorem 5.21 implies that $\pi_i(\mathbf{x})$ converges, say to x_i. But then by Theorem 5.23, $\mathbf{x}$ converges to $x = (x_i)$: done! □

This proof is due to J.L. Kelley [**Ke50**]. To my knowledge, it remains the outstanding application of universal nets.

EXERCISE 5.15 (Little Tychonoff). *: Let $\mathbf{x}_n$ be a sequence of metrizable spaces. Prove the Tychonoff theorem in this case by combining the following observations –*

(i) *A countable product of metrizable spaces is metrizable.*
(ii) *Sequential compactness is equivalent to quasi-compactness in metrizable spaces.*
(iii) *A sequence converges in a product space iff each projection converges –*

with a diagonalization argument. In particular, deduce the Heine-Borel theorem in $\mathbb{R}^n$ from the Heine-Borel theorem in $\mathbb{R}$.

EXERCISE 5.16. *Investigate to what extent the Axiom of Choice (AC) is used in the proof of Tychonoff's theorem. Some remarks:*
*a) The use of Zorn's Lemma in the proof that every net has a universal subnet is unavoidable in the sense that this assertion is known to be equivalent to the **Boolean Prime Ideal Theorem** (BPIT). BPIT is known to require AC (in the sense of being unprovable from Zermelo-Frankel set theory) but not to imply it (a similar meaning).*
b) A cursory look at the proof might then suggest that BPIT implies Tychonoff's theorem. However, it is a famous observation of Kelley that Tychonoff's theorem implies *AC,*[10] *so this cannot be the case. So AC must get invoked again in the proof of Tychonoff. Where?*
c) Hint: BPIT does imply that arbitrary products of quasi-compact Hausdorff spaces are quasi-compact Hausdorff!

5. Filters

5.1. Filters and ultrafilters on a set. Let X be a set. A **filter** on X is a nonempty family $\mathcal{F}$ of nonempty subsets of X satisfying

(F1) $A_1,\ A_2 \in \mathcal{F} \implies A_1 \cap A_2 \in \mathcal{F}$.
(F2) $A_1 \in \mathcal{F}$, $A_2 \supset A_1 \implies A_2 \in \mathcal{F}$.

EXAMPLE 5.11. *For any nonempty subset Y of X, the collection $\mathcal{F}_Y = \{A \mid Y \subset A\}$ of all subsets containing Y is a filter on X. Such filters are said to be **principal**.*

EXERCISE 5.17. *Show that every filter on a finite set is principal. (Hint: if $\mathcal{F}$ is a filter on the finite set X then $\cap_{A\in\mathcal{F}} A \in \mathcal{F}$.)*

EXAMPLE 5.12. *For any infinite set X, the family of all cofinite subsets of X is a filter on X, called the **Fréchet filter**.*

EXERCISE 5.18. *A filter $\mathcal{F}$ on X is **free** if $\cap_{A\in\mathcal{F}} A = \emptyset$.*

a) *Show that a principal filter is not free.*
b) *Show that a filter is free iff it contains the Fréchet filter.*

EXAMPLE 5.13. *If X is a topological space and $x \in X$, then the collection $\mathcal{N}_x$ of neighborhoods of x is a (nonfree) filter on X. It is principal iff x is an isolated point of X. More generally, if Y is a subset of X, then the collection $\mathcal{N}_Y$ of neighborhoods of Y (recall that we say that N is a neighborhood of Y is $Y \subset N^\circ$) is a nonfree filter on X, which is principal iff Y is an open subset.*

EXERCISE 5.19.
a) *Let $\{\mathcal{F}_i\}_{i\in I}$ be an indexed family of filters on a set X. Show that $\cap_{i\in I}\mathcal{F}_i$ is a filter on X, the largest filter which is contained in each $\mathcal{F}_i$.*
b) *Let X be a set with cardinality at least 2. Exhibit filters $\mathcal{F}_1$, $\mathcal{F}_2$ on X such that there is no filter containing both $\mathcal{F}_1$ and $\mathcal{F}_2$.*

---

[10] It is sometimes said that this is not surprising, since without AC the Cartesian product might be empty. But I have never understood this remark, since the empty set is of course quasi-compact. At any rate, the proof is not trivial.

The collection of all filters on a set X is partially ordered under set-theoretic containment. Exercise 5.19a) shows that in this poset arbitrary *joins* exist – i.e., any collection of filters admits a greatest lower bound – whereas Exercise 5.19b) shows that if $\#X > 1$ the collection of filters on X is not a directed set. If $\mathcal{F}_1 \subset \mathcal{F}_2$ we say that $\mathcal{F}_2$ **refines** $\mathcal{F}_1$, or is a **finer filter** than $\mathcal{F}_1$. An **ultrafilter** on X is a filter on X which is maximal with respect to this ordering, i.e., is not properly contained in any other filter.

Exercise 5.20. *Let Y be a nonempty subset of X. Then the principal filter $\mathcal{F}_Y$ is an ultrafilter iff $\#Y = 1$.*

If X is finite, this gives all the ultrafilters on X. More precisely, the ultrafilters on a finite set may naturally be identified with the elements x of X. However, if X is infinite (the case of interest to us here) there are a great many nonprincipal ultrafilters.

Lemma 5.25. *Any filter is contained in an ultrafilter.*

Proof. Since the union of a chain of filters is itself a filter, this follows from Zorn's Lemma. □

Proposition 5.26. *For a filter $\mathcal{F}$ on X, the following are equivalent:*

(i) *For every subset Y of X, $\mathcal{F}$ contains exactly one of Y and $X \setminus Y$.*
(ii) *$\mathcal{F}$ is an ultrafilter.*

Proof. If a filter $\mathcal{F}$ satisfies (i) and Y is any subset of X which is not an element of $\mathcal{F}$, then $X \setminus Y \in \mathcal{F}$, and since any finer filter $\mathcal{F}'$ would contain $X \setminus Y$, by (F1) it certainly cannot contain Y; i.e., $\mathcal{F}$ is not contained in any finer filter. Conversely, suppose that $\mathcal{F}$ is an ultrafilter and Y is a subset of X. Suppose first that for every $A \in \mathcal{F}$ we have $A \cap Y \neq \emptyset$. Then the family $\mathcal{F}'$ of all sets containing a set $A \cap Y$ with $A \in \mathcal{F}$ is easily seen to be a filter containing $\mathcal{F}$. Since $\mathcal{F}$ is an ultrafilter we have $\mathcal{F}' = \mathcal{F}$ and in particular $Y = Y \cap X \in \mathcal{F}$. Otherwise there exists an $A \in \mathcal{F}$ such that $A \cap Y = \emptyset$. Then $A \subset X \setminus Y$ and by (F2) $X \setminus Y \in \mathcal{F}$. □

Corollary 5.27. *A nonprincipal ultrafilter is free.*

Proof. If there exists $x \in \bigcap_{A \in \mathcal{F}} A$, then $X \setminus \{x\}$ is not an element of $\mathcal{F}$, so by Proposition 5.26 $\{x\} \in \mathcal{F}$ and $\mathcal{F} = \mathcal{F}_{\{x\}}$. □

Thus free ultrafilters exist on any infinite set: by Lemma 5.25 the Fréchet filter is contained in some ultrafilter, and any refinement of a free filter is free. To be sure, a free ultrafilter is a piece of set-theoretic devilry: it has the impressively decisive ability to, given any subset Y of X, select exactly one of Y and its complement $X \setminus Y$. A bit of thought suggests that even on $X = \mathbb{Z}^+$ this will be difficult or impossible to do in any constructive way. And indeed Lemma 5.25 is known to be equivalent to the Boolean Prime Ideal Theorem, so that it requires (but is not equivalent to) the Axiom of Choice.

Exercise 5.21. *Every filter is the intersection of the ultrafilters containing it.*

5.2. Another Characterization of Ultrafilters. Let us say that an indexed family $\{Y_i\}_{i\in I}$ (with $I \neq \varnothing$) of subsets of a set X is a **pseudo-partition** of X if for all $i \neq j \in I$ we have $Y_i \cap Y_j = \varnothing$ and $\bigcup_{i\in I} Y_i = X$.[11] A **finite pseudo-partition** is a pseudo-partition with finite index set I, and for $n \in \mathbb{Z}^+$, an **n-pseudo-partition** is a finite pseudo-partition with $\# I = n$.

Proposition 5.28. *For a nonempty set X and a set $\mathcal{F}$ of subsets of X, the following are equivalent:*

(i) *The set $\mathcal{F}$ is an ultrafilter on X.*
(ii) *For every finite pseudo-partition $\{Y_i\}_{i\in I}$ of X, there is exactly one $i \in I$ such that $Y_i \in \mathcal{F}$.*
(iii) *For every three pseudo-partition $\{Y_1, Y_2, Y_3\}$ of X, there is exactly one $1 \leq i \leq 3$ such that $Y_i \in \mathcal{F}$.*

Proof. (i) $\Longrightarrow$ (ii): let $\mathcal{F}$ be an ultrafilter on X. For any subset $Y \in \mathcal{F}$, if $Y = A \coprod B$ is a disjoint union of two subsets, then exactly one of A and B must lie in $\mathcal{F}$: indeed certainly not both, since A and B are disjoint, and if $A \notin \mathcal{F}$ then by Proposition 5.26 we have

$$(X \setminus A) \cap Y = B \in \mathcal{F}.$$

An induction argument extends this statement to: if for any $2 \leq n \leq \aleph_0$ we have $Y \in \mathcal{F}$ and $Y = \coprod_{i=1}^n A_i$, then there is exactly one $1 \leq i \leq n$ such that $A_i \in \mathcal{F}$. Applying this with $Y = X \in \mathcal{F}$ completes this implication.
(ii) $\Longrightarrow$ (iii): This is immediate: what holds for all positive integers n holds for $n = 2$. (iii) $\Longrightarrow$ (i): First consider the 3-pseudo-partition $Y_1 = X, Y_2 = Y_3 = \varnothing$. Since exactly one lies in $\mathcal{F}$ and $Y_2 = Y_3$, we have $Y_1 = X \in \mathcal{F}$ and $\varnothing \notin \mathcal{F}$. Now let $Y \subset X$ and take $Y_1 = Y$, $Y_2 = X \setminus Y$, $Y_3 = \varnothing$: it follows that $\mathcal{F}$ contains exactly one of Y and $X \setminus Y$.

We claim that if $Y_1, Y_2 \in \mathcal{F}$ then $Y_1 \cap Y_2 \neq \varnothing$. Indeed if $Y_1 \cap Y_2 = \varnothing$, then $Y_1, Y_2, Y_3 := X \setminus (Y_1 \cup Y_2)$ is a 3-pseudo-partition of X in which $\mathcal{F}$ contains at least two of the elements, a contradiction.

Next suppose that $A \in \mathcal{F}$ and $A \subset B \subset \mathcal{F}$. If $B \notin \mathcal{F}$ then $X \setminus B \in \mathcal{F}$ but $A \cap (X \setminus B) = \varnothing$, contradiction. So $B \in \mathcal{F}$.

Finally suppose that $Y_1, Y_2 \in \mathcal{F}$. Since Y_1, $Y_2 \setminus Y_1$, $X \setminus (Y_1 \cup Y_2)$ is a 3-pseudopartition of X and $Y_1 \in \mathcal{F}$ we have $Y_2 \setminus Y_1 \notin \mathcal{F}$. Since $Y_1 \cap Y_2$, $Y_2 \setminus Y_1$, $X \setminus Y_2$ is a 3-pseudopartition of X and neither $Y_2 \setminus Y_1$ nor $X \setminus Y_2$ lie in X, we must have $Y_1 \cap Y_2 \in \mathcal{F}$. So $\mathcal{F}$ is an ultrafilter. $\square$

In Propostion 5.28(iii) we can replace all instances of 3 with any other fixed $n \geq 4$, but we cannot replace it with $n = 2$: the family of subsets of size at least $N + 1$ of a set of finite order $2N + 1$ gives a counterexample.

5.3. Prefilters.

Proposition 5.29. *For a family F of nonempty subsets of a set X, the following are equivalent:*

[11] Why not just a partition? For minor reasons related to the empty set: first, we are allowing $Y_i = \varnothing$, which is not allowed in a partition. Second, in an indexed family the same element is allowed to appear multiple times. In this case a nonempty subset cannot appear more than once because then the pairwise disjointness condition would be violated. But the empty set is allowed to appear more than once.

(i) *For all* $A_1,\ A_2 \in F$, *there exists* $A_3 \in F$ *such that* $A_3 \subset A_1 \cap A_2$.
(ii) *The collection of all subsets that contain some element of* F *is a filter.*

EXERCISE 5.22. *Prove Proposition 5.29.*

We shall call a family F of nonempty subsets satisfying (i) a **prefilter**.[12] The collection $\mathcal{F}$ of all supersets of F is called the **filter generated by** F (or sometimes the **associated filter**). Note that the situation is reminiscent of the criterion for a family of subsets to be the base for a topology.

EXAMPLE 5.14. *Let* X *be a set and* $x \in X$. *Then* $F = \{\{x\}\}$ *is a prefilter on* X *(which might be called "constant"). The filter it generates is the principal ultrafilter* $\mathcal{F}_x$.

EXAMPLE 5.15. *Let* X *be a topological space and* Y *a subset of* X. *Then the collection* N_Y *of all open neighborhoods of* Y *(i.e., open sets containing* Y*) is a prefilter, whose associated filter is the neighborhood filter* $\mathcal{N}_Y$ *of* Y.

Our choice of terminology "prefilter" rather than "filter base" is motivated by the following principle: if we have in mind a certain property P of filters and we are seeking an analogous property for prefilters, then we need merely to define a prefilter to have property P if the filter it generates has property P. Then, if necessary, we unpack this definition more explicitly.

For instance, we can use this perspective to endow the collection of prefilters on X with a quasi-ordering: we say that a prefilter F_2 **refines** F_1 and write $F_1 \leq F_2$ if for the corresponding filters $\mathcal{F}_1$ and $\mathcal{F}_2$ we have $\mathcal{F}_1 \subset \mathcal{F}_2$. It is not hard to see that this holds iff for every $A_1 \in F_1$ there exists $A_2 \in F_2$ such that $A_1 \supset A_2$. If $F_1 \leq F_2 \leq F_1$ we say that F_1 and F_2 are **equivalent** prefilters and write $F_1 \sim F_2$.

EXERCISE 5.23. *If* $\#X \geq 2$, *show: there are prefilters* F_1 *and* F_2 *on* X *such that* $F_1 \sim F_2$ *but* $F_1 \neq F_2$.

Similarly we say a prefilter F on X is **ultra** if its associated filter is an ultrafilter. This amounts to saying that for any $Y \subset X$, there exists $A \in F$ such that either $A \subset Y$ or $A \subset (X \setminus Y)$.

EXERCISE 5.24. *(Filter subbases):*
a) Show that for a family I *of nonempty subsets of a set* X, *the following are equivalent:*
(i) I *has the finite intersection property: if* $A_1, \ldots, A_n \in I$, *then* $A_1 \cap \ldots A_n \neq \emptyset$.
(ii) There exists a prefilter F *such that* $I \subset F$.
(iii) There exists a filter $\mathcal{F}$ *such that* $I \subset \mathcal{F}$.
b) If I *satisfies the equivalent conditions of part a), show that there is a unique minimal filter* $\mathcal{F}(I)$ *containing* I, *called the* ***filter generated by*** I.

A family $\{F_i\}_{i \in I}$ of prefilters on a set X is **compatible** if there exists a prefilter $F \supset \bigcup_{i \in I} F_i$, i.e., if $\bigcup_{i \in I} F_i$ is a filter subbase. (It is equivalent to require that $\bigcup_{i \in I} F_i$ be refined by some prefilter.) In turn, this occurs iff for every finite subset $J \subset I$ and any assignment $j \mapsto A_j \in F_j$ we have $\bigcap_{j \in J} A_j \neq \emptyset$.

[12]The more traditional terminology is **filter base**. We warn that this terminology is often used in the literature for something else.

5.4. Convergence via filters.

Let F be a prefilter in a topological space X, and let x be a point of X. We say F **converges** to x – and write $F \to x$ – if F *refines* the neighborhood filter $\mathcal{N}_x$ of x. In This means that every neighborhood N of x contains an element A of F.

Let F be a prefilter in a topological space X, and let x be a point of X. We say that x is a **limit point**[13] of F if F *is compatible with* the neighborhood filter $\mathcal{N}_x$, or in plainer language, if every element of F meets every neighborhood of x.

Proposition 5.30. *Let $\mathcal{F}$ be a prefilter on X with associated filter $\mathcal{F}$, and let $F' \geq F$ be a finer prefilter.*

a) *If F converges to x, then x is a limit point of F.*
b) *F converges to $x \iff \mathcal{F}'$ converges to x.*
c) *x is a limit point of $F \iff x$ is a limit point of $\mathcal{F}$.*
d) *If F converges to x, then F' converges to x.*
e) *If x is a limit point of F', then x is a limit point of F.*
f) *X is Hausdorff $\iff$ every prefilter on X converges to at most one point.*

Exercise 5.25. *Prove Proposition 5.30.*

Exercise 5.26. *Show: for a topological space X, the following are equivalent:*
(i) *X has the trivial topology.*
(ii) *Every filter on X converges to every point of X.*

Exercise 5.27. *Let X be a topological space.*

a) *Show: the following are equivalent:*
   (i) *X is **Alexandroff**: $x \in X$ has a minimal neighborhood.*
   (ii) *For all $x \in X$, the neighborhood filter $\mathcal{N}_x$ is principal.*
b) *(E. Wofsey) Show: the following are equivalent:*
   (i) *Every convergent filter on X is principal.*
   (ii) *X is **locally finite**: every point of X has a finite neighborhood.*
c) *Show: finite implies locally finite implies Alexandroff, and neither implication can be reversed.*

Proposition 5.31. *Let F be a prefilter on X. The following are equivalent:*
(i) *x is a limit point of F.*
(ii) *There is a refinement F' of F such that F' converges to x.*

Proof. (i) $\implies$ (ii): If x is a limit point of F, there exists a prefilter F' refining both F and $\mathcal{N}_x$, and then F' is a finer prefilter converging to x.
(ii) $\implies$ (i): since $F' \to x$, x is a limit point of F' (Proposition 5.30a)), and since $F' \geq F$, x is a limit point of F (Proposition 5.30e)). $\square$

Proposition 5.32. *Let X be a topological space, Y a nonempty subset of X and x a point of x. The following are equivalent:*
(i) *x is a limit point of the prefilter $F_Y = \{Y\}$.*
(ii) *$x \in \overline{Y}$.*

Proof. Both (i) and (ii) say that every neighborhood of x meets Y. $\square$

[13]Alternate terminology: **cluster point**

A more traditional characterization of closure using filters is the following:

Corollary 5.33. *Let X be a topological space, Y a nonempty subset of X and x a point of x. The following are equivalent:*

(i) *We have $x \in \overline{Y}$.*
(ii) *There is a prefilter F on X consisting of subsets of Y such that $F \to x$.*
(iii) *There is a filter $\mathcal{F}$ on X such that $\mathcal{F} \to x$ and $Y \in \mathcal{F}$.*

Proof. (i) $\Longrightarrow$ (ii): We may take $F := \{N \cap Y \mid N \text{ is a neighborhood of } x\}$.
(ii) $\Longrightarrow$ (iii): We may take $\mathcal{F}$ to be the filter generated by F.
(iii) $\Longrightarrow$ (i): Since $\mathcal{F} \to x$, for every neighborhood N of x we have $N \in \mathcal{F}$. Since $Y \in \mathcal{F}$, we have $N \cap Y \in \mathcal{F}$ and thus $N \cap Y \neq \varnothing$. $\square$

Proposition 5.34. *Let X be a topological space, Y a nonempty subset of X and x a point of x. The following are equivalent:*
(i) The prefilter $F_Y = \{Y\}$ is compatible with the neighborhood filter $\mathcal{N}_x$ of x.
(ii) $x \in \overline{Y}$.

Proof. Each of (i) and (ii) says that every neighborhood of x meets Y. $\square$

Lemma 5.35. *If an ultra prefilter F has x as a limit point, then $F \to x$.*

Proof. As above, there is a prefilter F' refining both F and $\mathcal{N}_x$. But since F is ultra, it is equivalent to all of its refinements, so that F itself refines $\mathcal{N}_x$. $\square$

It may not come as a surprise that we can get further characterizations of quasi-compactness in terms of convergence / limit points of prefilters.

Theorem 5.36. *For a topological space X, the following are equivalent:*
(i) X satisfies the equivalent conditions of Theorem 5.21 ("X is quasicompact.")
(ii) Every prefilter on X has a limit point.
(iii) Every ultra prefilter on X is convergent.
The same equivalences hold with "'prefilter" replaced by "filter" in (ii) and (iii).

Proof. (i) $\Longrightarrow$ (ii): Let $F = \{A_i\}$ be a prefilter on X. The sets A_i satisfy the finite intersection property, hence *a fortiori* so do their closures. Appealing to condition e) in Theorem 5.21 there is an $x \in \bigcap_i \overline{A_i}$, and this means precisely that each A_i meets each neighborhood of x.
(ii) $\Longrightarrow$ (iii) follows immediately from Lemma 5.35.
(iii) $\Longrightarrow$ (i): Consider a family $I = \{F_i\}$ of closed subsets of X satisfying the finite intersection condition. Then I is a filter subbase, so that there exists some ultra prefilter refining I. By hypothesis, there exists $x \in X$ such that F converges to x, and *a fortiori* x is a limit point of F. So every element of F – and in particular each F_i – meets every neighborhood of x, so that $x \in \overline{F_i} = F_i$. Therefore $\cap_i F_i$ contains x and is thus nonempty.

The fact that the results hold also for filters instead of prefilters is easy and left to the reader. $\square$

Corollary 5.37. *Let F be a prefilter on the quasi-compact space X.*
a) If F does not *converge to a point $x \in X$, then F has a limit point $y \neq x$.*
b) If F has at most one limit point, it is convergent.
c) A filter on a compact space converges iff it has a unique limit point.

PROOF. a) If F does not converge to x, then there is an open neighborhood U of x which does not contain any element of F. Let $Y = X \backslash U$, and put $F_Y = \{A \cap Y \mid A \in F\}$. Then F_Y is a prefilter on Y: if $A \in F$ and $A \cap Y = \varnothing$ then $A \subset U$. For $A_1, A_2 \in F$, if $A_3 \in F$ is such that $A_3 \subset A_1 \cap A_2$ then $A_3 \cap Y \subset (A_1 \cap Y) \cap (A_2 \cap Y)$. Since Y is a closed subspace of the quasi-compact space X, F_Y has a limit point $y \in Y$. If now $\mathcal{N}$ is a neighborhood of y in X, then $\mathcal{N} \cap Y$ is a neighborhood of y in Y, so for all $A \in F$, $(A \cap Y) \cap (\mathcal{N} \cap Y) \neq \varnothing$, hence $A \cap Y \neq \varnothing$. It follows that y is a limit point of F. Since $x \in U$ and $y \in X \setminus U$, $y \neq x$.
b) Keeping in mind that by Theorem 5.36 F must have at least one limit point, this follows immediately from part a).
c) This follows from part b) and the uniqueness of limits in Hausdorff spaces. $\square$

Pushing forward filters: if $f : X \to Y$ is any map of sets and $I = \mathcal{A}_\rangle$ is a family of subsets of X, then by $f(I)$ we mean the family $\{f(A_i)\}_{i \in I}$.

PROPOSITION 5.38. *Let $f : X \to Y$ be a function and F a prefilter on X.*
a) $f(F)$ is a prefilter on Y.
b) If F is ultra, so is $f(F)$.

EXERCISE 5.28. *Prove Proposition 5.38.*

PROPOSITION 5.39. *Let $f : X \to Y$ be a function. The following are equivalent:*
(i) For every prefilter F on X with a limit point x, $f(F)$ has $f(x)$ as a limit point.
(ii) For every prefilter F on X converging to x, $f(F)$ converges to $f(x)$.
(iii) f is continuous.

PROOF. A function f between topological spaces is continuous iff for all $x \in X$, $f(\mathcal{N}_x)$ is a neighborhood base for Y. The result follows easily from this and is left to the reader. $\square$

Let $\{X_i\}_{i \in I}$ be an indexed family of topological spaces and suppose given a prefilter F_i on each X_i. We then define the **product prefilter** $\prod_i F_i$ to be the family of subsets of X of the form $\prod_{i \in I} M_i$, where there exists a finite subset $J \subset I$ such that $M_i = X_i$ for all $i \in I \setminus J$ and $M_i \in F_i$ for all $i \in J$. Since

$$(\prod_{i \in I} M_i) \cap (\prod_{i \in I} M_i') = \prod_{i \in I} (M_i \cap M_i') \supset \prod_{i \in I} M_i''$$

where M_i'' is an element of F_i contained in $M_i' \cap M_i''$ (or is X_i if $M_i = M_i'' = X_i$), this does indeed give a prefilter on X. Another way around is to say that F is the prefilter generated by taking finite intersections of the filter subbase $\pi_i^{-1}(M_i)$.

EXERCISE 5.29. *a) If for each i we are given equivalent prefilters $F_i \sim F_i'$ on X_i, then the product prefilter $\prod_i F_i$ is equivalent to $\prod_i F_i'$.*
b) (Remark): Because of part a), as far as convergence / limit points are concerned, it would be no loss of generality to assume that $X_i \in F_i$ for all i, and then we get a cleaner definition of the product prefilter.

THEOREM 5.40. *Let F be a prefilter on the product space $X = \prod_{i \in I} X_i$. The following are equivalent:*

(i) *F converges to $x = (x_i)$.*
(ii) *For all i, $\pi_i(F)$ converges to x_i.*

Proof. (i) $\implies$ (ii) is immediate from Proposition 5.39, so assume (ii). It is enough to show that for every $i \in I$ and every neighborhood N_{ij} of x_i in X_i there exists an element $A \in F$ with $\pi_i(A) \subset N_{ij}$, for then F will be a prefilter which is finer than the family $\pi_i^{-1}(N_{ij})$ which is a subbasis for the filter of neighborhoods of x in X. But this is tautological: since $\pi_i(F)$ converges to x_i, it contains an element, say $B = \pi_i(A)$, which is contained in N_{ij}, and then $A \subset \pi_i^{-1}(N_{ij})$. $\square$

Now for a proof of Tychonoff's Theorem (Theorem 5.24) using filters:

That b) implies a) follows from Exercise 4.1.1, since X_i is the image of X under the projection map X_i. Conversely, assume that each factor space X_i is quasi-compact. To show that X is quasi-compact, we shall use the notion of ultra prefilters: by Theorem 5.36 it suffices to show that every ultra prefilter F on X is convergent. Since F is ultra, by Proposition 5.38b) each projected prefilter $\pi_i(F)$ is ultra on X_i. Since X_i is quasi-compact, Theorem 5.36 implies that $\pi_i(F)$ converges, say to x_i. But then by Theorem 5.40, F converges to $x = (x_i)$: done!

This proof is due to H. Cartan [**Ca37**].

6. A characterization of quasi-compactness

Theorem 5.41. *For a topological space Y, the following are equivalent:*

(i) *The space Y is quasi-compact.*
(ii) *For all topological spaces X, the projection map $\pi_X : X \times Y \to X$ is closed.*

6.1. Proof of (i) $\implies$ (ii). Let $C \subset X \times Y$ be closed, and let $x_0 \in X \setminus \pi_X(C)$. Then $\mathcal{N} = (X \times Y) \setminus C$ is a neighborhood of $\{x_0\} \times Y$. By the Tube Lemma, there is a neighborhood U of x_0 in X such that $U \times Y \subset \mathcal{N}$, and then U is a neighborhood of x_0 in X which is disjoint from $\pi_X(C)$.

6.2. Proof of (ii) $\implies$ (i) using filters.

Let $\mathcal{F}$ be a filter on Y. Let $\star$ be a point which is not in Y, and let X be the set $Y \coprod \{\star\}$. We topologize X as follows: every subset not containing $\star$ is open; a subset $A \subset X$ containing $\star$ is open iff $A \setminus \star \in \mathcal{F}$. Since $\varnothing \notin \mathcal{F}$, $\{\star\}$ is not open in X and thus it lies in the closure of Y.

$$D = \{(y, y) \mid y \in Y\} \subset X \times Y,$$

and let

$$E = \overline{D}.$$

For any closed map $f : \mathcal{X} \to \mathcal{Y}$ of topological spaces and subset $A \subset \mathcal{X}$ we have $f(\overline{A})$ is closed and thus

$$\overline{f(A)} \subset f(\overline{A}) \subset \overline{f(A)},$$

so

$$f(\overline{A}) = \overline{f(A)}.$$

Since π_X is closed by assumption, we have

$$\pi_X(E) = \overline{\pi_X(D)} = \overline{Y} = X.$$

It follows that there is $y \in Y$ such that $(\star, y) \in E$. We claim that y is a limit point of $\mathcal{F}$. Indeed, let V be a neighborhood of y in Y, and let $M \in \mathcal{F}$. Then $\mathcal{N} = (M \coprod \{\star\}) \times V$ is a neighborhood of $(\star, y)$ in $Y \times X$. Since $(\star, y) \in E = \overline{D}$, there is $z \in Y$ such that $(z, z) \in \mathcal{N}$ and thus $z \in M \cap V$.

EXERCISE 5.30. *a) Observe that our proof of (i) $\implies$ (ii) in Theorem 5.41 used only that the conclusion of the Tube Lemma holds for Y. Combining with (ii) $\implies$ (i), observe that if a topological space satisfies the conclusion of the Tube Lemma, it is quasi-compact.*
b) The structure of the above argument was: quasi-compact $\implies$ Tube Lemma $\implies$ projections are closed $\implies$ quasi-compact, and part a) follows by going two steps around this triangle. Give a much shorter direct proof that closedness of projections implies the Tube Lemma.

7. The correspondence between filters and nets

Take a moment and compare Cartan's ultra prefilter proof with Kelley's universal net proof. By replacing every instance of "universal net" with "ultra prefilter" they become word for word identical! This, together with the other manifest parallelisms between §3 and §5, strongly suggests that nets and prefilters are not just different means to the same end but are somehow directly related: given a net, there ought to be a way to trade it in for a prefilter, and vice versa, in such a way as to preserve the concepts of: convergence, limit point, subnet / finer prefilter and universal net / ultra prefilter. This is exactly the correspondence that we now pursue.

If we search the preceding material for hints of how to pass from a net to a prefilter, sooner or later we will notice that we have already done so in the proof that b) $\implies$ e) in Theorem 5.21. We repeat that construction here, after introducing the following useful piece of notation.

If $\leq$ is a relation on a set I, for $i \in I$ we put $i^+ = \{i' \in I \mid i \leq i'\}$.

PROPOSITION 5.42. *Let $\mathbf{x} : I \to X$ be a net in the set X. Then the collection $\mathcal{P}(\mathbf{x}) := \{i^+\}_{i \in I}$ is a prefilter on X, the **prefilter of tails** of $\mathbf{x}$.*

PROOF. Indeed, for $i_1,\ i_2 \in I$, choose $i_3 \geq i_1,\ i_2$. Then $A_{i_3} \subset A_{i_1} \cap A_{i_2}$. $\square$

Conversely, suppose we are given a prefilter F on X: how to get a net? Evidently the first (and usually harder) task is to find the directed index set I and the second is to define the mapping $I \to X$. The key observation is that the condition $A_1,\ A_2 \in F \implies \exists A_3 \in F \mid A_3 \subset A_1 \cap A_2$ on a nonempty family of nonempty subsets of X says precisely that the elements of F are (like the neighborhoods of a point) directed under reverse inclusion. This suggests that we should take $I = F$. Then to get a net we are supposed to choose, for each $A \in F$, some element x_A of X. Other than to require $x_A \in A$, no condition presents itself. Making many arbitrary choices is dismaying, on the one hand for set-theoretic reasons but moreover because we shall inevitably have to worry about whether our choices are correct. So let's worry: once we have our net $\mathbf{x}(F)$, we can apply the previous construction to get another prefilter $\mathcal{P}(\mathbf{x}(F))$, and whether we dare to admit it out loud or not, we are clearly hoping that $\mathcal{P}(\mathbf{x}(F)) = F$.

Let us try our luck on the simplest possible example: let X be a set with more than one element, and let $F = \{X\}$, the unique *minimal* filter. A net $\mathbf{x}$ with index set F is just a choice of a point $x \in X$. The corresponding prefilter $\mathcal{P}(\mathbf{x})$ – namely the principal prefilter $F_x = \{x\}$ – is not only not equal to F, it is *ultra*: its associated filter is *maximal*. At least we don't have to worry about our choice of x_A in A: all choices fail equally.

We trust that we have now suitably motivated the correct construction:

PROPOSITION 5.43. *Let F be a prefilter on X. Let $I(F)$ be the set of all pairs (x, A) such that $x \in A \in F$. We endow $I(F)$ with the relation $(x_1, A_1) \leq (x_2, A_2)$ iff $A_1 \supset A_2$. Then $(I(F), \leq)$ is a directed set, and the assignment $(x, A) \mapsto x$ defines a net $x(F) : I \to X$.*

EXERCISE 5.31. *Prove Proposition 5.43.*

Coming back to our earlier example, if $F = \{X\}$, then $\mathbf{x}(F)$ has domain $I = \{X\} \times X$ and is just $(X, x) \mapsto x$. Note that the induced quasi-ordering on X makes $x \leq x'$ for any x, x': notice that it *is* directed and *is not* anti-symmetric (which at last justifies our willingness to entertain directed *quasi*-ordered sets). So for any $x \in X$, we have $(X, x)^+ = \{x' \geq x\} = X$, and we indeed get $\mathcal{P}(\mathbf{x}(F)) = \{X\} = F$. This was not an accident:

PROPOSITION 5.44. *For any prefilter G on X, we have $F(x(G)) = G$.*

PROOF. The index set of $x(G)$ consists of all pairs (x, A) for $x \in A \in F$, partially ordered under reverse inclusion. The associated prefilter consists of sets $A_{(x,A)} = \{\pi_1((x', A')) \,|(x', A') \geq (x, A)\}$. A moment's thought reveals this to be the set of all points x in filter elements $A' \subset A$, i.e., $A_{(x,A)} = A$. □

What about the relation $x(F(\mathbf{x})) = \mathbf{x}$? A moment's thought shows that this cannot possibly hold: the index set I of any net associated to a prefilter on X is a subset of $X \times 2^X$ hence has cardinality at most $\#(X \times 2^X)$ (i.e., $2^{\#X}$ is X is infinite), but every nonempty set admits nets based on index sets of arbitrarily large cardinality, e.g. constant nets. Indeed, if $\mathbf{x} : I \to X$ has constant value $x \in X$, then the associated prefilter $F(\mathbf{x})$ is just $\{x\}$, and then the associated net $x(F(\mathbf{x}))$ has $I = \{(x, \{x\})\}$, a one point set!

EXERCISE 5.32. *Suppose a net $\mathbf{x}$ is eventually constant, with eventual value $x \in X$.*
a) Show that the filter generated by $F(\mathbf{x})$ is the principal ultrafilter $\mathcal{F}_x$.
b) Suppose that F is a prefilter generating the principal ultrafilter $\mathcal{F}_x$
(i.e., $\{x\} \in F$!). Show that $\mathbf{x}(F)$ is eventually constant with eventual value x.

Nevertheless the nets $\mathbf{x}$ and $x(F(\mathbf{x}))$ are "pan-topologically equivalent" in the sense that they converge to the same points and have the same limit points for any topology on X. Indeed:

THEOREM 5.45. *Let X be a topological space, F be a prefilter on X, $\mathbf{x}$ a net on X and $x \in X$. Then:*

a) *F converges to $x \iff x(F)$ converges to x.*
b) *$\mathbf{x}$ converges to $x \iff F(\mathbf{x})$ converges to x.*
c) *x is a limit point of $F \iff x$ is a limit point of $x(F)$.*
d) *x is a limit point of $\mathbf{x} \iff x$ is a limit point of $F(\mathbf{x})$.*

e) *F is an ultra prefilter $\iff$ $x(F)$ is a universal net.*
f) *$\mathbf{x}$ is a universal net $\iff$ $F(\mathbf{x})$ is an ultra prefilter.*
g) *If $\mathbf{y}$ is a subnet of $\mathbf{x}$, then $F(\mathbf{y})$ refines $F(\mathbf{x})$.*

EXERCISE 5.33. *Prove Theorem 5.45.*

Were you expecting a part h)? Unfortunately it need not be the case that if $F' \geq F$ then the associated net $x(F')$ can be endowed with the structure of a subnet of $x(F)$. A bit of quiet contemplation reveals that a subnet structure is equivalent to the existence of a function $r : F' \to F$ satisfying $A' \subset r(A')$ for all $A' \in x(F')$ and $A'' \subset A' \implies r(A'') \subset r(A')$. To see that such a map need not exist, take $X = \mathbb{Z}^+$. For all $n \in \mathbb{Z}^+$, define let $A_n = \{1\} \cup \{n, n+1, \ldots\}$. Since $A_n \cap A_m = A_{\max m,n}$, $F = \{A_n\}$ is a prefilter on X. Let $F' = F \cup \{1\}$. The directed set I' on which $x(F')$ is based has an element which is larger than every element – namely $\{(1, \{1\}\}$ – but this does not hold for the directed set I on which $x(F)$ is based. (Indeed, I is order isomorphic to the positive integers, or the ordinal ω, whereas I' is order isomorphic to $\omega + 1$.) There is therefore no order homomorphism $I' \to I$ so that $x(F')$ cannot be given the structure of a subnet of $x(F)$.

This example isolates the awkwardness of the notion of subnet. Taking a step back, we see that we became satisfied that we had the right definition of a subnet only insofar as it fit into the theory of convergence as it should: i.e., it rendered true the facts that "x is a limit point of $\mathbf{x}$ $\iff$ some subnet $\mathbf{y}$ converges to x" and "every net $\mathbf{x}$ admits a subnet $\mathbf{y}$ which converges to each of its limit points." These two results are what subnets *are for*. Now that we have at our disposal the correspondence with the theory of filters, the extent of our leeway becomes clear. Any definition of "$\mathbf{y}$ is a subnet of $\mathbf{x}$" which satisfies the following requirements:

(SN1) If $\mathbf{y}$ is a subnet of $\mathbf{x}$, then $F(\mathbf{y}) \geq F(\mathbf{x})$;
(SN2) For every net $\mathbf{x} : I \to X$ and every prefilter $F' \geq F(\mathbf{x})$, there exists a subnet $\mathbf{y}$ of $\mathbf{x}$ with $F(\mathbf{y}) = F'$;

will render valid the above results and hence give an acceptable definition. Note that (SN1) is part g) of Theorem 5.45. The following establishes (SN2) (and a little more).

THEOREM 5.46. *(Smiley) Let $\alpha : I \to X$ be a net, and let F' be a prefilter on X which is compatible with $F(\mathbf{x})$. Let $\mathcal{I}$ be the set of all triples (x, i, A) with $i \in I$, $A \in F'$ and $x \in A$ such that there exists $j \geq i$ with $\alpha_j = x$. Let $\leq$ be the relation on $\mathcal{I}$ by $(x, i, A) \leq (x', i', A')$ if $i \leq i'$ and $A \supset A'$. Let $\gamma : \mathcal{I} \to X$ be the function $(x, i, A) \mapsto x$. Then:*
a) $\mathcal{I}$ is a directed set, and γ is a net on X.
b) Via the natural map $\mathcal{I} \to I$ given by $(x, i, A) \mapsto i$, γ is a subnet of I.
c) The associated prefilter $F(\gamma)$ is the prefilter generated by $F(\mathbf{x})$ and F'.
So if $F' \geq F(\mathbf{x})$, then γ is a subnet of $\mathbf{x}$ with $F(\gamma) = F'$.

EXERCISE 5.34. *Prove Theorem 5.46.*

Thus our definition of subnet is an acceptable one in the sense of (SN1) and (SN2). (In particular, the material of this section and §4 on filters gives independent proofs of the material of §3.) However, from the filter-theoretic perspective there is certainly a simpler definition of subnet that renders valid (SN1) and (SN2): just define

$\mathbf{y} : J \to X$ to be a subnet of $\mathbf{x} : I \to X$ if $F(\mathbf{y}) \geq F(\mathbf{x})$; or, in other words, that for all $i \in I$, there exists $j \in J$ such that $y(j^+) \subset x(i^+)$. That this should be the definition of a subnet was in fact suggested by Smiley.

8. Notes

The material of §1 ought to be familiar to every undergraduate student of mathematics. Among many references, we can recommend Kaplansky's elegant text [**Ka**]. That the key properties of metric spaces making the theory of sequential convergence go through are first countability and (to a lesser extent) Hausdorffness was first appreciated by Hausdorff himself. There is a very rich theory of the sequential closure operator, e.g. in set-theoretic topology (via the sequential order). Apparently there has been a recent interest in the general theory of operators satisfying the three Kuratowksi closure axioms (KC1), (KC2) and (KC4) but not (KC3) (idempotence): such an operator is called a **praclosure**.

The development of a repaired convergence theory via nets has a complicated history. In some form, the concept was first developed by E.H. Moore in his 1910 colloquium lectures [**Mo10**] and then in his 1915 note *Definition of limit in general integral analysis* [**?**]. A fuller treatment was given in the 1922 paper [**MS22**], written jointly with his student H.L. Smith. As the titles of these articles suggest, Moore and Smith were primarily interested in analytic applications: as in §3.2, the emphasis of their work was on a single notion of limit to which all the various complicated-looking limiting processes one meets in analysis can refer back to.[14] Thus their theory was (as I understand it; I have not had a chance to read their original paper) limited to "Moore-Smith sequences" (i.e., nets) with values in $\mathbb{R}$, $\mathbb{C}$, or some Banach space.

In 1937, Birkhoff published a paper *Moore-Smith Convergence in General Topology* whose point of departure is precisely the same as ours: to use mappings from a directed set to a topological space to generalize facts about neighborhoods, closure and continuous functions that hold using sequences only under the assumption of first countability (and to a lesser extent, Hausdorffness). He paper then goes on to discuss applications to the *completion* of various structures of mixed algebraic/topological character, e.g. topological vector spaces and topological algebras. In this aspect he goes beyond the material we have presented so far and competes with the work of André Weil, who in that same year introduced the seminal concept of **uniform space** as the correct generalization of special classes of spaces, notably metric spaces and topological groups, in which one can speak of one pair of points being as close together as another.

In 1940 Tukey published a short book which explored the interrelationships of Moore-Smith convergence and Weil's uniform spaces. Tukey's book is systematic and foundational, in particular employing a language which does not seem to have persuaded many to speak. (E.g. we find in his book that a **stack** is the directed set of finite subsets of a given set S – if only that's what stack meant today! – and a **phalanx** if a function from a stack to a topological space (cf. Example 3.2.1).) The book is probably most significant for its formulation of the notion of a uniform space in terms of **star refinements**, which is still useful today (e.g. [**?**]). Moreover the

[14] It is therefore a bit strange, is it not, that one does not learn about nets in basic real analysis courses? Admittedly the abstract Lebesgue integral plays a similar unifying role.

notion of uniform completion seems to appear here for the first time. We quote the first two sentences of Steenrod's review of Tukey's book: "The extension of metric methods to non-metrizable topological spaces has been a principal development in topology of the past few years. This has occurred in two directions: one through a rebirth of interest in Moore-Smith convergence due to results of Garrett Birkhoff, and the other through the concept of uniform structure due to André Weil." May it not even be the case that the emerging study of uniform spaces was the major cause of the rebirth of interest in Moore-Smith convergence?

Our treatment of nets in §3 closely follows Kelley's 1950 paper *Convergence in topology* [**Ke50**] and his text *General Topology* [**Ke**]. Apart from introducing the term "nets" for the first time, [**Ke50**] is the first to recognize the subnet as an essential tenet of the theory, to prove Proposition 5.16, to introduce the notion of universal net and apply it to give a strikingly simple proof of Tychonoff's theorem. On the other hand the idea of a universal net is motivated by that of an ultrafilter, and Kelley makes explicit reference to earlier work of H. Cartan.

Indeed, in 1937 Henri Cartan came up with the definition of a filter: apparently inspiration struck during a lull in a Séminaire Bourbaki (and Cartan stayed behind to think about his new idea rather than go hiking with the rest of the group). His ideas are written up briefly in [**Ca37**]. Evidently he had no trouble convincing André Weil (the *de facto* leader of Bourbaki at its inception in the 1930's) of the importance of this idea: Bourbaki's 1940 text *Topologie Generale* introduces filters and uses them systematically throughout. It may well be the case that this was the most influential of the many innovations introduced across Bourbaki's many books.

Bourbaki's treatment of filters is much more intensive than what we have given here. In particular Bourbaki rewrites the theory of convergent series and integrals in the filter-theoretic language. To my taste this becomes tiresome and serves as a *de facto* demonstration of the usefulness of nets in more analytic applications. One Bourbakism we have adopted here is the emphasis of the development of the theory at the level of prefilters (called there and elsewhere "filter bases"). It is not necessary to do so – at any stage, one can just pass to the associated filter – but seems to lead to a more precise development of the theory. We have emphasized the notion of compatible prefilters more than is typical (an exception is [**Sm57**]). The existence of free ultrafilters (due, so far as I know, to Cartan) even on a countably infinite set leads to what must be the single most striking application of set-theoretic machinery in general mathematics, the **ultraproduct**. The proof of Tychonoff's theorem via ultrafilters first appears in [**Bo**] and is one of Bourbaki's most celebrated results.

The material of §6 is distressingly absent from most standard treatments. Most texts choose to present either the results of §3 or the results of §4 but not both, and then give a few exercises on the convergence theory they did not develop. In terms of relating the two theories, standard is to drop the unhelpful remark "The equivalence of nets and filters is part of the folklore of the subject." Even Kelley's text [**Ke**] does this, although he gives the construction of a net from a filter and a filter from a net (the latter amounts to taking the associated filter of our prefilter of tails) and asks the reader to show our Proposition 5.44 (for filters). But this result is cited as "grounds for suspicion" that filters and nets are "equivalent", a phrasing which leads the careful reader to wonder whether things do in fact work out as they appear to. Of interest here is R.G. Bartle's 1955 paper *Nets and Filters in Topology* [**Ba55**]. Written at about the same time as [**Ke**], it aspires to make explicit the equivalence between the two theories. Unfortunately the paper is rather defective: the net that Bartle associates with a filter $\mathcal{F}$ is indexed by the elements of $\mathcal{F}$ (and one chooses arbitrarily a point in each element to define the net). As discussed in §6, this is inadequate: upon passing to the (pre)filter of tails, one gets a (pre)filter which may be strictly finer than the original one. (The correct definition is given in a footnote, following the suggestion of the referee!) As a result, instead of the equivalences of Theorem 5.45 Bartle gives only one-sided statements of the form "If the filter converges, then the net converges." Moreover, he erroneously claims [**Ba55**, Prop. 2.5] that given a net $\mathbf{x}$ and a finer prefilter $F' \geq F(\mathbf{x})$, there exists a subnet $\mathbf{y}$ of $\mathbf{x}$ with $F(\mathbf{y}) = F'$. (Interestingly, Kelley reviews this paper in MathReviews; his review is complimentary and finds nothing amiss.) There is a 1963 (eight years later!) erratum [**Ba55er**] to [**Ba55**] which replaces Prop. 2.5 by our (SN2). In between the paper and its erratum comes Smiley's 1957 paper [**Sm57**], whose results we have presented in §6. (Bartle's erratum does not make reference to [**Sm57**].) It is tempting to derive a moral about the dangers of leaving "folklore" unexamined; we will leave this to the interested reader.

9. Ultrafilters in Social Choice Theory

9.1. Preference Aggregation Functions, Arrow and Fishburn. Let X be a nonempty set, whose elements will be called voters, and let $\mathcal{P}$ be a set of outcomes, of size at least 2, that voters will be choosing among. To fix ideas, it may help to think of $\mathcal{P}$ as the set of candidates in an election. Let $\mathrm{Tot}(\mathcal{P})$ be the set of total orderings on $\mathcal{P}$. A **vote** is a function

$$v : X \to \mathrm{Tot}(\mathcal{P}).$$

That is, each voter ranks all the outcomes in order of preference. Let

$$\mathcal{V} := \mathrm{Tot}(\mathcal{P})^X$$

be the set of all possible votes. Each $v \in \mathcal{V}$ should be thought of as the data of one particular election – of a somewhat complicated sort, in which each voter does not just pick one candidate but linearly ranks all the candidates in order of preference.

Remark 5.47. *We are allowing the case in which $\mathcal{P}$, which may be thought of as the set of candidates, is infinite. On the one hand this seems a bit counterintuitive to me: the case of infinitely many voters feels somehow reasonable (it will be a very interesting case, as we will see). When $\mathcal{P}$ is finite of size n, we may as well identify it with $\{1, \ldots, n\}$, and total orderings are identified with permutations of n and may be viewed as a simple ranked list $\sigma_1 < \ldots < \sigma_n$. When $\mathcal{P}$ is infinite, to submit a total ordering means to submit answers the $\#\mathcal{P} \times \#P$ questions "Is $p \leq q$?" (and the check that the answers are consistent with a total ordering may be nontrivial!), so this involves an infinite amount of data. The number of order isomorphism types on a countably infinite set has continuum cardinality, so a single ballot becomes a rather complicated object. But these complications do not really make our analysis more complicated, so we might as well.*

We now want to come up with a reasonable set of "election rules," or in other words a procedure takes as input each possible "election datum" and outputs an *aggregate* total order on the candidates. In other words, this procedure examines as input a ranked list of preferences for each individual voter and outputs one final ranked list. Mathematically, this is a function

$$f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P}),$$

and we call such an f a **preference aggregation function**. For each $v \in \mathcal{V}$ and $x \in X$, $v(x)$ is a total ordering of $\mathcal{P}$, so for $p, q \in \mathcal{P}$ we write

$$p <_{v(x)} q$$

to mean that in the particular vote v, the voter x prefers candidate q to candidate p. Similarly, for a preference aggregation function f, a vote $v \in \mathcal{V}$ and $p, q \in \mathcal{P}$, we write

$$p <_{f(v)} q$$

to mean that in the vote v and according to the election rules f, candidate q beats (is preferred to) candidate p. For more technical reasons to come, for a subset $Y \subset X$ we also write

$$p <_{v(Y)}< q$$

to mean

$$\forall y \in Y,\ p <_{v(y)}< q,$$

or in other words that in the vote v, every voter $y \in Y$ prefers candidate q to candidate p.

Now there are lots and lots of preference aggregation functions f: indeed, there are $(\#\mathcal{P})^{(\#P)^{\#X}}$ of them. For instance, if we have $\mathcal{P} = \{1, \ldots, n\}$, then f could map every vote v to $1 < 2 < \ldots < n$. This is analogous to having the election be determined by the alphabetical order of the candidates' names. It is very orderly but has the defect of having no dependence whatsoever on the vote! So let us try to come up with some reasonable conditions that a preference aggregation function should satisfy: We will probably all agree on the following two:

(PAF1) (Unanimity) For all $v \in \mathcal{V}$ and all $p, q \in \mathcal{P}$, if for all $x \in X$ we have $p <_{v(x)}< q$, then $p <_{f(v)} q$.

In other words, if every single voter prefers q to p, then the aggregate preference should do so as well. (Notice that the silly constant preference function considered above does not satisfy this.)

(PAF2) (Independence) For all $p, q \in \mathcal{P}$ and all $v \in \mathcal{V}$, the ordering $f(v)$ on $\{p, q\}$ depends only on the orderings $v(x)$ on $\{p, q\}$ for all $x \in X$.

So for instances if a group is deciding whether to order Chinese food, Thai food or sushi, the aggregate outcome prefers Chinese food to Thai food, and then I realize that I like sushi more or less than I had said, that change should not effect the preference between Chinese food and Thai food, because sushi was not involved.

We introduce some notation that will become useful later. For each $p \neq q \in \mathcal{P}$ and each $v \in \mathcal{V}$, let

$$U^v_{p,q} := \{x \in X \mid p <_{v(x)} q\}$$

be the set of voters that prefer q to p in the vote v, and let

$$\mathcal{U}_{p,q} := \{U^v_{p,q} \mid v \in \mathcal{V}\}.$$

Then (PAF2) holds iff for all $p \neq q \in \mathcal{P}$ and all $v \in V$, whether $p <_{f(v)} q$ or not depends only on $U^v_{p,q}$. Thus a preference aggregation function satisfying (PAF2) is equivalent to selecting, for all $p \neq q \in \mathcal{P}$, a family of "winning subsets" $\mathcal{W}_{p,q}$ of X: then for all $v \in \mathcal{V}$, we put

$$U^v_{p,q} \in \mathcal{W}_{p,q} \iff p <_{f(v)} q.$$

A preference aggregation function is **admissible** if it satisfies (PAF1) and (PAF2).

There are other desirable properties for a preference aggregation function to have, but many of them turn out to be implied by (PAF1) and (PAF2). For instance, we now consider the property of "neutrality," which essentially means that viewing $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ as a set of "election rules," these rules should treat all the candidates symmetrically.

For the precise definition we introduce a modest set-theoretic formalism. Let $\mathcal{R}$ be a binary relation on a subset $\mathcal{P}$: that is, $\mathcal{R}$ is a subset of $\mathcal{P} \times \mathcal{P}$. Let $\mathrm{Sym}(\mathcal{P})$ be the set of all bijections of $\mathcal{P}$. For any function $\sigma : \mathcal{P} \to \mathcal{P}$, we can define a "pulled back relation" on $\mathcal{P}$ by

$$\sigma^*(\mathcal{R}) := (\sigma \times \sigma)^{-1}(\mathcal{R}).$$

That is, for $p, q \in \mathcal{P}$, we have

$$p\sigma^*(\mathcal{R})q \iff \sigma(p)\mathcal{R}\sigma(q).$$

EXERCISE 5.35. a) *If $\mathcal{R} \in \mathrm{Tot}(\mathcal{P})$ is a total ordering and $\sigma \in \mathrm{Sym}(\mathcal{P})$ is a bijection, show: $\sigma^*(\mathcal{R})$ is a total ordering.*

b) *There is therefore an induced map $\sigma^* : \mathrm{Tot}(\mathcal{P}) \to \mathrm{Tot}(\mathcal{P})$. Show that it is a bijection. (Hint: what is its inverse?)*

Denoting a total ordering on $\mathcal{P}$ by $\leq$, the exercise says that for a bijection $\sigma \in \mathrm{Sym}(\mathcal{P})$, we get a new total ordering:

$$p <_{\sigma^*} q \iff \sigma(p) < \sigma(q).$$

Henceforth for notational simplicity we will write σ in place of σ^*. For $v \in \mathcal{V}$, we will write $\sigma(v)$ for $\sigma \circ v : X \to \mathrm{Tot}(\mathcal{P})$.

LEMMA 5.48. *Let $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ be a preference aggregation function satisfying (PAF2). For all $p, q \in \mathcal{P}$, $v \in \mathcal{V}$ and $\sigma \in \mathrm{Sym}(\mathcal{P})$ we have*

$$U^v_{\sigma(p),\sigma(q)} = U^{\sigma(v)}_{p,q}.$$

PROOF. We have

$$U^v_{\sigma(p),\sigma(q)} = \{x \in X \mid \sigma(p) <_{v(x)} \sigma(q)\} = \{x \in X \mid p <_{\sigma(v(x))} q\} = U^{\sigma(v)}_{p,q}. \qquad \square$$

Now we can define

(PAF3) (Neutrality) For all $\sigma \in \mathrm{Sym}(\mathcal{P})$ we have $f \circ \sigma = \sigma \circ f$. In other words, for all $v \in \mathcal{V}$ we have $f(\sigma(v)) = \sigma(f(v))$.

When $\mathcal{P}$ is finite, a good way to think about this is: if for two candiates p and

q, if every single voter swaps p and q in their ranked list of all the candidates, then in the aggregate ordering p and q should be swapped. (The above formalism nails down what such a swapping means for infinite ordered sets as well.)

LEMMA 5.49. *Let f be a preference aggregation function satisfying (PAF2). Then f satisfies (PAF3) iff for all $p, q, p', q' \in \mathcal{P}$ with $p \neq q$ and $p' \neq q'$ we have*

$$\mathcal{W}_{p,q} = \mathcal{W}_{p',q'}.$$

PROOF. Suppose f is neutral, and let $Y \in \mathcal{W}_{p,q}$. To see that $Y \in \mathcal{W}_{p',q'}$ we must show that for all $v \in \mathcal{V}$, if $Y = U^v_{p',q'}$ then $p' <_{f(v)} q'$. Let $\sigma \in \operatorname{Sym}(\mathcal{P})$ be such that $\sigma(p) = p'$ and $\sigma(q) = q'$. By Lemma 5.48 we have

$$U^{\sigma(v)}_{p,q} = U^v_{\sigma(p),\sigma(q)} = Y \in \mathcal{W}_{p,q},$$

so $p <_{f(\sigma(v))} q$. Neutrality gives $p <_{\sigma(f(v))} q$, so $p' = \sigma(p) <_{f(v)} \sigma(q) = q'$.

Conversely, suppose that $\mathcal{W}_{p,q} = \mathcal{W}_{p',q'}$ for all $p, p'q, q' \in \mathcal{P}$ with $p \neq q$ and $p' \neq q'$. We must show that for all $v \in \mathcal{V}$ we have $f(\sigma(v)) = \sigma(f(v))$. Let $v \in \mathcal{V}$ and suppose that $p <_{f(\sigma(v))} q$. Then we have

$$U^v_{\sigma(p),\sigma(q)} = U^{\sigma(v)}_{p,q} \in \mathcal{W}_{p,q} = \mathcal{W}_{\sigma(p),\sigma(q)},$$

so $\sigma(p) <_{f(v)} \sigma(q)$, which means that $p <_{\sigma(f(v))} q$. This shows that as subsets of $\mathcal{P} \times \mathcal{P}$, the total ordering $f(\sigma(v))$ is contained in the total ordering $\sigma(f(v))$, and any containment of total orderings is an equality. $\square$

Let us now consider some simple cases.

• The case $\#X = 1$ – i.e., of one voter – is trivial. For a preference aggregation function f to satisfy (PAF1) in this case just means that for each vote v, the aggregate ordering $f(v)$ must be the ordering of the one voter, and this certainly satisfies (PAF2): if the one voter doesn't change his preference between two candidates p and q...then he doesn't change his preference between p and q, so the aggregate preference between p and q doesn't change.

• The case $\#\mathcal{P} = 2$ of two candidates is the familiar one. Here choosing a ranked list amounts to choosing a winner.

EXAMPLE 5.16. *It is tempting to take inspiration from real life: why don't we just take f to be the function that counts the number of voters who voted for each candidate and awards victory to the one with the most votes? This is a solid idea, and we analyze it in several cases.*

Suppose that $\#\mathcal{P} = 2$ and that X is finite of odd order. Then in any election v, one candidate will receive more votes than the other. If everyone votes for the same candidate, then yes, that candidate wins: (PAF1). The condition (PAF2) holds vacuously unless there are at least three candidates, so this preference aggregation function is admissible. It is also neutral: in an election with two candidates this just means that if every voter switches their vote then the outcome changes, which is indeed the case here, because whichever candidate has the majority of the votes will, after the switch, have the minority.

Next suppose that $\#\mathcal{P} = 2$ and that X is finite of even order $2n$. Now our proposed preference aggregation function is is incomplete: what do we do if the result of the vote is an n to n tie? One choice of f is to designate one voter $x_\bullet \in X$

as the tiebreaker: if the vote is a tie, the candidate preferred by $x_\bullet$ wins. This f is again admissible and neutral.

Finally suppose that $\#\mathcal{P} = 2$ and that X is infinite. Then we can interpret the above preference aggegration function in terms of cardinal arithmetic and again use a tiebreaker $x_\bullet$ in the case of ties (which feels much more likely than when X is finite: e.g. if X is countably infinite, then ties occur unless one candidate receives all but finitely many of the votes). This f is once again admissible and neutral.

EXAMPLE 5.17. *Suppose $\mathcal{P} = \{1, 2\}$, X is finite of even order $2n$, and consider the function f that awards the candidate with the most votes the victory but in case of a tie awards the victory to Candidate 1. This is admissible but not neutral: in the case of a tie, if we flip all the votes, then Candidate 1 still wins. Nevertheless this preference aggregation function seems not necessarily worse than the one of Example 5.16. In the 2020 American Presidential Election, there was an evening in which I was wondering what would happen if the Electoral College vote was a 269-269 tie. The answer is that, by some fine print concerning the present composition of the House of Representatives that need not concern us here, the Republican party candidate would win. Compare that to for instance the rule being that in the case of a tie, the winner is whoever a particular prearranged elector in Maine has voted for. The actual system seems better!*

EXAMPLE 5.18. *When $\#\mathcal{P} = 2$ and X is finite of odd order, we can award victory to the unique candidate with an odd number of votes. It is a bit distressing to observe that this function is admissible and neutral.*

EXERCISE 5.36. *Suppose $\#X = \mathcal{P} = 2$. There are $2^{2^2} = 16$ preference aggregation functions. Write them all out: how many are admissible? Neutral? Both?*

The case of at least three candidates is much more interesting. The first thing to see is that adding up the votes doesn't really make sense here, because each voter is now submitting an ordered list rather than a single candidate. There are still various ways reasonable ways of proceeding, and it is enlightening to try to come up with some and see that they fail to be admissible.

EXAMPLE 5.19. *Suppose that $2 \leq \#X < \infty$ is finite, $\mathcal{P} = \{1, 2, 3\}$ and that in any vote, first place goes to the candidate with the largest number of first place votes, second place goes to the remaining candidate with the largest number of second place votes, and third place goes to the remaining candidate. We have to break ties somehow, for instance by awarding the victory to the candidate with the larger number. But consider the election with 7 voters voting as follows (we list first place vote, followed by second place vote, followed by third place vote):*

1. 123
2. 123
3. 123
4. 231
5. 231
6. 321
7. 321

Then candidate 1 *has* 3 *first place votes, while the others have* 2*: no ties. Candidate* 2 *has* 5 *second place votes, while candidate* 1 *has* 2*. So the result of the election – or aggregate outcome – is*

$$123.$$

Now suppose that voters 4. and 5. change their mind about 2 *and* 3*:*

$$4'.\ 321$$

$$5'.\ 321$$

Now candidate 3 *has* 4 *first place votes and candidate* 7 *has* 7 *second place votes, so the aggregate outcome changes to*

$$321.$$

But no one changed any preferences involving candidate 1 *while some (in fact, both of the two) aggregate preferences for* 1 *changed, so this violates (PAF2). Notice that in this example there were no ties.*

Eventually one sees that for all X and $\mathcal{P}$ there still is an admissible preference aggregation function, perhaps taking inspiration from the idea of a designated tiebreaker. Namely, fix $x_\bullet$ in X and take for all $v \in \mathcal{V}$ $f(v) = v(x_\bullet)$. Such an $x_\bullet$ is called a **dictator**. Again, this is not so strange that we haven't seen it in the real world, but it does not seem very fair! We say that a preference aggregation function is **dictatorial** if it involves a dictator in this way, and thus the dictatorial preference aggregation functions are in bijection with the set of voters.

Modern social choice theory began with this 1951 result of K.A. Arrow.

Theorem 5.50 (Arrow Impossibility Theorem [**A**]). *If* $\#\mathcal{P} \geq 3$ *and* X *is finite, then every admissible preference aggregation function is dictatorial.*

What about an infinite set of voters? Fascinatingly, the situation reverses itself, as seems to have first been noted in the literature in a 1970 work of Fishburn.

Theorem 5.51 (Fishburn Possibility Theorem [**Fi70**]). *For all* $\#\mathcal{P} \geq 2$ *and all infinite* X*, there are preference aggregation functions that are admissible and nondictatorial.*

In the next section we state a result that implies both Theorems 5.50 and 5.51. That result is then proved in the following two sections.

9.2. The Connection With Ultrafilters. Let us take a step back: in the case of $\#\mathcal{P} \geq 3$, we see a striking dichotomy between the case in in which X is finite and the case in which X is infinite: in the former case, an admissible preference aggregation function is determined by a single element of X, while in the latter case such functions exist but there are also other functions that are not determined by a single element of X. To those familiar with the scent of ultrafilters, this reeks of them. To make the connection we need one more key definition. Let $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ be an admissible preference aggregation function. We say a subset $Y \subset X$ of voters is **decisive** if for all votes $v \in V$ and all $p, q \in \mathcal{P}$, if $p <_{v(Y)} q$, then $p <_{f(v)} q$. In words: a subset Y is decisive if in every vote, if all the voters in Y prefer candidate q to candidate p then the aggregate ordering prefers q to p. A dictator is precisely a one element decisive subset.

Here is our main result.

Theorem 5.52 (Kirman-Sondermann [**KS72**]). *Let $\#\mathcal{P} \geq 3$.*

a) *For every admissible preference aggregation function $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$, the set $\mathcal{F}_f$ of decisive subsets forms an ultrafilter on X.*
b) *For every ultrafilter $\mathcal{F}$ on X, define $f_{\mathcal{F}}$ by: for all $v \in \mathcal{V}$ and $p, q \in \mathcal{P}$, $p <_{f(v)} q$ if and only if $\{x \in X \mid p <_{v(x)} q\}$ lies in $\mathcal{F}$. Then $f_{\mathcal{F}}$ is an admissible preference aggregation function.*
c) *The assignments $f \mapsto \mathcal{F}_f$ and $\mathcal{F} \mapsto f_{\mathcal{F}}$ are mutually inverse bijections from the set of admissible preference aggregation functions to the set of ultrafilters on X.*

Let us see how Theorem 5.52 immediately implies Theorems 5.50 and 5.51. First, suppose X is finite and let $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ be a preference aggregation function. By Theorem 5.52a) the family of decisive subsets forms an ultrafilter $\mathcal{F}_f$ on X. Since X is finite, there is $x_\bullet \in X$ such that $\mathcal{F}_f = \mathcal{F}_{x_\bullet}$, the principal ultrafilter consisting of all subsets containing $x_\bullet$. This means that $\{x_\bullet\}$ is decisive, so $x_\bullet$ is a dictator: Arrow's Theorem. Now suppose that X is infinite. As we saw above, the Fréchet filter on X extends to a nonprincipal ultrafilter $\mathcal{F}$ on X, and by Theorem 5.52b) we get an admissible preference aggregation function $f_{\mathcal{F}}$ essentially by asking the ultrafilter whether it prefers p to q for each $p, q \in \mathcal{P}$. By Theorem 5.52c) the decisive subsets of $f_{\mathcal{F}}$ are precisely the elements of $\mathcal{F}$, so since $\mathcal{F}$ is nonprincipal, $f_{\mathcal{F}}$ is nondictatorial: Fishburn's Theorem.

In fact we get not just the existence of a nondictatorial preference aggregation function but an enormous supply of them: by the above discussion and Theorem 7.32, the number of such functions on an infinite set is $2^{2^{\#X}} \geq 2^{\mathfrak{c}}$.

9.3. Proof of Theorem 5.52, Parts b) and c). Let $\mathcal{F}$ be an ultrafilter on X, and let $f_{\mathcal{F}} : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ be the function that for all $p \neq q \in \mathcal{P}$, satisfies $p <_{f(v)} q$ if and only if $\{x \in X \mid p <_{v(x)} q\}$ lies in $\mathcal{F}$. First of all this function is well-defined: for all $p \neq q$, we have

$$X = U^v_{p,q} \coprod U^v_{q,p},$$

and since $\mathcal{F}$ is an ultrafilter on X, it contains exactly one of $U^v_{p,q}$ and $U^v_{q,p}$. This shows that exactly one of $p <_{f(v)} q$ and $q <_{f(v)} p$ holds. For $p, q, r \in \mathcal{P}$ and all $v \in \mathcal{V}$, we have

$$U^v_{p,r} \supset U^v_{p,q} \cap U^v_{q,r},$$

so if $p <_{f(v)} q$ and $q <_{f(v)} r$, then $U^v_{p,q}$, $U^v_{q,r} \in \mathcal{F}$, so $U^v_{p,r} \in \mathcal{F}$ and $p <_{f(v)} r$. This shows that $f(v)$ is a total ordering on $\mathcal{P}$.

For $v \in \mathcal{V}$, if $p <_{v(X)} q$ then $V_{pq} = X \in \mathcal{F}$, so $p <_{f(v)} q$: this is (PAF1). If we start with a vote v and modify it to a vote v' without changing any voter's preference between p and q, then $U^{v'}_{p,q} = U^v_{p,q}$, so

$$p <_{f(v)} q \iff U^v_{p,q} \in \mathcal{F} \iff U^{v'}_{p,q} \in \mathcal{F} \iff p <_{f(v')} q,$$

establishing (PAF2). This proves Theorem 5.52b).

Let f be an admissible preference aggregation function, let $\mathcal{F}_f$ be the ultrafilter of decisive subsets for f, and let $g := f_{\mathcal{F}_f}$ be the associated admissible preference aggregation function. Then for all $v \in \mathcal{V}$ and all $p, q \in \mathcal{P}$, $p <_{g(v)} q$ then $U^v_{p,q} \in \mathcal{F}_f$, so $U^v_{p,q}$ is decisive and by definition, in the vote v all voters in $U^v_{p,q}$ prefer q to p, so $p <_{f(v)} q$. Since $f(v)$ and $g(v)$ are both total orderings on $\mathcal{P}$, this implies that $f(v) = g(v)$ and thus $f = g = f_{\mathcal{F}_f}$.

Let $\mathcal{F}$ be an ultrafilter on X, let $f_{\mathcal{F}}$ be the associated admissible preference aggregation function, and let $\mathcal{G}$ be the ultrafilter $\mathcal{F}_{f_{\mathcal{F}}}$. If $Y \in \mathcal{F}$ then for all $v \in \mathcal{V}$ and all $p, q \in \mathcal{P}$, if $p <_{v(x)} q$ for all $x \in Y$ then $U^v_{p,q} \supset Y$, so $U^v_{p,q} \in \mathcal{F}$, so $p <_{f_{\mathcal{F}}} q$. This shows that Y is decisive for $f_{\mathcal{F}}$, so $Y \in \mathcal{F}_{f_{\mathcal{F}}}$, so $\mathcal{F} \subset \mathcal{F}_{f_{\mathcal{F}}}$. Since both are ultrafilters, we have $\mathcal{F} = \mathcal{F}_{f_{\mathcal{F}}}$. This proves Theorem 5.52c).

We observe that our deduction of Fishburn's Theorem (Theorem 5.51) from the Kirman-Sondermann Theorem (Theorem 5.52) only used parts b) and c) of Theorem 5.52 and thus we have already proved Fishburn's Theorem.

9.4. Proof of Theorem 5.52, Part a). It remains to prove that if $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ is an admissible preference aggregation function and $\#\mathcal{P} \geq 3$ then the decisive sets form an ultrafilter on X. Unlike the rest of the proof this will not follow simply by unwinding the definitions. In order to help see that we have something to do, we observe that this assertion fails if $\#\mathcal{P} = 2$: e.g. if $\#\mathcal{P} = 2$ and X is finite of odd cardinality $2n + 1$, then as above having the winner of the election be the candidate with the most votes is an admissible preference aggregation function. For this function the decisive subsets are the ones of cardinality greater than n. This family of sets is not closed under finite intersection so is not an ultrafilter (nor could it be, because this is a nondictatorial admissible function!).

We need one interesting preliminary result.

Lemma 5.53. *If $\#\mathcal{P} \geq 3$, every admissible preference aggregation function is neutral.*

Proof. By Lemma 5.49, it is enough to show that $\mathcal{W}_{p,q} = \mathcal{W}_{p',q'}$ for all p, q, p', q' with $p \neq q$ and $q \neq p$.
Step 1: Let $a, b, c \in \mathcal{P}$ be distinct. Suppose $Y \in \mathcal{W}_{a,b}$. There is $v \in \mathcal{V}$ such that

$$a <_{v(Y)} b <_{v(Y)} c, \ b <_{v(X \setminus Y)} c <_{v(X \setminus Y)} a,$$

Since $Y \in \mathcal{W}_{a,b}$ we have $a <_{f(v)} b$. For all $x \in X$ we have $b <_{v(x)} < c$, so unanimity gives $b <_{f(v)} c$. It follows that $a <_{f(v)} c$. Since $U^v_{a,c} = Y$, we have $Y \in \mathcal{W}_{a,c}$. It follows that $\mathcal{W}_{a,b} \subset \mathcal{W}_{a,c}$. Reversing the roles of b and c we get $\mathcal{W}_{a,b} = \mathcal{W}_{a,c}$.
Step 2: Similarly, if $Y \in \mathcal{W}_{a,b}$, considering a vote v with $c <_{v(Y)} a <_{v(Y)} b$ and $b <_{v(X \setminus Y)} c <_{v(X \setminus Y)} a$ shows that $\mathcal{W}_{a,b} = \mathcal{W}_{c,b}$.
Step 3: Finally we consider $p, q, p', q' \in \mathcal{P}$ with $p \neq q$ and $p' \neq q'$.
Case 1: If $p' = q$ and $q' = p$ then there is $r \in \mathcal{P} \setminus \{p, p', q, q'\}$, so

$$\mathcal{W}_{p',q'} = \mathcal{W}_{q,p} = \mathcal{W}_{q,r} = \mathcal{W}_{p,r} = \mathcal{W}_{p,q}.$$

Case 2: If $q' \neq p$, then

$$\mathcal{W}_{p,q} = \mathcal{W}_{p,q'} = \mathcal{W}_{p',q'}.$$

Case 3: If $q' = p$ and $p' \neq q$, then

$$\mathcal{W}_{p,q} = \mathcal{W}_{p',q} = \mathcal{W}_{p',q'}. \qquad \square$$

Now suppose $\#\mathcal{P} \geq 3$ and consider an admissible preference aggregation function f. By Lemmas 5.49 and 5.53 there is an associated family of subsets $\mathcal{W}$, namely those of the form $\mathcal{W}_{p,q}$ for any $p \neq q \in \mathcal{P}$. These are the "winning subsets": that is, in any vote $v \in V$ and for any $p, q \in \mathcal{P}$, we have $p <_{f(v)} q \iff U^v_{p,q} \in \mathcal{W}$. In the remainder of the proof we will show first that $\mathcal{W}$ is an ultrafilter on X and second that it coincides with the ultrafilter $\mathcal{F}$ of decisive subsets. Notice that what is clear directly from the definitions is that $\mathcal{F} \subset \mathcal{W}$: indeed, a decisive subset is precisely a subset Y such that every subset Z with $Y \subset Z \subset X$ is winning, so in particular decisive subsets are winning.

Proposition 5.54. *For $\mathcal{P} \geq 3$ and an admissible preference aggregation function f, the winning subsets $\mathcal{W}$ form an ultrafilter.*

Proof. Step 1: Unanimity implies $X \in \mathcal{W}$ and $\varnothing \notin \mathcal{W}$. For any subset Y of X, exactly one of Y and $X \setminus Y$ lies in $\mathcal{W}$: indeed, in an election between two candidates, the sets of voters for the first candiate and the set of voters for the second candidate are complementary subsets, and exactly one must be the winner!
Step 2: Next we claim that if $Y_1, Y_2 \in \mathcal{W}$, then $Y_1 \cap Y_2 \neq \varnothing$. Seeking a contradiction, suppose that Y_1 and Y_2 are disjoint, and put $Y_3 := X \setminus (Y_1 \cup Y_2)$. Let a, b, c be distinct elements of $\mathcal{P}$. Start with any vote v in which $Y_1 = U^v_{b,a}$. We alter v to a new vote v' as follows:
- For all $x \in Y_1 \cup Y_2$, we alter $v(x)$ by placing c directly below b.
- For $x \in Y_3$, we alter $v(x)$ by placing c directly below a.

Then $U^{v'}_{a,c} = Y_2 \in \mathcal{W}$, so $a <_{f(v')} c$. (PAF1) gives $c <_{f(v')<} b$, so $a <_{f(v')<} b$. However we also have $U^{v'}_{b,a} = Y_1 \in \mathcal{W}$, so $b <_{f(v')} a$: contradiction.
Step 3: Finally we will show that for any 3-pseudo-partition Y_1, Y_2, Y_3 of X exactly one of the sets lies in $\mathcal{W}$, which by Proposition 5.28 will show that $\mathcal{W}$ is an ultrafilter. Since Y_1, Y_2, Y_3 are pairwise disjoint, by Step 2 at most one of them can be winning, so seeking a contradiction we suppose that none of them are winning. By Step 1, this means that each of their complements $C_i := X \setminus Y_i$ is winning. Let a, b, c be distinct elements of $\mathcal{P}$. There is a vote v such that

$$c <_{v(C_1)} b <_{v(C_1)} a,$$

$$a <_{v(C_2)} c <_{v(C_2)} b,$$

$$b <_{v(C_3)} a <_{v(C_3)} c.$$

Now we have

$$U^v_{a,b} = C_2 \in \mathcal{W}, \text{ so } a <_{f(v)} b,$$

$$U^v_{b,c} = C_3 \in \mathcal{W}, \text{ so } b <_{f(v)} c,$$

$$U^v_{c,a} = C_1 \in \mathcal{W}, \text{ so } c <_{f(v)} a.$$

This is clearly impossible - a form of **Condorcet's Paradox** – and the contradiction completes the proof. $\square$

Corollary 5.55. *We have $\mathcal{W} = \mathcal{F}$: the decisive subsets are precisely the winning subsets, and thus the decisive subsets form an ultrafilter on X.*

Proof. Again, it is clear that $\mathcal{F} \subset \mathcal{W}$, so conversely suppose that $Y \in \mathcal{W}$ is a winning subset. As above, we need to show that if $Y \subset Z \subset X$, then Z is also winning. Well, if not, then since the winning subsets form an ultrafilter, we would have $X \setminus Z \in \mathcal{W}$ and then $Y \cap (X \setminus Z) = \varnothing \in \mathcal{W}$, a contradiction. $\square$

9.5. Anonymity. Is it worth asking why we find dictatorial preference aggregation functions unsatisfactory? I think so. Such a system feels "unfair," but after all fairness in voting algorithms is exactly what we are trying to model and study, so it would be a step forward if we could see this instance of unfairness as a violation of some more general principle.

This is not hard to do, using a further axiom on a preference aggregation function, the notion of anonymity. Whereas neutrality is a symmetry condition on the candidates, anonymity is a symmetry condition on the voters. For $\alpha \in \mathrm{Sym}(X)$ and a vote $v : X \to \mathrm{Tot}(\mathcal{P})$, we get a new vote

$$\alpha^*(v) := v \circ \alpha^{-1} : X \to \mathrm{Tot}(\mathcal{P}).$$

(Here we use α^{-1} instead of α only so as to get a left action of $\mathrm{Sym}(X)$ on $\mathcal{V}$. We have no particular reason to need this, but from an algebraically minded perspective it is the right way to set things up.) If we think of each voter as submitting a ballot, passing from v to $\alpha^*(v)$ simply corresponds to collecting the ballots in a different order. In an anonymous election it should not matter whose ballot was whose, so we saay that a preference aggregation function is **anonymous**, or satisfies (PAF4), if for all candidates $p, q \in \mathcal{P}$ we have

$$\forall \alpha \in \mathrm{Sym}(X),\ p <_{f(v)} q \iff p <_{f(\alpha^*(v))} q.$$

Now we observe that a dictatorial preference aggregation function violates anonymity: since $\mathrm{Sym}(X)$ acts transitively on X, under (PAF4) if I'm a dictator then you're a dictator and everyone is a dictator, which is clearly impossible (except in the trivial case $\#X = 1$).

Therefore Arrow's Theorem implies that if there are more than two candidates and finitely many voters, there is no preference aggregation function that is admissible and anonymous. What about in the other cases?

I am delighted to report that in the case of $\#\mathcal{P} = 2$ and $\#X$ finite and odd, having the winner of the election be the unique candidate that receives the most votes is admissible, neutral and anonymous. And I am both intrigued and horrified to report that declaring the winner to be the unique candidate with an odd number of voters as in Example 5.18 is also admissible, neutral and anonymous.

Suppose now that $\#\mathcal{P} = 2$ and $\#X$ is finite and even. There are admissible preference aggregation functions that are neutral but not anonymous: select one voter as the tie-breaker. (In the case $\#X = 2$ this is a dictatorship: we'll do what you want, unless we disagree, in which case we'll do what I want.) And there are admissible preference aggregation functions that are anonymous but not neutral: all ties go to the same candidate. It seems likely that we cannot have both at once. For instance, in the case where you and I are each voting on chocolate versus vanilla, suppose I vote chocolate and you vote vanilla: one of the flavors wins. Now suppose you and I switch: I vote vanilla and you vote chocolate. On the one hand, this vote is obtained from the previous one by switching the two candidates, so by neutrality the outcome must change. On the other hand, this vote is *also* obtained from the previous one by switching *us*, so by anonymity the outcome must not change. Contradiction!

To go further, it is helpful to have an alternate characterization of anonymity analogous to the characterization of neutrality given in Lemma 5.49. This time we leave it as an exercise.

Exercise 5.37. *Let $f : \mathcal{V} \to \mathrm{Tot}(\mathcal{P})$ be a preference aggregation function satisfying (PAF2).*

a) *Suppose X is finite. Show: f is anonymous iff for all $p, q \in \mathcal{P}$ and for all subsets $Y_1, Y_2 \in X$, if $Y_1 \in \mathcal{W}_{p,q}$ and $\#Y_2 = \#Y_1$, then $Y_2 \in \mathcal{W}_{p,q}$.*
b) *Suppose X is infinite. Show: f is anonymous iff for all $p, q \in \mathcal{P}$ and for all subsets Y_1, Y_2, if $Y_1 \in \mathcal{W}_{p,q}$, $\#Y_2 = \#Y_1$ and $\#(X \setminus Y_1) = \#(X \setminus Y_2)$, then $Y_2 \in \mathcal{W}_{p,q}$.*[15]

Exercise 5.37 says that under (PAF2), a preference aggregation function is anonymous if for any two candidates p, q, who is preferred depends only on the number of voters (in the sense of cardinal arithmetic, when X is infinite) that prefer each candidate to the other. This means that for any candidates $p, q \in \mathcal{P}$, in any vote v in which $\#U^v_{p,q} = \#U^v_{q,p}$, the aggregate preference for p versus q must be the same. Just as in our toy example above, this condition is incompatible with neutrality: again, we let the interested reader work out the details.

Exercise 5.38. *Suppose that X is either finite of even cardinality or infinite, and let $\mathcal{P}$ be any set with $\#\mathcal{P} \geq 2$. Then there is no preference aggregation function $f : \mathrm{Tot}(\mathcal{P})^X \to \mathrm{Tot}(\mathcal{P})$ that satisfies (PAF2), (PAF3) and (PAF4).*

9.6. Connections With Topology? In view of its inclusion in this text, it is reasonable to ask whether the theorems of Arrow and Fishburn are actually topological in some way. Our approach exposed the connection to ultrafilters, which we developed in this chapter to apply to study convergence in topological spaces, but ultrafilters make sense even without a topology, and indeed a topological space has yet to make an appearance in this section of the text.

The answer to this is that some connections between Arrow's Theorem (and related impossiblity theorems in social choice theory) and topology have been made, but how essential they are seems up for interpretation. In [**Ch82**] Chichilnisky develops choice functions in a differential topological setting: a preference p is a certain kind of vector field on the closed unit ball B in $\mathbb{R}^n$ and the space of preferences $\mathcal{P}$ is a fixed subspace of the Banach space of all C^1-vector fields on B. One of the main results in the paper [**Ch82**, Thm. 2] is "Let $\varphi : P^k \to P$ be a continuous W-Pareto aggregation rule." Then φ is homotopic to a dictatorial rule." Here $k \geq 2$ is an integer that it supposed to model the number of voters. A later paper of Baryshnikov [**Ba93**] employs techniques from algebraic topology – nerves, covers, singular homology groups – and proves a result [**Ba93**, Thm. 2] that implies both Arrow's Theorem and Chichilnisky's Theorem.

Neither of these are the *general* topology that we study in this text. In fact though, in [**KS72**] Kirman and Sondermann give a topological interpretation of their main result in the case of finitely many candidates and infinitely many voters: as we will

[15]The condition of part b) holds also when X is finite, but when X is finite, the cardinality of a subset determines the cardinality of its complement, so it is equivalent to the simpler condition given in part a).

see later on in this text, the set $\tilde{X}$ of ultrafilters on a set X can be given a natural compact topology in which the principal ultrafilters embed X as a discrete space. This is in fact the Stone-Cech compactification of the discrete space X. When $\mathcal{P}$ is finite of size at least 3 then we give $\mathrm{Tot}(\mathcal{P})$ the discret topology, the unique compact topology on a finite set. Then each vote has a unique continuous extension

$$\tilde{v} : \tilde{X} \to \mathrm{Tot}(\mathcal{P})$$

and Kirman-Sondermann interpret the element $\mathcal{F}_f \in \tilde{X}$ corresponding to the admissible preference aggregation function f as an "invisible dictator."

9.7. Final Remarks. Our statement of Theorem 5.52 is in fact a variant of [**KS72**, Thm. 1]. That paper as well as [**A**] and [**Fi70**] work not with total orderings but with "weak orderings," which in our terminology would be called total quasi-orderings. A total quasi-ordering $\prec$ on a set $\mathcal{P}$ is a binary relation on $\mathcal{P}$ that is reflexive, transitive and total: for all $p, q \in \mathcal{P}$ we have $p \prec q$ or $q \prec p$. However, since we are not imposing anti-symmetry, for distinct p and q we could have $p \prec q$ and $q \prec p$. As for any quasi-ordering, the relation $p \sim q$ iff $p \prec q$ and $q \prec p$ is an equivalence relation, and $\prec$ descends to give a partial ordering on the set of equivalence classes. The imposition of totality on $\mathcal{P}$ means that on $\mathcal{P}/\sim$ we get a total order. Thus total quasi-orderings model preferences in which "ties" are allowed.

As with several other more recent authors, we have elected to develop the theory with total orderings rather than total quasi-orderings. For starters, restricting to total orders seems simpler. Moreover, [**KS72**, Thm. 1] does not claim that the assignment of an ultrafilter to an admissible preference function is bijective but only surjective, and I believe the map need not be a bijection in the total quasi-ordered case. Finally, to me at least, allowing ties in an election does not seem especially natural, and although that is a more general setup, this raises the question of whether ties are crucially involved in Arrow's Impossibility phenomenon. The answer is that they are not, and this approach makes that clear.

Our exposition does not follow [**KS72**] directly but rather is a mix of [**T**] and [**Tr20**]. The main difference is that the notion of neutrality appears in [**T**] and [**Tr20**] but not in [**KS72**]. The use of neutrality does not really shorten the proof, but it seems to make it more interesting.

CHAPTER 6

Separation and Countability

1. Axioms of Countability

1.1. First Countable Spaces.

Let X be a topological space, and let x be a point of X. We say X is **first countable at x** if there is a countable neighborhood base at x. A space is **first countable** – or, more formally, **satisfies the first axiom of countability** – if it is first countable at each of its points.

Exercise 6.1. *Suppose that X has a countable neighborhood base at x. Show that there is a countable base of open neighborhoods $\mathcal{N} = \{U_n\}_{n=1}^{\infty}$ of x which is* nested*: $U_1 \supset U_2 \supset \ldots \supset U_n \supset \ldots$*

Proposition 6.1. *Metrizable spaces are first countable.*

Proof. Let d be a metric on (X, τ) inducing the topology τ. For $p \in X$, $\{B(p, \frac{1}{n})\}_{n=1}^{\infty}$ is a countable neighborhood base at p. □

Example 6.1. *Discrete spaces are first countable: this is a special case of the last result. Certainly any topological space with finitely many open sets is first countable. This includes any finite topological space and the indiscrete topology on any set. The cofinite topology on a set X is first countable iff X is countable. The cocountable topology on a set X is first countable iff X is countable (in which case it is discrete).*

Exercise 6.2. *Show: the Arens-Fort space is countable but not first countable.*

Proposition 6.2. *First countability is hereditary: a subspace of a first countable space is first countable.*

Proof. Let X be a topological space and Y a subspace. If $y \in Y$ and $\mathcal{N}$ is a neighborhood base for y in X, then $\mathcal{N} \cap Y = \{N \cap Y \mid N \in \mathcal{N}\}$ is a neighborhood base for y in Y. □

Theorem 6.3.
a) *If X is first countable and $f : X \to Y$ is a continuous surjection, then Y need not be first countable.*
b) *If X is first countable and $f : X \to Y$ is continuous, surjective and open, then Y is first countable.*

Proof. We leave this to the reader as an exercise. In the next section we will prove the analogous result with "first countable" replaced by "second countable". This is so similar that the reader who wants to prove this result for herself should do so before going on to the next section. □

THEOREM 6.4. *Let $\{X_i\}_{i\in I}$ be a nonempty family of nonempty topological spaces, let $X = \prod_{i\in I} X_i$, and let*

$$\kappa = \{i \in I \mid X_i \textit{ is } \text{not } \textit{indiscrete}\}.$$

The following are equivalent:

(i) *The space X is first countable.*
(ii) *For all $i \in I$ the space X_i is first countable, and moreover the set κ is countable.*

PROOF. (i) $\implies$ (ii) Suppose X is first countable, and for each $i \in I$ let $\pi_i : X \to X_i$ be the projection map. Then π_i is continuous open and surjective, so $X_i = \pi_i(X)$ is first countable by Theorem 6.3b).

Next, seeking a contradiction we assume that κ is uncountable and show that X is not first countable. For $i \in \kappa$ choose $x_i \in X_i$ and an open $N_i \subset X_i$ such that

$$x_i \in N_i \subsetneq X_i.$$

For $i \in I \setminus \kappa$, choose any $x_i \in X_i$, and put $x = (x_i)$. Let $\{U_n\}_{n=1}^{\infty}$ be any countable sequence of neighborhoods of x in X. We will construct an open neighborhood of x that does not contain U_n for any n. In this regard it it harmless to replace each U_n by any smaller open neighboorhood V_n of x and thereby we may assume without loss of generality that each $U_n = \prod_{i\in I} U_{i,n}$ is a Cartesian product of neighborhoods such that for each $n \in \mathbb{Z}^+$ and all but finitely many $i \in I$ we have $U_{i,n} = X_i$. It follows that the set K of $i \in I$ such that there exists $n \in \mathbb{Z}^+$ such that $U_{i,n} \subsetneq X_i$ is countable. Since κ is uncountable, there is $i \in \kappa \setminus K$, and then $N_i \times \prod_{j\neq i} X_j$ is an open neighborhood of $x \in X$ that does not contain U_n for any $n \in \mathbb{Z}^+$.
(ii) $\implies$ (i): Suppose that each X_i is first countable and that κ is countable and let $x = (x_i)_{i\in I}$ be a point of X. For each $i \in I$, let $\mathcal{N}_i$ be a countable neighborhood base at x_i in X_i. Let

$$\mathcal{N} := \left\{ \prod_{i\in I\setminus\kappa} X_i \times \prod_{i\in\kappa} U_i \;\middle|\; U_i \in \mathcal{N}_i \text{ and } U_i = X_i \text{ for all but finitely many } i \in I \right\}.$$

Then $\mathcal{N}$ is a countable neighborhood base at $x \in X$. □

Theorem 6.4 gives us a "no go" result on metrizability of large products:

COROLLARY 6.5. *Let I be an uncountable set, and for all $i \in I$ let X_i be a metric space with at least two points. Then $X = \prod_{i\in I} X_i$ is not metrizable.*

PROOF. Indeed, in the setting of Theorem 6.4 we have $\kappa = I$ is uncountable, so X is not first countable, hence certainly not metrizable. □

PROPOSITION 6.6. *Let X be a first countable space and $Y \subset X$. Then $\overline{Y}$ is the set of all limits of sequences from Y.*

PROOF. Suppose y_n is a sequence of elements of Y converging to x. Then every neighborhood N of x contains some $y_n \in Y$, so that $x \in \overline{Y}$. Conversely, suppose $x \in \overline{Y}$. If X is first countable at x, we may choose a nested collection $N_1 \supset N_2 \supset \ldots$ of open neighborhoods of x such that every neighborhood of x contains some N_n. Each N_n meets Y, so choose $y_n \in N_n \cap Y$, and y_n converges to y. □

PROPOSITION 6.7. *Let X be a first countable space, Y a topological space, and let $f : X \to Y$ be a function. The following are equivalent:*
(i) f is continuous.
(ii) If $\mathbf{x}_n \to x$, $f(\mathbf{x}_n) \to f(x)$.

PROOF. a) $\implies$ b): Let V be an open neighborhood of $f(x)$; by continuity there is an open neighborhood U of x with $f(U) \subset V$. Since $\mathbf{x}_n \to x$, there is $N \in \mathbb{Z}^+$ such that $n \geq N$ implies $\mathbf{x}_n \in U$, so $f(\mathbf{x}_n) \in V$. Therefore $f(\mathbf{x}_n) \to f(x)$.

b) $\implies$ a): Suppose f is not continuous, so that there exists an open subset V of Y with $U = f^{-1}(V)$ not open in X. More precisely, let x be a non-interior point of U, and let $\{N_n\}$ be a nested base of open neighborhoods of x. By non-interiority, for all n, choose $\mathbf{x}_n \in N_n \setminus U$; then $\mathbf{x}_n \to x$. By hypothesis, $f(\mathbf{x}_n) \to f(x)$. But V is open, $f(x) \in V$, and $f(\mathbf{x}_n) \in Y \setminus V$ for all n, a contradiction. $\square$

PROPOSITION 6.8. *A first countable space in which each sequence converges to at most one point is Hausdorff.*

PROOF. Suppose not, so there exist distinct points x and y such that every neighborhood of x meets every neighborhood of Y. Let U_n be a nested neighborhood basis for x and V_n be a nested neighborhood basis for y. By hypothesis, for all n there exists $\mathbf{x}_n \in U_n \cap V_n$. Then $\mathbf{x}_n \to x$, $\mathbf{x}_n \to y$, contradiction. $\square$

PROPOSITION 6.9. *Let $\mathbf{x}$ be a sequence in a first countable topological space, and let x be a point of X. The following are equivalent:*
 (i) *The point x is a limit point of the sequence $\mathbf{x}$.*
 (ii) *There is a subsequence converging to x.*

PROOF. (i) $\implies$ (ii): Take a nested neighborhood basis N_n of x, and for each $k \in \mathbb{Z}^+$ choose successively a term $n_k > n_{k-1}$ such that $x_{n_k} \in N_k$. Then $x_{n_k} \to x$.
(ii) $\implies$ (i): This direction holds in all topological spaces. $\square$

EXAMPLE 6.2. *We consider yet again the cocountable topology on an uncountable set (cf. Example 3.7). This is a non-Hasudorff topology in which a sequence $\mathbf{x}_n \to x$ iff $\mathbf{x}_n$ is eventually constant with eventual value x. Indeed, let $\mathbf{x}_n$ be a sequence for which the set of n such that $\mathbf{x}_n \neq x$ is infinite. Then $X \setminus \{x_n \neq x\}$ is a neighborhood of x which omits infinitely many terms $\mathbf{x}_n$ of the sequence, so $\mathbf{x}_n$ does not converge to x. This implies that the set of all limits of sequences from a subset Y is just Y itself, whereas for any uncountable Y, $\overline{Y} = X$.*

EXERCISE 6.3. *A point x of a topological space is **isolated** if $\{x\}$ is open.*
 a) *If x is isolated, and $\mathbf{x}_n \to x$, then $\mathbf{x}_n$ is eventually constant with limit x.*
 b) *Show that if X is first countable and x is not isolated, then there exists a non-eventually constant sequence converging to x. Must there exist an injective sequence – i.e., $\mathbf{x}_m \neq \mathbf{x}_n$ for all $m \neq n$ conveging to x?*

1.2. Second Countability, Separability and the Lindelöf Property.

A topological space is **second countable** – or, more formally, **satisfies the second axiom of countability** – if there is a countable base for the topology.

A topological space is **separable** if it admits a countable dense subset.

A topological space is **Lindelöf** if every open cover admits a countable subcover.

Proposition 6.10. *Let X be a topological space. Then:*

a) *If X is second countable, it is first countable, separable and Lindelöf.*
b) *If X is metrizable, then being second countable, separable and Lindelöf are all equivalent properties.*

Proof. a) Second countable implies first countable: base for the topology of a space is also a neighborhood base at each of its points.

Second countable implies separable: let $\mathcal{B} = \{U_n\}_{n=1}^{\infty}$ be a countable base for X. For each $n \in \mathbb{Z}^+$, choose $P_n \in U_n$, and let $Y = \{P_n\}_{n=1}^{\infty}$. We claim that $\overline{Y} = X$, which is sufficient. To see this, let $U \subset X$ be nonempty and open. Then $U \supset U_n$ for some n and thus $P_n \in U$.

Second countable implies Lindelöf: Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of X. For each positive integer n, if $V_n \subset U_i$ for some i, then choose one such index and call it i_n; if not, choose i_n to be any element of I. We claim that $\{U_{i_n}\}_{n=1}^{\infty}$ is a countable subcovering. Indeed, for any $x \in X$, $x \in U_i$ for some i and thus $x \in V_{n(i)} \subset U_i$ for some $n(i)$, and thus $x \in U_{i_{n(i)}}$.
b) This is Theorem 2.69. We recall it here for the sake of comparison. □

Example 6.3. a) *Let X be an uncountable set endowed with the discrete topology. Then X is first countable, but not separable or second countable.*
b) *The Sorgenfrey line is first countable, separable and Lindelöf, but not second countable.*

Exercise 6.4. a) *Prove Example 6.3a).*
b) *Prove Example 6.3b).*
c) *Try to prove Example 6.3c). (This is harder, and we'll come back to it.)*

Exercise 6.5. *The* ***weight*** *$w(X)$ of a topological space is the least cardinality of a base for the topology. (Thus second countable means $w(X) \leq \aleph_0$.) The* ***density*** *$d(X)$ of a topological space is the least cardinality of a dense subspace. (Thus separable means $d(X) \leq \aleph_0$.) Define the* ***packing number*** *$\operatorname{pn}(X)$ of a space X to be the maximum cardinality of a pairwise disjoint family of nonempty open subsets of X. These are* ***cardinal invariants****.*

a) *Show: for any space, we have $\max(d(X), \operatorname{pn}(X)) \leq w(X)$.*
b) *Show: for every cardinal number κ, there is a space X with*

$$w(X) = d(X) = \operatorname{pn}(X) = \#X = \kappa.$$

Exercise 6.6. a) *Let $\alpha \leq \beta$ be cardinal numbers. Show: there is a topological space of density α and cardinality β.*
b) *Let X be a first countable, Hausdorff topological space. Show: $\#X \leq 2^{d(X)}$.*
(Suggestion: use the interpretation of closure via sequences.)
c) *Let X be a Hausdorff topological space. Show: $\#X \leq 2^{2^{d(X)}}$.*
(Suggestion: use the interpretation of closure via prefilters.)

Exercise 6.7. [**Mu**, Exc. 4.1.4] *Let A be an uncountable subset of a second countable space. Recall that A' denotes the set of limit points of A in X. Show that $A \cap A'$ is uncountable.*

Proposition 6.11. *Second countability is hereditary: a subspace of a second countable space is second countable.*

Proof. Let X be a topological space and Y a subspace. If $\mathcal{B}$ is a base for the topology of X, then $\mathcal{B} \cap Y = \{B \cap Y \mid B \in \mathcal{B}\}$ is a base for the topology of Y. The result follows. $\square$

Proposition 6.12. *a) A subspace of a separable space need not be separable.*
b) An open subspace of a separable space is separable.
c) A subpace of a Lindelöf space need not be separable.
d) A closed subspace of a Lindelöf space is Lindelöf.

Proof. a) ... Moore-Nymetskii plane
b) Let $A \subset X$ be countable and dense, and let $U \subset Y$ be open. Then every open nonempty open subset V of U is also a nonempty open subset of X, so $A \cap V \neq \varnothing$. It follows that $A \cap U$ is dense in U. Certainly it is also countable, so U is separable.
c) ... Moore-Nymetskii plane
d) We leave it to the reader to check that the proof that a closed subspace of a quasi-compact space carries over easily to this context. $\square$

Exercise 6.8. *Show that for a topological space X, the following are equivalent:*

(i) *Every subset of X is Lindelöf.*
(ii) *Every open subset of X is Lindelöf.*

*(A space satisfying these properties is called **strongly Lindelöf**.)*

Exercise 6.9. *Let X be quasi-compact and Y be Lindelöf. Show: $X \times Y$ is Lindelöf. (Suggestion: adapt the proof of Corollary 3.39.)*

Proposition 6.13.
a) *A continuous image of a separable space is separable.*
(That is: If X is separable and $f : X \to Y$ is a continuous surjection, then Y is separable.)
b) *A continuous image of a Lindelöf space is Lindelöf.*

Proof. a) Let $A \subset X$ be countable and dense, let $f : X \to Y$ be a continuous surjection, and let $V \subset Y$ be nonempty and open. Then $f^{-1}(Y)$ is nonempty and open in Y, so there is $a \in A \cap f^{-1}(Y)$, so $f(a) \in f(A) \cap Y$. It follows that $f(A)$ is dense. Certainly $f(A)$ is countable, so Y is separable.
b) We leave it to the reader to check that the proof that a continuous image of a quasi-compact space is quasi-compact carries over easily to this context. $\square$

Proposition 6.14. *a) The continuous image of a second countable space need not be second countable.*
b) If X is second countable and $f : X \to Y$ is continuous, surjective and open, then Y is second countable.

Proof. a) Let X be $\mathbb{R}$ with its usual Euclidean topology, let Y be $\mathbb{R}$ with cofinite topology, and let $f : X \to Y$ be the identity map. We leave the verification of the properties as a nice exercise.
b) Let $\mathcal{B}$ be a countable base for the topology of X. Let $f(\mathcal{B}) = \{f(B) \mid B \in \mathcal{B}\}$. Since f is open, $f(\mathcal{B})$ is a family of open sets. If V is open in Y, then $f^{-1}(V)$ is open in X, so there is $\mathcal{B}' \subset \mathcal{B}$ such that $\bigcup_{B \in \mathcal{B}'} B = U$. Since f is surjective, $V = f(U) = \bigcup_{B \in \mathcal{B}'} f(B)$. So $f(\mathcal{B})$ is a countable base for the topology of Y. $\square$

REMARK 6.15. *The proof of part a) above shows that second countability is not a **coarsenable** property (recall that a property P of topological spaces is coarsenable if (X, τ_1) has property P and $\tau_2 \subset \tau_1$ is another topology on X, then (X, τ_2) has property P). Comparing $\mathbb{R}$ with the Euclidean topology to $\mathbb{R}$ with the discrete topology shows that second countability is not **refineable** either.*

THEOREM 6.16. *Let $\{X_i\}_{i \in I}$ be a nonempty family of nonempty topological spaces, let $X = \prod_{i \in I} X_i$, and let*

$$\kappa = \{i \in I \mid X_i \textit{ is } \text{not } \textit{indiscrete}\}.$$

The following are equivalent:

(i) *The space X is second countable.*
(ii) *For all $i \in I$ the space X_i is second countable, and moreover the set κ is countable.*

EXERCISE 6.10. *Prove Theorem 6.16.*
(Suggestion: adapt the proof of Theorem 6.4.)

PROPOSITION 6.17. *The product of two Lindelöf spaces need not be Lindelöf.*

PROOF. Sorgenfrey plane... $\square$

2. Density of Product Spaces

In Theorem 6.4 we saw that if $\{X_i\}_{i \in I}$ is a family of topological spaces, none of which has the indiscrete topology, then $\prod_{i \in I} X_i$ is first countable iff each X_i is first countable and I is countable. Theorem 6.16 shows that the same holds with "first countable" replaced everywhere by "second countable." What about separable spaces? Quite surprisingly, it turns out that uncountable products of separable spaces can be separable. The following is a standard (but rather difficult) exercise.

EXERCISE 6.11. *Let $\mathfrak{c} = \#\mathbb{R}$. Show: $[0,1]^{\mathfrak{c}}$ is separable.*

In fact if I is a set with $\#I \leq \mathfrak{c}$ and for each $i \in I$ we have a nonempty separable space X_i, then $\prod_{i \in I} X_i$ is separable, whereas $[0,1]^I$ will fail to be separable for sufficiently large I. We will now prove a cardinal generalization of these results that is due independently to Pondiczery [**Po44**], Hewitt [**He46b**] and Marczewski [**Ma47**]. First a simple preliminary.

LEMMA 6.18. *Let I be a nonempty set. For each $i \in I$ let X_i be a nonempty topological space, and let $Y_i \subset X_i$ be a subset.*

a) *In the product space $\prod_{i \in I} X_i$ we have*

$$\overline{\prod_{i \in I} Y_i} = \prod_{i \in I} \overline{Y_i}.$$

b) *The set $\prod_{i \in I} Y_i$ is dense in $\prod_{i \in I}$ iff we have that Y_i is dense in X_i for all $i \in I$.*

PROOF. a) First, if for all $i \in I$ we are given a closed subspace Z_i of X_i, then $\prod_{i \in I} Z_i$ is closed in $\prod_{i \in I} X_i$: indeed,

$$\prod_{i \in I} X_i \setminus \prod_{i \in I} Z_i = \bigcup_{i \in I} \left((X_i \setminus Z_i) \times \prod_{j \neq i} X_j \right)$$

is open in $\prod_{i\in I} X_i$. Thus $\prod_{i\in I} \overline{Y_i}$ is a closed subspace containing $\prod_{i\in I} Y_i$, so

$$\overline{\prod_{i\in I} Y_i} \subset \prod_{i\in I} \overline{Y_i}.$$

Conversely, if $x \in \prod_{i\in I} X_i \setminus \overline{\prod_{i\in I} Y_i}$ then there is an open neighborhood U of x that is disjoint from $\prod_{i\in I} Y_i$. By definition of the product topology, U contains an open neighborhood of x of the form $\prod_{i\in I} U_i$ with $U_i = X_i$ for all but finitely many i, and then since $\prod_{i\in I} U_i \cap \prod_{i\in I} Y_i = \varnothing$ there is $i \in I$ such that $U_i \cap Y_i = \varnothing$. Since U_i is open we have $U_i \cap \overline{Y_i} = \varnothing$ and thus $x \notin \prod_{i\in I} \overline{Y_i}$.
b) This follows immediately. $\square$

THEOREM 6.19 (Pondiczery-Hewitt-Marczewski). *Let κ an infinite cardinal, and let I be a set with $\#I \leq 2^\kappa$. For each $i \in I$, let X_i be a nonempty topological space of density at most κ: that is, each X_i has a dense subset of cardinality at most κ. Then $X := \prod_{i\in I} X_i$ also has density at most κ.*

PROOF. [**En**, p. 81] Step 1: We may as well assume that $\#I = 2^\kappa$: if $\#I < 2^\kappa$ we can enlarge I and take a product of one-point spaces. For $i \in I$, let Y_i be a dense subset of X_i with $\#Y_i \leq \kappa$. Since "being a dense subspace" is a transitive relation, by Lemma 6.18b) it is enough to show that $\prod_{i\in I} Y_i$ has a dense subspace of cardinality at most κ. Let D be a discrete space of cardinality κ. For each $i \in I$ there is a (necessarily continuous, since D is discrete) surjection $f_i : D \to Y_i$; then $f = \prod_{i\in I} f_i$ is a surjective continuous function from $D^{2^\kappa} \to \prod_{i\in I} Y_i$. So it suffices to show that the space D^{2^κ} has a dense subspace of cardinality at most κ.
Step 2: Let $T := \{0, 1\}^\kappa$, so $\#T = 2^\kappa$ and T has a base $\mathcal{B}$ of cardinality κ (namely, the set of all $\prod_{i\in\kappa} Y_i$ where $Y_i \subset \{0, 1\}$ and $Y_i = \{0, 1\}$ for all but finitely many i). Let $\mathcal{T}$ be the collection of all finite families $\{U_1, \ldots, U_k\}$ of pairwise disjoint members of $\mathcal{B}$, so $\#\mathcal{T} = \kappa$.

Since $\#T = 2^\kappa$ the space D^T is homeomorphic to D^{2^κ}. Let A be the subset of all functions $f : T \to D$ such that there is $\{U_1, \ldots, U_k\} \in \mathcal{T}$ such that f is constant on each U_i and also on $T \setminus \bigcup_{i=1}^k U_i$. For each $U_1, \ldots, U_k \in \mathcal{T}$ there are $\kappa^{k+1} = \kappa$ such functions, so overall there are $\kappa \cdot \#\mathcal{T} = \kappa^2 = \kappa$: that is, $\#A = \kappa$. We claim that A is dense in D^T, which will complete the proof. Let $V \subset D^T$ be nonempty and open. Choose distinct points $t_1, \ldots, t_k \in T$ and points $y_1, \ldots, y_k \in D$ such that $\bigcap_{i=1}^k \pi_{t_i}^{-1}(y_i) \subset V$. Since T is Hasudorff, there is $\{U_1, \ldots, U_k\} \in \mathcal{T}$ such that $t_i \in U_i$ for all $1 \leq i \leq k$. Define

$$f : T \to D,\ f(t) = \begin{cases} y_i & t \in U_i \\ y_1 & t \in T \setminus \bigcup_{i=1}^k U_i \end{cases}.$$

Then $f \in A \cap V$, so $A \cap V \neq \varnothing$ and A is dense in D^T. $\square$

COROLLARY 6.20. *Let κ be an infinite cardinal. Let I be a set. For each $i \in I$, let X_i be a Hasudorff space with $\#X_i \geq 2$. Let $X := \prod_{i\in I} X_i$. The following are equivalent:*

(i) *We have $d(X) \leq \kappa$: that is, there is a dense subset Y of X of cardinality at most κ.*
(ii) *For all $i \in I$ we have $d(X_i) \leq \kappa$, and we have $\#I \leq 2^\kappa$.*

PROOF. (i) $\implies$ (ii): For $i \in I$, let $\pi_i : X \to X_i$ be the projection map. Since Y is a dense subset of X of cardinality at most κ, the subset $\pi_i(Y)$ of X_i is a dense subset of cardinality at most κ. For $i \in I$, let U_i and V_i be disjoint nonempy open subsets of X_i, and put

$$Y_i := Y \cap \pi_i^{-1}(U_i).$$

Then for all $i \in I$, we have $\varnothing \neq Y_i$, and for all $i \neq j$ we have

$$\varnothing \subsetneq Y \cap \pi_i^{-1}(U_i) \cap \pi_j^{-1}(V_j) \subset Y_i \setminus Y_j,$$

so $Y_i \neq Y_j$. Thus we have defined an injection $I \hookrightarrow 2^Y$, so $\#I \leq 2^{\#Y} \leq 2^\kappa$.
(ii) $\implies$ (i): This holds by Pondiczery-Hewitt-Marczewski Theorem. □

Corollary 6.21. a) *Let X be a countably infinite separable Hausdorff space, and let κ be an infinite cardinal. Then X^{2^κ} has cardinality 2^{2^κ} and density at most κ.*

b) *Let κ be an infinite cardinal. The maximum cardinality of a compact space X of density at most α is 2^{2^α}.*

Proof. a) First of all X has density $\aleph_0$: indeed, otherwise it would have a finite dense set, but any finite subset of a Hausdorff space is closed, so a Hausdorff space of finite density is finite. Second, for any infinite cardinal κ we have

$$2^\kappa = 2^{\aleph_0 \cdot \kappa} = (2^{\aleph_0})^\kappa = \mathfrak{c}^\kappa.$$

It is easy to see that for any cardinals $\alpha_1 \leq \alpha_2$ and $\beta_1 \leq \beta_2$ we have $\alpha_1^{\beta_1} \leq \alpha_2^{\beta_2}$, so it follows that for any $2 \leq \alpha \leq \mathfrak{c}$ we have

$$2^\kappa = \alpha^\kappa.$$

In particular, for all infinite κ we have

$$\#X^{2^\kappa} = \aleph_0^{2^\kappa} = 2^{2^\kappa}.$$

The Pondiczery-Hewitt-Marczewski Theorem shows that the density of X^{2^κ} is at most κ.
b) The $X := \{\frac{1}{n} \mid n \in \mathbb{Z}^+\} \cup \{0\} \subset \mathbb{R}$ is countably infinite compact and separable, so by part a) and Tychonoff's Theorem the space X^{2^α} is compact of cardinality 2^{2^α} and density at most α.

It follows from Exercise 6.6c) that if X is a separable Hausdorff topological space, then $\#X \leq 2^{\mathfrak{c}}$. Having come this far, let us solve Exercise 6.6c): let $Y \subset X$ be a dense subset of a Hausdorff space X. By Corollary 5.33b), for each $x \in X$ there is a prefilter F_x on X consisting entirely of subsets of Y such that $F_x \to x$. Since X is Hausdorff the assignment $x \mapsto F_x$ is injective, so the cardinality of X is at most the number of prefilters on X consisting entirely of subsets of Y, which is at most the number of families of subsets of Y: thus $\#X \leq 2^{2^{\#Y}}$.[1] □

Notice that in Corollary 6.21 we did not assert that the density of X^{2^κ} is equal to κ, only that it is at most κ. In fact Theorem 6.19 implies that the density of X^{2^κ} is the least cardinal γ such that $2^\kappa \leq 2^\gamma$. It follows that the density of $X^{2^{\aleph_0}}$ is $\aleph_0$. Moreover, if one assumes the Generalized Continuum Hypothesis (GCH), then $2^\kappa \leq 2^\gamma \iff \kappa \leq \gamma$, in which case the density of X^{2^κ} is κ for all κ. Without

[1] A different proof of this occurs in [**He46b**]: for each $x \in X$, let $\mathcal{N}_x$ be the family of open neighborhoods of x. Since Y is dense in X and X is Hausdorff, for all $x \neq y$ we have

$$\{N \cap Y \mid N \in \mathcal{N}_x\} \neq \{N \cap Y \mid N \in \mathcal{N}_y\},$$

which again gives an injection $X \hookrightarrow 2^{2^Y}$.

GCH we can say much less: e.g. there are models of ZFC set theory in which $2^{\aleph_0} = 2^{\aleph_1} = \aleph_2$, in which case the density of $X^{2^{\aleph_1}}$ would be $\aleph_0$.

3. The Lower Separation Axioms

A general topological space need not be Hausdorff, but a metrizable space is necessarily Hausdorff. The Hausdorff axiom is an example – probably the single most important example – of a "separation axiom" for a topological space. Very roughly speaking, a separation axiom is one which guarantees that certain kinds of set-theoretic distinctnesses of points or subsets are witnessed by the topology. Exactly what this means we will now explore, but one motivation for studying separation axioms is that metric topologies satisfy very strong separation axioms, so if we are looking for necessary and/or sufficient conditions for metrizability, separation axioms are the first place to look. (We will see later that metrizability is not implied by separation axioms alone, but it is a good starting point.)

Let A and B be subsets of X. It may happen that A and B do not overlap in the set-theoretic sense – i.e., $A \cap B = \emptyset$ but they are "touching" in the topological sense: e.g., the intervals $(-\infty, 0]$ and $(0, \infty)$ are "just touching." More formally, we define two subsets A and B to be **separated** if

$$\overline{A} \cap B = A \cap \overline{B} = \emptyset.$$

For subsets A, B in a topological space, $\overline{A} \cap B = \varnothing$ means that for every $b \in B$, there is an open neighborhood N_b of b which is disjoint from A. Thus the condition that A and B are separated is a sort of "disjointness with insurance.

Exercise 6.12. *Suppose (X, d) is a metric space. Show that subsets A, B of X are separated iff every point in A has positive distance from B and conversely.*

Exercise 6.13.
a) *Show: separated subsets of a topological space are disjoint.*
b) *Find an open subset A and a closed subset B of $\mathbb{R}$ which are disjoint but not separated.*
c) *Let A and B be disjoint subsets of a topological space. Suppose that A and B are either both closed or both open. Show that A and B are separated.*
d) *Find open subsets A and B of $\mathbb{R}$ which are separated but for which $\overline{A}$ and $\overline{B}$ are not separated.*

3.1. Separated spaces.

We call a space **separated**, or **Fréchet**, if for any distinct points x and y, the one-point subsets $\{x\}$ and $\{y\}$ are separated.[2]

Proposition 6.22.
a) *For a topological space X, the following are equivalent:*
(i) *X is separated.*
(ii) *For all pairs x, y of distinct points of X, there is an open set U containing x and not y.*

[2]Another common name for this separation axiom is T_1. We will not use this terminology here.

(iii) *For all $x \in X$, the singleton set $\{x\}$ is closed. (Briefly: "points are closed".)*
b) Every Hausdorff space is separated.
c) There are spaces which are separated but not Hausdorff.

Proof. a) (i) $\implies$ (ii): Suppose X is separated, and let x, y be distinct points of x. The existence of an open set containing x and not y is equivalent to y not lying in the closure of x, which follows from the definition of the sets $\{x\}$ and $\{y\}$ being separated. (ii) $\implies$ (iii): If $\{x\}$ is not closed, then there are $y \neq x$ such that every open neighborhood of x contains y. (iii) $\implies$ (i) is immediate.
b) Suppose X is Hausdorff, and let $x, y \in X$. Then there are distinct open neighborhoods U_x and U_y of x and y respectively. In particular $y \notin U_x$, so $y \notin \overline{\{x\}}$. Therefore $\{x\}$ is closed.
c) The cofinite topology on an infinite set is separated but not Hausdorff. □

Exercise 6.14. *Show that being separated is a refineable property: if (X, τ_1) is separated and $\tau_2 \supset \tau_1$ is a finer topology on X, then (X, τ_2) is separated.*

Exercise 6.15. *Let X be a separated space and $q : X \to Y$ a quotient map. Show that Y is separated iff all the fibers of q are closed.*

3.2. Kolmogorov spaces and the Kolmogorov quotient.

In many branches of modern mathematics, a yet weaker separation axiom turns out to be more useful. One way to motivate it is by consideration of the following relation on a topological space X: we say that $x, y \in X$ are **topologically indistinguishable** if for all open sets U of x, $x \in U \iff y \in U$. We write $x \sim y$ iff x and y are topologically indistinguishable.

Exercise 6.16. *Let X be a topological space. Show: topological indistinguishability is an equivalence relation on X.*

A space X is **Kolmogorov**[3] if the relation of topological indistinguishability is simply equality: for all $x, y \in X$, $x \sim y \iff x = y$.

Proposition 6.23. *a) A topological space is Kolmogorov iff, for any two disinct points $x, y \in X$, either there is an open set U containing x and not y, or there is an open set V containing y and not x (or both).*
b) A separated space is Kolmogorov.
c) There are spaces which are Kolmogorov and not separated.

Proof. a) This is a simple unwinding of the definition and is left to the reader.
b)By Proposition 6.22, a space is separated iff for any distinct points $x, y \in X$, there is an open set U containing x and not y, hence by part a) X is Kolmogorov.
c) The Sierpinski space – a two-point set $X = \{\circ, \bullet\}$ with topology $\tau = \{\varnothing, \circ, X\}$ – is Kolomogorov but not separated. □

Lemma 6.24. *Let $f : X \to Y$ be a continuous map between topological spaces. If $x_1, x_2 \in X$ are topologically indistinguishable, then $f(x_1), f(x_2) \in Y$ are topologically indistinguishable.*

[3]It is common to call such spaces T_0.

PROOF. We argue by contraposition: suppose $y_1 = f(x_1)$ and $y_2 = f(x_2)$ are topologically distinguishable in Y; without loss of generality, we may assume that there is an open set V in Y containing y_1 but not y_2. Then $f^{-1}(V)$ is an open subset of X containing x_1 but not x_2. $\square$

Let X be a topological space and let $\sim$ be the equivalence relation of topological indistinguishability on X. Let $X_K = X/\sim$ be the set of $\sim$-equivalence classes and $q : X \to \overline{X}$ the quotient map. We endow X_K with the quotient topology – a subset of X_K is open iff its preimage in X is open – and then the space X_K and the continuous map $q : X \to X_K$ is called the **Kolmogorov quotient** of X.

PROPOSITION 6.25.
Let X be a topological space and $q : X \to X_K$ its Kolmogorov quotient.
a) The map q induces a bijection from the open sets of X to the open sets of X_K.
b) The space X_K is a Kolmogorov space.
c) The map q is universal *for continuous maps from X into a Kolmogorov space: i.e., for any Kolmogorov space Y and continuous map $f : X \to Y$, there is a unique continuous map $\overline{f} : X_K \to Y$ such that $f = \overline{f} \circ q$.*

PROOF. a) We claim that q (direct image) and q^{-1} (inverse image) are mutually inverse functions from the set of open sets of X to the set of open sets of X_K. For any quotient map $q : X \to Y$ and any open subset V of Y, one has $q(q^{-1}(V)) = V$. The other direction is more particular to the current situation: reall that a quotient map need not be open. But for any open subset U of X, $q^{-1}(q(U))$ is the set of all points which are topologically indistinguishable from some element of U. This set plainly contains U, and conversely if $x \in U$ and $y \in X \setminus U$, then U itself is an open set distinguishing x from Y, so $q^{-1}(q(U)) = U$.
b) Let $y_1 \neq y_2 \in X_K$, and choose $x_1 \in q^{-1}(y_1)$, $x_2 \in q^{-1}(y_2)$. Because $y_1 \neq y_2$, there is an open set U of X which either contains x_1 and not x_2 or contains x_2 and not x_1; relabelling if necessary, we suppose that $x_1 \in U$ and $x_2 \notin U$. By part a), $q(U)$ is open in Y, so $y_1 \in q(U)$. If we had $y_2 \in q(U)$, then we would have $x_2 \in q^{-1}(q(U)) = U$, contradiction.
c) By Lemma 6.24, f factors through q. The resulting map F is unique, and is continuous by the universal property of quotient maps. $\square$

The upshot is that, intuitively speaking, passing to the Kolmogorov quotient does not disturb the underlying topology – only the underlying set! That doesn't quite make sense in the standard set-theoretic setup for topology (to be sure, the only one we are considering!) but one can make sense of it via the theory of **locales**.

EXERCISE 6.17. *Show: Kolmogorov completion is a functor and is left adjoint to the forgetful functor from Kolmogorov spaces to topological spaces.*

EXERCISE 6.18. *Show: a space is quasi-compact iff its Kolmogorov quotient is quasi-compact.*

3.3. The specialization quasi-ordering.

We define a second relation on the points of a topological space X. Namely, for $x, y \in X$, we say that y **is a specialization** of x if $y \in \overline{\{x\}}$.

Many of the concepts we have been exploring in this section can be interpreted in terms of a specialization relation. In particular, a point is closed iff it does not specialize to any other point. Thus, a space is separated iff the specialization relation is equality. Moreover, two points x and y are topologically indistinguishable iff x specializes to y and y specializes to x.

In general, a binary relation R on a set X is a **preordering** if it satisfies the following axioms:

(PO1) For all $x \in X$, xRx (reflexivity).
(PO2) For all $x, y, z \in X$, xRy and yRz implies xRz (transitivity).

LEMMA 6.26. *Let R be any preorder on a set X, and define a new relation $\sim$ on X by $x \sim y$ if xRy and yRx. Then:*
a) The relation $\sim$ is an equivalence relation on X. Put $\overline{X} = X/\sim$.
b) The relation R descends to a partial ordering on $\overline{X}$.

EXERCISE 6.19. *Prove it.*

PROPOSITION 6.27. *Let X be a topological space.*

a) *X is Kolmogorov iff the specialization relation is a partial ordering (equivalently, if it antisymmetric).*
b) *For any space X, the quotient by the specialization relation is, as a partially ordered set, canonically isomorphic to the Kolmogorov quotient.*

EXERCISE 6.20. *Prove it.*

EXERCISE 6.21. *Let $f : X \to Y$ be a continuous map of topological spaces.*

a) *Show that the map is compatible with the specialization preorderings on X and Y, in the following sense: if $x_1 \preccurlyeq x_2$ in X, then $f(x_1) \preccurlyeq f(x_2)$ in Y.*
b) *Use part a) to define a functor $\mathcal{P}$ from the category of topological spaces and continuous maps to the category of preordered sets and preorder-preserving maps.*

It is natural to ask what the essential image of $\mathcal{P}$ is, i.e., which preordered sets, up to isomorphism, arise from the specialization preorder on a topological space? To answer this we will define a functor in the other direction.

If $(X, \preccurlyeq)$ is a quasi-ordered set, an **upward set** in X is a subset Y of X such that for all $y \in Y$ and $x \in X$, if $y \preccurlyeq x$, then $x \in Y$. Similarly, a subset Y of X is a **downward set** if for all $y \in Y$ and $x \in X$, if $x \preccurlyeq y$, then $x \in Y$.

Alexandroff space of a preordered set: let $(X, \preccurlyeq)$ be a preordered set. Let τ_X be the family of all downward sets in X. It is easy to see that τ_X contains $\varnothing$ and X and is closed under arbitrary unions and also arbitrary intersections. In particular τ_X is a topology on X, and (X, τ_X) is called the **Alexandroff topology** on $(X, \preccurlyeq)$.

EXERCISE 6.22. *Let X be any set.*

a) *Endow X with the **trivial** quasi-ordering – $x \preccurlyeq y \iff x = y$ – and show that the associated Alexandroff topology is the discrete topology.*
b) *Endow X with the **discrete** quasi-ordering – for all $x, y \in X$, $x \preccurlyeq y$ – and show that the associated Alexandroff topology is the trivial (or indiscrete) topology.*

c) *Endow X with a nontrivial partial ordering. Show that the associated Alexandroff topology is Kolmogorov but not separated.*

EXERCISE 6.23. *Show that $(X, \preccurlyeq) \mapsto (X, \tau_X)$ extends to a functor $\mathcal{T}$ from the category of topological spaces and continuous maps to the category of preordered sets and preorder-preserving maps.*

PROPOSITION 6.28. *Let $(X, \preccurlyeq)$ be a preordered set. Then the identity map $X \mapsto \mathcal{P}(\mathcal{T}(X))$ is an isomorphism of preordered spaces. It follows that every preordered space is, up to isomorphism, the specialization preordering on some topological space.*

PROOF. Let $x_1,\ x_2 \in X$. Suppose first that $x_1 \leq x_2$. Then every downward set which contains x_2 also contains x_1, i.e., every τ_X-open set containing x_2 also contains x_1, so x_2 is a specialization of x_1. Now suppose that x_1 is not less than or equal to x_2. Then the downward set $D(x_2)$ of all elements less than or equal to x_2 is a τ_X-open set containing x_2 but not x_1, so x_2 is not a specialization of x_1. $\square$

This answers the question of which preordered sets arise as a specialization preorder, but gives rise to another question: which topological spaces are the Alexandroff topology of some preorder on the underlying set? Note that here the answer is certainly not "all of them", because the Alexandroff topology on $(X, \preccurlyeq)$ has a property which most topologies lack: the family of open sets is closed under not just finite intersections but *arbitrary intersections*. This gives rise to interesting class of topological spaces which we study next.

3.4. Alexandroff Spaces.

PROPOSITION 6.29. *For a topological space X, the following are equivalent:*
(i) If $\{U_i\}_{i \in I}$ is any family of open sets of X, $\bigcap_{i \in I} U_i$ is open.
(ii) If $\{F_i\}_{i \in I}$ is any family of closed sets of X, $\bigcap_{i \in I} F_i$ is closed.
(iii) Every $x \in X$ has a unique minimal open neighborhood.
(iv) Every downward set in the specialization quasi-ordering is open.
(v) For every $S \subset X$ and $y \in \overline{S}$, there is $x \in S$ such that x specializes to y.
(vi) For every $S \subset X$ and $y \in \overline{S}$, there is a finite subset S' of S such that $y \in \overline{S}$.
A space satisfying these equivalent conditions is called an ***Alexandroff space****.*

PROOF. Obviously (i) $\iff$ (ii) by complementation.
(i) $\iff$ (iii): Note that (iii) amounts to: for every $x \in X$, the intersection of all open neighborhoods of x is open, say equal to $N(x)$. So certainly (i) $\implies$ (iii). Conversely, suppose (iii) holds, let $\{U_i\}_{i \in I}$ be a family of open sets, and let $x \in U = \bigcap_i U_i$. Then $N(x) \subset U_i$ for all i, so $N(x) \subset U$ and x is an interior point of U. Since x was arbitrary, U is open.
(iii) $\iff$ (iv): Let $x, y \in X$. Then $y \in N(x)$ iff y lies in every open neighborhood of x iff $x \in \overline{y}$ iff $y \preccurlyeq x$ in the specialization preorder. Thus $N(x)$ is precisely the principal downard set associated to x, and (iii) is equivalent to each of these sets being open. So (iv) $\implies$ (iii). Moreover, since any downard set is the union of its principal downward subsets, (iii) $\implies$ (iv).
(ii) $\implies$ (v): Since $y \in \overline{S}$, there is $x \in N(y) \cap S$.
(v) $\implies$ (vi) trivially.
(vi) $\implies$ (iu): Let $\{F_i\}_{i \in I}$ be a family of closed sets of X, put $F = \bigcup_{i \in I} F_i$, and let $x \in \overline{F}$. By assumption, there exist $x_1, \ldots, x_n \in F$ such that $x \in \overline{\{x_1, \ldots, x_n\}}$.

For each $1 \leq j \leq n$, x_j lies in some F_{i_j}, so that $\{x_1, \ldots, x_n\} \subset F' = \bigcup_{j=1}^n F_{i_j}$. Since F' is a finite union of closed sets, it is closed, and thus

$$x \in \overline{\{x_1, \ldots, x_n\}} \subset \overline{F'} \subset \overline{F}.$$

Since x was arbitrary, F is closed. □

Finite spaces, discrete and indiscrete spaces are all Alexandroff.

Exercise 6.24. *Show: an Alexandroff space is separated iff it is discrete.*

Exercise 6.25. *Show: the class of Alexandroff spaces is closed under passage to subspaces and finite products.*

Proposition 6.30. *A quotient of an Alexandroff space is Alexandroff.*

Proof. Let X be an Alexandroff space and $q : X \to Y$ be a quotient map. Let $\{V_i\}_{i \in I}$ is a family of open subsets of Y and put $V = \bigcap_i V_i$. Then

$$f^{-1}(V) = f^{-1}(\bigcap_i f^{-1}(V_i)) = \bigcap_i f^{-1}(V_i)$$

is open, since f is continuous and X is Alexandroff. By definition of the quotient topology, this implies that V is open in Y. □

Exercise 6.26. *Let X be an Alexandroff space and $f : X \to Y$ be continuous, open and surjective. Show that Y is an Alexandroff space.*

In particular, the Kolmogorov quotient of an Alexandroff space is Alexandroff and Kolmogorov. This is the topological analogue of passing from a quasi-order to its associated partial order. An Alexandroff spcace is Kolmogorov iff the assignment $x \in X \mapsto D(x)$ is injective.

Proposition 6.31. *Let X be an Alexandroff space and $x \in X$. Then the principal downset $D(x)$ is quasi-compact.*

Proof. Indeed, since $D(x)$ is the unique minimal open neighborhood of x, in any covering of $D(x)$ by open subsets of X, at least one of the elements U of the cover must contain $D(x)$, so $\{U\}$ is a finite subcovering. □

Note that this gives many examples of quasi-compact Alexandroff spaces, namely the Alexandroff topology on a quasi-ordered set X with a top element, i.e., an element x_T such that for all $x \in X$, $x \leq x_T$.

For any topological space X, we define its **Alexandroff completion** to be $\mathcal{T}(\mathcal{P}X)$, i.e., the topological space with the same underlying set as X but retopologized so that the open sets are precisely the downward sets for the specialization preordering on X. By Proposition 6.28, passage to the Alexandroff completion does not change the specialization preordering, so in particular a space is Kolmogorov (resp. separated) iff its Alexandroff completion is Kolmogorov (resp. separated). But of course most spaces are not Alexandroff, so the Alexandroff completion usually carries a different topology.

Example 6.4. *Let X be a set endowed with the cofinite topology. Then X is separated, so the specialization preorder is the trivial order, so the Alexandroff completion is discrete. On the other hand X is itself quasi-compact, so X coincides with its Alexandroff completion iff it is a finite space.*

EXAMPLE 6.5. *Let $Y = X \cup \{\eta\}$, where X is an infinite set. We topologize Y as follows: a nonempty subset of Y is open iff it contains η and is cofinite. In this topology, the points of X are each closed whereas the closure of η is all of Y. The specialization preordering on Y is as follows: no two distinct points of X specialize to each other, whereas η specializes to every point of X. In particular X is quasi-compact, Kolmogorov but not separated. In the Alexandroff completion of Y, the minimal open sets are the singleton set η and the pairs $\{\eta, x\}$ for $x \in X$. In other words, this is the topology – seen at the very beginning of our notes but not "in nature" until now – in which a subset of Y is open iff it contains η. This new topology is far from being quasi-compact.*

In both of these examples, passage to the Alexandroff completion resulted in a finer topology. The following result establishes this, and a little more.

PROPOSITION 6.32. *Let $(X, \preccurlyeq)$ be a preordered set. Then the Alexandroff topology (X, τ_X) is the finest topology τ on X such that the associated specialization preordering coincides with $\preccurlyeq$.*

PROOF. Let (X, τ) be a topological space with specialization preordering $\preccurlyeq$. It suffices to show: if $U \in \tau$ and $x \in U$, then the principal downset $D(x) = \{y \mid y \leq x\}$ is contained in U. But indeed, $y \in D(x)$ iff $x \in \overline{y}$ iff every open neighborhood N_x of x meets $\{y\}$. So in particular U meets y, i.e., $y \in U$. □

An equivalent phrasing of Proposition 6.32 is that, for any topological space X, the identity map $\mathcal{T}(\mathcal{P}X) \to X$ is continuous. It follows that every topological space is the continuous image of an Alexandroff space.

COROLLARY 6.33. *a)The functors $\mathcal{P}$ and $\mathcal{T}$ induce an equivalence between the category of Alexandroff topological spaces and the category of preordered sets.*

3.5. Irreducible spaces, Noetherian spaces, and sober spaces.

A topological space is **irreducible** if it is nonempty and if it cannot be expressed as the union of two proper closed subsets.

EXERCISE 6.27. *Show that for a Hausdorff topological space X, the following are equivalent:*

(i) *X is irreducible.*
(ii) *$\#X = 1$.*

PROPOSITION 6.34. *For a topological space X, the following are equivalent:*

(i) *X is irreducible.*
(ii) *Every finite intersection of nonempty open subsets (including the empty intersection!) is nonempty.*
(iii) *Every nonempty open subset of X is dense.*
(iv) *Every open subset of X is connected.*

EXERCISE 6.28. *Prove Proposition 6.34.*

PROPOSITION 6.35. *Let X be a nonempty topological space.*
a) If X is irreducible, every nonempty open subset of X is irreducible.
b) If a subset Y of X is irreducible, so is its closure $\overline{Y}$.
c) If $\{U_i\}$ is an open covering of X such that $U_i \cap U_j \neq \emptyset$ for all i, j and each U_i

is irreducible, then X is irreducible.
d) If $f : X \to Y$ is continuous and X is irreducible, then $f(X)$ is irreducible in Y.

PROOF. a) Let U be a nonempty open subset of X. By Proposition 6.34, it suffices to show that any nonempty open subset V of U is dense. But V is also a nonempty open subset of the irreducible space X.
b) Suppose $\overline{Y} = A \cup B$ where A and B are each proper closed subsets of $\overline{Y}$; since $\overline{Y}$ is itself closed, A and B are closed in X, and then $Y = (Y \cap A) \cup (Y \cap B)$. If $Y \cap A = Y$ then $Y \subset A$ and hence $\overline{Y} \subset \overline{A} = A$, contradiction. So A is proper in Y and similarly so is B, thus Y is not irreducible.
c) Let V be a nonempty open subset of X. Since the U_i's are a covering of X, there is at least one i such that $V \cap U_i \neq \emptyset$, and thus by irreducibility $V \cap U_i$ is a dense open subset of U_i. Therefore, for any index j, $V \cap U_i$ intersects the nonempty open subset $U_j \cap U_i$, so in particular V intersects every element U_j of the covering. Thus for all sets U_i in an open covering, $V \cap U_i$ is dense in U_i, so V is dense in X.
d) If $f(X)$ is not irreducible, there exist closed subsets A and B of Y such that $A \cap f(X)$ and $B \cap f(X)$ are both proper subsets of $f(X)$ and $f(X) \subset A \cup B$. Then $f^{-1}(A)$ and $f^{-1}(B)$ are proper closed subsets of X whose union is all of X. □

EXERCISE 6.29. *Show: the union of a chain of irreducible subspaces is irreducible.*

Let x be a point of a topological space, and consider the set of all irreducible subspaces of X containing x. (Since $\{x\}$ itself is irreducible, this set is nonempty.) Applying Exercise 6.29 and Zorn's Lemma, there is at least one maximal irreducible subset containing x. A maximal irreducible subset – which by Proposition 6.35b) is necessarily closed – is called an **irreducible component** of X. Since irreducible subsets are connected, each irreducible component lies in a unique connected component, and each connected component is the union of its irreducible components.

However, unlike connected components, it is possible for a given point to lie in more than one irreducible component. We will see examples shortly.

In the case of the Zariski topology $\operatorname{Spec} R$, there is an important algebraic interpretation of the irreducible components. Namely, the irreducible components Y of $\operatorname{Spec} R$ correspond to $V(\mathfrak{p})$ where $\mathfrak{p}$ ranges through the **minimal primes**.

PROPOSITION 6.36. *For an ideal I of R, the closed subset $V(I)$ is irreducible iff the radical ideal*
$$\operatorname{rad}(I) = \{x \in R \mid \exists n \in \mathbb{Z}^+ x^n \in I\}$$
is prime.

PROOF. See [**CA**, §9.1]. □

It follows that the irreducible components – i.e., the maximal irreducible subsets – are the sets of the form $V(\mathfrak{p})$ as $\mathfrak{p}$ ranges over the distinct minimal prime ideals.

PROPOSITION 6.37. *For a topological space X, the following are equivalent:*
(i) Every ascending chain of open subsets is eventually constant.
(ibis) Every descending chain of closed subsets is eventually constant.
(ii) Every nonempty family of open subsets has a maximal element.
(iibis) Every nonempty family of closed subsets has a minimal element.

(iii) Every open subset is quasi-compact.
(iv) Every subset is quasi-compact.
A space satisfying any (and hence all) of these conditions is called ***Noetherian****.*

Proof. The equivalence of (i) and (ibis), and of (ii) and (iibis) is immediate from taking complements. The equivalence of (i) and (ii) is a general property of partially ordered sets.

(i) $\iff$ (iii): Assume (i), let U be any open set in X and let $\{V_j\}$ be an open covering of U. We assume for a contradiction that there is no finite subcovering. Choose any j_1 and put $U_1 := V_{j_1}$. Since $U_1 \neq U$, there exists j_2 such that U_1 does not contain V_{j_2}, and put $U_2 = U_1 \cup V_{j_2}$. Again our assumpion implies that $U_2 \supsetneq U$, and continuing in this fashion we will construct an infinite properly ascending chain of open subsets of X, contradiction. Conversely, assume (iii) and let $\{U_i\}_{i=1}^{\infty}$ be an infinite properly ascending chain of subsets. Then $U = \bigcup_i U_i$ is not quasi-compact.

Obviously (iv) $\implies$ (iii), so finally we will show that (iii) $\implies$ (iv). Suppose that $Y \subset X$ is not quasi-compact, and let $\{V_i\}_{i \in I}$ be a covering of Y by relatively open subsets without a finite subcover. We may write each V_i as $U_i \cap Y$ with U_i open in Y. Put $U = \bigcup_i U_i$. Then, since U is quasi-compact, there exists a finite subset $J \subset I$ such that $U = \bigcup_{j \in J} U_j$, and then $Y = U \cap Y = \bigcup_{j \in J} U_j \cap Y = \bigcup_{j \in J} V_j$. □

Corollary 6.38. *A Noetherian Hausdorff space is finite.*

Exercise 6.30. *Prove Corollary 6.38.*

Proposition 6.39. *Let X be a Noetherian topological space.*
a) There are finitely many closed irreducible subsets $\{A_i\}_{i=1}^n$ such that $X = \bigcup_{i=1}^n A_i$.
b) Starting with any finite family $\{A_i\}_{i=1}^n$ as in part a) and eliminating all redundant sets – i.e., all A_i such that $A_i \subset A_j$ for some $j \neq i$ – we arrive at the set of irreducible components of X. In particular, the irreducible components of a Noetherian space are finite in number.

Proof. a) Let X be a Noetherian topological space. We first claim that X can be expressed as a finite union of irreducible closed subsets. Indeed, consider the collection of closed subsets of X which cannot be expressed as a finite union of irreducible closed subsets. If this collection is nonempty, then by Proposition 6.37 there exists a minimal element Y. Certainly Y is not itself irreducible, so is the union of two strictly smaller closed subsets Z_1 and Z_2. But Z_1 and Z_2, being strictly smaller than Y, must therefore be expressible as finite unions of irreducible closed subsets and therefore so also can Y be so expressed, contradiction.

b) So write

$$X = A_1 \cup \ldots \cup A_n$$

where each A_i is closed and irreducible. If for some $i \neq j$ we have $A_i \subset A_j$, then we call A_i **redundant** and remove it from our list. After a finite number of such removals, we may assume that the above finite covering of X by closed irreducibles is **irredundant** in the sense that there are no containment relations between distinct A_i's. Now let Z be any irreducible closed subset. Since $Z = \bigcup_{i=1}^n (Z \cap A_i)$ and Z is irreducible, we must have $Z = Z \cap A_i$ for some i, i.e., $Z \subset A_i$. It follows that the "irredundant" A_i's are precisely the maximal irreducible closed subsets, i.e., the irreducible components. □

We deduce the following important result, which is not so straightforward to prove using purely algebraic methods:

COROLLARY 6.40. *Let I be a proper ideal in a Noetherian ring R. The set of prime ideals $\mathfrak{p}$ which are minimal over I (i.e., minimal among all prime ideals containing I) is finite and nonempty.*

EXERCISE 6.31. *Prove Corollary 6.40.*

4. More on Hausdorff Spaces

Recall that a topological space X is **Hausdorff** if for each pair x, y of distinct points in X, there exist open neighborhoods U_x, U_y of x and y such that $U_x \cap U_y = \varnothing$.

PROPOSITION 6.41. *The Hausdorff property is hereditary.*

PROOF. Let Y be a subspace of the Hausdorff space X, and let $y_1 \neq y_2 \in Y$. Since X is Hausdorff there are disjoint open sets U_1 and U_2 of X with $y_1 \in U_1$ and $y_2 \in U_2$. Then $V_1 = U_1 \cap Y$ and $V_2 = U_2 \cap Y$ are dsijoint open sets of Y containing y_1 and y_2 respectively. $\square$

EXERCISE 6.32. *Let X be an infinite Hausdorff space.*

a) *Show: there is a nonempty open subset U of X such that $X \setminus \overline{U}$ is infinite.*
b) *Show: X admits a countably infinite discrete subspace.*

PROPOSITION 6.42. *The Hausdorff property is faithfully productive: that is, let $\{X_i\}_{i \in I}$ be a nonempty family of nonempty topological spaces, and let $X = \prod_{i \in I} X_i$, endowed with the product topology. Then X is Hausdorff iff for all $i \in I$, X_i is Hausdorff.*

PROOF. Suppose X is Hausdorff. Since Hausdorff is a hereditary property, it follows from Corollary 3.20 that each X_i is Hausdorff. Suppose X_i is Hausdorff for all $i \in I$ and let $x \neq y \in X$. Then there is $i \in I$ such that $x_i \neq y_i$. Let U_i and V_i be disjoint open subsets of X_i containing x_i and y_i respectively. Then $U = \pi_i^{-1}(U_i)$ and $V = \pi_i^{-1}(V_i)$ are disjoint open subsets of X containing x and y respectively. $\square$

PROPOSITION 6.43.
a) *The continuous open image of a Hausdorff space need not be Hausdorff.*
b) *If X is Hausdorff and $q : X \to Y$ is a closed quotient map, then Y need not be Hausdorff.*

PROOF. [**Wi**, p. 88]. $\square$

For a set X, we define the **diagonal map** $\Delta_X : X \hookrightarrow X \times X$ by $x \mapsto (x, x)$. It is plainly an injection. If X is a topological space, we claim that Δ_X is moreover an embedding, i.e., continuous and open. Indeed, let x in X. Then a neighborhood base of (x, x) in $X \times X$ is given by sets of the form $U \times V$, where U and V are both open neighborhoods of x in X. Then $\Delta_X^{-1}(U \times V) = U \times V$ is open in X, so Δ_X is continuous at x. Moreover, for any open subset U of X, $\Delta_X(U) = U \times U$ is open in $X \times X$.

EXAMPLE 6.6. *(The line with two origins): Let X be the union of two lines in $\mathbb{R}^2$, say $y = 0$ and $y = 1$. We define a quotient of X via the following equivalence relation: if $x \neq 0$, $(x, 0) \sim (x, 1)$, but $(0, 0)$ is not equivalent to $(0, 1)$. The quotient $Y = X/\sim$ is "almost" homeomorphic to the Euclidean line, except that it has "two origins". Y is locally Euclidean: for any $\epsilon > 0$, $((\epsilon, \epsilon) \times \{1\}) \cup ((-\epsilon, 0) \times \{0\}) \cup ((0, \epsilon) \times \{0\})$ is a neighborhood base at the image of $(0, 1)$ in Y each of whose* *elements is disjoint from* $(0,0)$*. In particular* Y *is separated. But it is evidently not Hausdorff.*

Proposition 6.44. *Let* $f : X \to Y$ *be a continuous map with* Y *a Hausdorff space. The set* $S = \{(x_1, x_2) \in X \times X \mid f(x_1) = f(x_2)\}$ *is closed in* $X \times X$*.*

Proof. If $(x_1, x_2) \in X \times X \setminus S$, then $f(x_1) \neq f(x_2)$. Since Y is Hausdorff, there exist disjoint open neighborhoods V_1 of $f(x_1)$ and V_2 of $f(x_2)$. Then $f^{-1}(V_1) \times f^{-1}(V_2)$ is an open neighborhood of (x_1, x_2) in $X \times X$ which is disjoint from S. $\square$

The following result gives a necessary and sufficient condition for the image under an open quotient map to be Hausdorff.

Theorem 6.45. *Let* $f : X \to Y$ *be an continuous, open and surjective. Then the following are equivalent:*
(i) Y *is Hausdorff.*
(ii) $S = \{(x_1, x_2) \in X \times X \mid f(x_1) = f(x_2)\}$ *is closed in* $X \times X$*.*

Proof. By Proposition 6.44, (i) $\implies$ (ii) (even without the hypothesis that f is an open quotient map). Conversely, assume that S is closed in $X \times X$, and let $f(x_1)$, $f(x_2)$ be distinct points of Y. Then $(x_1, x_2) \notin S$, so there exist open neighborhoods U_1, U_2 of x_1, x_2 in X such that $(U_1 \times U_2) \cap S = \varnothing$. Since f is open, $V_1 = f(U_1)$ and $V_2 = f(U_2)$ are open neighborhoods of $f(x_1)$, $f(x_2)$. If there existed a $y \in V_1 \cap V_2$, then there exist $x_1' \in U_1$ and $x_2' \in U_2$ such that $f(x_1') = y = f(x_2')$, contradicting the fact that $(U_1 \times U_2) \cap S = \varnothing$. $\square$

Exercise 6.33. [**Wi**, Exc. 13H] *Show that for every topological space* Y *there is a Hausdorff space* X *and a continuous, open surjection* $f : X \to Y$*.*

Proposition 6.46. *Let* X *be a space,* Y *a Hausdorff space and* $f, g : X \to Y$ *two continuous functions.*

a) *The set* $E(f, g) = \{x \in X \mid f(x) = g(x)\}$ *is closed in* X*.*
b) *If* f *and* g *agree on a dense subset of* X*, then* $f = g$*.*

Exercise 6.34. *Prove it.*

Exercise 6.35. *Recall that for any function* $f : X \to Y$*, the graph of* f *is*

$$G(f) = \{(x, f(x) \mid x \in X\} \subset X \times Y\}.$$

a) *Show that if* f *is continuous and* Y *is Hausdorff then* $G(f)$ *is closed.*
b) *Find a discontinuous function* $f : \mathbb{R} \to \mathbb{R}$ *for which* $G(f)$ *is closed.*

Exercise 6.36. *(Insel)*

a) *Suppose* X *is first countable and every quasi-compact subset of* X *is closed. Show:* X *is Hausdorff.*
b) *Give a counterexameple to part a) with the hypothesis of first countability omitted.*

5. Regularity and Normality

Let A, B be subsets of a topological space X. We say that A and B are **separated by open sets** if there are disjoint open subsets U, V of X with $A \subset U$, $B \subset V$.

A topological space X is **quasi-regular** if for every point $p \in X$ and every closed subset $A \subset X$, if $p \notin A$ then $\{p\}$ and A can be separated by open sets. A topological space is **regular** if it is quasi-regular and Hausdorff. A topological space X is **quasi-normal** if every pair of disjoint closed subsets can be separated by open sets. A topological space is **normal** if it is quasi-normal and Hausdorff.

EXERCISE 6.37. *Show: the Moore-Niemytzki plane is Hausdorff but not regular.*

The following exercise should help to explain the "quasi"s.

EXERCISE 6.38. *a) Show that normal spaces are regular.*
b) Show that the Sierpinski space is quasi-regular but not quasi-normal.

PROPOSITION 6.47. a) *For a topological space X, the following are equivalent:*
(i) *The space X is quasi-regular.*
(ii) *Every point of X admits a neighborhood base of closed neighborhoods.*

b) *For a topological space X, the following are equivalent:*
(i) *The space X is quasi-normal.*
(ii) *For all subsets $B \subset U \subset X$ with B closed and U open, there is an open subset V with*

$$B \subset V \subset \overline{V} \subset U.$$

PROOF. a) (i) $\implies$ (ii) Let $p \in X$, and let U be an open set containing p. Then $A = X \setminus U$ is closed and $p \notin A$, so by assumption there are disjoint open sets V containing p and W containing A. Then $\overline{V} \cap A = \varnothing$: indeed, if $x \in A$, then W is a neighborhood of x disjoint from V. So $p \in \overline{V} \subset U$.
(ii) $\implies$ (i): Let $A \subset X$ be closed, let $U = X \setminus A$, and let $p \in U$. By hypothesis, there is an open neighborhood V of p with $p \in \overline{V} \subset U$. Then V and $X \setminus \overline{V}$ are disjoint open sets with $p \in V$ and $A \subset X \setminus \overline{V}$.
b) (i) $\implies$ (ii): Let $B \subset U \subset X$ with A closed and U open. Let $A = X \setminus U$, so A is closed and $A \cap B = \varnothing$. By hypothesis there are disjoint open sets V containing B and W containing A. As above, we have $\overline{V} \cap A = \varnothing$, so $\overline{V} \subset U$.
(ii) $\implies$ (i): Let A and B be disjoint closed subsets of X. Let $U = X \setminus A$, so $B \subset U$. By hypothesis there is an open subset V with $B \subset V \subset \overline{V} \subset U$. Then V and $X \setminus \overline{V}$ are disjoint open sets containing B and A respectively. □

PROPOSITION 6.48. a) *A space is quasi-regular iff its Kolmogorov quotient is regular.*
b) *In particular, a Kolmogorov quasi-regular space is regular.*
c) *A space is quasi-normal iff its Kolmogorov quotient is normal.*
d) *In particular, a Kolmogorov quasi-normal space is normal.*

PROOF. It suffices to prove parts a) and c); parts b) and d) follow immediately. □

PROPOSITION 6.49. a) *Quasi-regularity and regularity are hereditary properies: subspaces of quasi-regular (resp. regular) spaces are regular.*
b) *Quasi-regularity and regularity are faithfully productive properties: if $\{X_i\}_{i\in I}$ is a family of nonempty topological spaces, then $X = \prod_{i\in I} X_i$ is quasi-regular (resp. regular) iff each X_i is quasi-regular (resp. regular).*

PROOF. It is enough to show parts a) and b) for quasi-regular spaces and combine with the analogous result for separated spaces.

a) Let X be a quasi-regular space, let $Y \subset X$, let $B \subset Y$ be closed in Y and let $y \in Y \setminus B$. Then there is a closed subset $A \subset X$ such that $B = A \cap Y$. Since $y \in Y$ and $y \notin B$ we have $y \notin A$. By quasi-regularity, there are disjoint open subsets U, V of X with $y \in U$ and $A \subset V$. The subsets $U \cap Y$ and $V \cap Y$ are disjoint, open in Y, and contain y and B respectively.
b) As usual, since each X_i is homeomorphic to a space of $X = \prod_{i \in I} X_i$, if X is quasi-regular, then it follows from part a) that each X_i is quasi-regular. Conversely, suppose each X_i is quasi-regular, let $x \in X$, and consider a basic neighborhood

$$U = \bigcap_{j=1}^{n} \pi_{i_j}^{-1}(U_{i_j})$$

of x in X. Then each U_{i_j} is a neighborhood of $x_{i_j} = \pi_{i_j}(x)$ in X_{i_j}, so by Proposition 6.47a) there is a closed neighborhood C_{i_j} of x_{i_j} contained in U_{i_j}. Then

$$C = \bigcap_{j=1}^{n} \pi_{i_j}^{-1}(C_{i_j})$$

is a closed neighborhood of x contained in U. So X is quasi-regular. $\square$

Theorem 6.50. *(Ubiquity of Normality)*

a) *Metrizable spaces are normal.*
b) *Compact spaces are normal.*
c) *(Tychonoff's Lemma) Regular Lindelöf spaces are normal.*
d) *Order spaces are normal.*

Proof. a) Let A, B be disjoint closed subsets of X. Since $A \cap \overline{B} = \emptyset$, for every $a \in A$, there exists $\epsilon_a > 0$ such that $B(a, \epsilon_a) \cap B = \emptyset$. Similarly, since $B \cap \overline{A} = \emptyset$, for every $b \in B$, there exists $\epsilon_b > 0$ such that $B(b, \epsilon_b) \cap A = \emptyset$. Put $U = \bigcup_{a \in A} B(a, \frac{\epsilon_a}{2})$ and $V = \bigcup_{b \in B} B(a, \frac{\epsilon_b}{2})$. Then $U \cap V = \emptyset$. Indeed, suppose $x \in U \cap V$; then there exist $a \in a$ and $b \in B$ such that $x \in B(a, \frac{\epsilon_a}{2}) \cap B(b, \frac{\epsilon_b}{2})$. Then

$$d(a, b) < \frac{\epsilon_a + \epsilon_b}{2} \leq \max\{\epsilon_a, \epsilon_b\}.$$

That is, either $d(a, b) < \epsilon_a$ – in which case there exists a point of B in $B(a, \epsilon_a)$, a contradiction – or $d(a, b) < \epsilon_b$, which is similarly contradictory.
b) Step 1: We will show that X is regular.
Let A be a closed subset of the compact space X and $x \in X \setminus A$. Since X is Hausdorff, each point $y \in A$ has an open neighborhood U_y such that $x \notin \overline{U_y}$. The closed subset A is itself compact, so we can extract a finite covering $\{U_{y_i}\}_{i=1}^N$ of A. Put $U = \bigcup_{i=1}^N U_{y_i} \supset A$.

Then

$$\overline{U} = \bigcup_{i=1}^{N} \overline{U_{y_i}}$$

does not contain p, so $X \setminus \overline{U}$, U are disjoint open subsets containing p and A.
Step 2: Now suppose A and B are disjoint closed subsets of X. Let $p \in B$, and apply the previous step to get disjoint open neighborhoods U_p of A and V_p of B. Because B is compact, there are $p_1, \ldots, p_n \in B$ such that $B \subset \bigcup_{i=1}^n V_{p_i}$. Let

$$U = \bigcap_{i=1}^{N} U_{p_i}, \ V = \bigcup_{i=1}^{n} V_{p_i}.$$

Then U and V are disjoint open subsets containing A and B.
c) Let X be regular Lindelöf, and let A and B be disjoint closed subsets of X. Because X is regular, for all $a \in A$ there is an open neighborhood U_a of a such that $\overline{U_a} \cap B = \varnothing$; and similarly for each $b \in B$ there is an open neighborhood V_b of b such that $A \cap \overline{V_b} = \varnothing$. Since A and B are closed in a Lindelöf space, they too are Lindelöf, so there are sequences $\{a_n\}_{n=1}^{\infty}$ in A and $\{b_n\}_{n=1}^{\infty}$ in B such that

$$A = \bigcup_n U_n, \; B = \bigcup_n V_n.$$

We now inductively construct two sequences of open sets:

$$S_1 = U_1, \; T_1 = V_1 \setminus \overline{S_1},$$

$$S_2 = U_2 \setminus \overline{T_1}, \; T_2 = V_2 \setminus (\overline{S_1} \cup \overline{S_2}),$$

$$S_3 = U_3 \setminus (\overline{T_1} \cup \overline{T_2}), \; T_3 = V_3 \setminus (\overline{S_1} \cup \overline{S_2} \cup \overline{S_3}),$$

and so forth. Put

$$S = \bigcup_{n=1}^{\infty} S_n, T = \bigcup_{n=1}^{\infty} T_n.$$

Then S and T are disjoint open subsets with $A \subset S$ and $B \subset T$.
d) **FIX ME!** $\square$

Theorem 6.51. *(Fragility of Normality)*

a) *A subspace of a normal space need not be normal.*
b) *The product of two normal spaces need not be normal.*
c) *(Noble's Theorem) Let X be a topological space such that for all cardinal numbers κ, the product X^{κ} is normal. Then X is compact.*

Proof. a,b) Our example (a very famous one) which establishes both of these facts will be the following: let ω_1 be the least uncountable ordinal, endowed with the order topology, in which a base is given by open intervals. Let $\omega_1 + 1 = \omega_1 \cup \{\omega_1\}$ be its successor ordinal. We claim ω_1 and $\omega_1 + 1$ are both normal; and indeed, that $\omega_1 + 1$ is compact. However, the product $\omega_1 \times (\omega_1 + 1)$ is not normal. Moreover, it is a subspace of the space $(\omega_1 + 1) \times (\omega_1 + 1)$, which is compact and hence normal.
c) **FIX ME!** See `https://dantopology.wordpress.com/2014/03/09/` $\square$

Theorem 6.50 gives an insight into the importance of normality: it gives a rather strong necessary condition for metrizability of a topological space. Unfortunately the same result shows that normality is not sufficient for metrizability.

Example 6.7. *Let X be a compact space containing more than one point, and let J be an uncountable set. By Tychonoff's Theorem, the product X^J is compact, hence normal by Theorem 6.50b). On the other hand, by Theorem 6.3 the space X^J is not first countable, so it cannot be metrizable.*

This suggests that we should add on some countability axiom in order to guarantee metrizability. Since metrizable spaces are necessarily first countable, it is natural to look at the class of normal, first-countable spaces. However, these need not be metrizable, even when compact. A counterexample is given by the space $[0,1] \times [0,1]$, topologized via the lexicographic ordering: $(x_1, y_1) < (x_2, y_2)$ iff $x_1 < x_2$ or $x_1 = x_2$ and $y_1 < y_2$.

It is then natural to ask whether a normal, second countable space must be metrizable. The answer to this question is one of the main goals of the following chapter.

6. An application to (dis)connectedness

THEOREM 6.52. *Let X be a compact space. Then the connected components and the quasi-components coincide: for all $x \in X$ we have $C(x) = C_Q(x)$.*

PROOF. Let $x \in X$. As above, we have $C(x) \subset C_Q(x)$. Since $C(x)$ is the maximal connected subset containing X, the equality $C(x) = C_Q(x)$ holds iff $C_Q(x)$ is connected. So suppose $C_Q(x) = Y_1 \coprod Y_2$ for disjoint closed subsets of $C_Q(x)$ with $x \in Y_1$. Since $C_Q(x)$ is closed in X and Y_1 and Y_2 are closed in $C_Q(x)$, we get that Y_1 and Y_2 are disjoint closed subsets in X. Being compact, X is thus normal, so there are disjoint open subsets $U_1 \supset Y_1$ and $U_2 \supset Y_2$. Since $C_Q(x)$ is the intersection of all clopen subsets containing x, $X \setminus C_Q(x)$ is a union of clopen subsets not containing x, hence $X \setminus (U_1 \cup U_2)$ is contained in a union of clopen subsets not containing x. By compactness, there are finitely many clopen subsets $B_1, \ldots, B_n$ not containing x such that

$$(X \setminus (U_1 \cup U_2)) \subset \bigcup_{i=1}^{n} B_i.$$

Then $F_i \coloneqq X \setminus B_i$ is a clopen subset containing x, hence

$$C_Q(X) \subset \bigcap_{i=1}^{n} F_i \subset U_1 \cup U_2.$$

Put $F \coloneqq \bigcap_{i=1}^{n} F_i$, a clopen subset. Since

$$\overline{U_1 \cap F} \subset \overline{U_1} \cap F \subset \overline{U_1} \cap (U_1 \cup U_2) \cap F = U_1 \cap F,$$

so $U_1 \cap F$ is clopen. Since $x \in U_1 \cap F$, we have $C_Q(x) \subset U_1 \cap F$ and thus $Y_2 \subset C_Q(x) \subset U_1$. It follows that $Y_2 \subset U_1 \cap U_2 = \varnothing$. Hence $C_Q(x)$ is connected. $\square$

A topological space is **zero-dimensional** if it admits a base of clopen sets.

EXERCISE 6.39. a) *Let X be a zero-dimensional topological space. Show that the following are equivalent:*
- (i) *The space X is Hausdorff.*
- (ii) *The space X is separated.*
- (iii) *The space X is Kolmogorov.*

b) *For every cardinal number κ find a connected zero-dimensional space of cardinality κ.*

THEOREM 6.53. *Let X be locally compact and totally disconnected. Then every point of x admits a neighborhood base of compact clopen neighborhoods. In particular, X is zero-dimensional.*

PROOF. Let $x \in X$, and let U be an open neighborhood of X. Since X is regular, there is an open neighborhood V of x such that $\overline{V}$ is compact and $\overline{V} \subset U$. Thus $\overline{V}$ is compact and totally disconnected, so by Theorem 6.52 the quasi-component of x in $\overline{V}$ is $\{x\}$. So for every $y \in \overline{V} \setminus V$, there is a clopen subset U_y disjoint from x. By compactness, $\overline{V} \setminus V$ has a finite covering by clopen subsets disjoint from x, and taking complements we get finitely many clopen subsets $F_1, \ldots, F_n$ such that

$$x \in \bigcap_{i=1}^{n} F_i \subset V.$$

Then $F := \bigcap_{i=1}^{n}$ is a compact clopen neighborhood of x contained in U. $\square$

COROLLARY 6.54. *For a compact metric space X, the following are equivalent:*
(i) X is totally disconnected.
(ii) X is zero-dimensional.
(iii) For all $\delta > 0$, X is a finite disjoint union of open subsets, each of diameter at most δ.

PROOF. (i) $\iff$ (ii) is a special case of Theorem 6.53.
(ii) $\implies$ (iii): Fix $\delta > 0$. For each $x \in X$, by Theorem 6.53 the open ball $B(x, \delta)$ contains a clopen subset F_x, and by compactness there are $x_1, \ldots, x_n \in X$ such that $X = \bigcup_{i=1}^{n} F_{x_i}$. Let $F_1 := F_{x_1}$, and for $2 \leq i \leq n$, let

$$F_i := F_{x_i} \setminus \bigcup_{j=1}^{i-1} F_f x_j.$$

This works.
(iii) $\implies$ (ii): Suppose that X is not totally disconnected. Then there is a connected subset $Y \subset X$ consisting of more than one point, thus of positive diameter δ. If then X is a disjoint union of finitely many open subsets $U_1, \ldots, U_n$, then for some U_i we have $U_i \cap Y = Y$ and thus the diameter of U_i is at least δ. $\square$

7. Further Exercises

EXERCISE 6.40. [**Wi**, Thm. 14.6] *Let $f : X \to Y$ be a continuous map of topological spaces. Show: if X is regular and f is open and closed, then Y is Hausdorff.*

EXERCISE 6.41. [**Wi**, Thm. 14.7] *Let X be a regular space, let $A \subset X$ be a closed subset, let $\sim$ be the equivalence relation on X in which A is an equivalence class and all singletons $x \in X \setminus A$ are equivalence classes, let $Y = X/\sim$ and let $q : X \to Y$ be the quotient map. Show that Y is Hausdorff.*

EXERCISE 6.42. *Show that a closed subspace of a quasi-normal (resp. normal) space is quasi-normal (resp. normal).*

EXERCISE 6.43. *Let $f : X \to Y$ be a closed map of topological spaces.*

a) *Show: if X is quasi-normal, so is Y.*
b) *Show: if X is normal, so is Y.*

EXERCISE 6.44. *A* ***Brown space*** *is a topological space X in which for all nonempty open subsets $U, V \subset X$ we have $\overline{U} \cap \overline{V} \neq \varnothing$.*

a) *Show: every nonempty Brown space is connected.*
b) *Show: a Brown space is quasi-regular iff it is indiscrete.*

EXERCISE 6.45. *For $a \in \mathbb{Z}^+$ and $b \in \mathbb{N}$, put*

$$U_{a,b} := \{an + b \mid n \in \mathbb{Z}^+\}.$$

a) *Show that* $\mathcal{B} = \{U_{a,b} \mid a, b \in \mathbb{Z}^+$ *with* $\gcd(a, b) = 1\}$ *forms a base for a topology* τ_G *on* $\mathbb{Z}^+$*. We call this the* ***Golomb topology*** *on* $\mathbb{Z}^+$*, following* [**Go59**].

b) *Show that the Golomb topology on* $\mathbb{Z}^+$ *is Hausdorff.*

c) *Show that* $(\mathbb{Z}^+, \tau_G)$ *is a Brown space. Deduce that it is connected and not regular.*

c) *Use the Golomb topology to show that there are infinitely many prime numbers.*
(Suggestion: Show that for any prime number p*, the set* $U_{p,0}$ *is closed. Suppose there are only finitely many prime numbers* $p_1 \ldots, p_n$*, and consider* $\bigcup_{i=1}^{n} U_{p,0}$*.)*

CHAPTER 7

Embedding, Metrization and Compactification

1. Completely Regular and Tychonoff Spaces

Two subsets A and B of a topological space X can be **separated by a continuous function** if there exists a continuous function $f : X \to [0,1]$ with $A \subset f^{-1}(0)$, $B \subset f^{-1}(1)$. This is indeed a strong separation axiom, for it follows immediately that A and B are separated by open neighborhoods, e.g. $f^{-1}([0, \frac{1}{2})$ and $f^{-1}((\frac{1}{2}, 1])$.

A space is **completely regular** if for every point x of X and every closed set A not containing x, $\{x\}$ and A can be separated by a continuous function. A separated completely regular space is called a **Tychonoff space**.[1]

Exercise 7.1. *Show that a completely regular space is quasi-regular but not necessarily regular.*

Theorem 7.1. a) *There is a regular space that is not completely regular.*
b) *(Hewitt) There is an infinite regular topological space X such that the only continuous functions $f : X \to \mathbb{R}$ are the constant functions. More precisely for every cardinal κ of uncountable cofinality, there is such a space of cardinality κ.*

Proof. a) I don't know a simple enough example to be worth our time. But see e.g. the **Deleted Tychonoff Corkscrew** [**SS**, pp. 109-11]. b) See [**He46a**]. □

Proposition 7.2. *Metric spaces are Tychonoff.*

Proof. Let (X, d) be a metric space, let $A \subset X$ be closed, and let $p \in X \setminus A$. Define $f : X \to \mathbb{R}$ by $f(x) = \min(\frac{d(x,A)}{d(p,A)}, 1)$. It works! □

Proposition 7.3. a) *Complete regularity and the Tychonoff property are hereditary (each passes from a space to all of its subspaces).*
b) *Complete regularity and the Tychonoff property are faithfully productive: if $\{X_i\}_{i \in I}$ is a family of nonempty topological spaces, then $X = \prod_{i \in I} X_i$ is completely regular (resp. Tychonoff) iff each X_i is completely regular (resp. Tychonoff).*
c) *A quotient of a Tychonoff space need not be Hausdorff, and even if it is, it need not be Tychonoff.*

Proof. Since we know that the Hausdorff property is hereditary and faithfully productive, it suffices to show parts a) and b) for complete regularity.
a) Suppose X is completely regular, let $Y \subset X$ be a subspace, let $A \subset Y$ be closed, and let $p \in Y \setminus A$. Then $A = B \cap Y$ for some closed $B \subset X$. Since p is in Y and

[1] It would be more consistent with our nomenclature to call completely regular spaces "quasi-Tychonoff". Unfortunately no one does this and the term "completely regular" is quite standard.

not in A, $p \in X \setminus B$, so there is a continuous function $f : X \to \mathbb{R}$ with $f(p) = 1$, $f(B) = \{0\}$. Then $f|_Y : Y \to [0,1]$ is a continuous function separating p from A.
b) Suppose each X_i is completely regular, let $A \subset X$ be closed and let $p \in X \setminus A$. Then there is a finite subset $J \subset I$ and for all $j \in J$ an open $U_j \subset X_j$ such that

$$p \in \prod_{j \in J} U_j \times \prod_{i \in I \setminus J} X_i \subset X \setminus A.$$

For each $j \in J$, choose $f_j : X_j \to [0,1]$ such that $f_j(p_j) = 1$ and $f_j(X_j \setminus U_j) = \{0\}$. Let $g : X \to I$ by

$$g(x) = \min_{j \in J} f_j(x_j) = \min_{j \in J} (f_j \circ \pi_j)(x).$$

The second description exhibits g as a minimum of finitely many continuous real-valued functions, hence g is continuous. Moreover we have $g(p) = 1$ and $g(x) = 0$ unless $\pi_j(x) \in U_j$ for all $j \in J$, so $g(A) = \{0\}$.

Being hereditary and productive, complete regularity is faithfully productive.
c) See [**Wi**, p. 96]. □

2. Urysohn and Tietze

Theorem 7.4. *(Tietze Extension Theorem)*
For a topological space X, the following are equivalent:
(i) X is quasi-normal.
(ii) If $A \subset X$ is closed and $f : A \to [0,1]$ is continuous, then there is a continuous map $F : X \to [0,1]$ with $F|_A = f$.
*(iii) For all disjoint closed subsets B_1, B_2 of X, there is a **Urysohn function**: a continuous function $f : X \to [0,1]$ with $B_1 \subset f^{-1}(0)$ and $B_2 \subset f^{-1}(1)$.*

Proof. (i) $\implies$ (ii): We directly follow an argument of M. Mandelkern [**Ma93**]. Let $A \subset X$ be a closed subset of a quasi-normal topological space, and let $f : A \to [0,1]$ be a continuous function. For $r \in \mathbb{Q}$, we put

$$A_r = f^{-1}([0,r]),$$

so $A_r \subset X$ is closed. For $s \in \mathbb{Q} \cap (0,1)$, we put

$$U_s = X \setminus (A \cap f^{-1}([s,1])),$$

so $U_s \subset X$ is open. Let

$$P = \{(r,s) \mid r,s \in \mathbb{Q}, 0 \leq r < s < 1\}.$$

The set P is countably infinite; let $P = \{(r_n, s_n)\}_{n=1}^{\infty}$ be an enumeration.
Let $n \in \mathbb{Z}^+$. Inductively, we suppose that for all $1 \leq k < n$ we have defined closed subsets $H_k \subset X$ such that

$$A_{r_k} \subset H_k^{\circ} \subset H_K \subset U_{s_k} \forall k < n \tag{19}$$

and

$$H_j \subset H_k^{\circ} \text{ when } j,k < n,\ r_j < r_k \text{ and } s_j < s_k. \tag{20}$$

We will define H_n. First put

$$J = \{j \mid j < n, r_j < r_n \text{ and } s_j < s_n\}$$

and

$$K = \{k \mid k < n, r_n < r_k \text{ and } s_n < s_k\}.$$

Since X is quasi-normal, there is a closed subset $H_n \subset X$ such that

$$A_{r_n} \cup \bigcup_{j \in J} H_j \subset H_n^\circ \subset H_n \subset U_{s_n} \cap \bigcap_{k \in K} H_k^\circ.$$

We write H_{rs} for H_n when $r = r_n$ and $s = s_n$. Inductively, we have defined a family $\{H_{(r,s)}\{_{(r,s)\in P}$ of closed subsets of X such that

$$\forall (r,s) \in P,\ A_r \subset H_{rs}^\circ \subset H_{rs} \subset U_s, \tag{21}$$

$$H_{rs} \subset H_{tu}^\circ \text{ when } r < t \text{ and } s < u. \tag{22}$$

For $r \in \mathbb{Q} \cap [0,1]$, put

$$X_r = \bigcap_{s>r} H_{rs}.$$

For $r < 0$, let $X_r = \varnothing$. For $r \geq 1$, let $X_r = X$. For $(r,s) \in P$, choose $t \in \mathbb{Q}$ such that $r < t < s$. Then

$$X_r \subset H_{rt} \subset H_{ts}^\circ \subset H_{ts} \subset \bigcap_{u>s} H_{su} = X_s.$$

For $r \in \mathbb{Q} \cap [0,1)$, we have

$$A_r \subset X_r \cap A = A \cap \bigcap_{s>r} H_{rs} \subset A \cap \bigcap_{s>r} U_s = A_r.$$

Thus we have constructed a family $\{X_r\}_{r\in\mathbb{Q}}$ of closed subsets of X such that

$$X_r \subset X_s^\circ \text{ when } r,s \in \mathbb{Q} \text{ and } r < s, \tag{23}$$

$$\forall r \in \mathbb{Q},\ X_r \cap A = A_r. \tag{24}$$

Finally, for $x \in X$ put $g(x) = \inf\{r \mid x \in X_r\}$. Then $g : X \to [0,1]$; since for all $x \in A$ we have $f(x) = \inf\{r \mid x \in A_r\}$, we have that $g|_A = f$. If $a < b \in \mathbb{R}$ then

$$g^{-1}((a,b)) = \bigcup \{X_s^\circ \setminus X_r : r,s \in \mathbb{Q} \text{ and } a < r < s < b\}$$

is open. Thus g is a continuous extension of f.
(ii) $\implies$ (iii): Let $B_1, B_2 \subset X$ be closed and disjoint; put $A = B_1 \cup B_2 = B_1 \coprod B_2$. The function $g : A \to [0,1]$ with $g|_{B_1} \equiv 0$ and $g|_{B_2} \equiv 1$ is locally constant, hence continuous. By assumption it extends to a continuous function $f : X \to [0,1]$.
(iii) $\implies$ (i): Let $B_1, B_2 \subset X$ be closed and disjoint. By our hypothesis, there is a continuous function $f : X \to [0,1]$ with $f(B_1) = \{0\}$, $f(B_2) = \{1\}$, let $U_1 = f^{-1}([0,\frac{1}{2}))$, $U_2 = f^{-1}(\frac{1}{2},1])$. Then $U_1, U_2 \subset X$ are disjoint and open with $U_1 \supset B_1$ and $U_2 \supset B_2$. □

Corollary 7.5. *(Urysohn's Lemma)*
Normal spaces are Tychonoff. In particular compact spaces, regular Lindelöf spaces and order spaces are Tychonoff.

Proof. Normal spaces are Hausdorff, so $\{p\}$ is closed for all $p \in X$. So according to Theorem 7.4 we can separate points from closed sets by continuous functions. □

The following variant of Theorem 7.4 is also useful.

Corollary 7.6. *Let X be quasi-normal, let $A \subset X$ be closed, and let $f : A \to \mathbb{R}$ be continuous. Then there is a continuous map $F : A \to \mathbb{R}$ such that $F|_A = f$.*

Proof. The obvious idea is the following: $\mathbb{R}$ is homeomorphic to $(0,1)$, so we may as well asume that $f(A) \subset (0,1)$. Then in particular $f(A) \subset [0,1]$, so by Tietze-Urysohn we may extend to a continuous function $F : X \to [0,1]$. However, this is not good enough, since we don't want F to take the values 0 or 1. (I.e.: we can extend $f : A \subset \mathbb{R}$ to a continuous function to the extended real line $[-\infty, \infty]$.)

We get around this as follows: first, for shallow reasons to be seen shortly, it will be better to work with the interval $(-1,1)$ instead of $\mathbb{R}$. Certainly Theorem 7.4 holds for functions with values in $[-1,1]$ in place of $[0,1]$, so let $F : X \to [-1,1]$ such that $F|_A = f$. Put

$$B = F^{-1}(0) \cup F^{-1}(1),$$

so $B \subset X$ is closed. Since F extends f and $f(A) \in (0,1)$, we have $A \cap B = \varnothing$. Let $\varphi : X \to [0,1]$ be a Urysohn function for B and A: $\varphi(B) = \{0\}$, $\varphi(A) = \{1\}$. Put

$$h : X \to [0,1],\ h(x) = F(x)\varphi(x).$$

This works: h is a continuous extension of f with values in $(-1,1)$. $\square$

Corollary 7.7. a) *A normal, connected topological space with more than one point has at least continuum cardinality.*

b) *[Urysohn] No topological space is countably infinite, connected and regular.*

Proof. a) Let X be normal and connected, and let x, y be distinct points of X. The subspace $\{x, y\}$ is discrete, so the function $f : \{x, y\} \to [0,1]$ by $f(x) = 0$, $f(y) = 1$ is continuous. By the Tietze Extension Theorem, there is a continuous function $f : X \to [0,1]$. Thus $f(X)$ is a connected subset of $[0,1]$ containing $\{0,1\}$, so $f(X) = [0,1]$. Since $\#[0,1] = \mathfrak{c}$, we're done.
b) Suppose not: let X be countable infinite, connected and regular. Like every countable space, X is Lindelöf, so by Tychonoff's Lemma (Theorem 6.50c) X is normal. Applying part a) gives a contradiction. $\square$

3. Perfect Normality

In this section we introduce a stronger separation axiom than normality and show that it is still satisfied by all metrizable spaces.

We recall that a subset of a topological space is a **G_δ-set** if it is a countable intersection of open sets.

A subset A of a topological space X is a **zero set** if there is a continuous function $f : X \to [0,1]$ with $A = f^{-1}(0)$. Thus zero sets are closed. A subset U of a topological space X is a **co-zero set** if its complement is a zero set; equivalently, if there is a continuous function $f : X \to [0,1]$ with $U = f^{-1}((0,1)]$. Thus co-zero sets are open.

Proposition 7.8. *For a topological space X, the following are equivalent:*

(i) *The space X is completely regular.*
(ii) *Every closed set is an intersection of zero sets.*
(iii) *The co-zero sets form a base for the topology.*

Proof. Taking complements shows (ii) $\iff$ (iii).
(i) $\implies$ (iii): Suppose X is completely regular, let U be an open subset of X, and let $x \in U$. By complete regularity, there is a continuous $g : X \to [0,1]$ such that $g(x) = 0$ and $g(X \setminus U) = 1$; then $f \coloneqq 1 - g : X \to [0,1]$ is continuous such that $f(x) = 1$ and $f(X \setminus U) = 0$. Then $V \coloneqq f^{-1}((0,1])$ is a co-zero set with $x \in V \subset U$.
(ii) $\implies$ (i): Let $A \subset X$ be closed and let $x \in X \setminus A$. Since A is an intersection of zero sets, there is a continuous function $g : X \to [0,1]$ such that $g(A) = 0$ and $g(x) > 0$. Then $f \coloneqq \frac{1}{g(x)} g : X \to [0,1]$ is continuous, $g(A) = 0$ and $g(x) = 1$. $\square$

Lemma 7.9. *Let X be a topological space.*

a) *A zero set in X is a closed G_δ-set.*
b) *A countable intersection of zero sets is a zero set.*
c) *If X is quasi-normal, then every closed G_δ-set is a zero set.*

Proof. a) Points are closed in the metric space $[0,1]$, so if $f : X \to [0,1]$ is continuous, then $f^{-1}(0)$ is closed. Moreover for all $n \in \mathbb{Z}^+$ the subset $f^{-1}([0, \frac{1}{n}))$ is open in X, so

$$f^{-1}(0) = f^{-1}(\bigcap_{n=1}^{\infty} [0, \frac{1}{n})) = \bigcap_{n=1}^{\infty} f^{-1}([0, \frac{1}{n}))$$

is a G_δ-set.
b) For $n \in \mathbb{Z}^+$, let $f_n : X \to [0,1]$ be continuous and let $A_n \coloneqq f_n^{-1}(0)$. Then we claim that

$$f \coloneqq \sum_{n=1}^{\infty} \frac{f_n}{2^n} : X \to [0,1]$$

is continuous. First of all, for $x \in X$, all $N \in \mathbb{Z}^+$ and all $m_1, m_2 \geq N$ we have

$$|\sum_{n=1}^{m_1} \frac{f_n(x)}{2^n} - \sum_{n=1}^{m_2} \frac{f_n(x)}{2^n}| \leq \sum_{n=N}^{\infty} \frac{f_n(x)}{2^n} \leq \sum_{n=N}^{\infty} \frac{1}{2^n} = 2^{-N},$$

so the series $\sum_{n=1}^{\infty} \frac{f_n(x)}{2^n}$ converges by the completeness of $\mathbb{R}$. Also we have

$$0 \leq \sum_{n=1}^{\infty} \frac{f_n(x)}{2^n} \leq \sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Moreover, for all $x \in X$, all $\epsilon > 0$ and all $N \in \mathbb{Z}^+$, since $f_1, \ldots, f_N$ are continuous at x there are neighborhoods $U_1(x), \ldots, U_N(x)$ such that for all $y \in U_i(x)$ we have $|f_i(y) - f_i(x)| < \frac{1}{2^{N+i+1}}$. Then $U = \bigcap_{i=1}^{N} U_i(x)$ is a neighborhood of x such that for all $y \in U$ we have

$$|f(y) - (x)| \leq \sum_{n=1}^{\infty} \frac{|f_n(y) - f_n(x)|}{2^n} < \frac{1}{2^{N+2}} + \ldots + \frac{1}{2^{2N+1}} + \frac{1}{2^{N+1}} < \frac{1}{2^N},$$

so f is continuous at x. Finally, for $x \in X$ we have $f(x) = 0$ iff $f_n(x) = 0$ for all $n \in \mathbb{Z}^+$, so

$$\bigcap_{n=1}^{\infty} A_n = f^{-1}(0)$$

is a zero set.
c) Let A be a closed G_δ-set: thus there is a sequence $\{U_n\}_{n=1}^{\infty}$ of open subsets of

X such that

$$A = \bigcap_{n=1}^{\infty} U_n.$$

Applying the Tieze Extension Theorem to the disjoint closed subsets A and $X \setminus U_n$ in the quasi-normal space X we get a continuous function $f_n : X \to [0,1]$ such that $f_n(A) = 0$, $f_n(X \setminus U_n) = 1$. Once again

$$f \coloneqq \sum_{n=1}^{\infty} \frac{f_n}{2^n} : X \to [0,1]$$

is well-defined and continuous, and we have $A = f^{-1}(0)$. □

Exercise 7.2. *Let X be a topological space. For each $n \in \mathbb{Z}^+$, let $f_n : X \to [0,1]$ be continuous, and let $a_n \in \mathbb{R}$. Show: if $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then*

$$f \coloneqq \sum_{n=1}^{\infty} a_n f_n : X \to \mathbb{R}$$

is well-defined and continuous.

Let A, B be disjoint closed subsets in a topological space X. We have seen that if X is quasi-normal, it admits a Urysohn function, i.e., a continuous function $f : X \to [0,1]$ with $A \subset f^{-1}(0)$ and $B \subset f^{-1}(1)$. It is natural to ask whether we can always find a Urysohn function with $A = f^{-1}(0)$ and $B = f^{-1}(1)$: let us call such an f a **perfect Urysohn function for A and B** and say that X is **perfectly normal** if it is Hausdorff and a perfect Urysohn function exists for all pairs of disjoint closed subsets.

Theorem 7.10. *For a topological space X, the following are equivalent:*

(i) *The space X is perfectly normal.*
(ii) *The space X is separated and every closed subset of X is a zero set.*
(iii) *The space X is normal and every closed subset is a G_δ-set.*

Proof. (i) $\implies$ (ii): If X is perfectly normal, then by definition it is Hausdorff, hence it is also separated. Let $A \subset X$ be closed. If $A = X$ then the constant function 0 shows that A is a zero set. Otherwise choose $y \in X \setminus A$. Then A and $\{y\}$ are disjoint closed subsets, so by perfect normality there is $f : X \to [0,1]$ such that $A = f^{-1}(0)$ and $\{y\} = f^{-1}(1)$ and thus A is a zero set.
(ii) $\implies$ (iii): Suppose X is separated and that every closed subset of X is a zero set. It suffices to show that X is quasi-normal, for then by Proposition 6.48d) X is normal and by Lemma 7.9b) every closed subset of X is a G_δ-set. Let A and B be disjoint closed subsets of X. Then there is a continuous function $f_1 : X \to [0,1]$ such that $A = f_1^{-1}(0)$ and a continuous function $g : X \to [0,1]$ such that $B = g^{-1}(0)$. Then $f_2 = 1 - g : X \to [0,1]$ is a continuous function such that $B = f_2^{-1}(1)$. Finally, $f \coloneqq f_1 + f_2 : X \to \mathbb{R}$ is a continuous function such that

$$\forall a \in A,\ f(a) = f_1(a) + f_2(a) = f_2(a) < 1,$$

$$\forall b \in B,\ f(b) = f_1(b) + f_2(b) = f_1(b) + 1 > 1.$$

Therefore $U_1 \coloneqq f^{-1}((-\infty,1)$ and $U_2 \coloneqq f^{-1}((1,\infty))$ are disjoint open subsets with $A \subset U_1$ and $B \subset U_2$.
(iii) $\implies$ (i): Suppose X is normal, and let A and B be disjoint closed subsets

of X. By Lemma 7.9c) there is a continuous function $g : X \to [0,1]$ such that $A = g^{-1}(0)$ and a continuous function $h : X \to [0,1]$ such that $B = h^{-1}(0)$. Then

$$f \coloneqq \frac{g}{g+h} : X \to [0,1]$$

is continuous, and we have $A = f^{-1}(0)$, $B = f^{-1}(1)$. □

COROLLARY 7.11. *Metrizable spaces are perfectly normal.*

PROOF. If X is metrizable, then it admits a compatible metric with diameter at most 1. Let Y be a closed subset of X. By Theorem 2.113 and Lemma 2.115, the function $f = d(\cdot, Y) : X \to [0,1]$ is continuous, and we have $Y = f^{-1}(0)$. □

COROLLARY 7.12. *Perfect normality is hereditary: if a topological space X is perfectly normal, then so is every subspace. In particular, perfectly normal spaces are hereditarily normal.*

PROOF. Let X be perfectly normal, and let Y be a subset of X. If Z is a closed subset of Y, then there is a closed subset $\tilde{Z}$ of X such that $Z = \tilde{Z} \cap Y$. By Theorem 7.10 there is a continuous function $f : X \to [0,1]$ such that $f^{-1}(0) = \tilde{Z}$. The restriction of f to Y is a continuous function such that $(f|_Y)^{-1}(0) = f^{-1}(0) \cap Y = \tilde{Z} \cap Y = Z$. Thus every closed subset of Y is a zero set, so Y is perfectly normal. □

4. The Tychonoff Embedding Theorem

By a **cube** we mean a topological space $[0,1]^\kappa$ for some cardinal κ.

THEOREM 7.13. *(Tychonoff Embedding Theorem)*
For a topological space X, the following are equivalent:
(i) X is homeomorphic to a subspace of a cube.
(ii) X admits a compactification, i.e., there is a compact space C and an embedding $\iota : X \hookrightarrow C$ with $\overline{\iota(X)} = C$.
(iii) X is Tychonoff.

PROOF. (i) $\implies$ (ii): Let $\iota : X \hookrightarrow [0,1]^\kappa$ be an embedding into a cube, and let $Y = \overline{\iota(X)}$. By Tychonoff's Theorem $[0,1]^\kappa$ is compact, hence so is the closed subspace Y. The map $\iota : X \hookrightarrow Y$ is an embedding of X into a compact space with dense image, i.e., a compactification of X. (ii) $\implies$ (iii): Compact spaces are normal (Theorem 6.50b), normal spaces are Tychonoff (Urysohn's Lemma: Theorem 7.5), and subspaces of Tychonoff spaces are Tychonoff (Theorem 7.3a)), so any space which is homeomorphic to a compact space is Tychonoff.
(iii) $\implies$ (i): Consider the **evaluation map**

$$e : X \to [0,1]^{C(X,[0,1])}, \ x \mapsto (f \mapsto f(x)).$$

By the universal property of the product topology, e is continuous. The map e is injective because X is completely regular and separated and thus for all $x \neq y \in X$ there is a continuous function $f : X \to [0,1]$ with $f(x) \neq f(y)$. Similarly, if $A \subset X$ is closed and $p \in X \setminus A$, then there is a continuous function $f : X \to [0,1]$ with $f|_A \equiv 0$, $f(p) = 1$ and thus $f(p) \notin \overline{f(A)}$. By Theorem 3.29c), e is an embedding. □

COROLLARY 7.14. *Locally compact spaces are Tychonoff.*

EXERCISE 7.3. *Prove Corollary 7.14.*

5. The Big Urysohn Theorem

PROPOSITION 7.15. *Let X be a Tychonoff space with a countable base $\mathcal{B}$. Then there exists a countable family $\mathcal{F}$ of continuous $[0,1]$-valued functions on X such that $\mathcal{F}$ separates points from closed subsets.*

PROOF. Consider the set $\mathcal{A}$ of all pairs (U,V) with $U,V \in \mathcal{B}$ and $\overline{U} \subset V$; evidently $\mathcal{A}$ is countable. Since X is regular and second countable, it is normal; hence for each such pair (U,V), choose a function $f : X \to [0,1]$ which is 0 on $\overline{U}$ and 1 on $X \setminus V$. This gives a countable family. Moreover let $x \in X$ and B be a closed set not containing x; we may choose an element V of $\mathcal{B}$ such that $x \in V \subset X \setminus B$ and U in $\mathcal{B}$ such that $x \in \overline{U} \subset V$. Then the continuous function f corresponding to the pair (U,V) separates x from B. □

THEOREM 7.16. *(Big Urysohn Theorem) For a separated, second countable topological space X, the following are equivalent:*

(i) *The space X can be embedded in the Hilbert cube $[0,1]^{\aleph_0} = \prod_{n=1}^{\infty}[0,1]$.*
(ii) *The space X is metrizable.*
(iii) *The space X is normal.*
(iv) *The space X is Tychonoff.*
(v) *The space X is regular.*

PROOF. (i) $\implies$ (ii): Metrizability is countably productive and hereditary.
(ii) $\implies$ (iii) $\implies$ (iv) $\implies$ (v) hold for all topological spaces.
(v) $\implies$ (i): Since X is regular and second countable, it is normal (Theorem 6.50c)) and thus Tychonoff by Urysohn's Lemma. Proposition 7.15 gives us a countable family $\{f_n : X \to [0,1]\}_{n=1}^{\infty}$ of continuous functions which separates points from closed subsets of X. By the Embedding Lemma (Theorem 3.29c), the restricted evaluation map

$$e_{\mathcal{F}} : X \to \prod_{n=1}^{\infty}[0,1],\ x \mapsto (n \mapsto f_n(x))$$

is an embedding. □

COROLLARY 7.17. *For a metrizable topological space X, the following are equivalent:*

(i) *The topology on X is induced by a totally bounded metric.*
(ii) *The space X is separable.*

PROOF. (i) $\implies$ (ii): Let d be a totally bounded metric on X. Then for all $n \in \mathbb{Z}^+$, the space (X,d) admits a finite $\frac{1}{n}$ net B_n, and then $\bigcup_{n=1}^{\infty} B_n$ is a countable dense subset.
(ii) $\implies$ (i): If X is separable and metrizable, it is separated and secound countable, so by Theorem 7.16 it can be embedded into the Hilbert cube, a compact metrizable space. Every compact metric space is totally bounded, and every subset of a totally bounded metric space is totally bounded. □

COROLLARY 7.18. *Let $f : X \to Y$ be continuous, with X compact metrizable and Y Hausdorff. Then $f(X)$ is metrizable.*

PROOF. The space $f(X)$ is compact, hence normal, so by Theorem 7.16 it suffices to show that $f(X)$ has a countable base. (In fact it is also necessary: by Theorem 2.69 separability, second countability and Lindelöfness are equivalent for metrizable spaces. So X, being compact metrizable, is separable. By Proposition 6.13 its continuous image $f(X)$ is separable. By Theorem 2.69 again, if $f(X)$ is metrizable then it has a countable base.)

Let $\mathcal{B}$ be a countable base for X, let $\mathcal{C}$ be the set of finite unions of sets from $\mathcal{B}$, and let

$$\mathcal{D} := \{Y \setminus f(X \setminus U) \mid U \in \mathcal{C}\}.$$

Since X is quasi-compact and Y is Hasudorff, f is a closed map, so $Y \setminus f(X \setminus U)$ is open, and thus $\mathcal{D}$ is a countable family of open sets in Y, so it suffices to show that $\mathcal{D}$ is a base for Y.

To see this, let $V \subset Y$ be open, and let $p \in V$. Then $f^{-1}(p)$ is a compact subset of the open subset $f^{-1}(V)$. So there are $B_1, \ldots, B_n \in \mathcal{B}$ such that

$$f^{-1}(p) \subset \bigcup_{i=1}^{n} B_i \subset f^{-1}(V).$$

Putting $C := \bigcup_{i=1}^{n} B_i$, we have that $C \in \mathcal{C}$ and, since f is surjective, we have

$$p \in Y \setminus f(X \setminus C) \subset f(C) \subset V. \qquad \square$$

Exercise 7.4. *Let X be locally compact and metrizable. Show that the following are equivalent:*

(i) *The space X is second countable.*
(ii) *The space X is σ-compact.*
(iii) *The Alexandroff extension $X^* = X \coprod \{\infty\}$ is metrizable.*

6. A Manifold Embedding Theorem

A **manifold** is a second countable Hausdorff topological space X such that for all $p \in X$, there is an open neighborhood U_p of p which is homeomorphic to $\mathbb{R}^{n(p)}$ for some positive integer $n(p)$. An **n-manifold** is a second countable Hausdorff topological space such that for all $p \in X$, there is an open neighborhood U_p of p which is homeomorphic to $\mathbb{R}^n$.

Exercise 7.5. a) *Show: a countable coproduct $\coprod_{i=1}^{\infty} M_i$ of manifolds is a manifold.*
b) *Let $d > 1$. Show: $\mathbb{R} \coprod \mathbb{R}^d$ is not a d-manifold for any $d \in \mathbb{Z}^+$.*

Let M be a manifold. Then M is locally connected and second countable, so is the coproduct of its connected components, which form a countable set. It is often the case that the study of manifolds reduces easily to the case of connected manifolds.

It is natural to suspect that a *connected* manifold must be an n-manifold for some positive integer n. And in fact it is true, but annoyingly difficult to prove. In particular, if this holds then for all $1 \leq m < n$ we must have that $\mathbb{R}^m$ is not homeomorphic to $\mathbb{R}^n$. This is easy to show when $m = 1$; for $m \geq 2$ it is most naturally approached using the methods of algebraic topology.

Exercise 7.6. a) *Let $m < n \in \mathbb{Z}^+$. Show: if $\mathbb{R}^m \cong \mathbb{R}^n$ then $S^m \cong S^n$.*
b) *(Exercise for a future course) Show that the mth homotopy of group of S^m is nontrivial and the mth homotopy group of S^n is trivial, so $S^m \not\cong S^n$.*
c) *(Exercise for a future course) Show that for a positive integer d, the dth homology group of S^m is nontrivial iff $d = m$. Deduce $S^m \not\cong S^n$.*

We can get what we want using the following result of L.E.J. Brouwer.

THEOREM 7.19. *(Invariance of Domain) Let $U \subset \mathbb{R}^n$ be open, and let $f : U \to \mathbb{R}^n$ be a continuous injection. Then f is an open map.*

EXERCISE 7.7. a) *Use Invariance of Domain to show that if $\mathbb{R}^m \cong \mathbb{R}^n$ then $m = n$.*

b) *Use Invariance of Domain to show that if a point p admits an open neighborhood $U_p \cong \mathbb{R}^m$ and an open neighborhood $V_p \cong \mathbb{R}^n$ then $m = n$. Thus there is a well-defined function* $\dim : M \to \mathbb{Z}^+$, *the* ***dimension at p****.*

c) *Show: the function* $\dim : M \to \mathbb{Z}^+$ *is locally constant.*

d) *Every connected manifold is an m-manifold for a unique $m \in \mathbb{Z}^+$.*

EXERCISE 7.8. *Let M be a manifold, and let $N \in \mathbb{Z}^+$. Suppose every connected component of M can be embedded in $\mathbb{R}^N$. Show: M can be embedded in $\mathbb{R}^N$.*

An open covering $\mathcal{U} = \{U_i\}$ of a topological space X is **locally finite** if for all $p \in X$, there is a neighborhood N_p such that $\{i \in \mathcal{U} \mid U_i \cap N_p \neq \varnothing\}$ is finite. Certainly any finite cover is locally finite.

For a function $f : X \to \mathbb{R}$, the **support of f** is

$$\operatorname{supp} f = \overline{f^{-1}(\mathbb{R} \setminus \{0\})}.$$

Thus p does not lie in the support of f iff there is a neighborhood N_p of p on which f is identically 0.

Let X be a topological space. A family of functions $\mathcal{F} = \{f : X \to [0, 1]\}$ is a **partition of unity** if:
(PU1) For all $x \in X$, there is a neighborhood U_x of x such that $\{f \in \mathcal{F} \mid \operatorname{supp} f \cap U_x \neq \varnothing\}$ is finite; and
(PU2) For all $x \in X$, $\sum_{f \in \mathcal{F}} f(x) = 1$.

Notice that because of (PU1), the sum in (PU2) amounts to a finite sum.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open covering of X. A partition of unity $\mathcal{F} = \{f_i : X \to [0, 1]\}_{i \in I}$ is **subordinate to the covering** if $\operatorname{supp} f_i \subset U_i$ for all $i \in I$.

THEOREM 7.20. *(Existence of Partitions of Unity) Let X be quasi-normal, and let $\mathcal{U} = \{U_i\}_{i=1}^n$ be a finite open cover of X. Then there is a partition of unity $\{f_i : X \to [0, 1]\}_{i=1}^n$ which is subordinate to $\mathcal{U}$.*

PROOF. Step 1: We show there are open subsets $V_1, \ldots, V_n$ of X with $X = \bigcup_{i=1}^n V_i$ and $\overline{V_i} \subset U_i$ for all $1 \leq i \leq n$. Let $A_1 = X \setminus \bigcup_{i=2}^n U_i$. Then A is closed, and since $\bigcup_{i=1}^n U_i = X$, we have $A_1 \subset U_1$. By quasi-normality, there is an open subset V_1 with $A_1 \subset V_1 \subset \overline{V_1} \subset U_1$, and thus $\{V_1, U_2, \ldots, U_n\}$ covers X. Let $2 \leq k \leq n$. Having constructed open subsets $V_1, \ldots, V_{k-1}$ such that $\overline{V_i} \subset U_i$ for all $1 \leq i \leq k-1$ and such that $\{V_1, \ldots, V_{k-1}, U_k, U_{k+1}, \ldots, U_n\}$ covers X, let

$$A_k = X \setminus (\bigcup_{i=1}^{k-1} V_i \cup \bigcup_{j=k+1}^{n} U_j).$$

Then A_k is closed in X and $A_k \subset U_k$, so by quasi-normality there is an open subset V_k with $A_k \subset V_k \subset \overline{V_k} \subset U_k$, and thus $\{V_1, \ldots, V_k, U_{k+1}, \ldots, U_n\}$ covers X and

$\overline{V_i} \subset U_i$ for all $1 \leq i \leq k$. We are done by induction: take $k = n$.
Step 2: Apply Step 1 to the finite open covering $\{U_i\}_{i=1}^n$ of X to get a finite open covering $\{V_i\}_{i=1}^n$ of X with $\overline{V_i} \subset U_i$ for all i. Then apply Step 1 again (!) to get a finite open covering $\{W_i\}_{i=1}^n$ of X with $\overline{W_i} \subset V_i$ for all i. By the Tietze Extension Theorem, for all $1 \leq i \leq n$ there is a continuous function $g_i : X \to [0,1]$ with $g_i|_{\overline{W_i}} \equiv 1$ and $g_i|_{X \setminus V_i} \equiv 0$. Thus for all $1 \leq i \leq n$ we have

$$\operatorname{supp} g_i \subset \overline{V_i} \subset U_i.$$

Define

$$g : X \to [0,1], \ g(x) = \sum_{i=1}^n g_i(x).$$

Because $X = \bigcup_{i=1}^n W_i$ we have $g(x) > 0$ for all $x \in X$. For $1 \leq i \leq n$, put

$$f_i : X \to [0,1], \ f_i(x) = \frac{g_i(x)}{g(x)}.$$

Then $\{f_i : X \to [0,1]\}_{i=1}^n$ is a partition of unity subordinate to $\{U_i\}_{i=1}^n$. $\square$

THEOREM 7.21. *(Manifold Embedding Theorem) Let M be a compact manifold. Then there is $N \in \mathbb{Z}^+$ and a continuous embedding $\iota : M \hookrightarrow \mathbb{R}^N$.*

PROOF. By compactness, M admits a finite covering $\mathcal{U}$ by open sets $U_1, \ldots, U_n$ such that each U_i is homeomorphic to $\mathbb{R}^{m(i)}$. Let $m = \max_{i=1}^n m(i)$. Then each U_i can be embedded in $\mathbb{R}^m$; choose such an embedding $\iota_i : U_i \to \mathbb{R}^m$. Since M is compact, it is normal, so by Theorem 7.20 there is a partition of unity $\{f_i : X \to [0,1]\}_{i=1}^n$ subordinate to $\mathcal{U}$. Let $A_i = \operatorname{supp} f_i$. For all $1 \leq i \leq n$, define $h_i : X \to \mathbb{R}^m$ by

$$\begin{aligned} h_i(x) &= f_i(x) \cdot \iota_i(x), \ x \in U_i \\ &= 0, \ x \in X \setminus A_i. \end{aligned}$$

This function is well-defined because the two prescriptions agree on the intersection and is continuous by the Pasting Lemma. Now consider the function

$$F : X \to \mathbb{R}^{n+mn}$$

given by

$$F(x) = (f_1(x), \ldots, f_n(x), h_1(x), \ldots, h_n(x)).$$

The characteristic property of the product topology shows that F is continuous. Suppose $F(x) = F(y)$. Since $\sum_{i=1}^n f_i(x) = 1$ we have $f_i(x) > 0$ for some i; thus $f_i(y) = f_i(x) > 0$, so $x, y \in U_i$. We have

$$f_i(x)\iota_i(x) = h_i(x) = h_i(y) = f_i(y)\iota_i(y),$$

so $\iota_i(x) = \iota_i(y)$. But $\iota_i : U_i \to \mathbb{R}^m$ is an embedding, so $x = y$. Thus F is injective. Being an injective continuous map from a compact space to a Hausdorff space, F is an embedding. $\square$

REMARK 7.22. *Theorem 7.21 can be improved in several ways (which are unfortunately beyond the scope of our ambitions).*

a) *The result continues to hold without the compactness hypothesis. See e.g.* [**Mu**, p. 315]. *The proof given there uses topological dimension theory and defines a class of "n-dimensional spaces" (including manifolds in which each connected component is an m-manifold for some $m \leq n$) that can be embedded in $\mathbb{R}^{2n+1}$.*

b) *The Whitney Embedding Theorem* [**Wh44**] *states that every* smooth *n-manifold can be smoothly embedded in $\mathbb{R}^{2n}$. This gives sharper results in small dimensions, since every manifold of dimension at most three admits a smooth structure. It follows that all surfaces can be embedded in $\mathbb{R}^4$. This can also be shown by completely classifying all topological surfaces and then finding explicit embeddings.*

c) *Only rather late in the day did Bryant-Mio* [**BM99**] *and Johnston* [**Jo99**] *each show that every topological n-manfiold can be embedded in $\mathbb{R}^{2n}$.*

7. The Stone-Cech Compactification

In this section we give a first look at the Stone-Cech compactification, by way of rings of continuous functions. Our treatment is inpired by that of [**Run**, §4.2].

7.1. $C(X,\mathbb{R})$ and $C_b(X,\mathbb{R})$. For a topological space X, let

$C(X,\mathbb{R}) \coloneqq \{\text{continuous } f : X \to \mathbb{R}\}$, $C_b(X,R) \coloneqq \{\text{bounded continuous } f : X \to \mathbb{R}\}$.

We have $C_b(X,\mathbb{R}) \subset C(X,R)$. We have equality if X is quasi-compact, since then $f(X)$ is a compact, hence closed and bounded, subset of $\mathbb{R}$. We define a topological space to be **pseudocompact** if every continuous function $f : X \to \mathbb{R}$ is bounded.

Exercise 7.9. *Let X be an infinite set, and choose $x_0 \in X$. Let τ_{X,x_0} be the set of subsets of X that are either empty or contain x_0. We call the space X equipped with τ_{X,x_0} the* ***particular point topology****. (Whenever we refer to the "topological space X" in this exercise, we will mean the topology τ_{X,x_0}.)*

a) *Show: τ_{X,x_0} is a topology on X. In fact, show that it is an Alexandroff topology: τ_{X,x_0} is closed under arbitrary intersections.*
b) *Show: X is neither quasi-compact nor separated.*
c) *Show: if Y is a Hausdorff space and $f : X \to Y$ is continuous, then f is constant. In particular $C_b(X,\mathbb{R}) = \mathbb{R}$ and X is pseudocompact.*

The sets $C(X,\mathbb{R})$ and $C_b(X,\mathbb{R})$ have a natural algebraic structure: for $f,g \in C(X,\mathbb{R})$ and $\alpha \in \mathbb{R}$, we define

$$(f+g)(x) \coloneqq f(x) + g(x),$$

$$(\alpha f)(x) \coloneqq \alpha f(x),$$

$$(f \cdot g)(x) \coloneqq f(x)g(x).$$

Exercise 7.10. *Let X be a topological space.*

a) *Show: for $f,g \in C(X,\mathbb{R})$ and $\alpha \in \mathbb{R}$, we have $f+g, \alpha f, f \cdot g \in C(X,\mathbb{R})$.*
b) *Show: for $f,g \in C_b(X,\mathbb{R})$ and $\alpha \in \mathbb{R}$, we have $f+g, \alpha f, f \cdot g \in C_b(X,\mathbb{R})$.*

It is easy to see that the operations of addition and scalar multiplication make $C(X,\mathbb{R})$ and $C_b(X,\mathbb{R})$ into vector spaces over $\mathbb{R}$ and $C_b(X,\mathbb{R})$ into a subspace of $C(X,\mathbb{R})$. Moreover, the operation of multiplication make $C(X,\mathbb{R})$ and $C_b(X,\mathbb{R})$ into commutative rings into which $\mathbb{R}$ embeds as the subring of constant functions. All this may be summed up as follows: for any topological space X we get commutative $\mathbb{R}$-algebras $C_b(X,\mathbb{R}) \subset C(X,\mathbb{R})$. This turns out to open the door for a deep interplay between general topology and commutative algebra.

REMARK 7.23. *From a topological perspective, the difference between a bounded continuous function $f : X \to \mathbb{R}$ and a continuous function $f : X \to [0,1]$ is immaterial. On the one hand, we have $C(X,[0,1]) \subset C_b(X,\mathbb{R})$. On other, for $f \in C_b(X,\mathbb{R})$ there is $M > 0$ such that $f(X) \subset [-M,M]$ and then $\frac{1}{M} f \in C(X,[0,1])$. This rescaling is generally harmless. However, $C_b(X,\mathbb{R})$ is better than $C(X,[0,1])$ algebraically: the former is closed under addition and scalar multiplication, while neither holds for the latter. In fact, $C_b(X,\mathbb{R})$ is precisely the $\mathbb{R}$-subspace of $\mathbb{R}^X$ spanned by $C(X,[0,1])$.*

If $F : X \to Y$ is a continuous map of topological spaces, there are induced maps

$$C(F) : C(Y,\mathbb{R}) \to C(X,\mathbb{R}),\ f : Y \to \mathbb{R} \mapsto f \circ F : X \to \mathbb{R},$$

$$C_b(F) : C_b(Y,\mathbb{R}) \to C_b(X,\mathbb{R}),\ f : Y \to \mathbb{R} \mapsto f \circ F : X \to \mathbb{R}.$$

The well-definedness of $C(F)$ just comes from the fact the composition of continuous functions is continuous. As for $C_b(F)$: if f is bounded, then $f(Y)$ is a bounded subset of $\mathbb{R}$, so $(f \circ F)(X) = f(F(X)) \subset f(Y)$ is also bounded.

EXERCISE 7.11. *Show: if $F : X \to Y$ is a continuous map, then the pullback maps*

$$C(F) : C(Y,\mathbb{R}) \to C(X,\mathbb{R}),\ C_b(F) : C_b(Y,\mathbb{R}) \to C_b(X,\mathbb{R})$$

are $\mathbb{R}$-algebra homomorphisms.

EXERCISE 7.12. *Show: if $F : X \to Y$ is a homeomorphism, then*

$$C(F) : C(Y,\mathbb{R}) \to C(X,\mathbb{R}), C_b(F) : C_b(Y,\mathbb{R}) \to C_b(X,\mathbb{R})$$

are $\mathbb{R}$-algebra isomorphisms.

Exercise 7.12 shows that $C(X,\mathbb{R})$ is an **algebraic invariant** of the topological space X. In other words, if the rings of continuous functions (or bounded continuous functions) are "different" (non-isomorphic), then the topological spaces are different (non-isomorphic). Whenever you attach an algebraic invariant to a topological or geometric object, you should at least entertain the hope that it might be a "complete invariant": that is, that if the invariants are the "same" (isomorphic), then the spaces are the same (isomorphic). This is not the case in full generality: evidently if $X = \{\bullet\}$ is a one-point space we have $C(X,\mathbb{R}) = C_b(X,\mathbb{R}) = \mathbb{R}$, while Exercise 7.9 exhibits an infinite topological space Y for which $C(Y,\mathbb{R}) = C_b(Y,\mathbb{R}) = \mathbb{R}$. So the ring of continuous functions does not always determine the space, but things work better on certain nicer classes of topological spaces: there will be a later spectacular result along these lines.

It is also interesting to explore how properties of a continuous function $F : X \to Y$ are reflected in the pullback maps $C(F)$ and $C_b(F)$. For instance, suppose $\iota : Y \hookrightarrow X$ is an embedding (which we may view as the inclusion of a subspace Y of X). Then $C(\iota)$ (resp. $C_b(\iota)$) is the map that restricts a continuous real-valued function (resp. a bounded continuous real-valued function) on X to a continuous real-valued function (resp. a bounded continuous real-valued function) on Y. So $C(\iota)$ is surjective iff every continuous function $f : Y \to \mathbb{R}$ extends to a continuous function $f : X \to \mathbb{R}$. Since $C(\iota)$ is a ring homomorphism, it is injective iff $C(\iota)(f) = 0 \implies f = 0$, i.e., the only function $f : X \to \mathbb{R}$ that restricts to the zero function on Y is the zero function on X. The following result restates what we already know about extending maps in these terms.

EXERCISE 7.13. *Let $\iota : Y \hookrightarrow X$ be the inclusion of a subspace Y into a topological space X.*

a) *Suppose X is quasi-normal and Y is closed in X. Show: $C(\iota)$ and $C_b(\iota)$ are surjective.*
c) *Suppose Y is dense in X. Show: $C(\iota)$ and $C_b(\iota)$ are injective.*

EXERCISE 7.14. a) *Show that for a topological space X, the following are equivalent:*
   (i) *For all $x, y \in X$ with $x \neq y$, there is $f \in C(X, \mathbb{R})$ with $f(x) \neq f(y)$.*
   (ii) *For all $x, y \in X$ with $x \neq y$ and all $\alpha, \beta \in \mathbb{R}$, there is $f \in C(X, \mathbb{R})$ with $f(x) = \alpha$, $f(y) = \beta$.*
   (iii) *For every finite subset $Y \subset X$ and any function $f : Y \to \mathbb{R}$, there is $F \in C(X, \mathbb{R})$ such that $F|_Y = f$.*

   A space satisfying these equivalent properties is ***functionally Hausdorff****.*

b) *Show: a Tychonoff space is functionally Hausdorff, and a functionally Hausdorff space is Hausdorff.*

LEMMA 7.24. *Let X be a topological space. For a point $x \in X$, put*

$$\mathfrak{m}_x := \{f \in C(X, \mathbb{R}) \mid f(x) = 0\}.$$

Then $\mathfrak{m}_x$ is a maximal ideal of $C(X, \mathbb{R})$, and the ring homomorphism

$$E_x : f \in C(X, \mathbb{R}) \mapsto f(x) \in \mathbb{R}$$

induces an isomorphism

$$C(X, \mathbb{R})/\mathfrak{m}_x \xrightarrow{\sim} \mathbb{R}.$$

PROOF. It is immediate that $E_x : C(X, \mathbb{R}) \to \mathbb{R}$ is a ring homomorphism:

$$E_x(f+g) = (f+g)(x) = f(x)+g(x),\ E_x(fg) = (fg)(x) = f(x)g(x),\ E_x(1_X) = 1_X(x) = 1.$$

Moreover E_x is surjective, since embedding $\iota : \mathbb{R} \hookrightarrow C(X, \mathbb{R})$ via the constant functions, we have

$$E_x \circ \iota = 1_{\mathbb{R}}.$$

Therefore by the fundamental isomorphism theorem for rings, we have

$$C(X, \mathbb{R})/(\operatorname{Ker} E_x) \xrightarrow{\sim} \mathbb{R}.$$

But by definition we have $\operatorname{Ker} E_x = \mathfrak{m}_x$, so certainly $\mathfrak{m}_x$ is an ideal of $C(X, \mathbb{R})$ and

$$C(X, \mathbb{R})/\mathfrak{m}_x \xrightarrow{\sim} \mathbb{R}.$$

Since an ideal I in a commutative ring R is maximal iff R/I is a field, this shows that $\mathfrak{m}_x$ is a maximal ideal of $C(X, \mathbb{R})$. □

Letting $\operatorname{MaxSpec} C(X, \mathbb{R})$ denote the set of maximal ideals of $\mathbb{R}$, we have thus defined a function

$$\mathcal{M} : X \to \operatorname{MaxSpec} C(X, \mathbb{R}),\ x \mapsto \mathfrak{m}_x.$$

We put

$$M(X) := \mathcal{M}(X),$$

so $M(X)$ is the set of maximal ideals that arise as the kernel of evaluation at x for some $x \in X$.

LEMMA 7.25. *For a topological space X, the following are equivalent:*

(i) *The space X is functionally Hausdorff.*

(ii) *The map $\mathcal{M} : X \to M(X)$ is a bijection.*

PROOF. (i) $\implies$ (ii) Suppose X is functionally Hausdorff, and let $x \neq y$ be distinct points of X. Then there is $f \in C(X, \mathbb{R})$ such that $f(x) = 0$ and $f(y) = 1$, so $f \in \mathfrak{m}_x \setminus \mathfrak{m}_y$.
(ii) $\implies$ (i): We go by contrapositive: suppose X is *not* functionally Hausdorff. Then there are $x \neq y$ in X such that $f(x) = f(y)$ for all $f \in C(X, \mathbb{R})$. In particular we have $f(x) = 0 \iff f(y) = 0$, so $\mathfrak{m}_x = \mathfrak{m}_y$. □

EXERCISE 7.15. *Let X be a topological space. For $f \in C(X, \mathbb{R})$, show that the following are equivalent:*

(i) *We have $f \in C(X, \mathbb{R})^{\times}$: that is, there is $g \in C(X, \mathbb{R})^{\times}$ such that $fg = 1$.*
(ii) *We have $0 \notin f(X)$.*

LEMMA 7.26. *If X is a quasi-compact topological space, then* $\operatorname{MaxSpec} C(X, \mathbb{R}) = M(X)$*: every maximal ideal of $C(X, \mathbb{R})$ is of the form $\mathfrak{m}_x$ for at least one $x \in X$.*

PROOF. Let $\mathfrak{m} \in \operatorname{MaxSpec} C(X, \mathbb{R})$ be a maximal ideal of $C(X, \mathbb{R})$. Seeking a contradiction, we suppose that for all $x \in X$ we have $\mathfrak{m} \neq \mathfrak{m}_x$. It follows that for no $x \in X$ is $\mathfrak{m}$ contained in $\mathfrak{m}_x$ (since a containment of maximal ideals must be an equality): that is, for all $x \in X$, there is $f_x \in \mathfrak{m}$ such that $f_x(x) \neq 0$; put

$$U_x := f_x^{-1}(\mathbb{R} \setminus \{0\}).$$

Then U_x is open containing x, so $\bigcup_{x \in X} U_x$ is an open cover of X. Since X is quasi-compact, there are $x_1, \ldots, x_n \in X$ such that $X = \bigcup_{i=1}^n U_{x_i}$. So

$$f := f_{x_1}^2 + \ldots + f_{x_n}^2 \in \mathfrak{m}$$

and for all $1 \leq i \leq in$ and $y \in U_{x_i}$ we have

$$f(y) = f_{x_1}^2(y) + \ldots + f_{x_n}^2(y) \geq f_{x_i}^2(y) > 0.$$

Since every $y \in X$ lies in U_{x_i} for some i, Exercise 7.15 shows that $f \in C(X, \mathbb{R})^{\times}$, which is a contradiction: no proper ideal of a ring can contain a unit of the ring. □

COROLLARY 7.27. *Let X be a compact space.*

a) *The map $\mathcal{M} : X \to \operatorname{MaxSpec} C(X, \mathbb{R})$ is a bijection.*
b) *Let $\varphi : C(X, \mathbb{R}) \to \mathbb{R}$ be an $\mathbb{R}$-algebra homomorphism. Then $\varphi = E_x$ for a unique $x \in X$.*

PROOF. a) Since X is compact, it is Tychonoff, hence functionally Hausdorff, so by Lemma 7.25 the map $\mathcal{M}$ is injective. Since X is compact, it is quasi-compact, so by Lemma 7.26 the map $\mathcal{M}$ is surjective. So $\mathcal{M}$ is a bijection.
b) Since φ is an $\mathbb{R}$-algebra homomorphism, the composite map $\mathbb{R} \hookrightarrow C(X, \mathbb{R}) \xrightarrow{\varphi} \mathbb{R}$ is the identity. In particular φ is surjective, so $\operatorname{Ker} \varphi$ is a maximal ideal of $C(X, R)$, and by part a) we have $\operatorname{Ker} \varphi = \mathfrak{m}_x$ for a unique $x \in X$. Now for any $f \in C(X, \mathbb{R})$ we have that $f - f(x) \in \mathfrak{m}_x = \operatorname{Ker} \varphi$, so

$$0 = \varphi(f - f(x)) = \varphi(f) - \varphi(f(x)) = \varphi(f) - f(x),$$

and thus $\varphi(f) = f(x)$. This x is unique since X is functionally Hausdorff. □

Corollary 7.27 is already a rather remarkable result: for a compact space, the set of maximal ideals of the ring $C(X, \mathbb{R})$ of continuous real-valued functions on X is naturally in bijection with X itself. We can push things further: for any topological space X we can endow $M(X)$ with a natural topology: namely, each $f \in C(X,\mathbb{R})$ naturally defines a function

$$E_f : M(X) \to \mathbb{R},\ \mathfrak{m}_x \mapsto f(x).$$

When X is not functionally Hausdorff, we must check that this definition makes since, since then there are points $x \neq y$ such that $\mathfrak{m}_x = \mathfrak{m}_y$. But we claim that if $\mathfrak{m}_x = \mathfrak{m}_y$ then we have $f(x) = f(y)$ for all $f \in C(X,\mathbb{R})$, which makes the map E_f well-defined. Indeed, if we had $f \in C(X,\mathbb{R})$ such that $f(x) \neq f(y)$, then $f - f(x) \in \mathfrak{m}_x \setminus \mathfrak{m}_y$. We endow $M(X)$ with the **initial topology** with respect to the family of maps $\{E_f : M(X) \to \mathbb{R} \mid f \in C(X,\mathbb{R})\}$: that is, the coarsest topology that makes each E_f continuous.

Lemma 7.28. *If X is functionally Hausdorff, then $M(X)$ is Hausdorff.*

Proof. Since X is functionally Hausdorff, the map $\mathcal{M} : X \to M(X),\ x \mapsto \mathfrak{m}_x$ is a bijection. Thus we may identify $M(X)$ with X as a set, but $M(X)$ is endowed with the initial topology, which is coarser than the given topology on X. (Remember that when we say one topology is coarser than another, we allow the possibility that the two topologies coincide!) Let $x \neq y$ be distinct points of X. Since X is functionally Hausdorff, there is $f \in C(X,\mathbb{R})$ such that $f(x) = 0$ and $f(y) = 1$. Put

$$U := f^{-1}((-1/2, 1/2)),\ V := f^{-1}((1/2, 3/2)).$$

Then U and V are disjoint subsets containing x and y respectively, and U and V are open (in the given topology on X, but also in) the initial topology on $M(X)$, so indeed $M(X)$ is Hausdorff. □

Theorem 7.29. *Let X be a compact space. Then the map $\mathcal{M} : X \to M(X)$ given by $x \mapsto \mathfrak{m}_x$ is a homeomorphism.*

Proof. After Lemma 7.28 we know that $M(X)$ may be viewed as a coarser Hausdorff topology, say τ_2, than the given topology, say τ_1, on X. This means that the identity map $1_X : (X,\tau_1) \to (X,\tau_2)$ is a continuous bijection. If Y is any closed subset of (X,τ_1) then Y is quasi-compact for τ_1, so $1_X(Y) = Y$ is quasi-compact for τ_2; since (X,τ_2) is Hausdorff, it follows that Y is closed in (X,τ_2), so 1_X is a continuous closed bijection and thus a homeomorphism.

All this really comes down to the following simple fact: if $\tau_2 \subset \tau_1$ are topologies on X with τ_1 quasi-compact and τ_2 Hausdorff, then $\tau_1 = \tau_2$. □

Theorem 7.30. *Let X and Y be compact spaces.*

a) *Let $\kappa : X \to Y$ be a continuous function. Then*

$$h_\kappa : C(Y,\mathbb{R}) \to C(X,\mathbb{R}),\ f \mapsto f \circ \kappa$$

is an $\mathbb{R}$-algebra homomorphism.

b) *Let $h : C(Y,\mathbb{R}) \to C(X,\mathbb{R})$ be an $\mathbb{R}$-algebra homomorphism. Then there is a unique continuous function $h : X \to Y$ such that $h = h_\kappa$.*

Proof. a) Certainly if $\kappa : X \to Y$ is continuous and $f : Y \to \mathbb{R}$ is continuous, then $f \circ \kappa : X \to \mathbb{R}$ is continuous. The rest of this part is similarly routine: if $f, g \in C(Y,\mathbb{R})$ then

$$h_\kappa(f + g) = (f + g)(\kappa) = f \circ \kappa + g \circ \kappa = h_\kappa(f) + h_\kappa(g),$$

and so forth.

b) Let $x \in X$. As above, we have an $\mathbb{R}$-algebra map $E_x : C(X,\mathbb{R}) \to \mathbb{R}$ obtained by evaluating at x. The map $E_x \circ h : C(Y, \mathbb{R}) \to \mathbb{R}$ is an $\mathbb{R}$-algebra map, which by Corollary 7.27b) is necessarily of the form $E_{y(x)}$ for a unique $y(x) \in Y$. So this defines a function $\kappa : X \to Y$. For $f \in C(Y, \mathbb{R})$ and $x \in X$, we have

$$h(f)(x) == E_x(h(f)) = (E_x \circ h)(f) = E_{\kappa(x)} \circ f = f(\kappa(x)) = (f \circ \kappa)(x),$$

so

$$h(f) = f \circ \kappa.$$

It remains to show that κ is continuous. However, for all $f \in C(Y, \mathbb{R})$ we have $f \circ \kappa = h(f) \in C(X, \mathbb{R})$ is continuous. By Theorem 7.29, the compact space X has the initial topology for the family of maps $f \in C(Y, \mathbb{R})$, so κ is continuous by Proposition 3.26. $\square$

7.2. The Stone-Cech Compactification.

Theorem 7.31 (Stone-Cech). *Let X be a Tychonoff space.*

a) *There is a compact space βX and a dense embedding $e : X \hookrightarrow \beta X$ such that for any compact space Y and continuous map $\kappa : X \to Y$, there is a unique continuous map $\hat{\kappa} : \beta X \to Y$ such that $\kappa = \hat{\kappa} \circ e$.*
b) *Every bounded continuous function $f : X \to \mathbb{R}$ admits a unique continuous extension $\hat{f} : \beta X \to \mathbb{R}$. Otherwise put: the natural map $C(\beta X, \mathbb{R}) \to C_b(X, \mathbb{R})$ is an $\mathbb{R}$-algebra isomorphism.*

Proof. Step 1: Put

$$\tilde{X} \coloneqq \prod_{f \in C_b(X,\mathbb{R})} \overline{f(X)} \subset \mathbb{R}^{C_b(X,\mathbb{R})}.$$

Each $\overline{f(X)}$ is a closed, bounded subset of $\mathbb{R}$, hence compact, so $\tilde{X}$ is compact by Tychonoff's Theorem. Moreover the natural map

$$e : X \to \tilde{X}, x \mapsto (f \in C_b(X, \mathbb{R}) \mapsto f(x))$$

is an embedding of topological spaces: in the proof of the Tychonoff Embedding Theorem, we showed that the similar map $e : X \to [0,1]^{C(X,[0,1])}$ was an embedding, and exactly the same argument shows that our map e is an embedding.

(In fact, since $C(X, [0,1]) \subset C_b(X, \mathbb{R})$, our present map e is the induced map into a larger product space than our previous map e. It follows from the Embedding Lemma (Theorem 3.29c) that if X is a separated space, I is an index set, and for each $i \in I$ we have a continuous map $f_i : X \to Y_i$, and for some subset $J \subset I$ we have that the map

$$e_J : X \to \prod_{i \in J} Y_i, \ x \mapsto (f_i(x))$$

is an embedding, then also the map

$$e_I : X \to \prod_{i \in I} Y_i, \ x \mapsto (f_i(x))$$

is an embedding.) We put

$$\beta X \coloneqq \overline{e(X)},$$

so

$$e : X \to \beta X$$

is indeed a dense embedding into a compact space.
Step 2: Let $f \in C_b(X,\mathbb{R})$, and let $\pi_f : \mathbb{R}^{C_b(X,\mathbb{R})} \to \mathbb{R}$ be the corresponding coordinate projection map. Since $\beta X \subset \tilde{X} \subset \mathbb{R}^{C_b(X,\mathbb{R})}$, its restriction to βX

$$\hat{f} = \pi_f|_{\beta X} : \beta X \to \mathbb{R}$$

is continuous, and we have

$$\hat{f} \circ e = f.$$

Thus $\hat{f} : \beta X \to \mathbb{R}$ is a continuous extension of $f : X \to \mathbb{R}$. In terms of the associated ring homomorphism

$$C_b(e) : C_b(\beta X,\mathbb{R}) \to C_b(X,\mathbb{R})$$

obatined by restricting a continuous real-valued function from βX to X, this shows that $C_b(e)$ is surjective. Since X is dense in βX, Exericse 7.13 shows that $C_b(e)$ is injective, so $C_b(e)$ is an isomorphism of rings. Since βX is compact, we have $C_b(\beta X,\mathbb{R}) = C(\beta X,\mathbb{R})$. This proves part b).
Step 3: Let Y be compact, and let $\kappa : X \to Y$ be a continuous function. Let

$$h_\kappa : C(Y,\mathbb{R}) \to C_b(X,\mathbb{R}) \text{ by } f : Y \to \mathbb{R} \mapsto f \circ \kappa : X \to \mathbb{R}.$$

(Since Y is compact and $f(\kappa(X)) \subset f(Y)$ is a bounded subset of $\mathbb{R}$.) Since $C_b(e)$ is an isomorphism, it has an inverse map; composing h_κ with the inverse gives a homomorphism

$$h \coloneqq C_b(e)^{-1} \circ h_\kappa : C(Y,\mathbb{R}) \to C(\beta X,\mathbb{R}).$$

By Theorem 7.30b), we have $h = h_{\hat{\kappa}}$ for a unique continuous function $\hat{\kappa} : \beta X \to Y$. In other words, we have

$$h_\kappa = C_b(e) \circ h_{\hat{\kappa}},$$

which means that for all continuous maps $f : Y \to \mathbb{R}$, we have

$$f \circ \kappa = f \circ \hat{\kappa} \circ e. \tag{25}$$

From this it follows that $\kappa = \hat{\kappa} \circ e$: indeed, if not, there is $x \in X$ such that

$$y_1 \coloneqq \kappa(x) \neq \hat{\kappa}(e(x)) \eqqcolon y_2.$$

Since Y is compact hence functionally Hausdorff, there is $f \in C(Y,\mathbb{R})$ such that $f(y_1) \neq f(y_2)$, and thus

$$(f \circ \kappa)(x) = f(y_1) \neq f(y_2) = (f \circ \hat{\kappa} \circ e)(x),$$

contradicting (25). Finally, $\hat{\kappa}$ is a continuous extension of a continuous map from the dense subspace X of βX to the Hausdorff space Y, so $\hat{\kappa}$ is unique. □

The Stone-Cech compactification of a reasonable space can be disconcertingly large. Here is a first result in this direction:

Corollary 7.32. *Let κ be an infinite cardinal, and let X_κ be a discrete space of cardinality κ. Then*

$$\#\beta X_\kappa = 2^{2^\kappa}.$$

Proof. By Corollary 6.21 there is a compact space Y of cardinality 2^{2^κ} and a dense subspace $Z \subset Y$ of cardinality at most κ. So there is a function $f : X_\kappa \to Y$ such that $f(X_\kappa) = Z$. Certainly f is continuous, since X_κ is discrete, so by Theorem 7.31 there is a continuous extension $\hat{f} : \beta X_\kappa \to Y$. Since $\hat{f}(\beta X_\kappa)$ is compact in the Hausdorff space Y, it is closed; since $\hat{f}(\beta X_\kappa) \supset f(X_\kappa) = Z$, also $\hat{f}(\beta X_\kappa)$ is dense in Y. Thus $\hat{f} : \beta X_\kappa \to Y$ is surjective, so $\#\beta X_\kappa \geq \#Y = 2^{2^\kappa}$. Since βX_κ

is a Hausdorff space with a dense subspace of cardinality κ, as seen in the proof of Corollary 6.21 we have $\#\beta X_\kappa \leq 2^{2^\kappa}$. $\square$

In particular, consider $\beta\mathbb{N}$, the Stone-Cech compactification of the natural numbers $\mathbb{N}$ (endowed with the topology they receive as a subspace of $\mathbb{R}$: the discrete topology). The space $\beta\mathbb{N}$ is a separable compact space of cardinality $2^{\mathfrak{c}}$. That is very large – in particular, too large to be first countable – but as the proof of Corollary 7.32 easily showed, $\beta\mathbb{N}$ has every separable compact space as a continuous image, so it must have cardinality $2^{\mathfrak{c}}$.

Corollary 7.33. *Let X be an infinite set of cardinality κ. Then the set of nonprincipal ultrafilters on X has cardinality 2^{2^κ}.*

Proof. Let βX be the Stone-Cech compactification of the discrete space X. For each $y \in \beta X$, let

$$F_y := \{N \cap X \mid N \text{ is a neighborhood of } y \text{ in } \beta X\}.$$

Then F_y is both a prefilter on βX that converges to y and a prefilter on X. Let $\mathcal{F}_y$ be any ultrafilter on X containing F_y. We claim that for $y_1 \neq y_2$ we have $\mathcal{F}_{y_1} \neq \mathcal{F}_{y_2}$. Indeed, since βX is Hausdorff, there are disjoint neighborhoods N_1 of y_1 and N_2 of y_2, and then

$$N_1 \cap X \in F_{y_1} \subset \mathcal{F}_{y_1},\ N_2 \cap X \in F_{y_2} \subset \mathcal{F}_{y_2},$$

so $N_1 \cap X$ lies in $\mathcal{F}_{y_1}$ and not in $\mathcal{F}_{y_2}$, for otherwise

$$\varnothing = (N_1 \cap X) \cap (N_2 \cap X) \in \mathcal{F}_{y_2}.$$

It follows that $y \mapsto \mathcal{F}_y$ is an injection from βX to the set of ultrafilters on X, so the latter set has cardinality at least $\#\beta X = 2^{2^\kappa}$. On the other hand every ultrafilter on X is an element of 2^{2^X}, so the number of ultrafilters on X is at most 2^{2^κ}. Finally, the principal ultrafilters on X are in bijection with X itself, so there are κ of them. Since $\kappa < 2^{2^\kappa}$, the number of nonprincipal ulrafilters must be 2^{2^κ}. $\square$

Exercise 7.16. *Let X be a separable Tychonoff space. Show: $\#(\beta X) \leq 2^{\mathfrak{c}}$.*

Let $f : X \to Y$ be a continuous map of Tychonoff spaces. Let $\beta_X : X \hookrightarrow \beta X$ and $\beta_Y : Y \hookrightarrow \beta Y$ be the respective Stone-Cech compactifications. Then

$$\beta_Y \circ f : X \to \beta Y$$

is a continuous map from the Tychonoff space X into the compact space βY, so by Theorem 7.31a) there is a unique continuous extension of $\beta_Y \circ f$ to βX. We call this continuous map $\beta f : \beta X \to \beta Y$.

Exercise 7.17. *Let $f : X \to Y$ be a continuous map of Tychonoff spaces. Suppose that $f(X)$ is dense in Y. Show: $\beta f : \beta X \to \beta Y$ is surjective.*

8. Complete Metrizability

Recall that a subset Y of a topological space X is a G_δ-set if there is a sequence $\{U_n\}_{n=1}^\infty$ of open subsets of X such that $Y = \bigcap_{n=1}^\infty U_n$.

8.1. Complete Metrizability I: Metric Spaces. Our treatment follows the very elegant presentation of [**Ox**, §11]. We need just one preliminary result, a sufficient condition for a metric space to be completely metrizable.

LEMMA 7.34 (Kuratowski). *Let (X, d) be a metric space. Suppose that there is a sequence $\{f_n : X \to \mathbb{R}\}_{n=1}^{\infty}$ of continuous functions on X with the following property: if $\{x_n\}$ is Cauchy in X and for all $m \in \mathbb{Z}^+$ the real sequence $\{f_m(x_n)\}_{n=1}^{\infty}$ is bounded, then $\{x_n\}$ converges. Then X is completely metrizable.*

PROOF. We define a function

$$\overline{d} : X \times X \to \mathbb{R},\ \overline{d}(x, y) := d(x, y) + \sum_{n=1}^{\infty} \frac{\min(1, |f_m(x) - f_m(y)|)}{2^m}.$$

It is immediate that $\overline{d}$ satisfies properties (M1) and (M2) of a metric function; that the triangle inequality (M3) holds is a straightforward calculation.

Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and let $x \in X$. Since $d \leq \overline{d}$ on $X \times X$, if $\{x_n\}$ converges to x in the $\overline{d}$-topology, it also converges to x in the d-topology. Conversely, suppose that $\{x_n\}$ converges to x in the d-topology, and fix $\epsilon > 0$. Choose $M \in \mathbb{Z}^+$ such that $2^{-M} < \frac{\epsilon}{3}$ and $0 < \delta < \frac{\epsilon}{3}$ such that

$$\forall y \in X, \forall 1 \leq m \leq M,\ d(x, y) < \delta \implies |f_m(x) - f_m(y)| < \frac{\epsilon}{3}.$$

Then if $d(x, y) < \delta$ we have

$$\overline{d}(x, y) < \frac{\epsilon}{3} + \sum_{m=1}^{M} \frac{|f_m(x) - f_m(y)|}{2^m} + 2^{-M} < \epsilon.$$

This shows that $\{x_n\}$ converges to x in the $\overline{d}$-topology. So d and $\overline{d}$ induce the same topology.

Let $\{x_n\}$ be a Cauchy sequence for the metric $\overline{d}$. Since $d \leq \overline{d}$, then sequence $\{x_n\}$ is also Cauchy for d. For $k \in \mathbb{Z}^+$ there is $N \in \mathbb{Z}^+$ such that

$$\forall m, n \geq N,\ \overline{d}(x_m, x_n) < 2^{-k}.$$

Thus for all $m, n \geq N$ we have

$$1 > 2^k \overline{d}(x_m, x_n) \geq \min(1, |f_k(x_m) - f_k(x_n)|),$$

so

$$|f_k(x_m) - f_k(x_n)| < 1.$$

It follows that for all $k \in \mathbb{Z}^+$ the sequence $\{f_k(x_n)\}$ is bounded, so by hypothesis the sequence $\{x_n\}$ converges (with respect to d but then also with respect to $\overline{d}$, since the two topologies coincide). So $(X, \overline{d})$ is complete. $\square$

THEOREM 7.35 (Alexandroff). *For a metric space (X, d) the following are equivalent:*

(i) *The space X is a G_δ-set in its completion $\tilde{X}$.*
(ii) *There is an isometric embedding $\iota : (X, d) \hookrightarrow (Y, d_Y)$ of X into a complete metric space Y such that $\iota(X)$ is a G_δ-subset of Y.*
(iii) *The topological space X is completely metrizable.*

PROOF. (i) $\implies$ (ii) is immediate: take $Y = \tilde{X}$.
(ii) $\implies$ (iii): Since X is completely metrizable iff $\iota(X)$ is, it is no loss of generality to assume that X is a G_δ-subset of the complete metric space Y: that is, suppose there is a sequence $\{U_n\}_{n=1}^\infty$ of open subsets of Y such that $X = \bigcap_{n=1}^\infty U_n$. We may assume that each U_n is a proper subset of Y. For $n \in \mathbb{Z}^+$, put $F_n := Y \setminus U_n$, a nonempty closed subset of Y. For $n \in \mathbb{Z}^+$, put

$$f_n : X \to \mathbb{R},\ f_n(x) := \frac{1}{d(x, F_n)}.$$

By Theorem 2.114 and Lemma 2.115 the function f_n is well-defined and continuous. We claim that the sequence $\{f_n : X \to \mathbb{R}\}_{n=1}^\infty$ satisfies the conditions of Lemma 7.34; thus, once we prove this claim, we get that X is completely metrizable.

To prove the claim, let $\{x_n\}$ be a Cauchy sequence in X such that $\{f_m(x_n)\}$ is bounded for all $m \in \mathbb{Z}^+$. Since Y is complete, there is $y \in Y$ such that $x_n \to y$. If for some $m \in \mathbb{Z}^+$ we had $y \in F_m$ then we would get

$$f_m(x_n) = \frac{1}{d(x_n, F_m)} \to \infty,$$

a contradiction, so $y \in \bigcap_{m=1}^\infty (Y \setminus F_m) = X$.
(iii) $\implies$ (i): Let $\overline{d}$ be a complete metric on X inducing the topology it gets as a subspace of its completion $\tilde{X}$. For $x \in X$ and $n \in \mathbb{Z}^+$, there is $\delta(x,n) \in (0, \frac{1}{n})$ such that for all $y \in X$ with $d(x,y) < \delta(x,n)$ we have $\overline{d}(x,y) < \frac{1}{n}$. For $n \in \mathbb{Z}^+$, let U_n be the union of all open balls in $\tilde{X}$ of radius $\frac{\delta(x,n)}{2}$ centered at x, as x ranges over points of X. Certainly we have $X \subset \bigcap_{n=1}^\infty U_n$, so to show that X is a G_δ in $\tilde{X}$ it suffices to show that opposite inclusion.

Let $z \in \bigcap_{n=1}^\infty U_n$. For all $n \in \mathbb{Z}^+$ there is $x_n \in X$ such that

$$d(z, x_n) < \frac{\delta(x_n, n)}{2} < \frac{1}{2n}.$$

It follows that $x_n \to z$. For all $m > n$ we have

$$d(x_m, x_n) \leq d(z, x_m) + d(z, x_n) < \frac{\delta(x_m, m)}{2} + \frac{\delta(x_n, n)}{2} \leq \max(\delta(x_m, m), \delta(x_n, n)),$$

and it follows that

$$\overline{d}(x_m, x_n) < \frac{1}{n}.$$

Thus the sequence $\{x_n\}$ is $\overline{d}$-Cauchy, so it is $\overline{d}$-convergent to an element of X. But it is also d-convergent to z and the metrics d and $\overline{d}$ induce the same topology on X, so $\{x_n\}$ is $\overline{d}$-convergent to z. It follows that $z \in X$ and thus $\bigcap_{n=1}^\infty U_n \subset X$. $\square$

The statement of Theorem 7.35 seems a bit hard to guess, because it includes the easier special case that closed subsets of a complete metric space are completely metrizable only rather indirectly: by Corollary 7.11 metric spaces are perfectly normal, hence by Theorem 7.10 in a metrizable space every closed subset is a G_δ. Theorem 7.35 also gives another route to part of Corollary 4.8: it shows once again that every G_δ-subspace in a complete metric space is a Baire space. In particular, we see again that the irrationals $\mathbb{I}$ form a Baire space. This theme will be pursued later on, when we we will put a particularly explicit complete metric on $\mathbb{I}$.

In the implication (iii) $\implies$ (i) we did not use that the ambient space $\tilde{X}$ was the completion of X: certainly any complete metric space would have served as well. A closer look shows that we did not even use the completeness of $\tilde{X}$. Later we will see that any completely metrizable subset of a Hausdorff topological space is a G_δ-subset.

8.2. Complete Metrizability II: Absolute G_δ Sets. Let X be a topological space, $A \subset X$ a dense subset, (Y, d) a metric space, and $f : A \to Y$ a continuous function. We are interested in continuously extending the domain of f. For this, the following definition is useful. For $x \in X$, we say that the **oscillation of f at x is zero** if for all $\epsilon > 0$ there is a neighborhood U of x such that $f(A \cap U)$ has diameter at most ϵ. Here are two simple observations:

• If for $x \in X$ the function f admits a continuous extension to $A \cup \{x\}$, then the oscillation of f at x is zero.
• Let $n \in \mathbb{Z}^+$. Then the locus

$$U_n := \{x \in X \mid \text{there is a neighborhood } U \text{ of } x \text{ with } \operatorname{diam}(f(A \cap U)) < \frac{1}{n}\}$$

is an open subset of X that contains A. It follows that the set

$$Z := \{x \in X \mid f \text{ has zero oscillation at x}\}$$

is a G_δ-subset of X containing A.

Theorem 7.36. *Let X be a topological space, let $A \subset X$ be a dense subset, let (Y, d) be a* complete *metric space, let $f : A \to Y$ be a continuous function, and let*

$$Z := \{x \in X \mid f \text{ has zero oscillation at } x\}$$

Then there is a unique continuous extension of f to a function $F : Z \to Y$. Thus every continuous function from a dense subset A of a topological space X into a completely metrizable space Y extends to a G_δ-subset of X that contains A.

Proof. For $x \in Z$, let $I(x)$ be the family of all open neighborhoods of x, and consider the family $\{\overline{f(A \cap U)}\}_{U \in I(x)}$ of closed subsets of Y. Since the intersection of finitely many open neighborhoods of x is also an open neighborhood of x and thus meets the dense subset A, this is a family of closed subsets of the complete metric space Y satisfying the finite intersection condition. Since $x \in Z$, for every $\epsilon > 0$, the family contains an element of diameter at most ϵ. It follows from Theorem 2.50 that $\bigcap_{U \in I(x)} \overline{f(A \cap U)}$ consists of a single point $F(x)$. Clearly if $x \in A$ then this single point is $f(x)$, so this defines a function $F : Z \to Y$ that extends f.

It remains to show that F is continuous. To see this, let $x \in Z$ and $\epsilon > 0$. Then there is an open neighborhood U of x such that $\operatorname{diam} \overline{f(A \cap U)} < \epsilon$. If $y \in Z \cap U$ then $U \in I(y)$, so $F(y) \subset \overline{f(A \cap U)}$. Since also $F(x) \in \overline{f(A \cap U)}$, we have

$$d(F(x), F(y)) \leq \operatorname{diam}(\overline{f(A \cap U)}) < \epsilon,$$

so F is continuous at x. □

Exercise 7.18. *Use Lemma 7.36 to give another proof of Theorem 2.85a).*

Theorem 7.37. *Let X be a Hausdorff space, and let A be a dense subset of X that is completely metrizable in the subspace topology. Then A is a G_δ-subset of X.*

Proof. Let $h : A \to M$ be a homeomorphism from A to a complete metric space M. By Theorem 7.36 there is a G_δ-subset $Y \supset A$ of X and a continuous extension $H : Y \to M$ of h. Consider

$$f : h^{-1} \circ H : Y \to A, \; g = 1_Y : Y \to Y.$$

Then $f|_A = g|_A = 1_A$. Since Y is Hausdorff and A is dense in Y, by Proposition 6.46 we have that $f = g$, which implies that $A = Y$ is a G_δ-set. □

8.3. Complete Metrizability III: Compactifications. So far we have established the following: a completely metrizable space is a G_δ-subset in any Hausdorff space in which it can be densely embedded. Conversely, a metric space that embeds as a G_δ in its metric completion is completely metrizable. This makes one wonder what other topological spaces can play the role of the completion here: namely, for which Hausdorff topological spaces X does it follow that if Y is metrizable and embeds as a dense G_δ in X then Y is completely metrizable? In particular any such space X must be completely metrizable if it is metrizable. Obviously the class of completely metrizable spaces satisfies this property! The other class of spaces that springs to mind are the compact spaces. And indeed this works:

Theorem 7.38. *Let X be a topological space, and let $\iota : X \hookrightarrow \tilde{X}$ be a compactification: that is, $\tilde{X}$ is compact and ι is an embedding with dense image. If X is metrizable and $\iota(X)$ is a G_δ-set in $\tilde{X}$, then X is completely metrizable.*

CHAPTER 8

Function Spaces

For topological spaces X and Y, we denote by $\mathrm{Map}(X,Y)$ the set of all functions $f: X \to Y$ and by $C(X,Y)$ the subset of all continuous functions $f: X \to Y$. A major aspect of general topology, with significant connections to analysis, lies in the study of topologies on the sets $\mathrm{Map}(X,Y)$ and $C(X,Y)$. In this study we we encounter a phenomenon that is familiar to us in the context of metric spaces but has not yet really arisen in the context of topological spaces: namely, even for fixed X and Y we do not wish to consider *just one* topology on $\mathrm{Map}(X,Y)$ or $C(X,Y)$. Rather, we will define several such topologies, each with a different intended purpose.

1. Pointwise Convergence

First of all, since $\mathrm{Map}(X,Y) = \prod_{x \in X} Y$ is a product of copies of Y indexed by X, we may endow $\mathrm{Map}(X,Y)$ with the product topology and thus $C(X,Y)$ with the subspace topology it receives from the product topology on X. A subbase for the product topology consists of all sets of the form

$$[x,U] := U \times \prod_{x' \neq x} X$$

for $x \in X$ and U an open subset of Y. In other words, $[x,U]$ is the set of all $f: X \to Y$ such that $f(x) \in U$. Moreover we have $[x,U] = \pi_x^{-1}(U)$, where $\pi_x : \mathrm{Map}(X,Y) \to Y$ is the xth coordinate projection map. This description shows that our topology on $\mathrm{Map}(X,Y)$ is the coarsest for which all the projection maps π_x are continuous (a special case of a previous observation: see Example 3.18). It is easy to see that the converse is also true, which is part a) of the following result.

Proposition 8.1. *Maintain the notation as above.*

a) *If* $\mathbf{x}$ *is a sequence in* $\mathrm{Map}(X,Y)$ *and* $p \in \mathrm{Map}(X,Y)$*, then* $\mathbf{x} \to p$ *if and only if* $\pi_x(\mathbf{x}) \to \pi_x(p)$ *for all* $x \in X$*.*
b) *If* $\mathbf{x}$ *is a net in* $\mathrm{Map}(X,Y)$ *and* $p \in \mathrm{Map}(X,Y)$*, then* $\mathbf{x} \to p$ *if and only if* $\pi_x(\mathbf{x}) \to \pi_x(p)$ *for all* $x \in X$*.*
c) *If* F *is a prefilter on* $\mathrm{Map}(X,Y)$ *and* $p \in X$*, then* $F \to p$ *if and only if* $\pi_x(F) \to \pi_x(F)$ *for all* $x \in X$*.*

Proof. Indeed part a) is a special case of part b), which is a special case of Theorem 5.23. Part c) is a special case of Theorem 5.40. $\square$

In light of Proposition 8.1, we call the product topology on $\mathrm{Map}(X,Y)$ the **topology of pointwise convergence**. Use of the definite article is warranted here: Proposition **??**b) tells us precisely which nets converge to which points, which by

Proposition 5.13 determines the closure operator and hence the topology. Alternately, we could use part c) of Proposition 8.1 together with Corollary 5.33.

A couple of remarks: first, although X is a topological space, we haven't used the topology on X yet. Second, from Theorem 6.4 we deduce that $\mathrm{Map}(X, Y)$ is first countable if and only if Y is first countable and moreover either Y is indiscrete or X is countable. In practice, these hypotheses are unlikely to be satisfied, e.g. they are not for $\mathbb{R}^{[0,1]}$. In particular, even when X and Y are, say, compact metric spaces, the topology of pointwise convergence need not be metrizable.

Example 8.1. *In general, the subspace $C(X, Y)$ of* $\mathrm{Map}(X, Y)$ *need* not *be closed in the topology of pointwise convergence. This is a fancy way of saying that a pointwise limit of continuous functions need not be continuous. Indeed, suppose e.g. that $X = [0, 1]$, $Y = \mathbb{R}$ and we take the standard example $f_n : [0, 1] \to \mathbb{R}$ by $f_n(x) = x^n$. For each $x \in [0, 1)$ we have $x^n \to 0$, while $1^n \to 0$, so if we put*

$$f : [0,1] \to \mathbb{R}, f(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases},$$

then $f_n \to f$ in our topology on $\mathrm{Map}(X, Y)$*. Since each f_n is continuous and f is not, the point f is a limit point of $C(X, Y)$ that does not lie in $C(X, Y)$, so $C(X, Y)$ is not closed in* $\mathrm{Map}(X, Y)$*.*

A major theme in the study of topologies on function spaces is to derive conditions for subsets to be quasi-compact – or slightly more generally, to have quasi-compact closure. The space $C(X, \mathbb{R})$ of continuous functions from X to $\mathbb{R}$ is an $\mathbb{R}$-vector space; if X is infinite and normal, then it follows from the Tietze Extension Theorem that this vector space is infinite-dimensional. So it is too much to expect to have a characterization as simple as the fact that a subset of $\mathbb{R}^N$ has compact closure if and only if it is bounded.

However, in the case of pointwise convergence, such a characterization is essentially an immediate consequence of Tychonoff's Theorem.

Theorem 8.2. *Let X be a set, let Y be a topological space, and give* $\mathrm{Map}(X, Y)$ *the topology of pointwise convergence. Let $\mathcal{F}$ be a subset of* $\mathrm{Map}(X, Y)$*.*

- a) *Suppose that $\mathcal{F}$ is closed and for all $x \in X$, the set $\pi_x(\mathcal{F})$ has quasi-compact closure. Then $\mathcal{F}$ is quasi-compact.*
- b) *Suppose that Y is Hausdorff. If $\mathcal{F}$ is quasi-compact, then $\mathcal{F}$ is closed and for all $x \in X$, the set $\pi_x(\mathcal{F})$ is compact.*

Proof. a) We have $\mathcal{F} \subseteq \prod_{x \in X} \overline{\pi_x(\mathcal{F})}$. By assumption each factor is quasi-compact, so the product is quasi-compact by Tychonoff's Theorem. Since $\mathcal{F}$ is closed in $\prod_{x \in X} \overline{\pi_x(\mathcal{F})}$, so $\mathcal{F}$ is quasi-compact by Proposition 3.34a).
b) If Y is Hausdorff, then so are all subspaces of $\mathrm{Map}(X, Y)$, so quasi-compactness and compactness coincide. If $\mathcal{F}$ is a compact subset of a Hausdorff space then it is closed (Proposition 3.34b)). Since π_x is continuous and $\mathcal{F}$ is quasi-compact, its image $\pi_x(\mathcal{F})$ is quasi-compact and Hausdorff, hence compact. $\square$

Let $\underline{X}$ be a subset of X. Then there is a natural restriction map

$$R : \mathrm{Map}(X, Y) \to \mathrm{Map}(\underline{X}, Y)$$

that is continuous (e.g. use the universal property of the product topology). For a subset $\mathcal{F} \subseteq \mathrm{Map}(X, Y)$, we may consider the "restricted restriction map"

$$R_{\mathcal{F}} : \mathcal{F} \to \mathrm{Map}(\underline{X}, Y).$$

This map is injective if and only if for any $f \neq g$ in $\mathcal{F}$ there is $\underline{x} \in \underline{X}$ such that $f(\underline{x}) \neq g(\underline{x})$. When this holds we say that $\underline{X}$ **distinguishes members** of $\mathcal{F}$.

We define on $\mathcal{F}$ **the topology of pointwise convergence on** $\underline{X}$: this is the topology with subbase $[\underline{x}, U] \cap \mathcal{F}$ for $\underline{x} \in \underline{X}$ and $U \in \tau_Y$. This is the coarsest topology on $\mathcal{F}$ that makes the map $R_{\mathcal{F}}$ continuous.

THEOREM 8.3. *Let X be a set, let $\underline{X}$ be a subset of X, let Y be a Hausdorff topological space, and let $\mathcal{F} \subseteq \mathrm{Map}(X, Y)$. Let τ be the topology of pointwise convergence restricted to $\mathcal{F}$ and let $\tau_{\underline{X}}$ be $\mathcal{F}$ with the topology of pointwise convergence on $\underline{X}$.*

a) *$\tau_{\underline{X}}$ is Hausdorff if and only if $\underline{X}$ distinguishes members of $\mathcal{F}$.*
b) *If $\underline{X}$ distinguishes members of $\mathcal{F}$ and τ is quasi-compact, then $\tau = \tau_{\underline{X}}$.*

PROOF. a) The topology $\tau_{\underline{X}}$ is the initial topology from the map $R_{\mathcal{F}} : \mathcal{F} \to \mathrm{Map}(\underline{X}, Y)$. If this map is injective, then this intiial topology is nothing else than the subspace topology that $\mathcal{F}$ receives from $\mathrm{Map}(\underline{X}, Y)$ (Example 3.17). Since Y is Hausdorff, so is $\mathrm{Map}(\underline{X}, Y)$ and thus so is $\tau_{\underline{X}}$. Conversely, the initial topology received from a non-injective map $f : W \to Z$ is never a Kolmogorov space: indeed the open sets of W are unions of fibers of f, so if $f(p) = f(q)$ then p and q have exactly the same open neighborhoods.
b) It is clear from the definition that $\tau_{\underline{X}} \subseteq \tau$ in all cases. If τ is quasi-compact, then so is the coarser topology $\tau_{\underline{X}}$. Because Y is Hausdorff, τ is also compact, so by part a), if $\underline{X}$ distinguishes members of $\mathcal{F}$ then $\tau_{\underline{X}} \subseteq \tau$ are compact topologies on $\mathcal{F}$. This forces the two topologies to coincide: indeed it suffices to show that every τ-closed subset is $\tau_{\underline{X}}$-closed. If a subset is τ-closed, then it is τ-quasicompact, so it is $\tau_{\underline{X}}$- quasi-compact in the Hausdorff topology $\tau_{\underline{X}}$, so it is $\tau_{\underline{X}}$-closed. □

2. Uniform Convergence

Let X be a nonempty set, and let Y be a nonempty metric space. For $f, g \in \mathrm{Map}(X, Y)$, we define

$$d(f, g) := \sup_{x \in X} d(f(x), g(x)).$$

EXERCISE 8.1. *For a set X and a metric space (Y, d), show that the following are equivalent:*

(i) *For all $f, g \in \mathrm{Map}(X, Y)$ we have $d(f, g) < \infty$.*
(ii) *Either Y is bounded or X is finite.*

Thus when X is infinite and Y is unbounded, there will be functions $f, g : X \to Y$ with $d(f, g) = \infty$, so the function d is an "emetric" in the sense of Exercise 2.56. As described there, the induced topology on $\mathrm{Map}(X, Y)$ is nevertheless metrizable in all cases: it is the same topology as the one induced by the honest metric $d_1(f, g) := \min(1, d(f, g))$. We call this the **topology of uniform convergence**. In an emetric space (X, d) it makes sense to speak of Cauchy sequences and thus of completeness, and again the notions are the same as with respect to the associated bounded metric d_1.

The following result is similar enough to part of Theorem 2.92 that we leave the proof as an exercise.

PROPOSITION 8.4. *Let X be a nonempty set, and let (Y, d) be a metric space. The following are equivalent:*

(i) *The metric space Y is complete.*
(ii) *The emetric space $\mathrm{Map}(X, Y)$ is complete.*
(iii) *The metric space $\mathrm{Map}_b(X, Y)$ is complete.*

EXERCISE 8.2. *Prove Proposition 8.4*

THEOREM 8.5 (Uniform Limit Theorem). *Let X be a topological space, let (Y, d) be a metric space, and let $f_n : X \to Y$ be a sequence of continuous functions converging uniformly to a function $f : X \to Y$. Then f is continuous.*

PROOF. To show that f is continuous, it suffices to show: for all open subsets V of Y and all $x_0 \in f^{-1}(V)$, there is a neighborhood U of x_0 such that $f(U) \subseteq V$. Put $y_0 := f(x_0)$ and choose $\epsilon > 0$ such that $B^\circ(y_0, \epsilon) \subseteq V$. Now choose $n \in \mathbb{Z}^+$ such that for all $x \in X$, we have

$$d(f_n(x), f(x)) < \frac{\epsilon}{3}.$$

Finally, choose a neighborhood U of x_0 such that $f_n(U) \subseteq B^\circ(f_n(x_0), \frac{\epsilon}{3})$. Then for all $x \in U$, we have

$$d(f(x), f(x_0)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(x_0)) + d(f_N(x_0), f(x_0)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Thus $f(U) \subseteq B^\circ(y_0, \epsilon) \subseteq V$. $\square$

COROLLARY 8.6. *Let X be a nonempty topological space, and let Y be a metric space.*

a) *The subset $C(X, Y)$ of continuous functions is closed in $\mathrm{Map}(X, Y)$ in the uniform topology.*
b) *The emetric space $C(X, Y)$ is complete if and only if the metric space Y is complete.*

EXERCISE 8.3. *Prove Corollary 8.6.*

Once again we are most interested in the question of when a subspace $\mathcal{F} \subseteq C(X, Y)$ is compact, or has compact closure. In this section we will give a classical answer to this question, namely to compactness in $C(X, \mathbb{R})$ when X is itself quasi-compact. In this case $C(X, \mathbb{R}) = C_b(X, \mathbb{R})$ is a metric space, so if $\mathcal{F} \subseteq C(X, \mathbb{R})$ to be compact, certainly $\mathcal{F}$ must be closed and bounded. If X is finite, these necessary conditions are also sufficient: Heine-Borel. When X is infinite, this may not be the case:

EXAMPLE 8.2. *Let X be a topological space that is infinite, compact and metrizable. Then X is limit point compact, so that in X there is a sequence $\{x_n\}_{n=1}^\infty$ of distinct terms converging to a point x_0 that is not equal to x_n for any $n \in \mathbb{Z}^+$. The space $Y := \{x_n \mid n \geq 0\}$ is then a closed subspace of X, hence also a compact metrizable space. Consider in Y the following sequence of functions:*

$$f_n(x_m) := \begin{cases} 0 & \text{if } 1 \leq m \leq n \\ 1 & \text{if } m = 0 \text{ or } m > n \end{cases}.$$

Then each function f_n is continuous: the only nonisolated point in the domain is x_0, and $f_n(x) = 1 = f(x_0)$ for all x sufficiently close to x_0. The sequence converges pointwise on Y to the function

$$f(x_m) = \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{if } m \geq 1. \end{cases}$$

This function is discontinuous, so the convergence cannot be uniform: also on $Y \setminus \{0\}$ we have pointwise convergence to 0 and $||f_n||_{Y\setminus\{0\}} = 1$, again showing the failure of uniform convergence. Nothing changes upon passage to a subsequence, so the subset $\mathcal{F} \coloneqq \{f_n \mid n \in \mathbb{Z}^+\}$ of $C(Y, \mathbb{R})$ has no convergent subsequence in the uniform metric. Therefore $\overline{\mathcal{F}}$ is not compact, though it is closed and bounded.

By the Tietze Extension Theorem, we can extend each f_n to a continuous function $f_n : X \to [0, 1]$ and thereby consider the extended family $\mathcal{F} \coloneqq \{f_n \mid n \in \mathbb{Z}^+\}$ inside $C(X, \mathbb{R})$. Since no subsequence converges uniformly on Y, certainly no subsequence converges uniformly on X. Again $\overline{\mathcal{F}}$ is closed, bounded but not compact.

EXERCISE 8.4. *Exhibit an infinite quasi-compact topological space X with the property that a subset $\mathcal{F}$ of $C(X, \mathbb{R})$ is compact in the uniform topology if and only if $\mathcal{F}$ is closed and bounded.*

Reflecting on Example 8.2 we may come to realize "what is wrong" with the family $\{f_n\}$. To see it more clearly, let us suppose that $Y = \{\frac{1}{n} \mid n \in \mathbb{Z}^+\} \cup \{0\}$, so $x_n = \frac{1}{n}$ and $x_0 = 0$. Then each f_n is continuous at 0, but taking $\epsilon = \frac{1}{2}$, if δ_n is such that

$$|x_m - x_0| < \delta_n \implies |f_n(\frac{1}{m}) - 1| = |f(x_m) - f(0)| < \frac{1}{2},$$

then $\delta_n \leq \frac{1}{n}$: any larger δ allows us to take $m = n$: $|x_n - x_0| = \frac{1}{n} < \delta_n$ but $|f_n(\frac{1}{n}) - 1| = 1$. Another way to say this is that for each $\epsilon > 0$ and $n \in \mathbb{Z}^+$ there is a neighborhood U_n of 0 such that $f_n(U_n) \subseteq B^\circ(0, \epsilon)$, but there is no neighborhood U of 0 such that

$$\forall n \in \mathbb{Z}^+, f_n(U) \subseteq B^\circ(0, \frac{1}{2}).$$

So although each function in the family is continuous, in this sense the continuity fails to be "uniform in the family." This motivates us to make the following key definition:

Let X be a topological space, let Y be a metric space, and let $\mathcal{F} \subseteq \operatorname{Map}(X, Y)$ be a family of functions from X to Y. We say that $\mathcal{F}$ is **equicontinuous at x in X** if for all $\epsilon > 0$, there is a neighborhood U of x in X such that:

$$\forall f \in \mathcal{F},\ \forall x' \in U,\ d(f(x), f(x')) < \epsilon.$$

We say that $\mathcal{F}$ is **equicontinuous** if it is equicontinuous at every point of X. Notice that if a family $\mathcal{F}$ is equicontinuous at x then so is every subfamily $\mathcal{F}'$. Taking he case of $\mathcal{F}' = \{f\}$, equicontinuity at x reduces to the function f being continuous at x, so if $\mathcal{F}$ is an equicontinuous family, then $\mathcal{F} \subseteq C(X, Y)$.

Again, a good way to think about equicontinuity is as "the other kind of uniform continuity," in which the uniformity is not across points of X but across members of $\mathcal{F}$. If X is also a metric space, it indeed makes sense to combine these two notions of uniformity: a family $\mathcal{F} \subseteq \mathrm{Map}(X,Y)$ is **uniformly equicontinuous** if for all $\epsilon > 0$ there is $\delta > 0$ such that for all $x, x' \in X$ and all $f \in \mathcal{F}$, we have

$$d(x,x') < \delta \implies d(f(x), f(x')) < \epsilon.$$

The following result shows a connection between equicontinuity and compactness:

LEMMA 8.7. *Let X be a quasi-compact topological space, and let (Y,d) be a metric space. Consider the following conditions on a family $\mathcal{F} \subseteq C(X,Y)$:*

(i) *The family $\mathcal{F}$ is equicontinuous.*
(ii) *The family $\mathcal{F}$ is totally bounded in the uniform metric on $C(X,Y)$.*

Then:

a) *If Y is totally bounded, then (i) $\implies$ (ii).*
b) *We always have (ii) $\implies$ (i).*

PROOF. a) Suppose that Y is totally bounded and $\mathcal{F}$ is equicontinuous, and fix $\epsilon > 0$. Using the equicontinuity of $\mathcal{F}$ and the quasi-compactness of X, we may cover X by finitely many open neighborhoods $U_1, \ldots, U_k$ of points $x_1, \ldots, x_k$ such that

$$\forall f \in \mathcal{F},\ \forall x \in U_i, d(f(x), f(x_i)) < \frac{\epsilon}{3}.$$

Since Y is totally bounded, we may cover it by finitely many open sets $V_1, \ldots, V_m$ each of diameter less than $\frac{\epsilon}{3}$.

For each $\alpha \in \mathrm{Map}(\{1,\ldots,k\},\{1,\ldots,m\})$, *if* there is $f \in \mathcal{F}$ such that $f(x_i) \in V_{\alpha(i)}$ for all $1 \le i \le k$, choose one and call it f_α. The set of such functions is indexed by a subset J of the finite set $\mathrm{Map}(\{1,\ldots,k\},\{1,\ldots,m\})$ and is thus finite. We claim that the finite set of open balls $\{B^\circ(f_\alpha,\epsilon)\}_{\alpha\in J}$ covers $\mathcal{F}$. Indeed, let $f \in \mathcal{F}$. For $1 \le i \le k$, choose $\alpha(i) \in \{1,\ldots,m\}$ such that $f(x_i) \in V_{\alpha(i)}$. Now let $x \in X$, and choose U_i containing x. then

$$\begin{aligned} d(f(x), f_\alpha(x)) &\le d(f(x), f(x_i) + d(f(x_i), f_\alpha(x_i)) + d(f_\alpha(x_i), f_\alpha(x)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

It follows that $d(f, f_\alpha) \le \epsilon$.
b) Suppose that $\mathcal{F}$ is totally bounded in the uniform metric, let $x_0 \in X$ and let $\epsilon > 0$. Because $\mathcal{F}$ is totally bounded, we may cover it by finitely many balls $\{B^\circ(f_i, \frac{\epsilon}{3})\}_{i=1}^n$. Since each f_i is continuous at x_0, there is a neighborhood U_i of x_0 such that for all $x \in U_i$ and all $1 \le i \le n$ we have

$$d(f_i(x), f_i(x_0)) < \frac{\epsilon}{3}.$$

Take $U \coloneqq \bigcap_{i=1}^n U_i$. Let $f \in \mathcal{F}$, and choose i such that $f \in B^\circ(f_i, \frac{\epsilon}{3})$. Then for all $x \in U$, we have

$$\begin{aligned} d(f(x), f(x_0)) &\le d(f(x), f_i(x)) + d(f_i(x), f_i(x_0)) + d(f_i(x_0), f(x_0)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$ □

THEOREM 8.8 (Classical Arzelà-Ascoli). *Let X be a quasi-compact topological space, and let $\mathcal{F} \subseteq C(X,\mathbb{R})$. Then $\mathcal{F}$ is compact in the uniform topology if and only if $\mathcal{F}$ is closed, bounded and equicontinuous.*

PROOF. Suppose $\mathcal{F}$ is compact in the uniform topology. Then being compact in the metric space $C(X,\mathbb{R})$, the subset $\mathcal{F}$ must be closed in $C(X,\mathbb{R})$ and bounded. Indeed, compactness forces $\mathcal{F}$ to be totally bounded, so by Lemma b) the family $\mathcal{F}$ is equicontinuous.

Conversely, suppose that $\mathcal{F}$ is closed, bounded and equicontinuous. Being bounded in the uniform metric means there is $M > 0$ such that $|f(x)| \leq M$ for all $f \in \mathcal{F}$ and all $x \in X$. Thus we have $\mathcal{F} \subseteq C(X,[-M,M])$. Since $[-M,M]$ is a totally bounded metric space, by Lemma 2a) the family $\mathcal{F}$ is totally bounded in the uniform metric. Moreover $C(X,\mathbb{R}) = C_b(X,\mathbb{R})$ is complete, so $\mathcal{F}$, being closed in $C(X,\mathbb{R})$, is also complete. Therefore $\mathcal{F}$ is compact by Theorem 2.75. $\square$

EXERCISE 8.5. *Show that Theorem 8.8 continues to hold with $\mathbb{R}$ replaced by any metric space in which all closed, bounded subsets are compact.*

3. Quasi-compact Convergence

Let X be a topological space, and let (Y,d) be a metric space. For a subset C of X and $f,g \in \operatorname{Map}(X,Y)$, we put

$$d_C(f,g) := \sup_{x\in C} d(f(x),g(x)) \in [0,\infty].$$

If moreover $\epsilon > 0$, we put

$$U(C,f,\epsilon) := \{g \in \operatorname{Map}(X,Y) \mid \sup_{x\in C} d_C(f,g) < \epsilon\}.$$

If $g \in U(C,f,\epsilon)$, then

$$U(C,g,\epsilon - d_C(f,g)) \subseteq U(C,f,\epsilon).$$

Now suppose

$$g \in U(C_1,f_1,\epsilon_1) \cap U(C_2,f_2,\epsilon_2).$$

Then

$$\begin{gathered} g \in U(C_1 \cup C_2, g, \min(\epsilon_1 - d_{C_1}(f_1,g), \epsilon_2 - d_{C_2}(f_2,g_2)) \\ U(C_1,g,\epsilon_1 - d_{C_1}(f_1,g)) \cap U(C_2,g,\epsilon_2 - d_{C_2}(f_2,g_2)) \\ \subseteq U(C_1,f_1,\epsilon_1) \cap U(C_2,f_2,\epsilon_2). \end{gathered}$$

Thus if $\mathcal{C}$ is a family of subsets of X that is closed under finite unions, then the set

$$U(C,f,\epsilon) \mid C \in \mathcal{C},\ f \in \operatorname{Map}(X,Y),\ \epsilon > 0\}$$

form the base for a topology on $\operatorname{Map}(X,Y)$. In this topology a neighborhood base of $f \in \operatorname{Map}(X,Y)$ is the set $\{U(C,f,\epsilon) \mid C \in \mathcal{C},\ \epsilon > 0\}$, so a net $\mathbf{f}$ converges to f in this topology if and only if $\mathbf{f}|_C$ is uniformly convergent to $f|_C$ for all $C \in \mathcal{C}$.

So:
• When $\mathcal{C}$ is the set of finite subsets of X, we recover the topology of pointwise convergence.
• When $\mathcal{C} = \{X\}$, we recover the topology of uniform convergence.
• When $\mathcal{C}$ is the set of all quasi-compact subsets of X, we get the **topology of uniform convergence on quasi-compact subsets** or, for short, the **topology of compact convergence**.

EXERCISE 8.6. *Let X be a topological space and let (Y,d) be a metric space. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two families of subsets of X, each closed under finite intersection. In this exercise, all topologies are on the set* $\operatorname{Map}(X,Y)$.

a) *Suppose that for each $C_1 \in \mathcal{C}_1$ there is $C_2 \in \mathcal{C}_2$ with $C_1 \subseteq C_2$. Show that the topology of uniform convergence on elements of $\mathcal{C}_1$ is coarser than the topology of uniform convergence on elements of $\mathcal{C}_2$.*
b) *Deduce: if $\mathcal{C}_1 \subseteq \mathcal{C}_2$, then the topology of uniform convergence on elements of $\mathcal{C}_1$ is coarser than the topology of uniform convergence on subsets of $\mathcal{C}_2$.*
c) *Deduce: the topology of pointwise convergence is coarser than the topology of compact convergence, which is coarser than the topology of uniform convergence.*

Example 8.3. *Let $\sum_{n=0}^{\infty} a_n z^n$ be a complex power series (you can make it a real power series if you prefer; the example will hold verbatim). Suppose the radius of convergence is $R \in (0, +\infty)$. Using the Root Test and the Weierstrass M-Test, one shows in complex analysis that for all $0 < r < R$, the convergence is uniform on the closed ball $B^\bullet(0, r)$, whereas the convergence need not be uniform on the open ball $B^\circ(0, R)$. For instance, for the power series $\sum_{n=0}^{\infty} z^n$ the radius of convergence is 1, and if the convergence were uniform on $B^\circ(0, 1)$ then the limit function*

$$f(x) = \sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$$

would be a uniform limit of bounded functions, hence bounded on $B^\circ(0, 1)$, which is clearly not the case.

First we observe that the above convergence property implies that the power series $\sum_n a_n z^n$ converges uniformly on compact subsets of $B^\circ(0, R)$. Indeed, if $C \subseteq B^\circ(0, R)$ is a compact subset, then the modulus function $z \mapsto |z|$ assumes a maximum $r < R$, so $C \subseteq B^\bullet(0, r)$.

Next we recall that this weaker version of uniform convergence is still enough to ensure that the corresponding function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is continuous on $B^\circ(0, R)$. Indeed, for any $z \in B^\circ(0, R)$, the point z has a compact neighborhood C contained in $B^\circ(0, R)$, and since f_C is a uniform limit of polynomial functions on C, $f|_C$ is continuous, so f is continuous at z.

For any space X that is weakly locally quasi-compact, the argument of the preceding example (using nets instead of sequences) shows that $C(X, Y)$ is closed in $\operatorname{Map}(X, Y)$ in the topology of compact convergence. However, by proceding more carefully we can derive the same conclusion under a weaker hypothesis on X.

For this we need a new definition that we shall pause to motivate a little. A subset of a topological space X is open if and only if its intersection with every open subset U of X is open in U. Similarly, a subset of X is closed if and only if its intersection with every closed subset A of X is closed in A. Indeed an open (resp. closed) subset has open (resp. closed) intersection with any subset, but the other way around is more interesting. In $\mathbb{R}^N$ it is easy to see that a subset A whose intersection with every closed ball is closed must itself be closed: otherwise, take a limit point p of A that lies outside A; then for any $R > 0$, the point p is a limit point of $A \cap B^\bullet(p, R)$ that does not lie in $A \cap B^\bullet(p, R)$.

A topological space X is **quasi-compactly generated** if a subset A of X such that $A \cap C$ is closed in C for every quasi-compact subset C of X is closed. Taking complements, we get that a space is compactly generated if and only if a subset tha thas open intersection with each quasi-compact subset is itself open.

PROPOSITION 8.9. *Let X be a topological space. If X is either weakly locally compact or is a sequential space, then X is quasi-compactly generated.*

PROOF. First suppose that X is weakly locally compact. In this case we can argue essentially as we did above in the case of $\mathbb{R}^N$: if A is a subset of X that is not closed, let $p \in X \setminus A$ be a limit point of X, and let C be a quasi-compact neighborhood of p. Then for every neighborhood N of p, also $C \cap N$ is a neighborhood of p in C, so $C \cap N$ contains a point of A, so p is a limit point of C that does not lie in $A \cap C$, so $A \cap C$ is not closed in C.

Next suppose that X is a sequential space. Let A be a subset of X that is not closed. Because X is sequential, A is then not sequentially closed: there is a sequence $x_\bullet$ in A converging to a point $x \in X \setminus A$. The set

$$C \coloneqq \{x_n \mid n \in \mathbb{Z}^+\} \cup \{x\}$$

is quasi-compact (any open neighborhood of x contains all but finitely many other elements of C), and x lies in the closure of $A \cap C$ in C but not in $A \cap C$ itself, so $A \cap C$ is not closed in C. $\square$

EXERCISE 8.7. *Let X be an uncountable set endowed with the cocountable topology: i.e., the open subsets U of X are those such that $X \setminus U$ is countable. This space is separated but not Hausdorff. Show: X is not quasi-compactly generated.*

EXERCISE 8.8. *Show that the Arens-Fort space (cf. Example 3.9) is a Hausdorff space that is not quasi-compactly generated.*

The following result is not surprising, but it is important.

PROPOSITION 8.10. *Let X be a quasi-compactly generated space, and let $f : X \to Y$ be a function. The following are equivalent:*

(i) *The function f is continuous.*
(ii) *For all quasi-compact subsets C of X, the restricted function $f|_C : C \to Y$ is continuous.*

PROOF. (i) $\implies$ (ii): For this direction no hypothesis on X is needed: if f is continuous, then its restriction $f|_C : C \to Y$ to *any* subspace C is continuous.
(ii) $\implies$ (i): Let V be an open subset of Y. Then for all quasi-compact subsets C of X we have that $(f|_C)^{-1}(V) = f^{-1}(V) \cap C$ is open in C. Since X is quasi-compactly generated, this means that $f^{-1}(V)$ is open in X. So f is continuous. $\square$

THEOREM 8.11. *Let X b a quasi-compactly generated space, and let (Y, d) be a metric space. Then $C(X, Y)$ is closed in* $\mathrm{Map}(X, Y)$ *in the topology of quasi-compact convergence.*

PROOF. Let $f \in \mathrm{Map}(X, Y)$ be a limit point of $C(X, Y)$. Since X is quasi-compactly generated, it suffices to show that $f|_C : C \to Y$ is continuous for all quasi-compact subsets C of X, for then f is continuous by Proposition 8.10. For $n \in \mathbb{Z}^+$, choose $f_n \in U(C, f, \frac{1}{n}) \cap C(X, Y)$. Then (f_n) converges uniformly to f on C, so by the Uniform Limit Theorem, $f_n|_C$ is continuous. $\square$

4. The Compact-open Topology

4.1. The $\mathcal{F}$-open topology. We now wish to move on to topologies on $\mathrm{Map}(X, Y)$ defined in terms of a generalization of the $[x, U]$ construction. Namely, for nonempty sets X and Y, let $A \subseteq X$ and $B \subseteq Y$ and put

$$[A, B] \coloneqq \{f : X \to Y \mid f(A) \subseteq B\}.$$

The following exercise collects some very basic facts about these sets of functions.

Exercise 8.9. *Let X and Y be nonempty sets.*

a) *Let $A_1 \subseteq A_2 \subseteq X$ and $B \subseteq Y$. Show: $[A_2, Y] \subseteq [A_1, Y]$.*
b) *Let $A \subseteq X$ and $B_1 \subseteq B_2 \subseteq Y$. Show: $[A, B_1] \subseteq [A, B_2]$.*
c) *Let $\{A_i\}_{i\in I}$ be a family of subsets of X and let $B \subseteq Y$. Show:*

$$[\bigcup_{i\in I} A_i, B] = \bigcap_{i\in I} [A_i, B].$$

d) *Let $A \subseteq X$ and let $\{B_i\}_{i\in I}$ be a family of subsets of Y. Show:*

$$[A, \bigcup_{i\in I} B_i] = \bigcup_{i\in I} [A, B_i].$$

Now let $\mathcal{F} \subseteq 2^X$ be a family of subsets of X. We define the **$\mathcal{F}$-open topology** on $\operatorname{Map}(X, Y)$ to be the topology with subbase $[A, U]$ for $A \in \mathcal{F}$ and $U \in \tau_Y$ (i.e., U is an open subset of Y). Because we are going to take finite intersections of the sets $[A, U]$ to form a base for the $\mathcal{F}$, it follows from Execise 8.9c) that if $\mathcal{F}'$ is the set of all finite unions of elements of $\mathcal{F}$, then the $\mathcal{F}$-topology and the $\mathcal{F}'$-topology on $\operatorname{Map}(X, Y)$ coincide. So we may as well assume that $\mathcal{F}$ is closed under finite intersections. Moreover, it is reasonable to only consider topologies on $\operatorname{Map}(X, Y)$ making the projection maps $\pi_x : \operatorname{Map}(X, Y) - > X$ continuous, which as above means that we want the sets $[x, U]$ to be open for all $x \in X$ and $U \in \tau_Y$. So as to ensure this, we will require that $\mathcal{F}$ contain all the singleton sets $\{x\}$, and thus that $\mathcal{F}$ contain all finite sets. Thus of all topologies we are now considering, the topology of pointwise convergence – a.k.a. the **finite-open topology** – is the coarsest.

Exercise 8.10. *Let $X = Y = \mathbb{R}$, and let $\mathcal{F}$ be the family of all finite subsets of $\mathbb{R}$ with at least 2 elements.*

a) *Show that in the $\mathcal{F}$-topology on $\mathbb{R}^R$, the subset $[0, (-1, 1)]$ is not open.*

b) *Show that after restricting to the set $C(\mathbb{R}, R)$ of continuous functions, all subsets $[x, U]$ are open in the $\mathcal{F}$-topology.*

4.2. The quasi-compact-open topology. Let $\mathcal{C}$ be the family of quasi-compact subsets of X. The **quasi-compact-open topology** on $\operatorname{Map}(X, Y)$ is the $\mathcal{C}$-open topology, i.e., the topology with subbasis the sets $[C, U]$ for C a quasi-compact subset of X and Y an open subset of Y.

Theorem 8.12. *Let X be a topological space, and let (Y, d) be a metric space. On the subset $C(X, Y)$ of continuous functions from X to Y, the compact-open topology and the topology of quasi-compact convergence coincide.*

Proof. Let τ_1 be the compact-open topology on $C(X, Y)$ and let τ_2 be the topology of quasi-compact convergence on $C(X, Y)$.

Step 1: For a subset A of Y and $\epsilon > 0$, let

$$N_\epsilon(A) \coloneqq \{y \in Y \mid d(A, y) < \epsilon\}$$

be the ϵ-neighborhood of A. If A is compact and $V \supseteq A$ is open, then by Theorem 2.113, for all $a \in A$ the sets $\{a\}$ and $X \setminus V$ are distance-separated; since the function $a \mapsto d(a, X \setminus V)$ is a continuous function on the compact set A, we have

$$\min_{a \in A} d(a, X \setminus V) > 0$$

and

$$U(A, \min_{a \in A} d(a, X \setminus V)) \subseteq V.$$

Step 2: We show that $\tau_1 \subseteq \tau_2$. To see this, let $C \subseteq X$ be quasi-compact, let $V \subseteq Y$ be open, and let $f \in [C, V] \cap C(X, Y)$. Then $f(C)$ is a compact subset of the open set V, so by Step 1, there is $\epsilon > 0$ such that $N_\epsilon(f(C)) \subseteq V$, and thus

$$U(C, f, \epsilon) \subseteq [C, V],$$

so

$$U(C, f, \epsilon) \cap C(X, Y) \subseteq [C, V] \cap C(X, Y).$$

Step 3: We show that $\tau_2 \subseteq \tau_1$. If $\mathcal{U}$ is is a τ_2-open set and $f \in \mathcal{U}$, then there is a quasi-compact subset $C \subseteq X$ and $\epsilon > 0$ such that $U(C, f, \epsilon) \subseteq \mathcal{U}$, so it will suffice to find an element $\mathcal{V}$ of τ_1 such that $f \in \mathcal{V} \subseteq U(C, f, \epsilon)$.

For $x \in X$, choose an open neighborhood N_x of x such that

$$f(N_x) \subseteq B^\circ(f(x), \frac{\epsilon}{4}).$$

Then

$$f(\overline{N_x}) \subseteq \overline{B^\circ(f(x), \frac{\epsilon}{4})} \subseteq B^\bullet(f(x), \frac{\epsilon}{4}) \subseteq B^\circ(f(x), \frac{\epsilon}{3}) =: U_x$$

and U_x has diameter less than ϵ. Since C is quasi-compact, we can cover it by finitely many neighborhoods $N_{x_1}, \ldots, N_{x_n}$. For $1 \leq i \leq n$, put

$$C_{x_i} := \overline{N_{x_i}} \cap C.$$

Each C_{x_i} is a closed in the quasi-compact space C, so is quasi-compact. We have

$$f \in \bigcap_{i=1}^{n} [C_{x_i}, U_{x_i}] \subseteq U(C, f, \epsilon).$$

Thus we may take $\mathcal{V} := \bigcap_{i=1}^{n} [C_{x_i}, U_{x_i}]$. □

5. Joint Continuity

Let X and Y be a topological spaces, and let $\mathcal{F} \subseteq \mathrm{Map}(X, Y)$ be a family of maps from X to Y. There is a natural evaluation map

$$e : \mathcal{F} \times X \to Y, \ e : (f, x) \mapsto f(x).$$

If $\mathcal{F}$ is endowed with a topology τ, then we can give $\mathcal{F} \times X$ the product topology. Then we say that $(\mathcal{F}, \tau)$ is **jointly continuous** if the map e is continuous.

LEMMA 8.13. *Let $\mathcal{F} \subseteq \mathrm{Map}(X, Y)$ and let τ be any topology on $\mathcal{F}$. If $(\mathcal{F}, \tau)$ is jointly continuous, then $\mathcal{F} \subseteq C(X, Y)$. So however we topologize a family of maps, it can only be jointly continuous if each member of the family is continuous.*

PROOF. Suppose that $e : \mathcal{F} \times X \to Y$ is continuous. Let $f \in \mathcal{F}$. Then the restriction of e to $\{f\} \times X$ is continuous, but this map $\{f\} \times X \to Y$ maps $x \in X$ to $f(x)$: in other words, it is none other than the map f. □

EXAMPLE 8.4. *We claim that the family* $\mathcal{C}([0,1],\mathbb{R})$ *of all continuous functions* $f:[0,1]\to\mathbb{R}$ *is* not *jointly continuous when given the topology of pointwise convergence. More precisely, let* f *be the zero function. We will show that* e *is not continuous at the point* $(f,0)$. *If it were, there would be an open neighborhood* $\mathcal{U}$ *of* X *in the product topology* $\mathbb{R}^{[0,1]}$ *and* $\delta>0$ *such that for every continuous function* $g\in\mathcal{U}$, $g([0,\delta))\subset(-1,1)$. *However, membership in* $\mathcal{U}$ *only imposes conditions on* g *at finitely many points of* $[0,1]$, *and clearly there are continuous functions* $g:[0,\delta)\to\mathbb{R}$ *with presecribed values at any finite number of points that do not have range contained in* $(-1,1)$.

EXERCISE 8.11. *Show that the map* $e:\mathcal{C}([0,1],\mathbb{R})\times[0,1]\to\mathbb{R}$ *of Example 8.4 is not continuous at* any *point of its domain.*

EXERCISE 8.12. *Let* X *and* Y *be topological spaces, let* $\mathcal{F}\subseteq C(X,Y)$, *and give* $\mathcal{F}$ *the discrete topology. Show:* $\mathcal{F}$ *is jointly continuous.*

Let X and Y be topological spaces, let $\mathcal{F}\subseteq C(X,Y)$ be a family of continuous functions, let τ be a topology on $\mathcal{F}$, and let $\underline{X}$ be a subset of X. We say that $\mathcal{F}$ is **jointly continuous on** $\underline{X}$ if the restricted evaluation map

$$e|_{\underline{X}}:\mathcal{F}\times\underline{X}\to Y,\ e:(f,\underline{x})\mapsto f(\underline{x})$$

is continuous. Since $e|_{\underline{X}}$ is indeed the restriction of e to $\mathcal{F}\times\underline{X}$, family that is jointly continuous is jointly continuous on every subset $\underline{X}$ of X.

THEOREM 8.14. *Let* X *and* Y *be topological spaces, let* $\mathcal{F}\subseteq C(X,Y)$ *be a family of continuous functions, and let* τ *be a topology on* $\mathcal{F}$. *If* $(\mathcal{F},\tau)$ *is jointly continuous on every quasi-compact subset of* X, *then* τ *is finer than the quasi-compact-open topology on* $\mathcal{F}$.

PROOF. Let C be a quasi-compact subset of X and let U be an open subset of Y. We must show that $[C,U]\in\tau$. Because $\mathcal{F}$ is jointly continuous on C, the restricted evaluation map $e_C:\mathcal{F}\times C\to Y$ is continuous, so $V\coloneqq e_C^{-1}(U)$ is open in $\mathcal{F}\times C$. Notice that for a subset $S\subseteq\mathcal{F}$, we have

$$S\subseteq[C,U]\iff S\times C\subseteq V.$$

Now let $f\in[C,U]$, so by our observation we have $\{f\}\times C\subseteq V$. By the Tube Lemma (Theorem 3.38), there is an open neighborhood $\mathcal{N}$ of f in $\mathcal{F}$ such that $\mathcal{N}\times C\subseteq V$, so $\mathcal{N}\subseteq[C,U]$. It follows that $[C,U]$ is open. $\square$

Theorem 8.14 implies in particular that any topology on $\mathcal{F}\subseteq C(X,Y)$ that makes $\mathcal{F}$ jointly continuous must contain the quasi-compact-open topology. The next result gives an important case in which the quasi-compact-open topology is conversely fine enough to make any such family jointly continuous.

THEOREM 8.15. *Let* X *be a locally quasi-compact topological space and let* Y *be a topological space. Let* $\mathcal{F}\subset C(X,Y)$ *be a family of continuous functions, given the quasi-compact-open topology. Then* $\mathcal{F}$ *is jointly continuous.*

PROOF. Let $f\in\mathcal{F}$, let $x\in X$ and let U be an open neighborhood of $f(x)$ in Y. Let $e:\mathcal{F}\times X\to Y$ be the evaluation map. We must find an open neighborhood $\mathcal{N}$ of (f,x) in $\mathcal{F}\times X$ such that $e(\mathcal{N})\subset U$. Since $f^{-1}(U)$ is a neighborhood of x and X is locally quasi-compact, there is a compact neighborhood C of x that is contained in $f^{-1}(U)$. We may therefore take

$$\mathcal{N}\coloneqq[C,U]\times C^{\circ}.\qquad\square$$

Let X, Y, Z be sets. Then there is a canonical bijection (sometimes called **currying**)

$$c : \mathrm{Map}(X \times Z, Y) \to \mathrm{Map}(X, \mathrm{Map}(Z, Y))$$

in which we send $f : X \times Z \to Y$ to the function $F : x \mapsto f(x, \cdot)$. If X, Y, Z are moreover topological spaces and $f : X \times Z \to Y$ is continuous, then clearly so is $f(x, \cdot)$ for all $x \in X$. That is, we can restrict the currying map to get

$$c : C(X \times Z, Y) \to \mathrm{Map}(X, C(Z, Y)).$$

It is natural to ask whether the map $c(f)$ is necessarily continuous, and if so whether the induced map

$$c : C(X \times Z, Y) \to C(X, C(Z, Y))$$

is a bijection.

THEOREM 8.16. *Let X, Y, Z be topological spaces, and give $C(X, Y)$ the quasi-compact-open topology. Let $f \in \mathrm{Map}(X \times Z, Y)$.*

a) *If f is continuous, so is $c(f) : Z \to C(X, Y)$.*
b) *If X is locally quasi-compact and $c(f)$ is continuous, then f is continuous.*
c) *If X is locally quasi-compact, c restricts to a bijection*

$$c : C(X \times Z, Y) \to C(X, C(Z, Y)).$$

PROOF. a) Suppose that f is continuous. Let $z \in Z$, let $C \subseteq X$ be quasi-compact and $U \subseteq Y$ be open, and suppose that $c(f)(z) \subseteq [C, U]$. The latter means that $f(C, z) \subseteq U$. It suffices to find a neighborhood $\mathcal{N}$ of z such that $c(f)(\mathcal{N}) \subseteq [C, U]$. Since f is continuous, $f^{-1}(U)$ is an open set in $X \times Z$ containing $C \times \{z\}$, so $f^{-1}(U) \cap (C \times Z)$ is open in $C \times Z$ and contains the slice $C \times \{z\}$. By the Tube Lemma, there is a neighborhood $\mathcal{N}$ of z in Z such that $C \times \mathcal{N} \subseteq f^{-1}(U)$. Thus for all $z \in \mathcal{N}$ and $x \in C$ we have $f(x, z) \in U$, so $c(f)(\mathcal{N}) \subseteq [C, U]$.
b) Suppose that $c(f)$ is continuous and X is locally quasi-compact. Then f is the composite map

$$X \times Z \stackrel{c(f) \times 1_X}{\to} C(X, Y) \times X \stackrel{e}{\to} Y,$$

so it is continuous by Theorem 8.15.
c) By part a), the restriction of c to $C(X \times Z, Y)$ lands in $C(X, C(Z, Y))$, and being the restriction of an injective map, is also injective. Part b) says that for any element of $C(X, C(Z, Y))$, its unique preimage under c is continuous, so the restriction of c to $C(X \times Z, Y)$ is a bijection. $\square$

Let X and Y be topological spaces, and let $f, g : X \to Y$ be continuous functions. A **homotopy** from f to g is a continuous function $H : X \times [0, 1] \to Y$ such that

$$\forall x \in X, H(x, 0) = f(x) \text{ and } H(x, 1) = g(x).$$

One thinks of H as a "continuous interpolation" from f to g. The previous result implies that a homotopy H induces a continuous map

$$c(H) : [0, 1] \to C(X, Y),$$

so we may think of homotopies beween f and g as certain *paths* in the space of continuous functions from X to Y. Moreover, when X is locally quasi-compact the converse is true: every path from f to g in the space of continuous functions is a homotopy from f to g.

EXERCISE 8.13. *Let X be a locally quasi-compact topological space, and let Y be a topological space. We define a relation on $C(X,Y)$ be $f \sim g$ if there is a homoptopy $H : X \times [0,1] \to Y$ from f to g.*

a) *Show that $\sim$ is an equivalence relation on $C(X,Y)$.*
b) *Show: the currying map c induces a bijection from the set of $\sim$-equivalence classes to the set of path-components of the space $C(X,Y)$ in the quasi-compact-open topology.*[1]

6. Arzelà-Ascoli

LEMMA 8.17. *Let X be a topological space, let (Y,d) be a metric space, and let $\mathcal{F} \subseteq C(X,Y)$ be a family that is equicontinuous at $x \in X$. Then the closure of $\mathcal{F}$ with respect to the topology of pointwise convergence is also equicontinuous at x.*

PROOF. Let $f_\bullet$ be a net (indexed by a directed set I, say) in $\mathcal{F}$ that is pointwise convergent to f. For all $\epsilon > 0$, there is an open neighborhood $N(x,\epsilon)$ of x such that for all $x' \in N(x,\epsilon)$ and all $g \in \mathcal{F}$, we have $d(g(x),g(x')) < \epsilon$. Let $x' \in N(x,\frac{\epsilon}{3})$. Because of the pointwise convergence, there is $i = i(x,x') \in I$ such that $d(f(x),f_i(x))$ and $d(f(x'),f_i(x'))$ are each less than $\frac{\epsilon}{3}$. Thus

$$\begin{aligned} d(f(x),f(x')) &\leq d(f(x),f_i(x)) + d(f_i(x),f_i(x')) + d(f_i(x'),f(x')) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

□

THEOREM 8.18. *Let X be a topological space, and let (Y,d) be a metric space. Let $\mathcal{F} \subseteq C(X,Y)$ be an equicontinuous family. Then on $\mathcal{F}$ the topologies of pointwise and quasi-compact convergence coincide.*

PROOF. The topology of quasi-compact convergence is always finer than the topology of pointwise convergence, so it suffices to prove the converse: let $f \in \mathcal{F}$. A neighborhood base at f for the topology of quasi-compact convergence is furnished by sets of the form $U(C,f,\epsilon)$ for $C \subseteq X$ a quasi-compact subset and $\epsilon > 0$. For each $x \in C$, let $N(x_i,\frac{\epsilon}{3})$ be an open neighborhood of x_i such that for all $g \in \mathcal{F}$, if $x' \in N(x_i,\frac{\epsilon}{3})$ then $d(g(x),g(x_i)) < \frac{\epsilon}{3}$. Since C is quasi-compact, there are $x_1,\ldots,x_n$ such that $C \subseteq \bigcup_{i=1}^n N(x_i,\frac{\epsilon}{3})$. Now let $\mathcal{N}$ be the set of all $g : X \to Y$ such that for all $1 \leq i \leq n$ we have $d(f(x_i),g(x_i)) < \frac{\epsilon}{3}$. Then $\mathcal{N} \cap \mathcal{F}$ is an open neighborhood of $\mathcal{F}$ in the topology of pointwise convergence. Let $g \in \mathcal{N} \cap \mathcal{F}$, let $x' \in C$, and choose i such that $x' \in N(x_i,\frac{\epsilon}{3})$. Then

$$d(g(x'),f(x')) \leq d(g(x'),g(x_i)) + d(g(x_i),f(x_i)) + d(f(x_i),f(x')) \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

It follows that

$$\mathcal{N} \cap \mathcal{F} \subseteq U(C,f,\epsilon) \cap \mathcal{F}.$$

□

THEOREM 8.19 (Arzelà-Ascoli). *Let X be a topological space that is Hausdorff or quasi-regular. Let (Y,d) be a metric space, and let $\mathcal{F} \subseteq C(X,Y)$ be a family of continuous functions. The following are equivalent:*

(i) *The family $\mathcal{F}$ is compact in the compact-open topology.*
(ii) *All of the following conditions hold:*

[1] This is not a very good exercise in the sense that there is almost nothing to show...but it's quite an important result!

(a) *The family $\mathcal{F}$ is closed in* $\operatorname{Map}(X,Y)$ *in the topology of pointwise convergence.*
(b) *For all $x \in X$, the set $\pi_x(\mathcal{F})$ has compact closure in Y.*
(c) *The restriction of X to each quasi-compact subset of X is equicontinuous.*

PROOF. (i) $\implies$ (ii): Suppose that $\mathcal{F}$ is compact in the compact-open topology. Then it is also compact in the coarser topology of pointwise convergence, a.k.a. the product topology. Since Y is Hausdorff, so is $\operatorname{Map}(X,Y)$ (in the product topology hence also in the compact-open topology), it follows that $\mathcal{F}$ is closed in $\operatorname{Map}(X,Y)$. Moreover, $\pi_x(\mathcal{F})$ is indeed compact in Y.

Now let C be a quasi-compact subset of X. Since the restriction map $r_C : \operatorname{Map}(X,Y) \to \operatorname{Map}(C,Y)$ is continuous, $\mathcal{F}_C \coloneqq r_C(\mathcal{F})$ is compact. The separation hypotheses on X imply that C is quasi-regular. Let $x \in X$ and $\epsilon > 0$. By the quasi-regularity of C and the continuity of f, there is a neighborhood $U(f)$ of x such that $f(\overline{U(f)}) \subseteq B^\circ(f(x), \frac{\epsilon}{2})$. Then $[\overline{(U(f)}, B^\circ(f(x), \frac{\epsilon}{2})]$ is a neighborhood of f in the compact-open topology, so $\{[\overline{(U(f)}, B^\circ(f(x), \frac{\epsilon}{2})]\}_{f\in\mathcal{F}_C}$ is an open cover of $\mathcal{F}_C$. Since $\mathcal{F}_C$ is compact, there are finitely many $f_1, \ldots, f_n \in \mathcal{F}_C$ such that

$$\mathcal{F}_C \subseteq \bigcup_{i=1}^{N} [\overline{(U(f)}, B^\circ(f(x), \frac{\epsilon}{2})].$$

Put $U \coloneqq \bigcap_{i=1}^n U(f_i)$. Then for each $f \in \mathcal{F}_C$, for some $1 \leq i \leq n$ we have $f \in [\overline{(U(f_i)}, B^\circ(f_i(x), \frac{\epsilon}{2})]$ and thus

$$f(U) \subseteq f(\overline{U(f_i)}) \subseteq B^\circ(f_i(x), \frac{\epsilon}{2}) \subseteq B^\circ(f(x), \epsilon),$$

showing that $\mathcal{F}$ is equicontinuous at x.
(ii) $\implies$ (i): We have

$$\mathcal{F} \subseteq \prod_{x\in X} \pi_x(\mathcal{F}) \subseteq \prod_{x\in X} \overline{\pi_x(\mathcal{F})}.$$

Since each $\overline{\pi_x(\mathcal{F})}$ is compact, by Tychonoff's Theorem $\prod_{x\in X} \overline{\pi_x(\mathcal{F})}$ is closed in the product topology. Since $\mathcal{F}$ is closed in the product topology, it is therefore compact in the product topology. So it suffices to show that the product topology on $\mathcal{F}$ coincides with the compact-open topology on $\mathcal{F}$. For any quasi-compact subset C of X, let $\mathcal{F}_C$ be the family restricted to C; by hypothesis this is equicontinuous. So by Theorem 8.18, for any open subset U of Y, we have that $\{f|_C \mid f \in [C,U] \cap \mathcal{F}\}$ is open in $\mathcal{F}_C$ in the product topology. However the restriction map $r : \operatorname{Map}(X,Y) \to \operatorname{Map}(C,Y)$ is continuous for the product topology, and

$$[C,U] \cap \mathcal{F} = r^{-1}(\{f|_C \mid f \in [C,U] \cap \mathcal{F}\}),$$

so $[C,U] \cap \mathcal{F}$ is also open in $\mathcal{F}$ in the product topology. Thus the product topology on $\mathcal{F}$ coincides with the compact-open topology on $\mathcal{F}$, completing the proof. $\square$

EXERCISE 8.14. *Discuss to what extent Theorem 8.8 can be deduced from Theorem 8.19.*

CHAPTER 9

Topological Characterization Theorems

1. Continuous Images of the Cantor Set

1.1. The Alexandroff-Hausdorff Theorem.

THEOREM 9.1. *(Alexandroff-Hausdorff) Let C be the Cantor set. For a Hausdorff space X, the following are equivalent:*

(i) *There is a continuous surjective map $f : C \to X$.*
(ii) *The space X is nonempty, compact and metrizable.*

PROOF. In this proof, for a metric space X, we denote $\prod_{i=1}^{\infty} X$ by X^{∞} and endow it with a good metric using Corollary 2.37.
(i) $\implies$ (ii): Since C is compact metrizable, f is continuous and $X = f(C)$ is Hausdorff, Corollary 7.18 implies that $f(C)$ is nonempty, compact and metrizable.
(ii) $\implies$ (i): We may as well assume that X is a nonempty compact metric space. Our proof closely follows a lovely note of I. Rosenholtz [**Ro76**]. Without changing the underlying topology on X we may (and shall) assume that $\operatorname{diam} X \leq 1$. We break the argument up into several steps.
Step 1: Since X is compact metric, it is separable by Corollary 2.71. For every separable metric space X, we will construct a continuous injection $f : X \to [0,1]^{\infty}$.

Replacing the metric on X by a topologically equivalent one, we may assume that $\operatorname{diam} X \leq 1$. Let $\{x_n\}_{n=1}^{\infty}$ be a countable dense subset, and put $f(x) = \{d(x, x_n)\}_{n=1}^{\infty}$: that is, the nth component is the function $d(\cdot, x_n)$. We know that each distance function $d(\cdot, x_n)$ is continuous, so by Proposition 2.38, the function f iscontinuous. Suppose that $x, y \in X$ are such that $f(x) = f(y)$. We may choose a subsequence $\{x_{n_k}\}$ converging to x, so that

$$0 = \lim_{k \to \infty} d(x, x_{n_k}) = \lim_{k \to \infty} d(y, x_{n_k})$$

and thus $\{x_{n_k}\}$ also converges to y. Since the limit of a sequence in a metric space is unique, we conclude $x = y$.
Step 2: There is a continuous surjection $f : C \to [0,1]$.
As above, we may write an element of C uniquely as $\sum_{n=1}^{\infty} \frac{a_n}{3^n}$ with $a_n \in \{0,2\}$, and then the ternary-to-binary expansion map

$$\sum_{n=1}^{\infty} \frac{a_n}{3^n} \mapsto \sum_{n=1}^{\infty} \frac{a_n/2}{2^n}$$

works.
Step 3: There is a homeomorphism $C \cong C^{\infty}$.
Indeed, since $\mathbb{Z}^+ \times \mathbb{Z}^+$ is countably infinite, by Lemma 2.100 we have

$$C^{\infty} \cong (\{0,1\}^{\infty})^{\infty} \cong \{0,1\}^{\infty} \cong C.$$

Step 4: There is a continuous surjection $C \to [0,1]^\infty$.
By Step 2, there is a surjection $f : C \to [0,1]$, which induces a surjection $f^\infty : C^\infty \to [0,1]^\infty$, $\{x_n\}_{n=1}^\infty \mapsto \{f(x_n)\}_{n=1}^\infty$. Precomposing this with a homemorphism $C \to C^\infty$ from Step 3 gives the result.
Step 5: If $K \subset C$ is nonempty and closed, there is a continuous surjection $C \to K$.
Let C' be the set of $x \in [0,1]$ of the form $\sum_{n=1}^\infty \frac{b_n}{6^n}$ with $b_n \in \{0,5\}$. (C' is constructed much as is C but by iteratively removing the open middle *two thirds* of each subinterval.) The proof of Lemma 2.100 immediately adapts to show that C' is homeomorphic to $\{0,1\}^\infty$ and thus also to C So we may work with C' instead of C. However C' has the following property: if $x \neq y \in C'$, then $\frac{x+y}{2} \notin C'$. It follows that for every nonempty closed subset K' of C' and every $x' \in C'$, there is a unique element $k' \in K'$ such that $d(k',x') = d(K',x')$. The map $x' \in C' \mapsto k' \in K'$ is continuous; moreover it restricts to the identity on K', so is surjective.
Step 6: By Step 1, there is a continuous injection $\iota : X \hookrightarrow [0,1]^\infty$, and by Step 4 there is a continuous surjection $F : C \to [0,1]^\infty$. Let $K = F^{-1}(X)$, which is a closed subset of C. By Step 5, there is a continuous surjection $f : C \to K$. Then $F \circ f : C \to [0,1]^\infty$ is continuous and

$$(F \circ f)(C) = F(f(C)) = F(K) = X. \qquad \square$$

EXERCISE 9.1. a) *Exhibit a nonempty compact space that is not the continuous image of the Cantor set.*

b) *Show: every nonempty countable compact space is a continuous image of the Cantor set.*
c) *[Dreher-Samuel [**DS14**]] Show: there is a countably infinite quasi-compact space that is not a continuous image of the Cantor set.*

1.2. Applications to Space Filling Curves.

COROLLARY 9.2 (Peano's Theorem). *For all $N \in \mathbb{Z}^+$ there is a continuous surjection $f : [0,1] \to [0,1]^N$.*

PROOF. Step 1: Indeed $[0,1]^N$ is a nonempty compact metric space, so there is a continuous surjective map $f : C \to [0,1]^N$.
Step 2: The complement of C in $[0,1]$ is a countable disjoint union of open intervals. On each such interval we may extend f linearly. A little thought shows that this gives a continuous surjective map $f : [0,1] \to [0,1]^N$.
Step 3: Alternately to Step 2, let $f = (f_1, \ldots, f_n) : C \to [0,1]^N$ and apply the Tietze Extension Theorem (Theorem 2.86) to f_i for all $1 \leq i \leq N$. We get a continuous function $F : [0,1] \to [0,1]^N$ which extends f, so is surjective. $\square$

EXERCISE 9.2. *Let $N \geq 2$.*

a) *Show that there is no continuous surjection $f : [0,1] \to \mathbb{R}^N$.*
b) *Show that there is a continuous surjection $f : \mathbb{R} \to \mathbb{R}^N$.*
c) *If X and Y are sets such that there are surjections $f : X \to Y$ and $g : Y \to X$, then there is a bijection ("isomorphism of sets") $\Phi : X \to Y$: this is one formulation of the Schröder-Bernstein Theorem. Deduce that this is far from the case for topological spaces: e.g. we have continuous surjections $\mathbb{R} \to \mathbb{R}^N$ and $\mathbb{R}^N \to \mathbb{R}$, but $\mathbb{R}$ is not homeomorphic to $\mathbb{R}^N$.*

It is very interesting and initially surprising that $[0,1]^N$ is a continuous image of $[0,1]$. This of course raises the question: which spaces are continuous images of $[0,1]$? Our next result answers this for Hausdorff spaces.

2. Continuous Images of $[0, 1]$

THEOREM 9.3 (Hahn-Mazurkiewicz). *For a Hausdorff space X, the following are equivalent:*

(i) *There is a continuous surjective map $f : [0, 1] \to X$.*
(ii) *The space X is nonempty, compact, connected and locally connected.*

Notice that Peano's Theorem (Corollary 9.2) is an immediate consequence of the Hahn-Mazurkiewicz Theorem (Theorem 9.3). Using Theorem 3.50 and Proposition 3.59 we also get that the Hilbert cube $\prod_{i=1}^{\infty}[0, 1]$ is a continuous image of $[0, 1]$: a one-dimensional space can map continuously to an infinite-dimensional space!

3. Topological Characterization of the Cantor Set

THEOREM 9.4. *Let X be a topological space that is nonempty, compact, second countable totally disconnected and perfect (i.e., without isolated points). Then X is homeomorphic to the Cantor set.*

PROOF. Step 0: A second countable compact space is normal, hence metrizable by the Big Urysohn Theorem. Henceforth we assume X is a compact metric space.
Step 1: Let X be a compact metric space. We suppose given a sequence of successive separations on X: we separate $X = \coprod X_0 \cup X_1$, we separate $X_0 = X_{0,0} \cup X_{0,1}$, $X_1 = X_{1,0} \cup X_{0,1}$, and so forth: at the nth stage we have partitioned X into 2^n nonempty clopen sets $X_{\epsilon_1,\ldots,\epsilon_n}$, $\epsilon_i \in \{0, 1\}$. Suppose also that for all $\epsilon > 0$, there is $n \in \mathbb{Z}^+$ such that for all $\epsilon_1, \ldots, \epsilon_n \in \{0, 1\}$, the diameter of $X_{\epsilon_1,\ldots,\epsilon_N}$ is at most ϵ. We claim that X is homeomorphic to $\prod_{n=1}^{\infty}\{0, 1\}$, and thus, by Lemma 2.100, to the Cantor set. Indeed, for $n \in \mathbb{Z}^+$, define $\Phi_n : X \to \{0, 1\}$ by $\Phi_n(x) = 0$ if $x \in X_{\epsilon_1,\ldots,\epsilon_{n-1},0}$ and $\Phi_n(x) = 1$ if $x \in X_{\epsilon_1,\ldots,\epsilon_{n-1},1}$, and let $f : X \to \{0, 1\}^n$ by $x \mapsto \{f_n(x)\}_{n=1}^{\infty}$. The map f is surjective by assumption, and it is injective because of the shrinking diameters condition, which implies that if x_1, x_2 are distinct points of X then for sufficiently large n they cannot lie in the same set $X_{\epsilon_1,\ldots,\epsilon_n}$, i.e., $f_n(x_1) \neq f_n(x_2)$ for some n. Each map f_n is locally constant, hence continuous, hence f is continuous by the universal property of the product topology. Thus $f : X \to \{0, 1\}^n$ is a continuous bijection from a compact topological space to a Hausdorff space, hence a homeomorphism.
Step 2: Let X be a nonempty compact, totally disconnected perfect metric space. We claim that X admits a sequence of successive separations as in Step 1, which will complete the proof. For this we will use Corollary 6.54. First, we can partition X into $2 \leq N_1 < \aleph_0$ clopen sets $\{U_i\}$, each of which has diameter at most $\frac{1}{2}$. Each U_i is again a nonempty compact totally disconnected metric space. Moreover, because X is perfect, so is each U_i (for an isolated point of U_i would be an isolated point of X). Observe that by further separating some of the U_i's if necessary, we may assume that $N_1 - 1 = 2^{n_1}$ for some $n_1 \in \mathbb{Z}^+$. We put $X_0 := U_1$ and we take the $X_{1,\epsilon_2,\ldots,\epsilon_{n_1}}$ to be the remaining U_i's in some order. We now repeat this procedure on each U_i, requiring the diameter of each subset of the partition into clopen sets to be at most $\frac{1}{4}$, and so forth.

It may take a little thought to see that the bookkeeping can be made to work here, and we leave this to the reader. For instance, perhaps it is cleaner to further partition U_1 into 2^{n_1} clopen subsets, so that after the first stage of the process we have partitioned X into 2^{n_1+1} clopen subsets each of diameter at most $\frac{1}{2}$. $\square$

Recall that a compact space is metrizable iff it is second countable. We observe that $\{0,1\}^{\mathbb{R}}$ with the product topology is nonempty, compact, totally disconnected and perfect. However it is not even first countable by Theorem 6.4. This shows that the second countability hypothesis in Theorem 9.4 is necessary.

4. Topological Characterization of $[0,1]$

A point x of a topological space x is a **cut point** if X is connected but $X \setminus \{x\}$ is not connected. In the space $[0,1]$, all points except 0 and 1 are cut points.

THEOREM 9.5. *Let X be a topological space that is compact, connected and second countable and has precisely two points that are not cut points. Then X is homeomorphic to* $[0,1]$.

5. Topological Characterization of $\mathbb{Q}$

THEOREM 9.6 (Sierpinski). *Let X be a metrizable space that is countably infinite and perfect (i.e., without isolated points). Then X is homeomorphic to $\mathbb{Q}$.*

PROOF. We follow an argument of Ciesielski [**Ci20**]. The first idea is that Sierpinski's Theorem is equivalent to the fact that any two countably infinite perfect metric spaces are homeomorphic, and to show this we can just as well take any fixed countably infinite metrizable space X_0 and show that any other countably infinite perfect metric space is homeomorphic to X_0. Our choice for X_0 will be not $\mathbb{Q}$ but the space $\mathcal{S}_0$ of inifnite sequences of natural numbers that are eventually 0. The space $\mathcal{S}_0$ lives naturally as a subset of $\mathcal{S} \coloneqq \prod_{n=1}^{\infty} \mathbb{N}$ of sequences of natural numbers, where we endow each factor $\mathbb{N}$ with its usual discrete topology and then take the product topology on $\mathcal{S}$. Being a countable product of metrizable spaces, we know that $\mathcal{S}$ is metrizable, hence so is its subspace $\mathcal{S}_0$. It is easy to see that $\mathcal{S}_0$ is countably infinite. Indeed, for all $k \in \mathbb{Z}^+$ the set of all sequences that are zero starting with the $(k+1)$st term is naturally in bijection with $\mathbb{N}^k$ so is countable, and $\mathcal{S}_0$ is the countable union of these subsets, hence is countable. In fact any infinite product of discrete spaces, with each factor consisting of more than one point, has no isolated points, since every nonempty open subset of the product projects onto an infinite product of spaces each consisting of more than one point, so nonempty open subsets are uncountable.

Let X be a countably infinite metric space with no isolated points, and enumerate X as $\{x_n\}_{n=1}^{\infty}$. Let $D \coloneqq \{d(x,y) \mid x,y \in X\}$ be the set of distances between points of X; since X is countable, so is D. For any $r \in (0,\infty) \setminus D$ and any $x \in X$ we have

$$B^{\circ}(x,r) = B^{\bullet}(x,r).$$

Henceforth in this proof we will *only* consider balls with radii $r \notin D$, so we write $B(x,r)$ only: this is a clopen set.
Step 1: We observe that every nonempty open subset of X is again countably infinite and without isolated points: indeed, in a metric space without isolated points, all nonempty open subsets are infinite, and an isolated point of an open subset would be an isolated point of the entire space.

Let U be a nonempty open subset of X; we enumerate its elements as $\{u_n\}_{n=1}^{\infty}$. Let $k \in \mathbb{Z}^+$. We **claim** there is a sequence $s_k(U) = \{B_n\}_{n=0}^{\infty}$ of pairwise disjoint clopen balls, each of radius at most $\frac{1}{2^{k+1}}$, such that $U = \bigcup_{n=0}^{\infty} B_n$. We may order them so that B_0 contains the least element of X that lies in U (with respect to the ordering on X we get from the enumeration $X = \{x_n\}_{n=1}^{\infty}$).

This can be shown by an inductive argument. Namely, first take a ball B_0 centered at x_1, with a radius $r \in (0, 2^{-k-1}] \setminus D$ and sufficiently small so that some point of X lies at outside it. The complement $U_1 := U \setminus B_1$ is a nonempty open subset of X_1, so as above it is again countably infinite and without isolated points. Let n_1 be the least index such that x_{n_1} lies in U_1, and let B_2 be a ball centered at x_{n_1}, with radius not in D, at most 2^{-k-1}, and small enough so that B_2 is a proper subset of U_1. Repeating this process generates an infinite disjoint sequence of clopen balls. By construction we have $\{u_1, \ldots, u_n\} \subset \bigcup_{i=1}^{n} B_i$, so $U = \bigcup_{n=1}^{\infty} B_n$.
Step 2: Let $\mathfrak{s}$ denote the set of all finite sequences in $\mathbb{N}$. For each $s \in \mathfrak{s}$ we will construct a nonempty clopen subset C_s. Our definition goes by induction on the length $|s|$ of the finite sequence s. For the unique sequence $\varnothing$ of length 0, we put

$$C_{\varnothing} := X.$$

If $s = (s_1, \ldots, s_k) \in \mathbb{N}^k$ is a sequence of length k and $n \in \mathbb{N}$, we denote by $s.n$ the sequence $(s_1, \ldots, s_k, n) \in \mathbb{N}^{k+1}$ of length $k+1$. (To be sure, $\varnothing.n$ is the length 1 sequence (n).) Inductively, having defined C_s for all sequences of length $k \geq 0$, for all $n \in \mathbb{N}$, we put $C_{s.n}$ to be the the nth term B_n in the sequence $s_k(C_s)$. Otherwise put: to get from stage 0 to stage 1, we partition the entire space X into a sequence of clopen balls of radius at most $\frac{1}{2}$. To get from stage 1 to stage 2, we take each such ball and partition it into a sequence of clopen balls of radius at most $\frac{1}{4}$, and so forth. At the kth stage, our sequence of length k represents this process of breaking a clopen subset up into infinitely many clopen balls k times, and to get from this to the $k+1$st stage we do this one more time.
Step 3: For a sequence $x \in S^{\mathbb{Z}^+}$ in any set and $k \in \mathbb{Z}^+$, denote by $\tau_k(x)$ its truncation to a sequence of length k:

$$\tau_k : (s_1, \ldots, s_n, \ldots) \mapsto (s_1, \ldots, s_k).$$

For each $x \in X$, there is a unique sequence s_x of natural numbers such that

$$x \in \bigcap_{k=1}^{\infty} C_{\tau_k(s_x)}.$$

We claim that the map

$$h : X \to \mathcal{S}_0, \ x \mapsto s_x$$

is a homeomorphism, which would complete the proof. The first order of business is to show that for all $x \in X$ we indeed have $s_x \in \mathcal{S}_0$, i.e., that the terms of the sequence are eventually zero. To see this, write $x = x_j$, and choose n such that $d(x_i, x_j) > \frac{1}{2^n}$ for all $1 \leq i < j$. Then for all $k > n$, the kth term of the sequence gives the index of the point x_j in a partition of $C := C_{\tau^{k-1}(h(x_j))}$, a clopen subset of diameter at most $2^{-k+1} \leq 2^{-n}$, into countably many disjoint clopen subsets. If for some $i < j$ we had $x_i \in C$, then x_j could not also lie in C, since $d(x_i, x_j) > \operatorname{diam} C$, but x_j lies in C by construction. It follows that j is the smallest index such that x_j lies in C, hence it is also the smallest index such that x_j lies in $C_{\tau^k h(x_j)}$, so the kth component of $h(x_j)$ is 0. Thus $h(x_j) \in \mathcal{S}_0$.

The injectivity of h follows from the fact that $\operatorname{diam} C_{\tau^k(s_x)} \leq 2^{-k} \to 0$. For any $s \in \mathcal{S}_0$, this amounts to successively selecting one clopen ball from each partition so that after a certain point we always select the 0th ball. By construction, the selection corresponding to each finite sequence of length k does give a nonempty clopen ball. Suppose that for all $k \geq K$ the kth term of s is 0, and let j be the least index such that $x_j \in C_{\tau^K(s)}$. Then we have $x_j \in C_{\tau^k(s)}$ for all $k > k$, hence $x_j \in \bigcap_{k=1}^{\infty} C_{\tau_k(s)}$, so $h(x_j) = s$.

Let $t \in \mathbb{N}^k$ be a finite sequence of length k, and let

$$[t] := \{s \in \mathcal{S}_0 \mid \tau^k(s) = t\}.$$

Then $\mathcal{B} := \{[t] \mid \ t \text{ is a finite sequence in } \mathbb{N}\}$ is a base for the topology on $\mathcal{S}_0$, and for any finite sequence t in $\mathbb{N}$ we have

$$h^{-1}([t]) = C_t,$$

which is open. So h is continuous. Similarly, the set $\{C_t \mid \ t \text{ is a finite sequence in } \mathbb{N}\}$ is a base for the topology on X: if $x \in X$ and $B^\circ(x, \delta)$ is an open ball centered at x, then there is some t such that C_t contains x and has diameter less than δ, so $x \in C_t \subset B^\circ(x, \delta)$. Since $h(C_t) = [t]$, the map h is open. Therefore it is a homeomorphism, which completes the proof. □

6. Topological Characterization of $\mathbb{I}$

Theorem 9.7 (Alexandroff-Urysohn). *For a topological space X, the following are equivalent:*

(i) *The space X is nonempty, separable, completely metrizable, and every compact subset of X has empty interior.*
(ii) *The space X is homeomorphic to the irrational numbers $\mathbb{I}$.*
(iii) *The space X is homeomorphic to a countably infinite Cartesian product of countably infinite discrete spaces.*

CHAPTER 10

Appendix: Very Basic Set Theory

1. The Basic Trichotomy: Finite, Countable and Uncountable

1.1. Introducing equivalence of sets, countable and uncountable sets.

We assume known the set $\mathbb{Z}^+$ of positive integers, and the set $\mathbb{N} = \mathbb{Z}^+ \cup \{0\}$ of natural numbers. For any $n \in \mathbb{Z}^+$, we denote by $[n]$ the set $\{1, \ldots, n\}$. We take it as obvious that $[n]$ has n elements, and also that the empty set $\emptyset$ has 0 elements. Just out of mathematical fastidiousness,[1] let's define $[0] = \emptyset$ (why not?).

It is pretty clear what it means for an arbitrary set S to have 0 elements: it must be the empty set. That is – and this is a somewhat curious property of the empty set – $\emptyset$ as a set is uniquely characterized by the fact that it has 0 elements.

What does it mean for an arbitrary set S to have n elements? By definition, it means that there exists a bijection $\iota : S \to [n]$, i.e., a function which is both injective and surjective; or, equivalently, a function for which there exists an inverse function $\iota' : [n] \to S$.[2]

Let us call a set *finite* if it has n elements for some $n \in \mathbb{N}$, and a set *infinite* if it is not finite.

Certainly there are some basic facts that we feel should be satisfied by these definitions. For instance:

Fact 10.1. *The set $\mathbb{Z}^+$ is infinite.*

Proof. It is certainly nonempty, so we would like to show that for no $n \in \mathbb{Z}^+$ is there a bijection $\iota : [n] \to \mathbb{Z}^+$. This seems obvious. Unfortunately, sometimes in mathematics we must struggle to show that the obvious is true (and sometimes what seems obvious is not true!). Here we face the additional problem of not having formally axiomatized things, so it's not completely clear what's "fair game" to use in a proof. But consider the following: does $\mathbb{Z}^+$ have one element? Absolutely not: for any function $\iota : [1] = \{1\} \to \mathbb{Z}^+$, ι is not surjective because it does not hit $\iota(1) + 1$. Does $\mathbb{Z}^+$ have two elements? Still, no: if ι is not injective, the same argument as before works; if ι is injective, its image is a 2 element subset of $\mathbb{Z}^+$. Since $\mathbb{Z}^+$ is totally ordered (indeed well-ordered), one of the two elements in the image is larger than the other, and then that element plus one is not in the image of

[1] Well, not really: this will turn out to be quite sensible.

[2] I am assuming a good working knowledge of functions, injections, surjections, bijections and inverse functions. This asserts at the same time (i) a certain amount of mathematical sophistication, and (ii) a certain amount of metamathematical informality.

our map. We could prove it for 3 as well, which makes us think we should probably work by induction on n. How to set it up properly? Let us try to show that for all n and all $\iota : [n] \to \mathbb{Z}^+$, there exists $N = N(\iota)$ such that $\iota([n]) \subset [N]$. If we can do this, then since $[N]$ is clearly a proper subset of $\mathbb{Z}^+$ (it does not contain $N+1$, and so on) we will have shown that for no n is there a surjection $[n] \to \mathbb{Z}^+$ (which is in fact stronger than what we claimed). But carrying through the proof by induction is now not obvious but (much better!) very easy, so is left to the reader. □

What did we use about $\mathbb{Z}^+$ in the proof? Some of the Peano axioms for $\mathbb{Z}^+$, most importantly that it satisfies the principle of mathematical induction (POMI). Since it is hard to imagine a rigorous proof of a nontrivial statement about $\mathbb{Z}^+$ that does not use POMI, this is a good sign: things are proceeding well so far.

What about $\mathbb{Z}$: is it too infinite? It should be, since it contains an infinite subset. This is logically equivalent to the following fact:

FACT 10.2. *A subset of a finite set is finite.*

PROOF. More concretely, it suffices to show that for any $n \in \mathbb{N}$ and and subset $S \subset [n]$, then for some $m \in \mathbb{N}$ there exists a bijection $\iota : S \to [m]$. As above, for any specific value of n, it straightforward to show this, so again we should induct on n. Let's do it this time: assume the statement for n, and let $S \subset [n+1]$. Put $S' = S \cap [n]$, so by induction there exists a bijection $\iota' : [m] \to S'$ for some $m' \in \mathbb{N}$. Composing with the inclusion $S' \subset S$ we get an injection $\iota : [m] \to S$. If $n+1$ is not an element of S, then $S' = S$ and ι is a bijection. If $n+1 \in S$, then extending ι' to a map from $[m+1]$ to S by sending $m+1$ to $n+1$ gives a bijection. □

Again, by contraposition this shows that many of our most familiar sets of numbers – e.g. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ – are infinite.

There is one more thing we should certainly check: namely, we have said that a set S has n elements if it can be put in bijection with $[n]$ for some n. But we have not shown that this n is unique: perhaps a set can have n elements and also $n + 691$ elements? Of course not:

FACT 10.3. *For natural numbers $n \neq n'$, there is no bijection from $[n]$ to $[n']$.*

Of course, we even know a more precise result:

FACT 10.4. *Let S be a set with m elements and T a set with n elements.*
a) If there exists a surjection $\varphi : S \to T$, then $m \geq n$.
b) If there exists an injection $\varphi : S \to T$, then $m \leq n$.

EXERCISE 10.1. *Give a proof of Fact 10.4 that is rigorous enough for your taste.*

Remark: For instance, part b) is the famous "Pigeonhole" or "Dirichlet's box" principle, and is usually regarded as obvious. Of course, if we play the game of formalized mathematics, then "obvious" means "following from our axioms in a way which is so immediate so as not to deserve mention," and Fact 10.4 is not obvious in this sense. (But one can give a proof in line with the above induction proofs, only a bit longer.)

EXERCISE 10.2. *Show: for sets S and T, the following are equivalent:*

(i) *There is a surjection $S \to T$.*
- *(ii)] There is an injection $T \to S$.*

Let us press on to study the properties of *infinite* sets.

The following definition is due to Cantor: we say that sets S and T as *equivalent*, and write $S \cong T$ if there exists a bijection $\iota : S \to T$.

Historical Remark: When there exists a bijection between S and T, Cantor first said that S and T have the same *power*.[3] As is often the case in mathematics, this forces us to play a linguistic-grammatical game – given that a definition has been made to have a certain part of speech, write down the cognate words in other parts of speech.[4] Thus a faithful rendition of Cantor's definition in adjectival form would be something like *equipotent*. The reader should be warned that it would be more common to use the term *equinumerous* at this point.

However, we have our reasons for choosing to use "equivalent." The term "equinumerous," for instance, suggests that the two sets have the same number of elements, or in other words that there is some numerical invariant we are attaching to a single set with the property that two sets can be put in bijection exactly when both have the same value of this numerical invariant. But we would like to view things in exactly the opposite way. Let us dilate a bit on this point.

It was Cantor's idea that we should regard two sets as "having the same size" iff they are equivalent, i.e., iff their elements can be paired off via a one-to-one correspondence. Certainly this is consistent with our experience from finite sets. There is, however, a brilliant and subtle twist: colloquially one thinks of counting or measuring something as a process which takes as input one collection of objects and outputs a "number." We therefore have to have names for all of the "numbers" which measure the sizes of things: if you like, we need to count arbitrarily high. Not every civilization has worked out such a general counting scheme: I have heard tell that in a certain "primitive tribe" they only have words for numbers up to 4 and anything above this is just referred to as "many." Indeed we do not have proper names for arbitrarily large numbers in the English language (except by recourse to iteration, e.g., million million for a trillion).

But notice that we do not have to have such an elaborate "number knowledge" to say whether two things have the same size or not. For instance, one may presume that shepherding predates verbal sophistication, so the proto-linguistic shepherd needs some other means of making sure that when he takes his sheep out to graze in the countryside he returns with as many as he started with. The shepherd can do this as follows: on his first day on the job, as the sheep come in, he has ready some sort of sack and places stones in the sack, one for each sheep. Then in the future he counts his sheep, not in some absolute sense, but in relation to these stones. If one day he runs out of sheep before stones, he knows that he is missing some sheep (at least if he has only finitely many sheep!).

Even today there are some situations where we test for equivalence rather than count in an absolute sense. For instance, if you come into an auditorium and everyone is sitting in a (unique!) seat then you know that there are at least as many

[3]Or rather, he said something in German that gets translated to this. Such pedantic remarks will be omitted from now on!

[4]This is a game that some play better than others, viz.: generization, sobrification, unicity.

seats as people in the room without counting both quantities.

What is interesting about infinite sets is that these sorts of arguments break down: the business of taking away from an infinite set becomes much more complicated than in the finite case, in which, given a set S of n elements and any element $x \in S$, then $S \setminus x$ has $n-1$ elements. (This is something that you can establish by constructing a bijection and is a good intermediate step towards Fact 10.4.) On the other hand, $\mathbb{Z}^+$ and $\mathbb{N}$ are equivalent, since the map $n \mapsto n-1$ gives a bijection between them. Similarly $\mathbb{Z}^+$ is equivalent to the set of even integers ($n \mapsto 2n$). Indeed, we soon see that much more is true:

FACT 10.5. *For any infinite subset $S \subset \mathbb{Z}^+$, S and $\mathbb{Z}^+$ are equivalent.*

PROOF. Using the fact that $\mathbb{Z}^+$ is well-ordered, we can define a function from S to $\mathbb{Z}^+$ by mapping the least element s_1 of S to 1, the least element s_2 of $S \setminus \{s_1\}$ to 2, and so on. If this process terminates after n steps then S has n elements, so is finite, a contradiction. Thus it goes on forever and clearly gives a bijection. $\square$

It is now natural to wonder which other familiar infinite sets are equivalent to $\mathbb{Z}^+$ (or $\mathbb{N}$). For this, let's call a set equivalent to $\mathbb{Z}^+$ *countable*.[5] A slight variation of the above argument gives

FACT 10.6. *Every infinite set has a countable subset.*

PROOF. Indeed, for infinite S just keep picking elements to define a bijection from $\mathbb{Z}^+$ to some subset of S; we can't run out of elements since S is infinite! $\square$

As a first example:

FACT 10.7. *The two sets $\mathbb{Z}$ and $\mathbb{Z}^+$ are equivalent.*

PROOF. We define an explicit bijection $\mathbb{Z} \to \mathbb{Z}^+$ as follows: we map $0 \mapsto 1$, then $1 \mapsto 2$, $-1 \mapsto 3$, $2 \mapsto 4$, $-2 \mapsto 5$ and so on. (If you are the kind of person who thinks that having a formula makes something more rigorous, then we define for positive n, $n \mapsto 2n$ and for negative n, $n \mapsto 2|n|+1$.) $\square$

FACT 10.8. *Suppose that S_1 and S_2 are two countable sets. Then $S_1 \bigcup S_2$ is countable.*

Indeed, we can make a more general splicing construction:

FACT 10.9. *Let $\{S_i\}_{i \in I}$ be an indexed family of pairwise disjoint nonempty sets; assume that I and each S_i is at most countable (i.e., countable or finite). Then $S := \bigcup_{i \in I} S_i$ is at most countable. Moreover, S is finite iff I and all the S_i are finite.*

PROOF. We sketch the construction: since each S_i is at most countable, we can order the elements as s_{ij} where either $1 \leq j \leq \infty$ or $1 \leq j \leq N_j$. If everything in sight is finite, it is obvious that S will be finite (a finite union of finite sets is finite). Otherwise, we define a bijection from $\mathbb{Z}^+$ to S as follows: $1 \mapsto s_{11}$, $2 \mapsto s_{12}$, $3 \mapsto s_{22}$, $4 \mapsto s_{13}$, $5 \mapsto s_{23}$, $6 \mapsto s_{33}$, and so on. Here we need the convention that when s_{ij} does not exist, we omit that term and go on to the next element in the codomain. $\square$

[5] Perhaps more standard is to say "countably infinite and reserve "countable" to mean countably infinite or finite. Here we suggest simplifying the terminology.

Fact 10.9 is used very often. As one immediate application:

FACT 10.10. *The set of rational numbers $\mathbb{Q}$ is countable.*

PROOF. Each nonzero rational number α can be written uniquely as $\pm\frac{a}{b}$, where $a,\ b \in \mathbb{Z}^+$. We define the height $h(\alpha)$ of α to be $\max a, b$ and also $h(0) = 0$. It is clear that for any height $n > 0$, there are at most $2n^2$ rational numbers of height n,[6] and also that for every $n \in \mathbb{Z}^+$ there is at least one rational number of height n, namely the integer $n = \frac{n}{1}$. Therefore taking $I = \mathbb{N}$ and putting some arbitrary ordering on the finite set of rational numbers of height n, Fact 10.9 gives us a bijection $\mathbb{Z}^+ \to \mathbb{Q}$. $\square$

In a similar way, one can prove that the set $\overline{\mathbb{Q}}$ of algebraic numbers is countable.

FACT 10.11. *If A and B are countable, then the Cartesian product $A \times B$ is countable.*

EXERCISE 10.3. *Prove Fact 10.11. (Strategy 1: Reduce to the case of $\mathbb{Z}^+ \times \mathbb{Z}^+$ and use the diagonal path from the proof of Fact 10.9. Strategy 2: Observe that $A \times B \cong \bigcup_{a\in A} B$ and apply Fact 10.9 directly.)*

The buck stops with $\mathbb{R}$. Let's first prove the following theorem of Cantor, which is arguably the single most important result in set theory. Recall that for a set S, its power set 2^S is the set of all subsets of S.

THEOREM 10.12. *(First Fundamental Theorem of Set Theory)*
There is no surjection from a set S to its power set 2^S.

PROOF. It is short and sweet. Suppose that $f : S \to 2^S$ is any function. We will produce an element of 2^S which is not in the image of f. Namely, let T be the set of all $x \in S$ such that x is not an element of $f(x)$, so T is some element of 2^S. Could it be $f(s)$ for some $s \in S$? Well, suppose $T = f(s)$ for some $s \in S$. We ask the innocent question, "Is $s \in T$?" Suppose first that it is: $s \in T$; by definition of T this means that s is not an element of $f(s)$. But $f(s) = T$, so in other words s is not an element of T, a contradiction. Okay, what if s is not in T? Then $s \in f(s)$, but again, since $f(s) = T$, we conclude that s *is* in T. In other words, we have managed to define, in terms of f, a subset T of S for which the notion that T is in the image of f is logically contradictory. So f is not surjective! $\square$

What does this have to do with $\mathbb{R}$? Let us try to show that the interval $(0, 1]$ is uncountable. By Fact 10.5 this implies that $\mathbb{R}$ is uncountable. Now using binary expansions, we can identify $(0, 1]$ with the power set of $\mathbb{Z}^+$. Well, almost: there is the standard slightly annoying ambiguity in the binary expansion, that

$$.a_1a_2a_3 \cdots a_n 01111111111\ldots = .a_1a_2a_2 \cdots a_n 1000000000\ldots.$$

There are various ways around this: for instance, suppose we agree to represent every element of $(0, 1]$ by an element which does not terminate in an infinite string of zeros. Thus we have identified $(0, 1]$ with a certain subset T of the power set of $\mathbb{Z}^+$, the set of *infinite* subsets of $\mathbb{Z}^+$. But the set of finite subsets of $\mathbb{Z}^+$ is countable (Fact 10.9 again), and since the union of two countable sets would be countable (and again!), it must be that T is uncountable. Hence so is $(0, 1]$, and so is $\mathbb{R}$.

[6] I will resist the temptation to discuss how to replace the 2 with an asymptotically correct constant.

There are many other proofs of the uncountability of $\mathbb{R}$. For instance, we could contemplate a function $f : \mathbb{Z}^+ \to \mathbb{R}$ and, imitating the proof of Cantor's theorem, show that it cannot be surjective by finding an explicit element of $\mathbb{R}$ not in its image. We can write out each real number $f(n)$ in its decimal expansion, and then construct a real number $\alpha \in [0,1]$ whose nth decimal digit α_n is different from the nth decimal digit of $f(n)$. Again the ambiguity in decimal representations needs somehow to be addressed: here we can just stay away from 9's and 0's. Details are left to the reader.

Example 10.1. *Since $\mathbb{R}$ is by definition the completion of $\mathbb{Q}$ with respect to the standard Euclidean metric, and has no isolated points, $\mathbb{R}$ must be uncountable. For that matter, even $\mathbb{Q}$ has no isolated points (which is strictly stronger: no element of the completion of a metric space minus the space itself can be isolated, since this would contradict the density of a space in its completion), so since we know it is countable, we deduce that it is incomplete without having to talk about $\sqrt{2}$ or any of that sort of thing. Indeed, the same argument holds for $\mathbb{Q}$ endowed with a p-adic metric: there are no isolated points, so $\mathbb{Q}_p$ is uncountable and not equal to $\mathbb{Q}$.*

The above was just one example of the importance of distinguishing between countable and uncountable sets. Let me briefly mention some other examples:

Example 10.2 (Measure Theory). *A measure is a $[0,\infty]$-valued function defined on a certain family of subsets of a given set; it is required to be countably additive but not uncountably additive. For instance, this gives us a natural notion of size on the unit circle, so that the total area is π and the area of any single point is 0. The whole can have greater measure than the sum of the measures of the parts if there are uncountably many parts!*

Example 10.3. *Given a differentiable manifold M of dimension n, then any submanifold of dimension $n-1$ has, in a sense which is well-defined independent of any particular measure on M, measure zero. In particular, one gets from this that a countable family of submanifolds of dimension at most $n-1$ cannot "fill out" an n-dimensional manifold. In complex algebraic geometry, such stratifications occur naturally, and one can make reference to a "very general" point on a variety as a point lying on the complement of a (given) countable family of lower-dimensional subvarieties, and be confident that such points exist!*

Example 10.4. *Model theory is a branch of mathematics that tends to exploit the distinction between countable and uncountable in rather sneaky ways. Namely, there is the Lowenheim-Skolem theorem, which states in particular that any theory (with a countable language) that admits an infinite model admits a countable model. Moreover, given any uncountable model of a theory, there is a countable submodel which shares all the same "first order" properties, and conversely the countable/uncountable dichotomy is a good way to get an intuition on the difference between first-order and second-order properties.*

1.2. Some further basic results.

Fact 10.13. *A set S is infinite iff it is equivalent to a proper subset of itself.*

Proof. One direction expresses an obvious fact about finite sets. Conversely, let S be an infinite set; as above, there is a countable subset $T \subset S$. Choose some bijection ι between T and $\mathbb{N}$. Then there is a bijection ι' between $T' := T \setminus \iota^{-1}(0)$ and T (just because there is a bijection between $\mathbb{N}$ and $\mathbb{Z}^+$. We therefore get a bijection between $S' := S \setminus \iota^{-1}(0$ and S by applying ι' from T' to T and the identity on $S \setminus T$. $\square$

This characterization of infinite sets is due to Dedekind. What is ironic is that in some sense it is cleaner and more intrinsic than our characterization of finite sets, in which we had to compare against a distinguished family of sets $\{[n] \mid n \in \mathbb{N}\}$. Thus perhaps we should define a set to be finite if it cannot be put in bijection with a proper subset of itself! (On the other hand, this is not a "first order" property, so is not in reality that convenient to work with.)

Notice that in making the definition "uncountable," i.e., an infinite set which is not equivalent to $\mathbb{Z}^+$, we have essentially done what we earlier made fun of the "primitive tribes" for doing: giving up distinguishing between very large sets. In some sense, set theory begins when we attempt to classify uncountable sets up to equivalence. This turns out to be quite an ambitious project – we will present the most basic results of this project in the next installment – but there are a few further facts that one should keep in mind throughout one's mathematical life.

Let us define a set S to be *of continuum type* (or, more briefly, a continuum[7]) if there is a bijection $\iota : S \to \mathbb{R}$. One deserves to know the following:

FACT 10.14. *There exists an uncountable set not of continuum type, namely* $2^{\mathbb{R}}$.

PROOF. By Theorem 10.12 there is no surjection from $\mathbb{R}$ to $2^{\mathbb{R}}$, so $2^{\mathbb{R}}$ is certainly not of continuum type. We must however confirm what seems intuitively plausible: that $2^{\mathbb{R}}$ is indeed uncountable. It is certainly infinite, since via the natural injection $\iota : \mathbb{R} \to 2^{\mathbb{R}}$, $r \mapsto \{r\}$, it contains an infinite subset. But indeed, this also shows that $2^{\mathbb{R}}$ is uncountable, since if it were countable, its subset $\iota(\mathbb{R}) \cong \mathbb{R}$ would be countable, which it isn't. $\square$

1.3. Some sets of continuum type.

For any two sets S and T, we define T^S as the set of all functions $f : S \to T$. When $T = [2]$, the set of all functions $f : S \to [2]$ is naturally identified with the power set 2^S of S (so the notation is *almost* consistent: for full consistency we should be denoting the power set of S by $[2]^S$, which we will not trouble ourselves to do).

FACT 10.15. *The sets* $(0, 1]$, $2^{\mathbb{Z}^+}$ *and* $\mathbb{R}^{\mathbb{Z}^+}$ *are of continuum type.*

Earlier we identified the unit interval $(0, 1]$ in $\mathbb{R}$ with the infinite subsets of $\mathbb{Z}^+$ and remarked that, since the finite subsets of $\mathbb{Z}^+$ form a countable set, this implies that $(0, 1]$ hence $\mathbb{R}$ itself is uncountable. Let us refine this latter observation slightly:

LEMMA 10.16. *Let S be an uncountable set and $C \subset S$ an at most countable subset. Then $S \setminus C \cong S$.*

---

[7]This has a different meaning in general topology, but no confusion should arise.

Proof. Suppose first that C is finite, say $C \cong [n]$. Then there exists an injection $\iota : \mathbb{Z}^+ \to S$ such that $\iota([n]) = C$ (as follows immediately from Fact 6). Let $C_\infty = \iota(\mathbb{Z}^+)$. Now we can define an explicit bijection β from $S \setminus C$ to S: namely, we take β to be the identity on the complement of C_∞ and on C_∞ we define $\beta(\iota(k)) = \iota(k-n)$.

Now suppose C is countable. We do something rather similar. Namely, taking $C_1 = C$, since $S \setminus C_1$ is uncountable, we can find a countably infinite subset $C_2 \subset S \setminus C_1$. Proceeding in this way we can find a family $\{C_i\}_{i \in \mathbb{Z}^+}$ of pairwise disjoint countable subsets of S. Let us identify each of these subsets with $\mathbb{Z}^+$, getting a doubly indexed countable subset $C_\infty := \bigcup_i C_i = \{c_{ij}\}$ – here c_{ij} is the jth element of C_i. Now we define a bijection β from $S \setminus C_1$ to S by taking β to be the identity on the complement of C_∞ and by putting $\beta(c_{ij}) = c_{(i-1)j}$. This completes the proof of the lemma. $\square$

Thus the collection of infinite subsets of $\mathbb{Z}^+$ – being a subset of $2^{\mathbb{Z}^+}$ with countable complement – is equivalent to $2^{\mathbb{Z}^+}$, and hence $(0,1] \cong 2^{\mathbb{Z}+}$. So let us see that $(0,1]$ is of continuum type. One way is as follows: again by the above lemma, $[0,1] \cong (0,1)$, and $\mathbb{R}$ is even homeomorphic to $(0,1)$: for instance, the function

$$\arctan(\pi(x - \frac{1}{2})) : (0,1) \xrightarrow{\sim} \mathbb{R}.$$

For the case of $(\mathbb{Z}^+)^{\mathbb{R}}$: since $\mathbb{R} \cong 2^{\mathbb{Z}^+}$, it is enough to find a bijection from $(\mathbb{Z}^+)^{2^{\mathbb{Z}^+}}$ to $2^{\mathbb{Z}^+}$. This is in fact quite easy: we are given a sequence a_{ij} of binary sequences and want to make a single binary sequence. But we can do this just by choosing a bijection $\mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{Z}^+$.

A little more abstraction will make this argument seem much more reasonable:

Lemma 10.17. *Suppose A, B and C are sets. Then there is a natural bijection*

$$(A^B)^C \cong A^{C \times B}.$$

Proof. Indeed, given a function F from C to A^B and an ordered pair $(c,b) \in C \times B$, $F(c)$ is a function from B to A and so $F(c)(b)$ is an element of a. Conversely, every function from $C \times B$ to A can be viewed as a function from C to the set A^B of functions from B to A, and these correspondences are evidently mutually inverse.[8] So what we said above amounts to

$$2^{\mathbb{Z}^+} \cong 2^{\mathbb{Z}^+ \times \mathbb{Z}^+} \cong (2^{\mathbb{Z}^+})^{\mathbb{Z}^+}.$$

$\square$

Exercise 10.4. *Show: A subinterval of $\mathbb{R}$ containing more than one point is of continuum type.*

It is also the case that $(\mathbb{Z}^+)^{\mathbb{Z}^+}$ is of continuum type. I do not see a proof of this within the framework we have developed so far. What we can show is that there exists an injection $\mathbb{R} \hookrightarrow (\mathbb{Z}^+)^{\mathbb{Z}^+}$ – indeed, since $\mathbb{R} \cong 2^{\mathbb{Z}^+}$, this is obvious – and also that there exists an injection $(\mathbb{Z}^+)^{\mathbb{Z}^+} \hookrightarrow 2^{\mathbb{Z}^+} \cong \mathbb{R}$.

To see this latter statement: given any sequence of positive integers, we want to return a binary sequence – which it seems helpful to think of as "encoding" our

[8]This is canonical bijection is sometimes called "adjunction."

original sequence – in such a way that the decoding process is unambiguous: we can always reconstruct our original sequence from its coded binary sequence. The first thought here is to just encode each positive integer a_i in binary and concatenate them. Of course this doesn't quite work: the sequence 2, 3, 1, 1, 1 ... gets coded as 1011 followed by an infinite string of ones, as does the sequence 11, 1, 1, 1 .... But this can be remedied in many ways. One obvious way is to retreat from binary notation to *unary* notation: we encode a_i as a string of i ones, and in between each string of a_i ones we put a zero to separate them. This clearly works (it seems almost cruelly inefficient from the perspective of information theory, but no matter).

Roughly speaking, we have shown that $(\mathbb{Z}^+)^{\mathbb{Z}+}$ is "at least of continuum type" and "at most of continuum type," so if equivalences of sets do measure some reasonable notion of their size, we ought to be able to conclude from this that $(\mathbb{Z}^+)^{\mathbb{Z}+}$ is itself of continuum type. This is true, a special case of the important Schröder-Bernstein theorem whose proof we defer until the next installment.

1.4. Many inequivalent uncountable sets.

From the fundamental Theorem 10.12 we first deduced that not all infinite sets are equivalent to each other, because the set $2^{\mathbb{Z}^+}$ is not equivalent to the countable infinite set $\mathbb{Z}^+$. We also saw that $2^{\mathbb{Z}^+} \cong \mathbb{R}$ so called it a set of continuum type. Then we noticed that Cantor's theorem implies that there are sets not of continuum type, namely $2^{\mathbb{R}} \cong 2^{2^{\mathbb{Z}+}}$. By now one of the most startling mathematical discoveries of all time must have occurred to the reader: we can keep going!

To simplify things, let us use (and even slightly abuse) an obscure[9] but colorful notation due to Cantor: instead of writing $\mathbb{Z}^+$ we shall write $\beth_0$. For $2^{\mathbb{Z}+}$ we shall write $\beth_1$, and in general, for $n \in \mathbb{N}$, having defined $\beth_n$ (informally, as the n-fold iterated power set of $\mathbb{Z}^+$), we will define $\beth_{n+1}$ as $2^{\beth_n}$. Now hold on to your hat:

FACT 10.18. *The infinite sets $\{\beth_n\}_{n\in\mathbb{N}}$ are pairwise inequivalent.*

PROOF. Let us first make the preliminary observation that for any nonempty set S, there is a surjection $2^S \to S$. Indeed, pick your favorite element of S, say x; for every $s \in S$ we map $\{s\}$ to s, which is "already" a surjection; we extend the mapping to all of 2^S by mapping every other subset to x.

Now we argue by contradiction: suppose that for some $n > m$ there exists even a surjection $s : \beth_m \to \beth_n$. We may write $n = m+k$. By the above, by concatenating (finitely many) surjections we get a surjection $\beta : \beth_{m+k} \to \beth_{m+1}$. But then $\beta \circ s : \beth_m \to \beth_{m+1} = 2^{\beth_m}$ is a surjection, contradicting Cantor's theorem. □

Thus there are rather a lot of inequivalent infinite sets. Is it possible that the $\beth_n$'s are all the infinite sets? In fact it is *not*: define $\beth_\omega := \bigcup_{n\in\mathbb{N}} \beth_n$. This last set $\beth_\omega$ is certainly not equivalent to $\beth_n$ for any n, because it visibly surjects onto $\beth_{n+1}$. Are we done yet? No, we can keep going, defining $\beth_{\omega+1} := 2^{\beth_\omega}$.

To sum up, we have a two-step process for generating a mind-boggling array of equivalence classes of sets. The first step is to pass from a set to its power set, and the second stage is to take the union over the set of all equivalence classes of

[9] At least, *I* didn't know about it until recently; perhaps this is not your favorite criterion for obscurity.

sets we have thus far considered. Inductively, it seems that each of these processes generates a set which is not surjected onto by any of the sets we have thus far considered, so it gives a new equivalence class. Does the process ever end?!?

Well, the above sentence is an example of the paucity of the English language to describe the current state of affairs, since even the sequence $\beth_0, \beth_1, \beth_2 \ldots$ does not end in the conventional sense of the term. Better is to ask whether or not we can reckon the equivalence classes of sets even in terms of infinite sets. At least we have only seen countably many equivalence classes of sets[10] thus far: is it possible that the collection of all equivalence classes of sets is countable?

No again, and in fact that's easy to see. Suppose $\{S_i\}_{i \in \mathbb{N}}$ is any countable collection of pairwise inequivalent sets. Then – playing both of our cards at once! – one checks immediately that there is no surjection from any S_i onto $2^{\bigcup_{i \in \mathbb{N}} S_i}$. In fact it's even stranger than this:

FACT 10.19. *For no set I does there exists a family of sets $\{S_i\}_{i \in I}$ such that every set S is equivalent to S_i for at least one i.*

PROOF. Again, take $S_{\text{bigger}} = 2^{\bigcup_{i \in I} S_i}$. There is no surjection from $\bigcup_{i \in I} S_i$ onto S_{bigger}, so for sure there is no surjection from any S_i onto S_{bigger}. $\square$

2. Order and Arithmetic of Cardinalities

Here we pursue Cantor's theory of cardinalities of infinite sets a bit more deeply. We also begin to take a more sophisticated approach in that we identify which results depend upon the Axiom of Choice and strive to give proofs which avoid it when possible. However, we defer a formal discussion of the Axiom of Choice and its equivalents to a later installment, so the reader who has not encountered it before can ignore these comments and/or skip ahead to the next installment.

We warn the reader that the main theorem in this installment – Theorem 10.22 (which we take the liberty of christening "The Second Fundamental Theorem of Set Theory") – will not be proved until the next installment, in which we give a systematic discussion of well-ordered sets.

For More Advanced Readers: We will mostly be content to use the Axiom of Choice (AC) in this handout, despite the fact that we will not discuss this axiom until Handout 3. However, whereas in [**?**] we blithely used AC without any comment whatsoever, here for a theorem whose statement requires AC we indicate that by calling it **AC-Theorem**. (If a theorem holds without AC, we sometimes still gives proofs which use AC, if they are easier or more enlightening.)

2.1. The fundamental relation $\leq$.

Let's look back at what we did in the last section. We introduced a notion of equivalence on sets: namely $S_1 \equiv S_2$ if there is a bijection $f : S_1 \to S_2$. This sets up a project of classifying sets up to equivalence. Looking at finite sets, we found that each equivalence class contained a representative of the form $[n]$ for a unique

[10]The day you ever "see" uncountably many things, let me know.

natural number n. Thus the set of equivalence classes of finite sets is $\mathbb{N}$. Then we considered whether all infinite sets were equivalent to each other, and found that they are not.

If we look back at finite sets (it is remarkable, and perhaps comforting, how much of the inspiration for some rather recondite-looking set-theoretic constructions comes from the case of finite sets) we can't help but notice that $\mathbb{N}$ has so much more structure than just a set. First, it is a semiring: this means that we have operations of $+$ and $\cdot$, but in general we do not have $-$ or $/$. Second it has a natural ordering $\leq$ which is indeed a *well-ordering*: that is, $\leq$ is a linear ordering on x in which every non-empty subset has a least element. (The well-ordering property is easily seen to be equivalent to the principle of mathematical induction.)

Remarkably, *all* of these structures generalize fruitfully to equivalence classes of sets! What does this mean? For a set S, let $|S|$ stand for its equivalence class. (This construction is commonplace in mathematics but has problematic aspects in set theory since the collection of sets equivalent with any nonempty set S does not form a set. Let us run with this notion for the moment, playing an important mathematician's trick: rather than worrying about what $|S|$ *is*, let us see how it *behaves*, and then later we can attempt to define it in terms of its behavior.)

We write $S_1 \leq S_2$ if there exists an injection $\iota : S_1 \hookrightarrow S_2$.

PROPOSITION 10.20. *Let S_1 be a nonempty set and S_2 a set. If there is an injection from S_1 to S_2, then there is a surjection from S_2 to S_1.*

PROOF. Let $\iota : S_1 \to S_2$ be an injection. We define $s : S_2 \to S_1$ as follows. Let $x_1 \in S_2$. If $y \in \iota(S_1)$, then since ι is injective there is exactly one $x \in S_1$ with $\iota(x) = y$, and we set $s(y) = x$. If $y \notin \iota(S_1)$, we set $s(y) = x_1$. This is a surjection. $\square$

THEOREM 10.21. *Let S_1 be a nonempty set and S_2 a set. If there is a surjection from S_2 to S_1, then there is an injection from S_1 to S_2.*

PROOF. Let $s : S_2 \to S_1$ be a surjection. We define $\iota : S_1 \to S_2$ as follows. For each $x \in S_1$, we **choose** $y \in S_2$ with $s(y) = x$ and define $\iota(x) = y$. If for $x_1, x_2 \in S_1$ we have $\iota(x_1) = \iota(x_2)$, then $x_1 = s(\iota(x_1)) = s(\iota(x_2)) = x_2$, so ι is an injection. $\square$

EXERCISE 10.5. *Suppose $S_1 = \varnothing$. Under what conditions on S_2 does Proposition 10.20 remain valid? What about Theorem 10.21?*

Let $\mathcal{F}$ be any family (i.e., set!) of sets S_α. Then our $\leq$ gives a relation on $\mathcal{F}$; what properties does it have? It is of course reflexive and transitive, which means it is (by definition) a *quasi-ordering*. On the other hand, it is generally not a partial ordering, because $S_{\alpha_1} \leq S_{\alpha_2}$ and $S_{\alpha_2} \leq S_{\alpha_1}$ does not in general imply that $S_{\alpha_1} = S_{\alpha_2}$: indeed, suppose have two distinct, but equivalent sets (say, two sets with three elements apiece). However, given a quasi-ordering we can formally associate a partial ordering, just by taking the quotient by the equivalence relation $x \leq y, y \leq x$. However, exactly how the associated partial ordering relates to the given quasi-ordering is in general unclear.

Therefore we can try to do something less drastic. Namely, let us write $|S_1| \leq |S_2|$

if $S_1 \leq S_2$. We must check that this is well-defined, but no problem: indeed, if $S_i \equiv T_i$ then choosing bijections $\beta_i : S_i \to T_i$, we get an injection

$$\beta_2 \circ \iota \circ \beta_1^{-1} : T_1 \to T_2.$$

Thus we can pass from the quasi-ordered set $(\mathcal{F}, \leq)$ to the quasi-ordered set of equivalence classes $(|\mathcal{F}, \leq)$. Since we removed an obvious obstruction to the quasi-ordering being a partial ordering, it is natural to wonder whether or not this partial ordering on equivalence classes is better behaved. If $\mathcal{F}$ is a family of finite sets, then $|\mathcal{F}|$ is a subset of $\mathbb{N}$ so we have a well-ordering. The following stunning result asserts that this remains true for infinite sets:

AC-Theorem 10.22. *(Second Fundamental Theorem of Set Theory) For any family $\mathcal{F}$ of sets, the relation $\leq$ descends to $|\mathcal{F}|$ and induces a well-ordering.*

In its full generality, Theorem 10.22 is best derived in the course of a systematic development of the theory of well-ordered sets, and we shall present this theory later on. However, the following special case can be proved now:

Theorem 10.23 (Schröder-Bernstein). *If $M \leq N$ and $N \leq M$, then $M \equiv N$.*

Proof. Certainly we may assume that M and N are disjoint. Let $f : M \hookrightarrow N$ and $g : N \hookrightarrow M$. Consider the following function B on $M \cup N$: if $x \in M$, $B(x) = f(x) \in N$; if $x \in N$, $B(x) = g(x) \in M$. Now we consider the B orbits on $M \cup N$. Put $B^m = B \circ \ldots \circ B$ (m times). There are three cases:
Case 1: The forward B-orbit of x is finite. Equivalently, for some m, $B^m(x) = x$. Then the backwards B-orbit is equal to the B-orbit, so the full B-orbit is finite.
Otherwise the B-orbit is infinite, and we consider the backwards B-orbit.
Case 2: The backwards B-orbit also continues indefinitely, so for all $m \in \mathbb{Z}$ we have pairwise disjoint elements $B^m(x) \in M \cup N$.
Case 3: For some $m \in \mathbb{Z}^+$, $B^{-m}(x)$ is not in the image of f or g.
As these possibilities are exhaustive, we get a partition of $M \cup N$ into three types of orbits: (i) finite orbits, (ii) $\{B^m \mid m \geq m_0\}$, and (iii) $\{B^m \mid m \in \mathbb{Z}\}$. We can use this information to define a bijection from M to N. Namely, f itself is necessarily a bijection from the Case 1 elements of M to the Case 1 elements of N, and the same holds for Case 3. f need not surject onto every Case 2 element of N, but the Case 2 element of $M \cup N$ have been partitioned into sets isomorphic to $\mathbb{Z}^+$, and pairing up the first element occurring in M with the first element occurring in N, and so forth, we have defined a bijection from M to N! $\square$

Theorem 10.22 asserts that $|S|$ is measuring, in a reasonable sense, the *size* of the set S: if two sets are inequivalent, it is because one of them is larger than the other. This motivates a small change of perspective: we will say that $|S|$ is the *cardinality* of the set S. Note well that we have not made any mathematical change: we have not defined cardinalities in an absolute sense – i.e., we have not said what sort of object $|\mathbb{N}|$ is – but only in a relational sense: i.e., as an invariant of a set that measures whether a set is bigger or smaller than another set.

Notation: For brevity we will write

$$\aleph_0 = |\mathbb{N}|$$

and

$$\mathfrak{c} = |\mathbb{R}|.$$

Here $\aleph$ is the Hebrew letter "aleph", and $\aleph_0$ is usually pronounced "aleph naught" or "aleph null" rather than "aleph zero". Exactly why we are choosing such a strange name for $|\mathbb{N}|$ will not be explained until the third handout. In contrast, we write $\mathfrak{c}$ for $|\mathbb{R}|$ simply because "c" stands for *continuum*, and in Handout 1 we said that a set S if **of continuum type** if $S \equiv \mathbb{R}$. In our new notation, [**?**, Fact 16] is reexpressed as

$$|2^{\aleph_0}| = \mathfrak{c}. \tag{26}$$

2.2. Addition of cardinalities.

For two sets S_1 and S_2, define the disjoint union $S_1 \coprod S_2$ to be $S_1' \cup S_2'$, where $S_i' = \{(s,1) \mid s \in S_i\}$. Note that there is an obvious bijection $S_i \to S_i'$; the point of this little artifice is that even if S_1 and S_2 are not disjoint, S_1' and S_2' will be.[11] Now consider the set $S_1 \coprod S_2$.

FACT 10.24. *The equivalence class $|S_1 \coprod S_2|$ depends only on the equivalence classes $|S_1|$ and $|S_2|$.*

PROOF. All this means is that if we have bijections $\beta_i : S_i \to T_i$, then there is a bijection from $S_1 \coprod S_2$ to $T_1 \coprod T_2$, which is clear: there is indeed a canonical bijection, namely $\beta_1 \coprod \beta_2$: by definition, this maps an element $(s,1)$ to $(\beta_1(s),1)$ and an element $(s,2)$ to $(\beta_2(s),2)$. $\square$

The upshot is that it makes formal sense to define $|S_1| + |S_2|$ as $|S_1 \coprod S_2|$: our addition operation on sets descends to equivalence classes. Note that on finite sets this amounts to

$$m + n = |[m]| + |[n]| = |[m] \coprod [n]| = |[m+n]| = m+n.$$

THEOREM 10.25. *Let $S \leq T$ be sets, with T infinite. Then $|S| + |T| = |T|$.*

There is a fairly quick and proof of Theorem 10.25, which however uses Zorn's Lemma (which is equivalent to the Axiom of Choice). At this stage in the development of the theory the reader might like to see such a proof, so we will present it now (certainly Zorn's Lemma is well known and used in "mainstream mathematics"). We begin with the following preliminary result which is of interest in its own right.

AC-THEOREM 10.26. *Any infinite set S is a disjoint union of countable subsets.*

PROOF. Consider the partially ordered set each of whose elements is a pairwise disjoint family of countable subsets of S, and with $\leq$ being set-theoretic inclusion. Any chain $\mathcal{F}_i$ in this poset has an upper bound: just take the union of all the elements in the chain: this is certainly a family of countable subsets of S, and if any element of $\mathcal{F}_i$ intersects any element of $\mathcal{F}_j$, then $\mathcal{F}_{\max(i,j)}$ contains both of these elements so is not a pairwise disjoint family, contradiction. By Zorn's Lemma we are entitled to a maximal such family $\mathcal{F}$. Then $S \setminus \bigcup_{i \in \mathcal{F}} S_i$ must be finite, so the remaining elements can be added to any one of the elements of the family. $\square$

AC-THEOREM 10.27. *For any infinite set A, there are disjoint subsets B and C with $A = B \cup C$ and $|A| = |B| = |C|$.*

[11] This in turn raises canonicity issues, which we will return to later.

Proof. Express $A = \bigcup_{i \in \mathcal{F}} A_i$, where each $A_i \cong \mathbb{Z}^+$. So partition S_i into $B_i \cup C_i$ where B_i and C_i are each countable, and take $B = \bigcup_{i \in \mathcal{F}} B_i$, $C = \bigcup_{i \in \mathcal{F}} C_i$. $\square$

Proof of Theorem 10.25: Let S and T be sets; by Theorem 10.22 we may assume $|S| \leq |T|$. Then clearly $|S| + |T| \leq |T| + |T|$, but the preceding result avers $|T| + |T| = |T|$. So $|S| + |T| \leq |T|$. Clearly $|T| \leq |S| + |T|$, so by the Schröder-Bernstein Theorem we conclude $|S| + |T| = |T|$.

Exercise 10.6. *Let T be an infinite set and S a nonempty subset of T. Show that S can be expressed as a disjoint union of subsets of cardinality $|T|$.*

AC-Theorem 10.28. *For all infinite sets S and T, $|S| + |T| = \max |S|, |T|$.*

Exercise 10.7. *Deduce Theorem 10.28 from Theorem 10.22 and Theorem 10.25.*

2.3. Subtraction of cardinalities.

It turns out that we cannot formally define a subtraction operation on infinite cardinalities, as one does for finite cardinalities using set-theoretic subtraction: given sets S_1 and S_2, to define $|S_1| - |S_2|$ we would like to find sets $T_i \equiv S_i$ such that $T_2 \subset T_1$, and then define $|S_1| - |S_2|$ to be $|T_1 \setminus T_2|$. Even for finite sets this only makes literal sense if $|S_2| \leq |S_1|$; in general, we are led to introduce negative numbers through a formal algebraic process, which we can recognize as the group completion of a monoid (or the ring completion of a commutative semiring).

However, here the analogy between infinite and finite breaks down: given $S_2 \subset S_1$, $T_2 \subset T_1$ and bijections $\beta_i : S_i \to T_i$, we absolutely do not in general have a bijection $S_1 \setminus S_2 \to T_1 \setminus T_2$. For instance, take $S_1 = T_1 = \mathbb{Z}^+$ and $S_2 = 2\mathbb{Z}^+$, the even numbers. Then $|S_1 \setminus S_2| = |\mathbb{N}|$. However, we could take $T_2 = \mathbb{Z}^+$ and then $T_2 \setminus T_1 = \emptyset$. For that matter, given any $n \in \mathbb{Z}^+$, taking T_2 to be $\mathbb{Z}^+ \setminus [n]$, we get $T_1 \setminus T_2 = [n]$. Thus when attempting to define $|\mathbb{N}| - |\mathbb{N}|$ we find that we get all conceivable answers, namely all equivalence classes of at most countable sets. This phenomenon does generalize:

Proposition 10.29. *(Subtraction theorem) For any sets $S_1 \subset S_2 \subset S_3$, there are bijections $\beta_1 : S_1 \to T_1$, $\beta_3 : S_3 \to T_3$ such that $T_1 \subset T_3$ and $|T_3 \setminus T_1| = |S_2|$.*

Proof. If S_1 and S_2 are disjoint, we may take $T_1 = S_1$, $T_2 = S_2$ and $T_3 = S_1 \cup S_2$. Otherwise we may adjust by bijections to make them disjoint. $\square$

2.4. Multiplication of cardinalities.

Let S_1 and S_2 be sets. We define

$$|S_1| \times |S_2| = |S_1 \times S_2|.$$

Exercise 10.8. *Check that this is well-defined.*

At this point, we have what appears to be a very rich structure on our cardinalities: suppose that $\mathcal{F}$ is a family of sets which is, up to bijection, closed under $\coprod$ and $\times$. Then the family $|\mathcal{F}|$ of cardinalities of these sets has the structure of a well-ordered semiring.

Example 10.5. *Take $\mathcal{F}$ to be any collection of finite sets containing, for all $n \in \mathbb{N}$, at least one set with n elements. Then $|\mathcal{F}| = \mathbb{N}$ and the semiring and (well)-ordering are the usual ones.*

EXAMPLE 10.6. *Take $\mathcal{F}$ to be a family containing finite sets of all cardinalities together with $\mathbb{N}$. Then, since $\mathbb{N}\coprod\mathbb{N} \cong \mathbb{N}$ and $\mathbb{N}\times\mathbb{N} \cong \mathbb{N}$, the corresponding family of cardinals $|\mathcal{F}|$ is a well-ordered semiring. It contains $\mathbb{N}$ as a subring and one other element, $|\mathbb{N}|$; in other words, as a set of cardinalities it is $\mathbb{N}\cup\{\mathbb{N}\}$, a slightly confusing-looking construction which we will see much more of later on. As a well-ordered set we have just taken $\mathbb{N}$ and added a single element (the element $\mathbb{N}$!) which is is larger than every other element. It is clear that this gives a well-ordered set; indeed, given any well-ordered set $(S,\leq)$ there is another well-ordered set, say $s(S)$, obtained by adding an additional element which is strictly larger than every other element (check and see that this gives a well-ordering). The semiring structure is, however, not very interesting: every $x \in \mathbb{N}\cup\{\mathbb{N}\}$, $x+\mathbb{N} = x\cdot\mathbb{N} = \mathbb{N}$. In particular, the ring completion of this semiring is the 0 ring. (It suffices to see this on the underlying commutative monoid. Recall that the group completion of a commutative monoid M can be represented by pairs (p,m) of elements of M with $(p,m)\sim(p',m')$ iff there exists some $x \in M$ such that $x+p+m' = x+p'+m$. In our case, taking $x = \mathbb{N}$ we see that all elements are equivalent.)*

However, like addition, multiplication of infinite cardinalities turns out not to be very interesting.

THEOREM 10.30. *Let T be infinite and S a nonempty subset of T. Then $|S|\times|T| = |T|$.*

The same remarks are in order here as for the addition theorem (Theorem 10.25): combining with cardinal trichotomy, we conclude that $|S|\times|T| = \max(|S|,|T|)$ for any infinite sets. This deduction uses the Axiom of Choice, whereas the theorem as stated does not. However, it is easier to give a proof using Zorn's Lemma, which is what we will do. Moreover, as for the additive case, it is convenient to first establish the case of $S = T$. Indeed, assuming that $T\times T\cong T$, we have

$$|S|\times|T| \leq |T|\times|T| = |T| \leq |S|\times|T|.$$

So let us prove that for any infinite set T, $T\times T\cong T$.

Consider the poset consisting of pairs (S_i,f_i), where $S_i\subset T$ and f_i is a bijection from S_i to $S_i\times S_i$. Again the order relation is the natural one: $(S_i,f_i)\leq(S_j,f_j)$ if $S_i\subset S_j$ and $f_j|_{S_i} = f_i$. Now we apply Zorn's Lemma, and, as is often the case, the verification that every chain has an upper bound is immediate because we can just take the union over all elements of the chain. Therefore we get a maximal element (S,f).

Now, as for the case of the addition theorem, we need not have $S = T$; put $S' = T\setminus S$. What we *can* say is that $|S'| < |S|$. Indeed, otherwise we have $|S'|\geq|S|$, so that inside S' there is a subset S'' with $|S''| = |S|$. But we can enlarge $S\times S$ to $(S\cup S'')\times(S\cup S'')$. The bijection $f : S\to S\times S$ gives us that

$$|S''| = |S| = |S|\times|S| = |S''|\times|S''|.$$

Thus using the addition theorem, there is a bijection $g : S\cup S''\to(S\cup S'')\times(S\cup S'')$ which can be chosen to extend $f : S\to S\times S$; this contradicts the maximality of (S,f).

Thus we have that $|S'| < |S|$ as claimed. But then we have $|T| = |S\cup S'| = \max(|S|,|S'|) = |S|$, so

$$|T|\times|T| = |S|\times|S| = |S| = |T|,$$

completing the proof.

Exercise 10.9. *Prove the analogue of Proposition 10.29 for cardinal division.*

Exercise 10.10. *Verify that* $+$ *and* $\cdot$ *are commutative and associative operations on cardinalities, and that multiplication distributes over addition. (There are two ways to do this. One is to use the fact that* $|S|+|T| = |S|\cdot|T| = \max(|S|,|T|)$ *unless* S *and* T *are both finite. On the other hand one can verify these identities directly in terms of identities on sets.)*

2.5. Cardinal Exponentiation.

For two sets S and T, we define S^T to be the set of all functions $f : T \to S$. Why do we write S^T instead of T^S? Because the cardinality of the set of all functions from $[m]$ to $[n]$ is n^m: for each of the m elements of the domain, we must select one of the n elements of the codomain. As above, this extends immediately to infinite cardinalities:

For any sets S and T, we put $|S|^{|T|} = |S^T|$.

Exercise 10.11. *Check that this is well-defined.*

Exercise 10.12. *Suppose* X *has at most one element. Compute* $|X|^{|Y|}$ *for any set* Y*.*

Henceforth we may as well assume that X has at least two elements.

Proposition 10.31. *For any sets* X*,* Y*,* Z *we have*

$$(|X|^{|Y|})^{|Z|} = |X|^{|Y|\cdot|Z|}.$$

Proof. By [?, Lemma 18] we have $(X^Y)^Z \equiv X^{YZ}$. The result follows immediately. □

Proposition 10.32. *For any sets* X*,* Y*,* Z*, we have*

$$|X|^{|Y|+|Z|} = |X|^{|Y|}\cdot|X|^{|Z|}$$

and

$$(|X|\cdot|Y|)^{|Z|} = |X|^{|Z|}\cdot|Y|^{|Z|}.$$

Exercise 10.13. *Prove Proposition 10.32.*

Theorem 10.33. *Let* X_1, X_2, Y_1, Y_2 *be sets with* $Y_1 \neq \varnothing$*. If* $|X_1| \leq |X_2|$ *and* $|Y_1| \leq |Y_2|$ *then* $|X_1|^{|Y_1|} \leq |X_2|^{|Y_2|}$*.*

Proof. Let $\iota_X : X_1 \to X_2$ be an injection. By Proposition 10.20, there is a surjection $s_Y : Y_2 \to Y_1$. There is an induced injection $X_1^{Y_1} \to X_2^{Y_1}$ given by

$$f : Y_1 \to X_1 \mapsto \iota_X \circ f : Y_1 \to X_2$$

and an induced injection $X_2^{Y_1} \to X_2^{Y_2}$ given by

$$f : Y_1 \to X_2 \mapsto f \circ s_Y : Y_2 \to X_2.$$

Composing these gives an injection from $X_1^{Y_1}$ to $X_2^{Y_2}$. □

If Y is finite, then $|X|^{|Y|} = |X|\cdot\ldots\cdot|X|$ so is nothing new. The next result evaluates, in a sense, $|X|^{|Y|}$ when $|Y| = \aleph_0$.

AC-Theorem 10.34. *Let S be a set with $|\{1,2\}| \leq |S| \leq \mathfrak{c}$. Then $|S|^{\aleph_0} = \mathfrak{c}$.*

Proof. There is an evident bijection from the set of functions $\mathbb{N} \to \{1,2\}$ to the power set $2^{\mathbb{N}}$, so $|\{1,2\}|^{\aleph_0} = |2^{\aleph_0}| = \mathfrak{c}$. Combining this with Theorem 10.33 and Proposition 10.32 we get

$$\mathfrak{c} = |\{1,2\}|^{\aleph_0} \leq |S|^{\aleph_0} \leq \mathfrak{c}^{\aleph_0} = (|\{1,2\}|^{\aleph_0})^{\aleph_0} = |\{1,2\}|^{\aleph_0 \times \aleph_0} = |\{1,2\}|^{\aleph_0} = \mathfrak{c}. \quad \square$$

What about $|X|^{|Y|}$ when Y is uncountable? By Cantor's Theorem we have

$$|X|^{|Y|} \geq |\{0,1\}|^{|Y|} = |2^Y| > |Y|.$$

Thus the first order of business seems to be the evaluation of $|2^Y|$ for uncountable Y. This turns out to be an extremely deep issue with a very surprising answer.

What might one expect $2^{|S|}$ to be? The most obvious guess seems to be the minimalist one: since any collection of cardinalities is well-ordered, for any cardinality κ, there exists a smallest cardinality which is greater than κ, traditionally called κ^+. Thus we might expect $2^{|S|} = |S|^+$ for all infinite S.

But comparing to finite sets we get a little nervous about our guess, since 2^n is very much larger than $n^+ = n + 1$. On the other hand, our simple formulas for addition and multiplication of infinite cardinalities do not hold for finite cardinalities either – in short, we have no real evidence so are simply guessing.

Notice that we did not even "compute" $|2^{\mathbb{N}}|$ in any absolute sense but only showed that it is equal to the cardinality $\mathfrak{c}$ of the real numbers. So already it makes sense to ask whether $\mathfrak{c}$ is the *least* cardinality greater than $\aleph_0$ or whether it is larger. The minimalist guess $\mathfrak{c} = \aleph_0^+$ was made by Cantor, who was famously unable to prove it, despite much effort: it is now called the **Continuum Hypothesis** (CH). Moreover, the guess that $2^S = |S|^+$ for all infinite sets is called the **Generalized Continuum Hypothesis** (GCH).

Will anyone argue if I describe the continuum hypothesis (and its generalization) as the most vexing problem in all of mathematics? Starting with Cantor himself, some of the greatest mathematical minds have been brought to bear on this problem. For instance, in his old age David Hilbert claimed to have proved CH and he even published his paper in *Crelle*, but the proof was flawed. Kurt Gödel proved in 1944 that CH is relatively consistent with the ZFC axioms for set theory – in other words, assuming that the ZFC axioms are consistent (if not, all statements in the language can be formally derived from them!), it is not possible to deduce CH as a formal consequence of these axioms. In 1963, Paul Cohen showed that the negation of CH is also relatively consistent with ZFC, and for this he received the Fields Medal. Cohen's work undoubtedly revolutionized set theory, and his methods ("forcing") have since become an essential tool. But where does this leave the status of the Continuum Hypothesis?

The situation is most typically summarized by saying that Gödel and Cohen showed the undecidability of CH – i.e., that it is neither true nor false in the same way that Euclid's parallel postulate is neither true nor false. However, to accept this as the end of the story is to accept that what we know about sets and set theory is exactly what the ZFC axiom scheme tells us, but of course this is a position that would require (philosophical as well as mathematical) justification – as well as a position that seems to be severely undermined by the very issue at hand! Thus, a more honest admission of the status of CH would be: we are not even sure whether or not the problem is open. From a suitably Platonistic mathematical perspective – i.e., a belief that what is true in mathematics is different from what we are able (in practice, or even in principle) to prove – one feels that either there exists some infinite subset of $\mathbb{R}$ which is equivalent to neither $\mathbb{Z}^+$ nor $\mathbb{R}$, or there doesn't, and the fact that none of the ZFC axioms allow us to decide this simply means that the ZFC axioms are not really adequate. It is worth noting that this position was advocated by both Gödel and Cohen.

In recent years this position has begun to shift from a philosophical to a mathematical one: the additional axioms that will decide CH one way or another are no longer hypothetical. The only trouble is that they are themselves very complicated, and "intuitive" mostly to the set theorists that invent them. Remarkably – considering that the Axiom of Choice and GCH are to some extent cognate (and indeed GCH implies AC) – the consensus among experts seems to be settling towards *rejecting* CH in mathematics. Among notable proponents, we mention the leading set theorist Hugh Woodin. His and other arguments are vastly beyond the scope of these notes.

To a certain extent, cardinal exponentation reduces to the problem of computing the cardinality of 2^S. Indeed, one can show the following result.

AC-Theorem 10.35. *If X has at least 2 elements and Y has at least one element,*

$$\max(|X|, |2^Y|) \leq |X|^{|Y|} \leq \max(|2^X|, |2^Y|).$$

We omit the proof for now.

2.6. Note on sources.

Most of the material of this installment is due to Cantor, with the exception of the Schröder-Bernstein theorem (although Cantor was able to deduce the Second Fundamental Theorem from the fact that every set can be well-ordered, which we now know to be equivalent to the Axiom of Choice). Our proofs of Theorems 10.25 and 10.30 follow Kaplansky's *Set Theory and Metric Spaces.* Gödel's views on the Continuum Problem are laid out with his typical (enviable) clarity in *What Is Cantor's Continuum Problem?* It is interesting to remark that this paper was first written before Cohen's work – although a 1983 reprint in Benacerraf and Putnam's *Philosophy of Mathematics* contains a short appendix acknowledging Cohen – but the viewpoint that it expresses (anti-formalist, and favoring the negation of CH) is perhaps *more* accepted today than it was at the time of its writing.

3. The Calculus of Ordinalities

3.1. Well-ordered sets and ordinalities.

The discussion of cardinalities in Chapter 2 suggests that the most interesting thing about them is their order relation, namely that any set of cardinalities forms a well-ordered set. So in this section we shall embark upon a systematic study of well-ordered sets. Remarkably, we will see that the problem of classifying sets up to bijection is literally contained in the problem of classifying well-ordered sets up to order-isomorphism.

EXERCISE 10.14. *Show that for a linearly ordered set X, the following are equivalent:*

(i) *The set X satisfies the descending chain condition: there are no infinite strictly descending sequences $x_1 > x_2 > \ldots$ in X.*
(ii) *The set X is well-ordered.*

We need the notion of "equivalence" of of well-ordered sets. A mapping $f : S \to T$ between partially ordered sets is **order preserving** (or an **order homomorphism**) if $s_1 \leq s_2$ in S implies $f(s_1) \leq f(s_2)$ in T.

EXERCISE 10.15. *Let $f : S \to T$ and $g : T \to U$ be order homomorphisms of partially ordered sets.*

a) *Show that $g \circ f : S \to U$ is an order homomorphism.*
b) *Note that the identity map from a partially ordered set to itself is an order homomorphism.*
*(It follows that there is a **category** whose objects are partially ordered sets and whose morphisms are order homomorphisms.)*

An **order isomorphism** between posets is a mapping f which is order preserving, bijective, and whose inverse f^{-1} is order preserving. (This is the general – i.e., categorical – definition of isomorphism of structures.)

EXERCISE 10.16. *Give an example of an order preserving bijection f such that f^{-1} is not order preserving.*

However:

LEMMA 10.36. *An order-preserving bijection whose domain is a totally ordered set is an order isomorphism.*

EXERCISE 10.17. *Prove Lemma 10.36.*

LEMMA 10.37 (Rigidity Lemma). *Let S and T be well-ordered sets and $f_1,\ f_2 : S \to T$ two order isomorphisms. Then $f_1 = f_2$.*

PROOF. Let f_1 and f_2 be two order isomorphisms between the well-ordered sets S and T, which we may certainly assume are nonempty. Consider S_2, the set of elements s of S such that $f_1(s) \neq f_2(s)$, and let $S_1 = S \setminus S_2$. Since the least element of S must get mapped to the least element of T by any order-preserving map, S_1 is nonempty; put $T_1 = f_1(S_1) = f_2(S_1)$. Supposing that S_2 is nonempty, let s_2 be its least element. Then $f_1(s_2)$ and $f_2(s_2)$ are both characterized by being the least element of $T \setminus T_1$, so they must be equal, a contradiction. □

EXERCISE 10.18. *Let S be a partially ordered set.*

a) *Show that the order isomorphisms $f : S \to S$ form a group, the **order automorphism group** $\mathrm{Aut}(S)$ of S. (The same holds for any object in any category.)*
b) *Notice that Lemma 10.37 implies that the automorphism group of a well-ordered set is the trivial group.*[12]

[12]One says that a structure is **rigid** if it has no nontrivial automorphisms.

c) *Suppose S is linearly ordered and f is an order automorphism of S such that for some positive integer n we have $f^n = 1_S$, the identity map. Show that $f = 1_S$. (Thus the automorphism group of a linearly ordered set is* ***torsionfree****.)*
d) *For any infinite cardinality κ, find a linearly ordered set S with $|\operatorname{Aut}(S)| \geq \kappa$. Can we always ensure equality?*
e) *(Harder!) Show: every group G is isomorphic to the automorphism group of some partially ordered set.*

Let us define an **ordinality** to be an order-isomorphism class of well-ordered sets, and write $o(X)$ for the order-isomorphism class of X. The intentionally graceless terminology will be cleaned up later on. Since two-order isomorphic sets are equipotent, we can associate to every ordinality α an "underlying" cardinality $|\alpha|$: $|o(X)| = |X|$. It is natural to expect that the classification of ordinalities will be somewhat richer than the classification of cardinalities – in general, endowing a set with extra structure leads to a richer classification – but the reader new to the subject may be (we hope, pleasantly) surprised at how much richer the theory becomes.

From the perspective of forming "isomorphism classes" (a notion the ontological details of which we have not found it profitable to investigate too closely) ordinalities have a distinct advantage over cardinalities: according to the Rigidity Lemma, any two representatives of the same ordinality are *uniquely* (hence canonically!) isomorphic. This in turn raises the hope that we can write down a *canonical* representative of each ordinality. This hope has indeed been realized, by von Neumann, as we shall see later on: the canonical representatives will be called "ordinals." While we are alluding to later developments, let us mention that just as we can associate a cardinality to each ordinality, we can also – and this is much more profound – associate an ordinality $o(\kappa)$ to each cardinality κ. This assignment is *one-to-one*, and this allows us to give a canonical representative to each cardinality, a "cardinal." At least at the current level of discussion, there is no purely mathematical advantage to the passage from cardinalities to cardinals, but it has a striking ontological[13] consequence, namely that, up to isomorphism, we may develop all of set theory in the context of "pure sets", i.e., sets whose elements (and whose elements' elements, and ...) are themselves sets.

But first let us give some basic examples of ordinalities and ways to construct new ordinalities from preexisting ones.

3.2. Algebra of ordinalities.

Example 10.7. *Trivially the empty set is well-ordered, as is any set of cardinality one. These sets, and only these sets, have unique well-orderings.*

Example 10.8. *Our "standard" example $[n]$ of the cardinality n comes with a well-ordering. Moreover, on a finite set, the concepts of well-ordering and linear ordering coincide, and it is clear that there is up to order isomorphism a unique linear ordering on $[n]$. Informally, given any two orderings on an n element set, we define an order-preserving bijection by pairing up the least elements, then the second-least elements, and so forth. (For a formal proof, use induction.)*

[13] I restrain myself from writing "ontological" (i.e., with quotation marks), being like most contemporary mathematicians alarmed by statements about the reality of mathematical objects.

EXAMPLE 10.9. *The usual ordering on* $\mathbb{N}$ *is a well-ordering. Notice that this is isomorphic to the ordering on* $\{n \in \mathbb{Z} \mid n \geq n_0\}$ *for any* $n_0 \in \mathbb{Z}$*. As is traditional, we write* ω *for the ordinality of* $\mathbb{N}$*.*

EXERCISE 10.19. *For any ordering* $\leq$ *on a set* X*, we have the opposite ordering* $\leq'$*, defined by* $x \leq' y$ *iff* $y \leq x$*.*

a) *If* $\leq$ *is a linear ordering, so is* $\leq'$*.*
b) *If both* $\leq$ *and* $\leq'$ *are well-orderings, then* X *is finite.*

For a partially ordered set X, we can define a new partially ordered set $X^+ := X \cup \{\infty\}$ by adjoining some new element ∞ and decreeing $x \leq \infty$ for all $x \in X$.

PROPOSITION 10.38. *If* X *is well-ordered, so is* X^+*.*

PROOF. Let Y be a nonempty subset of X^+. Certainly there is a least element if $|Y| = 1$; otherwise, Y contains an element other than ∞, so that $Y \cap X$ is nonempty, and its least element will be the least element of Y. $\square$

If X and Y are order-isomorphic, so too are X^+ and Y^+, so the passage from X to X^+ may be viewed as an operation on ordinalities. We denote $o(X^+)$ by $o(X)+1$, the **successor ordinality** of $o(X)$.

Note that all the finite ordinalities are formed from the empty ordinality 0 by iterated successorship. However, not every ordinality is of the form $o' + 1$, e.g. ω is clearly not: it lacks a maximal element. (On the other hand, it is obtained from *all* the finite ordinalities in a way that we will come back to shortly.) We will say that an ordinality o is a **successor ordinality** if it is of the form $o' + 1$ for some ordinality o' and a **limit ordinality** otherwise. Thus 0 and ω are limit ordinals.

EXAMPLE 10.10. *The successor operation allows us to construct from* ω *the new ordinals* $\omega+1$*,* $\omega+2 := (\omega+1)+1$*, and for all* $n \in \mathbb{Z}^+$*,* $\omega+n := (\omega+(n-1))+1$*: these are all distinct ordinals.*

For partially ordered sets $(S_1, \leq_1)$ and $(S_2, \leq_2)$, we define $S_1 + S_2$ to be the set $S_1 \coprod S_2$ with $s \leq t$ if either of the following holds:
(i) For $i = 1$ or 2, s and t are both in S_i and $s \leq_i t$;
(ii) $s \in S_1$ and $s \in S_2$.

Informally, we may think of $S_1 + S_2$ as "S_1 followed by S_2."

PROPOSITION 10.39. *If* S_1 *and* S_2 *are linearly ordered (resp. well-ordered), so is* $S_1 + S_2$*.*

EXERCISE 10.20. *Prove Proposition 10.39.*

Again the operation is well-defined on ordinalities, so we may speak of the **ordinal sum** $o + o'$. By taking $S_2 = [1]$, we recover the successor ordinality: $o + [1] = o + 1$.

EXAMPLE 10.11. *The ordinality* $2\omega := \omega + \omega$ *is the class of a well-ordered set which contains one copy of the natural numbers followed by another. Proceeding inductively, we have* $n\omega := (n-1)\omega + \omega$*, with a similar description.*

Tournant dangereux: We can also form the ordinal sum $1 + \omega$, which amounts to adjoining to the natural numbers a smallest element. But this is still order-isomorphic to the natural numbers: $1 + \omega = \omega$. In fact the identity $1 + o = o$ holds for every infinite ordinality (as will be clear later on). In particular $1 + \omega \neq \omega + 1$, so beware: the ordinal sum is not commutative! (To my knowledge it is the only non-commutative operation in all of mathematics which is invariably denoted by "+".) It is however immediately seen to be associative.

The notation 2ω suggests that we should have an ordinal product, and indeed we do:

For posets $(S_1, \leq_1)$ and $(S_2, \leq_2)$ we define the **lexicographic product** to be the Cartesian product $S_1 \times S_2$ endowed with the relation $(s_1, s_2) \leq (t_1, t_2)$ if(f) either $s_1 \leq t_1$ or $s_1 = t_1$ and $s_2 \leq t_2$. If the reasoning behind the nomenclature is unclear, I suggest you look up "lexicographic" in the dictionary.[14]

PROPOSITION 10.40. *If S_1 and S_2 are linearly ordered (resp. well-ordered), so is $S_1 \times S_2$.*

EXERCISE 10.21. *Prove Proposition 10.40.*

As usual this is well-defined on ordinalities so leads to the **ordinal product** $o \cdot o'$.

EXAMPLE 10.12. *For any well-ordered set X, $[2] \cdot X$ gives us one copy $\{(1, x) \mid x \in X\}$ followed by another copy $\{(2, x) \mid x \in X\}$, so we have a natural isomorphism of $[2] \cdot X$ with $X + X$ and hence $2 \cdot o = o + o$. (Similarly for $3o$ and so forth.) This time we should be prepared for the failure of commutativity: $\omega \cdot n$ is isomorphic to ω. This allows us to write down $\omega^2 := \omega \times \omega$, which we visualize by starting with the positive integers and then "blowing up" each positive integer to give a whole order isomorphic copy of the positive integers again. Repeating this operation gives $\omega^3 = \omega^2 \cdot \omega$, and so forth. Altogether this allows us to write down ordinalities of the form $P(\omega) = a_n \omega^n + \ldots + a_1 \omega + a_0$ with $a_i \in \mathbb{N}$, i.e., polynomials in ω with natural number coefficients. It is in fact the case that (i) distinct polynomials $P \neq Q \in \mathbb{N}[T]$ give rise to distinct ordinalities $P(\omega) \neq Q(\omega)$; and (ii) any ordinality formed from $[n]$ and ω by finitely many sums and products is equal to some $P(\omega)$ – even when we add/multiply in "the wrong order", e.g. $\omega * 7 * \omega^2 * 4 + \omega * 3 + 11 = \omega^3 + \omega + 11$ – but we will wait until we know more about the ordering of ordinalities to try to establish these facts.*

EXAMPLE 10.13. *Let $\alpha_1 = o(X_1), \ldots, \alpha_n = o(X_n)$ be ordinalities.*

a) *Show that $\alpha_1 \times (\alpha_2 \times \alpha_3)$ and $(\alpha_1 \times \alpha_2) \times \alpha_3$ are each order isomorphic to the set $X_1 \times X_2 \times X_3$ endowed with the ordering $(x_1, x_2, x_3) \leq (y_1, y_2, y_3)$ if $x_1 < y_1$ or ($x_1 = y_1$ and ($x_2 < y_2$ or ($x_2 = y_2$ and $x_3 \leq y_3$))). In particular ordinal multiplication is associative.*
b) *Give an explicit definition of the product well-ordering on $X_1 \times \ldots \times X_n$, another "lexicographic ordering."*

In fact, we also have a way to exponentiate ordinalities: let $\alpha = o(X)$ and $\beta = o(Y)$. Then by α^β we mean the order isomorphism class of the set $Z = Z(Y, X)$ of all functions $f : Y \to X$ with $f(y) = 0_X$ (0_X denotes the minimal element of X) for all but finitely many $y \in Y$, ordered by $f_1 \leq f_2$ if $f_1 = f_2$ or, for the greatest element $y \in Y$ such that $f_1(y) \neq f_2(y)$ we have $f_1(y) < f_2(y)$.

Some helpful terminology: one has the zero function, which is 0 for all values.

[14] Ha ha.

For every other $f \in W$, we define its **degree** $y_{\deg}$ to be the largest $y \in Y$ such that $f(y) \neq 0$ and its **leading coefficient** $x_l := f(y_{\deg})$.

PROPOSITION 10.41. *For ordinalities α and β, α^β is an ordinality.*

PROOF. Let Z be the set of finitely nonzero functions $f : Y \to X$ as above, and let $W \subset Z$ be a nonempty subset. We may assume 0 is not in W, since the zero function is the minimal element of all of Z. Thus the set of degrees of all elements of W is nonempty, and we may choose an element of minimal degree y_1; moreover, among all elements of minimal degree we may choose one with minimal leading coefficient x_1, say f_1. Suppose f_1 is not the minimal element of W, i.e., there exists $f' \in W_2$ with $f' < f_1$. Any such f' has the same degree and leading coefficient as f_1, so the last value y' at which f' and f_1 differ must be less than y_1. Since f_1 is nonzero at all such y' and f_1 is finitely nonzero, the set of all such y' is finite and thus has a *maximal* element y_2. Among all f' with $f'(y_2) < f(y_2)$ and $f'(y) = f(y)$ for all $y > y_2$, choose one with $x_2 = f'(y_2)$ minimal and call it f_2. If f_2 is not minimal, we may continue in this way, and indeed get a sequence of elements $f_1 > f_2 > f_3 \ldots$ as well as a descending chain $y_1 > y_2 > \ldots$. Since Y is well-ordered, this descending chain must terminate at some point, meaning that at some point we find a minimal element f_n of W. $\square$

EXAMPLE 10.14. *The ordinality ω^ω is the set of all finitely nonzero functions $f : \mathbb{N} \to \mathbb{N}$. At least formally, we can identify such functions as polynomials in ω with $\mathbb{N}$-coefficients: $P_f(\omega) = \sum_{n \in \mathbb{N}} f(n)\omega^n$. The well-ordering makes $P_f < P_g$ if the at the largest n for which $f(n) \neq g(n)$ we have $f(n) < g(n)$, e.g. $\omega^3 + 2\omega^2 + 1 > \omega^3 + \omega^2 + \omega + 100$.*

It is hard to ignore the following observation: ω^ω puts a natural well-ordering relation on all the ordinalities we had already defined. This makes us look back and see that the same seems to be the case for all ordinalities: e.g. ω itself is order isomorphic to the set of all the finite ordinalities $[n]$ with the obvious order relation between them. Now that we see the suggested order relation on the ordinalities of the form $P(\omega)$ one can check that this is the case for them as well: e.g. ω^2 can be realized as the set of all linear polynomials $\{a\omega + b \mid a, b \in \mathbb{N}\}$.

This suggests the following line of inquiry:

(i) Define a natural ordering on ordinalities (as we did for cardinalities).
(ii) Show that this ordering *well-orders* any set of ordinalities.

EXERCISE 10.22. *Let α and β be ordinalities.*

a) *Show: $0^\beta = 0$, $1^\beta = 1$, $\alpha^0 = 1$, $\alpha^1 = \alpha$.*
b) *Show: the correspondence between finite ordinals and natural numbers respects exponentiation.*
c) *For an ordinal α, the symbol α^n now has two possible meanings: exponentiation and iterated multiplication. Show that the two ordinalities are equal. (The proof requires you to surmount a small left-to-right lexicographic difficulty.) In particular $|\alpha^n| = |\alpha|^n = |\alpha|$.*
d) *For any infinite ordinal β, show that $|\alpha^\beta| = \max(|\alpha|, |\beta|)$.*

Tournant dangereux: In particular, it is generally *not* the case that $|\alpha^\beta| = |\alpha|^{|\beta|}$: e.g. 2^ω and ω^ω are both countable ordinalities. In fact, we have not yet seen any uncountable well-ordered sets, and one cannot construct an uncountable ordinal from ω by any finite iteration of the ordinal operations we have described (nor by a countable iteration either, although we have not yet made formal sense of that). This leads us to wonder: are there any uncountable ordinalities?

3.3. Ordering ordinalities. Let S_1 and S_2 be two well-ordered sets. In analogy with our operation $\leq$ on sets, it would seem natural to define $S_1 \leq S_2$ if there exists an order-preserving injection $S_1 \to S_2$. This gives a relation $\leq$ on ordinalities which is clearly symmetric and transitive.

However, this is *not* the most useful definition of $\leq$ for well-ordered sets, since it gives up the rigidity property. In particular, recall Dedekind's characterization of infinite sets as those which are in bijection with a proper subset of themselves, or, equivalently, those which *inject* into a proper subset of themselves. With the above definition, this will still occur for infinite ordinalities: for instance, we can inject ω properly into itself just by taking $\mathbb{N} \to \mathbb{N}$, $n \mapsto n+1$. Even if we require the least elements to be preserved, then we can still inject $\mathbb{N}$ into any infinite subset of itself containing 0.

So as a sort of mild *deus ex machina* we will work with a more sophisticated order relation. First, for a linearly ordered set S and $s \in S$, we define

$$I(s) = \{t \in S \mid t < s\},$$

an **initial segment** of S. Note that every initial segment is a proper subset. Let us also define

$$I[s] = \{t \in S \mid t \leq s\}.$$

Now, given linearly ordered sets S and T, we define $S < T$ if there exists an order-preserving embedding $f : S \to T$ such that $f(S)$ is an initial segment of T (say, an **initial embedding**). We define $S \leq T$ if $S < T$ or $S \cong T$.

Exercise 10.23. *Let $f : S_1 \to S_2$ and $g : T_1 \to T_2$ be order isomorphisms of linearly ordered sets.*

a) *Suppose $s \in S_1$. Show that $f(I(s)) = I(f(s))$ and $f(I[s])) = I(f[s])$.*
b) *Suppose that $S_1 < T_1$ (resp. $S_1 \leq T_1$). Show that $S_2 < T_2$) (resp. $S_2 \leq T_2$).*
c) *Deduce that $<$ and $\leq$ give well-defined relations on any set $\mathcal{F}$ of ordinalities.*

Exercise 10.24.
a) *Show that if $i : X \to Y$ and $j : Y \to Z$ are initial embeddings of linearly ordered sets, then $j \circ i : X \to Z$ is an initial embedding.*
b) *Deduce that the relation $<$ on any set of ordinalities is transitive.*

In a partially ordered set X, a subset Z is an **order ideal** if for all $z \in Z$ and $x \in X$, if $x < z$ then $x \in Z$. For example, the empty set and X itself are always order ideals. We say that X is an **improper** order ideal of itself, and all other order ideals are **proper**. For instance, $I[s]$ is an order ideal, which may or may not be an initial segment.

LEMMA 10.42 ("Principal ideal lemma"). *Any proper order ideal in a well-ordered set is an initial segment.*

Proof: Let Z be a proper order ideal in X, and s the least element of $X \setminus Z$. Then a moment's thought gives $Z = I(s)$.

The following is a key result:

THEOREM 10.43. *(Ordinal trichotomy) For any two ordinalities $\alpha = o(X)$ and $\beta = o(Y)$, exactly one of the following holds: $\alpha < \beta$, $\alpha = \beta$, $\beta < \alpha$.*

COROLLARY 10.44. *Any set of ordinalities is linearly ordered under $\leq$.*

EXERCISE 10.25. *Deduce Corollary 10.44 from Theorem 10.43. Is the Corollary equivalent to the Theorem?*

Proof of Theorem 10.43: Part of the assertion is that no well-ordered set X is order isomorphic to any initial segment $I(s)$ in X (we would then have both $o(I(s)) < o(X)$ and $o(I(s)) = o(X)$); let us establish this first. Suppose to the contrary that $\iota : X \to X$ is an order embedding whose image is an initial segment $I(s)$. Then the set of x for which $\iota(x) \neq x$ is nonempty (otherwise ι would be the identity map, and no linearly ordered set is *equal* to any of its initial segments), so let x be the least such element. Then, since ι restricted to $I(x)$ is the identity map, $\iota(I(x)) = I(x)$, so we cannot have $\iota(x) < x$ – that would contradict the injectivity of ι – so it must be the case that $\iota(x) > x$. But since $\iota(X)$ is an initial segment, this means that x is in the image of ι, which is seen to be impossible.

Now if $\alpha < \beta$ and $\beta < \alpha$ then we have initial embeddings $i : X \to Y$ and $j : Y \to X$. By Exercise 10.24 their composite $j \circ i : X \to X$ is an initial embedding, which we have just seen is impossible. It remains to show that if $\alpha \neq \beta$ there is either initial embedding from X to Y or vice versa. We may assume that X is nonempty. Let us try to build an initial embedding from X into Y. A little thought convinces us that we have no choices to make: suppose we have already defined an initial embedding f on a segment $I(s)$ of X. Then we *must* define $f(s)$ to be the least element of $Y \setminus f(I(s))$, and we *can* define it this way exactly when $f(I(s)) \neq Y$. If however $f(I(s)) = Y$, then we see that f^{-1} gives an initial embedding from Y to X. So assume Y is not isomorphic to an initial segment of X, and let Z be the set of x in X such that there exists an initial embedding from $I(z)$ to Y. It is immediate to see that Z is an order ideal, so by Lemma 10.42 we have either $Z = I(x)$ or $Z = X$. In the former case we have an initial embedding from $I(z)$ to Y, and as above, the only we could not extend it to x is if it is surjective, and then we are done as above. So we can extend the initial embedding to $I[x]$, which – again by Lemma 10.42 is either an initial segment (in which case we have a contradiction) or $I[x] = X$, in which case we are done. The last case is that $Z = X$ has no maximal element, but then we have $X = \bigcup_{x \in X} I(x)$ and a uniquely defined initial embedding ι on each $I(x)$. So altogether we have a map on all of X whose image $f(X)$, as a union of initial segments, is an order ideal. Applying Lemma 10.42 yet again, we either have $f(X) = Y$ – in which case f is an order isomorphism – or $f(X)$ is an initial segment of Y, in which case $X < Y$: done.

EXERCISE 10.26. *Let α and β be ordinalities. Show that if $|\alpha| > |\beta|$, then $\alpha > \beta$. (Of course the converse does not hold: there are many countable ordinalities.)*

Corollary 10.45. *Any set $\mathcal{F}$ of ordinalities is* well-ordered *with respect to $\leq$.*

Proof. It suffices to prove that there is no infinite descending chain in $\mathcal{F} = \{o_\alpha\}_{\alpha \in I}$. So, seeking a contradiction, suppose that we have a sequence of well-ordered sets S_1, $S_2 = I(s_1)$ for $s_1 \in S_1$, $S_3 = I(s_2)$,...,$S_{n+1} = I(s_n)$ for $s_n \in S_n$,.... But all the S_n's live inside S_1 and we have produced an infinite descending chain $s_1 > s_2 > s_3 > \ldots > s_n > \ldots$ inside the well-ordered set S_1, a contradiction. $\square$

Thus any set $\mathcal{F}$ of ordinalities itself generates an ordinality $o(\mathcal{F})$, the ordinality of the well-ordering that we have just defined on $\mathcal{F}$!

Now: for any ordinality o, it makes sense to consider the set $I(o)$ of ordinalities $\{o' \mid o' < o\}$: indeed, these are well-orderings on a set of cardinality at most the cardinality of o, so there are at most $2^{|o| \times |o|}$ such well-orderings. Similarly, define

$$I[o] = \{o' \mid o' \leq o\}.$$

Corollary 10.46. *$I(o)$ is order-isomorphic to o itself.*

Proof. We shall define an order-isomorphism $f : I(o) \to o$. Namely, each $o' \in I(o)$ is given by an initial segment $I(y)$ of o, so define $f(o') = y$. That this is an order isomorphism is essentially a tautology which we leave for the reader to unwind. $\square$

3.4. The Burali-Forti "Paradox". Do the ordinalities form a set? As we have so far managed to construct only countably many of them, it seems conceivable that they might. However, Burali-Forti famously observed that the assumption that there is a set of all ordinalities leads to a paradox. Namely, suppose $\mathbb{O}$ is a set whose elements are the ordinalities. Then by Corollary 10.45, $\mathbb{O}$ is itself well-ordered under our initial embedding relation $\leq$, so that the ordinality $o = o(\mathbb{O})$ would itself be a member of $\mathbb{O}$.

This is already curious: it is tantamount to saying that $\mathbb{O}$ is an element of itself, but notice that we are not necessarily committed to this: $(\mathbb{O}, \leq)$ is order isomorphic to one of its members, but maybe it is not *the same* set. (Anyway, is $o \in o$ paradoxical, or just strange?) Thankfully the paradox does not depend upon these ontological questions, but is rather the following: if $o \in \mathbb{O}$, then consider the initial segment $I(o)$ of $\mathbb{O}$: we have $\mathbb{O} \cong o \cong I(o)$, but this means that $\mathbb{O}$ is order-isomorphic to one of its initial segments, in contradiction to the Ordinal Trichotomy Theorem (Theorem 10.43).

Just as the proof of Cantor's *paradox* (i.e., that the cardinalities do not form a set) can be immediately adapted to yield a profound and useful *theorem* – if S is a set, there is no surjection $S \to 2^S$, so that $2^{|S|} > |S|$ – in turn the proof of the Burali-Forti paradox immediately gives the following result, which we have so far been unable to establish:

Theorem 10.47. *(Burali-Forti's Theorem) For any cardinal κ, the set $\mathcal{O}_\kappa$ of ordinalities o with $|o| \leq \kappa$ has cardinality greater than κ.*

Proof. Indeed, $\mathcal{O}_\kappa$ is, like any set of ordinalities, well-ordered under our relation $\leq$, so if it had cardinality at most κ it would contain its own ordinal isomorphism class o as a member and hence be isomorphic to its initial segment $I(o)$ as above. $\square$

It follows that there are uncountable ordinalities. There is therefore a *least* uncountable ordinality, traditionally denoted ω_1. This least uncountable ordinality is a truly remarkable mathematical object: mere contemplation of it is fascinating and a little dizzying. For instance, the minimality property implies that all of its initial segments are countable, so it is not only very large as a set, but it is extremely difficult to traverse: for any point $x \in \omega_1$, the set of elements less than x is countable whereas the set of elements greater than x is uncountable! (This makes Zeno's Paradox look like kid stuff.) In particular it has no largest element so is a limit ordinal.[15]

On the other hand its successor ω_1^+ is also of interest.

Exercise 10.27. *Let S be a totally ordered set, endowed with the order topology.*

a) *Show that the order topology on an ordinal o is discrete iff $o \leq \omega$. What is the order topology on $\omega + 1$? On 2ω?*
b) *Show that order topologies are Hausdorff.*
c) *Show that an ordinality is compact iff it is a successor ordinality. In particular $I[o]$ is the one-point compactification of $I(o) \cong o$; deduce that the order topology on an ordinality is Tychonoff.*
d) *(Harder) Show that, in fact, any linearly ordered space is normal, and moreover all subspaces are normal.*
e) *A subset Y of a linearly ordered set X can be endowed with two topologies: the subspace topology, and the order topology for the ordering on X restricted to Y. Show that the subspace topology is always finer than the order topology; by contemplating $X = \mathbb{R}$, $Y = \{-1\} \cup \{\frac{1}{n}\}_{n\in\mathbb{Z}^+}$ show that the two topologies need not coincide.*
f) *Show that it may happen that a subspace of a linearly ordered space need not be a linearly ordered space (i.e., there may be* no *ordering inducing the subspace topology). Suggestion: take $X = \mathbb{R}$, $Y = \{-1\} \cup (0, 1)$. One therefore has the notion of a* ***generalized order space****, which is a space homeomorphic to a subspace of a linearly ordered space. Show that no real manifold of dimension greater than one is a generalized order space.*
g) *Let X be a well-ordered set and Y a nonempty subset. Show that the embedding $Y \to X$ may be viewed as a* net *on X, indexed by the (nonempty well-ordered, hence directed) set Y. Show that for any ordinality o the net $I(o)$ in $I[o]$ converges to o.*

Exercise 10.28. *Let $\mathcal{F}$ be a set of ordinalities. As we have seen, $\mathcal{F}$ is well-ordered under our initial embedding relation $<$ so gives rise to an ordinality $o(\mathcal{F})$. In fact there is another way to attach an ordinality to $\mathcal{F}$.*

a) *Show that there is a least ordinality s such that $\alpha \leq s$ for all $\alpha \in \mathcal{F}$. (Write $\alpha = o(X_\alpha)$, apply the Burali-Forti theorem to $|2^{\coprod_{\alpha\in\mathcal{F}} X_\alpha}|$, and use Exercise 10.26.) We call this s the* ***ordinal supremum*** *of the ordinalities in $\mathcal{F}$.*
b) *Show that an ordinality is a limit ordinality iff it is the supremum of all smaller ordinalities.*

[15]In fact this only begins to express ω_1's "inaccessibility from the left"; the correct concept, that of **cofinality**, will be discussed later.

c) *Recall that a subset T of a partially ordered set S is* ***cofinal*** *if for all $s \in S$ there exists $t \in T$ such that $s \leq t$. Let α be a limit ordinality, and $\mathcal{F}$ a subset of $I(\alpha)$. Show that $\mathcal{F}$ is cofinal iff $\alpha = \sup \mathcal{F}$.*

d) *For any ordinality α, we define the* ***cofinality*** *$\mathrm{cf}(\alpha)$ to be the minimal ordinality of a cofinal subset $\mathcal{F}$ of $I(\alpha)$. E.g., an ordinality is a successor ordinality iff it has cofinality 1. Show that $\mathrm{cf}(\omega) = \omega$ and $\mathrm{cf}(\omega_1) = \mathrm{cf}(\omega_1)$. What is $\mathrm{cf}(\omega^2)$?*

e) *(Harder) An ordinality is said to be* ***regular*** *if it is equal to its own cofinality. Show that for every cardinality κ, there exists a regular ordinality o with $|o| > \kappa$.*

f) *(For D. Lorenzini) For a cardinality κ, let o be a regular ordinality with $|o| > \kappa$. Show that any linearly ordered subset of cardinality at most κ has an upper bound in o, but $I(\kappa)$ does not have a maximal element.*[16]

3.5. Von Neumann ordinals.

Here we wish to report on an idea of von Neumann, which uses the relation $I(o) \cong o$ to define a canonical well-ordered set with any given ordinality. The construction is often informally defined as follows: "we inductively define o to be the set of all ordinals less than o." Unfortunately this definition is circular, and not for reasons relating to the induction process: step back and see that it is circular in the most obvious sense of using the quantity it purports to define!

However, it is quite corrigible: rather than building ordinals out of nothing, we consider the construction as taking as input a well-ordered set S and returning an order-isomorphic well-ordered set $vo(S)$, the **von Neumann ordinal** of S. The only property that we wish it to have is the following: if S and T are order-isomorphic sets, we want $vo(S)$ and $vo(T)$ to be not just order-isomorphic but *equal*. Let us be a bit formal and write down some axioms:

(VN1) For all well-ordered sets S, we have $vo(S) \cong S$.
(VN2) For well-ordered S and T, $S \cong T \implies vo(S) = vo(T)$.

Consider the following two additional axioms:

(VN3) $vo(\emptyset) = \emptyset$.
(VN4) For $S \neq \emptyset$, $vo(S) = \{vo(S') \mid S' < S\}$.

The third axiom is more than reasonable: it is forced upon us, by the fact that there is a unique empty well-ordered set. The fourth axiom is just expressing the order-isomorphism $I(o) \cong o$ in terms of von Neumann ordinals. Now the point is that these axioms determine all the von Neumann ordinals:

THEOREM 10.48. *(von Neumann) There is a unique correspondence $S \mapsto vo(S)$ satisfying (VN1) and (VN2).*

Before proving this theorem, let's play around with the axioms by discussing their consequences for finite ordinals. We know that $vo(\emptyset) = \emptyset = [0]$. What is $vo([1])$?

[16]This shows that one must allow chains of arbitrary cardinalities, and not simply ascending sequences, in order for Zorn's Lemma to hold.

Well, it is supposed to be the set of von Neumann ordinals strictly less than it. There is in all of creation exactly one well-ordered set which is strictly less than [1]: it is $\emptyset$. So the axioms imply

$$vo([1]) = \{\emptyset\}.$$

How about $vo([2])$? The axioms easily yield:

$$vo([2]) = \{vo[0], vo[1]\} = \{\emptyset, \{\emptyset\}\}.$$

Similarly, for any finite number n, the axioms give:

$$v0([n]) = \{vo[0], vo[1], \ldots, vo[n-1]\},$$

or in other words,

$$vo([n]) = \{vo[n-1], \{vo[n-1]\}\}.$$

More interestingly, the axioms tell us that the von Neumann ordinal ω is precisely the set of all the von Neumann numbers attached to the natural numbers. And we can track this construction "by hand" up through the von Neumann ordinals of 2ω, ω^2, ω^ω and so forth. But how do we know the construction works (i.e., gives a unique answer) for every ordinality?

The answer is simple: by induction. We have seen that the axioms imply that at least for sufficiently small ordinalities there is a unique assignment $S \mapsto vo(S)$. If the construction does not always work, there will be a smallest ordinality o for which it fails. But this cannot be, since it is clear how to define $vo(o)$ given definitions of all von Neumann ordinals of ordinalities less than o: indeed, (VN4) tells us exactly how to do this.

This construction is an instance of **transfinite induction**.

EXERCISE 10.29. *Show: for any well-ordered set S, we have $vo(S^+) = \{vo(S), \{vo(S)\}\}$.*

This is not a foundationalist treatment of von Neumann ordinals. It would also be possible to define a von Neumann ordinal as a certain type of set, using the following exercise.

EXERCISE 10.30. *Show that a set S is a von Neumann ordinal iff both of the following hold:*

(i) *if $x \in S$ implies $x \subset S$;*
(ii) *the relation $\subset$ is a well-ordering on elements of S.*

For the rest of these notes we will drop the term "ordinality" in favor of "ordinal." The reader who wants an ordinal to be something in particular can thus take it to be a von Neumann ordinal. This convention has to my knowledge no real mathematical advantage, but it has some very convenient notational consequences, as for instance the following definition of "cardinal."

3.6. A definition of cardinals. Here we allow ourselves the following result, which we will discuss in more detail later on.

THEOREM 10.49. *(Well-ordering theorem) Assuming the Axiom of Choice, every set S can be well-ordered.*

We can use this theorem ("theorem"?) to reduce the theory of cardinalities to a special case of the theory of ordinalities, and thus, we can give a concrete definition of cardinal numbers in terms of Von Neumann's ordinal numbers.

Namely, for any set S, we define its cardinal $|S|$ to be the smallest von Neumann ordinal o such that o is equivalent to (i.e., in bijection with) S.

In particular, we find that the finite cardinals and the finite ordinals are the same: we have changed our standard n element set from $[1, n]$ to the von Neumann ordinal n, so for instance $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. On purely mathematical grounds, this is not very exciting. However, if you like, we can replace our previous attitude to what the set $[n] = \{1, \ldots, n\}$ "really is" (which was, essentially, "Why are you bothering me with such silly questions?") by saying that, in case anyone asks (we may still hope that they do not ask), we identify the non-negative integer n with its von Neumann ordinal. Again, this is not to say that we have discovered what 3 really is. Rather, we noticed that a set with three elements exists in the context of **pure set theory**, i.e., we do not have to know that there exist 3 objects in some box somewhere that we are basing our definition of 3 on (like the definition of a meter used to be based upon an actual meter stick kept by the Bureau of Standards). In truth 3 is not a very problematic number, but consider instead $n = 10^{10^{10^{10}}}$; the fact that n is (perhaps) greater than the number of distinct particles in the universe is, in our account, no obstacle to the existence of sets with n elements.

Let's not overstate the significance of this for finite sets: with anything like a mainstream opinion on mathematical objects[17] this is completely obvious: we could also have defined 0 as $\emptyset$ and n as $\{n-1\}$, or in infinitely many other ways. It becomes more interesting for infinite sets, though.

That is, we can construct a theory of sets without *individuals* – in which we never have to say what we mean by an "object" as an element of a set, because the only elements of a set are other sets, which ultimately, when broken up enough (but possibly infinitely many) times, are lots and lots of braces around the empty set. This is nice to know, most of all because it means that in practice we don't have to worry one bit about what the elements of are sets are: we can take them to be whatever we want, because each set is equivalent (bijective) to a *pure set*. If you would like (as I would) to take a primarily Bourbakistic view of mathematical structure – i.e., that the component parts of any mathematical object are of no importance whatsoever, and that mathematical objects matter only as they relate to each other – then this is very comforting.

Coming back to the mathematics, we see then that any set of cardinals is in particular a set of ordinals, and the notion of $<$ on cardinals induced in this way is the same as the one we defined before. That is, if α and β are von Neumann cardinals, then $\alpha < \beta$ holds in the sense of ordinals iff there exists an injection from α to β but not an injection from β to α.

EXERCISE 10.31. *Convince yourself that this is true.*

[17]The only contemporary mathematician I know who would have problems with this is Doron Zeilberger.

Thus we have now, at last, proved the Second Fundamental Theorem of Set Theory, modulo our discussion of Theorem 10.49.

3.7. Introducing the Axiom of Choice.

Now we come clean. Many of the results of Chapter II rely on the following "fact":

FACT 10.50. *(Axiom of Choice (AC)): For any nonempty family I of nonempty sets S_i, the product $\prod_{i \in I} S_i$ is nonempty.*

Remark: In other words, any product of nonzero cardinalities is itself nonzero. This is the version of the axiom of choice favored by Bertrand Russell, who called it the "multiplicative axiom." Aesthetically speaking, I like it as well, because it seems so simple and self-evident.

EXERCISE 10.32. *Show that if (AC) holds for all families of pairwise disjoint sets S_i, it holds for all nonempty families of nonempty sets.*

However, in applications it is often more convenient to use the following reformulation of (AC) which spells out the connection with "choice".

(AC′): If S is a set and $I = \{S_i\}$ is a nonempty family of nonempty subsets of S, then there exists a **choice function**, i.e., a function $f : I \to S$ such that for all $i \in I$, $f(S_i) \in S_i$.

Let us verify the equivalence of (AC) and (AC′).

(AC) $\implies$ (AC′): By (AC), $\mathcal{S} = \prod_{i \in I} S_i$ is nonempty, and an element f of $\mathcal{S}$ is precisely an assignment to each $i \in I$ of an element $f(i) \in S_i \subset S$. Thus f determines a choice function $f : I \to S$.

(AC′) $\implies$ (AC): Let $I = \{S_i\}$ be a nonempty family of nonempty sets. Put $S = \bigcup_{i \in I} S_i$. Let $f : I \to S$ be a choice function: for all $i \in I$, $f(S_i) \in S_i$. Thus $\{f(i)\}_{i \in I} \in \prod_{i \in I} S_i$.

The issue here is that if I is infinite we are making infinitely many choices – possibly with no coherence or defining rule to them – so that to give a choice function f is in general to give an infinite amount of information. Have any of us in our daily lives ever made infinitely many independent choices? Probably not. So the worry that making such a collection of choices is not possible is not absurd and should be taken with some seriousness.

Thus the nomenclature *Axiom* of Choice: we are, in fact, asserting some feeling about how infinite sets behave, i.e., we are doing exactly the sort of thing we had earlier averred to try to avoid. However, in favor of assuming AC, we can say: (i) it is a fairly basic and reasonable axiom – if we accept it we do not, e.g., feel the need to justify it in terms of something simpler; and (ii) we are committed to it, because most of the results we presented in Chapter II would not be true without it, nor would a great deal of the results of mainstream mathematics.

Every student of mathematics should be aware of some of the "facts" that are equivalent to AC. The most important two are as follows:

FACT 10.51. *(Zorn's Lemma) Let S be a partially ordered set. Suppose that every chain C – i.e., a totally ordered subset of S – has an upper bound in S. Then S has a maximal element.*

THEOREM 10.52. *The axiom of choice (AC), Zorn's Lemma (ZL), and the Well-Ordering Theorem (WOT) are all equivalent to each other.*

Remark: The fact that we are asserting the logical equivalence of an axiom, a lemma and a theorem is an amusing historical accident: according to the theorem they are all on the same logical footing.

WOT $\implies$ AC: It is enough to show WOT $\implies$ AC$'$, which is easy: let $\{S_i\}_{i\in I}$ be a nonempty family of nonempty subsets of a set S. Well-order S. Then we may define a choice function $f : I \to S$ by mapping i to the least element of S_i.

AC $\implies$ ZL: Strangely enough, this proof will use transfinite induction (so that one might initially think WOT would be involved, but this is absolutely not the case). Namely, suppose that S is a poset in which each chain C contains an upper bound, but there is no maximal element. Then we can define, for every ordinal o, a subset $C_0 \subset S$ order-isomorphic to o, in such a way that if $o' < o$, $C_{o'} \subset C_o$. Indeed we define $C_\emptyset = \emptyset$, of course. Assume that for all $o' < o$ we have defined $C_{o'}$. If o is a limit ordinal then we define $C_o := \bigcup_{o'<o} C_{o'}$. Then necessarily C_0 is order-isomorphic to o: that's how limit ordinals work. If $o = o' + 1$, then we have $C_{o'}$ which is assumed not to be maximal, so we choose an element x of $S \setminus C_{o'}$ and define $x_o := x$. Thus we have inside of S well-ordered sets of all possible order-isomorphism types. This is clearly absurd: the collection $o(|S|)$ of ordinals of cardinality $|S|$ is an ordinal of cardinality greater than the cardinality of S, and $o(|S|) \hookrightarrow S$ is impossible.

But where did we use AC? Well, we definitely made some choices, one for each non-successor ordinal. To really nail things down we should cast our choices in the framework of a choice function. Suppose we choose, for each well-ordered subset W of X, an element $x_W \in X \setminus W$ which is an upper bound for W. (This is easily phrased in terms of a choice function.) We might worry for a second that in the above construction there was some compatibility condition imposed on our choices, but this is not in fact the case: at stage o, any upper bound x for C_o in $S \setminus C_o$ will do to give us $C_{o+1} := C_o \cup \{x\}$. This completes the proof.

Remark: Note that we showed something (apparently) slightly stronger: namely, that if every well-ordered subset of a poset S has an upper bound in S, then S has a maximal element. This is mildly interesting but apparently useless in practice.

ZL $\implies$ WOT: Let X be a non-empty set, and let $\mathcal{A}$ be the collection of pairs $(A, \leq)$ where $A \subset X$ and $\leq$ is a well-ordering on A. We define a relation $<$ on $\mathcal{A}$: $x < y$ iff x is equal to an initial segment of y. It is immediate that $<$ is a strict partial ordering on $\mathcal{A}$. Now for each chain $C \subset \mathcal{A}$, we can define x_C to be the union of the elements of C, with the induced relation. x_C is itself well-ordered with the induced relation: indeed, suppose Y is a nonempty subset of x_C which is not well-ordered. Then Y contains an infinite descending chain $p_1 > p_2 > \ldots > p_n > \ldots$. But taking an element $y \in C$ such that $p_1 \in y$, this chain lives entirely inside y (since otherwise $p_n \in y'$ for $y' > y$ and then y is an initial segment of y', so $p_n \in y'$,

$p_n < p_1$ implies $p_n \in y$), a contradiction.

Thus, applying Zorn's Lemma, we are entitled to a maximal element $(M, \leq_M)$ of $\mathcal{A}$. It remains to see that $M = X$. If not, take $x \in X \setminus M$; adjoining x to $(M, \leq_M)$ as the maximum element we get a strictly larger well-ordering, a contradiction.

Remark: In the proof of AC $\implies$ ZL we made good advantage of our theory of ordinal arithmetic. It is possible to prove this implication (or even the direct implication AC $\implies$ ZL) directly, but this essentially requires proving some of our lemmata on well-ordered sets on the fly.

3.8. The Teichmüller-Tukey Lemma. Let X be a set. A family $\mathcal{P}$ of subsets of X is said to be **of finite character** if $\varnothing \in \mathcal{P}$ and for all $A \subset X$ we have that $A \in \mathcal{P}$ iff every finite subset $B \subset A$ is an element of $\mathcal{P}$.

LEMMA 10.53 (Teichmüller-Tukey). *If $\mathcal{P} \subset 2^X$ has finite character, then every $A \in \mathcal{P}$ is contained in a maximal element $B \in \mathcal{P}$: i.e., B is not properly contained in any element of $\mathcal{P}$.*

PROOF. Let $\mathcal{P}_A$ be the set of $\{B \in \mathcal{P} \mid B \supset A\}$, partially ordered under inclusion. If we can show that every chain in $\mathcal{P}_A$ has an upper bound, then by Zorn's Lemma the set $\mathcal{P}_A$ has a maximal element B, which is the desired conclusion.

Let $\mathcal{C}$ be a chain in $\mathcal{P}_A$. We claim that $\bigcup \mathcal{C} \in \mathcal{P}_A$: if not, there is a finite subset $C \subset \bigcup \mathcal{C}$ such that $C \notin \mathcal{P}_A$, but since $\mathcal{C}$ is a chain, the finite set C is a subset of some element of $\mathcal{C}$, contradicting the fact that $\mathcal{P}$ has finite character. Thus $\bigcup \mathcal{C}$ is an upper bound for $\mathcal{P}_A$ in $\mathcal{P}_A$. $\square$

Conversely:

LEMMA 10.54. *The Teichmüller-Tukey Lemma implies Zorn's Lemma.*

PROOF. Let $(X, \leq)$ be a partially ordered set in which each chain has an upper bound. Let $\mathcal{C} \subset 2^X$ be the set of all chains in X. Since the empty set is a chain and a subset of a partially ordered set is a chain iff all of its finite subsets is a chain (iff all of its 2 element subsets is a chain), the family $\mathcal{C}$ is of finite character, so by Teichmüller-Tukey, there is a maximal element $C \in \mathcal{C}$. By hypothesis, C has an upper bound x in X. If $x \notin C$, then $C \cup \{x\}$ would be a strictly larger chain in X, contradicting the maximality of C, so $x \in C$. Similarly, the element $x \in C$ must be a maximal element of X, because otherwise there would be $x < y$ in X and then $C \cup \{y\}$ would be a strictly larger chain. $\square$

It follows from Lemmas 10.53 and 10.54 that the Teichmüller-Tukey Lemma is equivalent to Zorn's Lemma and thus also to the Axiom of Choice.

3.9. Some equivalents and consequences of the Axiom of Choice. Although disbelieving AC is a tenable position, mainstream mathematics makes this position slightly unpleasant, because Zorn's Lemma is used to prove many quite basic results. One can ask which of these uses are "essential." The strongest possible case is if the result we prove using ZL can itself be shown to imply ZL or AC. Here are some samples of these results:

FACT 10.55. *For any infinite set A, we have $|A| = |A \times A|$.*

FACT 10.56. *For sets A and B, there is an injection $A \hookrightarrow B$ or an injection $B \hookrightarrow A$.*

FACT 10.57. *Every surjective map of sets has a section.*

FACT 10.58. *For any field k, every k-vector space V has a basis.*

FACT 10.59. *Every proper ideal in a commutative ring is contained in a maximal proper ideal.*

FACT 10.60. *The product of any number of compact spaces is itself compact.*

Even more commonly one finds that one can make a proof work using Zorn's Lemma but it is not clear how to make it work without it. In other words, many statements seem to require AC even if they are not equivalent to it. As a simple example, try to give an explicit well-ordering of $\mathbb{R}$. Did you succeed? In a precise formal sense this is impossible. But it is intuitively clear (and also true!) that being able to well-order a set S of any given infinite cardinality is not going to tell us that we can well-order sets of all cardinalities (and in particular, how to well-order 2^S), so the existence of a well-ordering of the continuum is not equivalent to AC.

Formally, speaking one says that a statement *requires* AC if one cannot prove that statement in the Zermelo-Fraenkel axiomation of set theory (ZF) which excludes AC. (The Zermelo-Fraenkel axiomatization of set theory including the axiom of choice is abbreviated ZFC; ZFC is essentially the "standard model" for sets.) If on the other hand a statement requires AC in this sense but one cannot deduce AC from ZF and this statement, we will say that the statement *merely requires* AC. There are lots of statements that merely require AC:[18]

THEOREM 10.61. *The following facts merely require AC:*

a) *The countable union of countable sets is countable.*
b) *An infinite set is Dedekind infinite.*
c) *There exists a non(-Lebesgue-)measurable subset of* $\mathbb{R}$.
d) *The Banach-Tarski paradox.*
e) *Every field has an algebraic closure.*
f) *Every field extension has a relative transcendence basis.*
g) *Every Boolean algebra has a prime ideal (BPIT).*
h) *Every Boolean algebra is isomorphic to a Boolean algebra of sets (Stone representation theorem).*
i) *Every subgroup of a free group is free.*
j) *The Hahn-Banach theorem (on extension of linear functionals), the open mapping theorem, the closed graph theorem, the Banach-Alaoglu theorem.*
k) *Baire's theorem.*
l) *The existence of a Stone-Cech compactification of every Tychonoff space.*

Needless to say the web of implications among all these important theorems is a much more complicated picture; for instance, it turns out that the BPIT is an interesting intermediate point (e.g. Tychonoff's theorem for Hausdorff spaces is equivalent to BPIT). Much contemporary mathematics is involved in working out the various dependencies.

[18]This list was compiled with the help of the Wikipedia page on the Axiom of Choice.

Bibliography

[A] K.A. Arrow, *Social Choice and Individual Values.* Cowles Commission Monograph No. 12. John Wiley & Sons, Inc., New York, N. Y.; Chapman & Hall, Ltd., London, 1951.

[Az18] Y. Azouzi, *Cluster Points of Slow Sequences.* Amer. Math. Monthly 125 (2018), 457–461.

[Ba22] S. Banach, *Sur les opérations dans les ensembles abstraits et leur applications aux équations intégrales.* Fund. Math. 3 (1922), 133–181.

[Ba55] R.G. Bartle, *Nets and Filters in Topology*, Amer. Math. Monthly 62 (1955), 551–557.

[Ba55er] R.G. Bartle, *A Correction for "Nets and Filters in Topology"*, Am. Math. Monthly, Vol. 70, No. 1. (Jan., 1963), 52-53.

[Ba93] Y.M. Baryshnikov, *Unifying impossibility theorems: a topological approach.* Adv. in Appl. Math. 14 (1993), 404–415.

[BC10] D. Burton and J. Coleman, *Quasi-Cauchy sequences.* Amer. Math. Monthly 117 (2010), 328–333.

[Be12] S.J. Bernstein, *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités.* Communications of the Kharkov Mathematical Society 13 (1912), 1–2.

[Be67] L. Bers, *Classroom Notes: On Avoiding the Mean Value Theorem.* Amer. Math. Monthly 74 (1967), 583.

[Be82] H. Bereková, *The principle of induction in continuum and related methods.* Acta Math. Univ. Comenian. 40(41) (1982), 97–100.

[Be98] Y. Benyamini, *Applications of the universal surjectivity of the Cantor set.* Amer. Math. Monthly 105 (1998), 832–839.

[Be06] A.F. Beardon, *Contractions of the Real Line.* Amer. Math. Monthly 113 (2006), 557–558.

[Bi37] G. Birkhoff, *Moore-Smith convergence in general topology.* Ann. of Math. (2) 38 (1937), 39–56.

[BJ75] J. Berman and S.L. Jordan, *The Kuratowski closure-complement problem.* Amer. Math. Monthly 82 (1975), 841–842.

[Bl19] A. Blaszczyk, *A simple proof of Sierpiński's theorem.* Amer. Math. Monthly 126 (2019), 464–466.

[BM99] J.L. Bryant and W. Mio, *Embeddings of homology manifolds in codimension* ≥ 3. Topology 38 (1999), 811–821.

[Bo58] J. Bosák, *Generalization of the method of complete induction.* Acta Fac. Nat. Univ. Comenian. Math. 2 (1958), 255–256.

[Bo] N. Bourbaki, *Éléments de Mathématique. Livre III: Topologie Générale*, Hermann, Paris, 1940, 1942, 1947, 1948, 1949.

[BPW17] M. Balcerzak, M. Poplawski and J. Wodka, *Local extrema and nonopenness points of continuous functions.* Amer. Math. Monthly 124 (2017), 436–443.

[CA] P.L. Clark, *Commutative Algebra.* `http://pete.math.uga.edu/~pete/integral.pdf`

[Ca37] H. Cartan, *Théorie des filtres.* C.R. Acad. Paris, 205, (1937) 595–598. *Filtres et ultrafiltres.* 777–779.

[Ca73] D.E. Cameron, *The mini-max property of the Tychonoff product topology.* Amer. Math. Monthly 80 (1973), 925–927.

[CdC] K. Conrad, *The contraction mapping theorem.* `http://www.math.uconn.edu/∼kconrad/blurbs/analysis/contractionshort.pdf`

[CHN74] P. Cameron, J.G. Hocking and S.A. Naimpally, *Nearness–A Better Approach to Continuity and Limits.* Amer. Math. Monthly 81 (1974), 739–745.

[Ch76] P.E. Cohen, *Products of Baire spaces.* Proc. Amer. Math. Soc. 55 (1976), 119–124.

[Ch19] Y. R. Chao, *A note on "Continuous mathematical induction."* Bull. Amer. Math. Soc. Volume 26, (1919), 17–18.

[Ch82] G. Chichilnisky, *The topological equivalence of the Pareto condition and the existence of a dictator.* J. Math. Econom. 9 (1982), 223–233.

[Ch92] P.R. Chernoff, *A simple proof of Tychonoff's theorem via nets.* Amer. Math. Monthly 99 (1992), 932—934.

[Ci20] K.C. Ciesielski, *Sierpiński's topological characterization of* $\mathbb{Q}$. Math. Mag. 93 (2020), 136–138.

[Cl-HC] P.L. Clark, *Honors Calculus.* `http://math.uga.edu/~pete/2400full.pdf`.

[Cl73] R. Cleveland, *A Global Characterization of Uniform Continuity.* Amer. Math. Monthly 80 (1973), 64–66.

[Cl19] P.L. Clark, *The instructor's guide to real induction.* Math. Mag. 92 (2019), 136–150.

[CLP19] P.L. Clark, N. Lebowitz-Lockard and P. Pollack, *A note on Golomb topologies.* Quaest. Math. 42 (2019), 73–86.

[CL20] D. Cao Labora, *When Is a Continuous Bijection a Homeomorphism?* Amer. Math. Monthly 127 (2020), 547–553.

[Co67] L.W. Cohen, *Classroom Notes: On Being Mean to the Mean Value Theorem.* Amer. Math. Monthly 74 (1967), 581–582.

[DF92] J. Dydak and N. Feldman, *Major theorems on compactness: a unified exposition.* Amer. Math. Monthly 99 (1992), 220–227.

[Di05] J.J. Dijkstra, *On homeomorphism groups and the compact-open topology.* Amer. Math. Monthly 112 (2005), 910–912.

[Do00] D.E. Dobbs, *When is an ordered field a metric space?* Tsukuba J. Math. 24 (2000), 325–336.

[Do03] G. Dowek, *Preliminary investigations on induction over real numbers*, preprint, 2003.

[DS84] V. Drobot and J. Sawka, *The Teaching of Mathematics: Why the Product Topology?* Amer. Math. Monthly 91 (1984), 137–138.

[DS14] F. Dreher and T. Samuel, *Continuous images of Cantor's ternary set.* Amer. Math. Monthly 121 (2014), 640–643.

[Du] J. Dugundji, *Topology.* Allyn and Bacon, Inc., Boston, Mass. 1966.

[Du57] W.L. Duren, Jr., *Mathematical induction in sets.* Amer. Math. Monthly 64 (1957), 19–22.

[Du64] R.M. Dudley, *On sequential convergence.* Trans. Amer. Math. Soc. 112 (1964), 483–507.

[Eb77] C. Eberhart, *Some remarks on the irrational and rational numbers.* Amer. Math. Monthly 84 (1977), 32–35.

[Ed62] M. Edelstein, *On fixed and periodic points under contractive mappings.* J. London Math. Soc. 37 (1962), 74–79.

[Eh94] P. Ehrlich, *Dedekind cuts of Archimedean complete ordered abelian groups.* Algebra Universalis 37 (1997), 223–234.

[El09] M. Elekes, *On a converse to Banach's fixed point theorem.* Proc. Amer. Math. Soc. 137 (2009), 3139-–3146.

[En] R. Engelking, *General topology.* Monografie Matematyczne, Tom 60. PWN—Polish Scientific Publishers, Warsaw, 1977.

[Fi70] P.C. Fishburn, *Arrow's impossibility theorem: concise proof and infinite voters.* J. Econom. Theory 2 (1970), 103–106.

[Fi91] J.H. Fife, *The Kuratowski closure-complement problem.* Math. Mag. 64 (1991), 180–182.

[Fl70] L. Flatto, *On Groups of Motions Generated by Two Rotations.* Amer. Math. Monthly 77 (1970), 733–735.

[Fo57] L.R. Ford, *Interval-additive propositions.* Amer. Math. Monthly 64 (1957), 106–108.

[Fo10] G.B. Folland, *A tale of topology.* Amer. Math. Monthly 117 (2010), 663–672.

[Fu55] H. Furstenberg, *On the Infinitude of Primes.* Amer. Math. Monthly 62 (1955), 353.

[Ga71] T.E. Gantner, *A Regular Space on which on every Continuous Real-Valued Function is Constant.* Amer. Math. Monthly 78 (1971), 52–53.

[GJ08] B.J. Gardner and M. Jackson, *The Kuratowski closure-complement theorem.* New Zealand J. Math. 38 (2008), 9–44.

[GMGS11] J. L. Gámez-Merino, G.A. Muñoz-Fernández, J.B. Seoane-Supúlveda, *A characterization of continuity revisited.* Amer. Math. Monthly 118 (2011), 167–170.

[Go59] S.W. Golomb, *A Connected Topology for the Integers.* Amer. Math. Monthly 66 (1959), 663–665.

[Go15] W.B. Goldbloom, *Open discontinuous maps from* $\mathbb{R}^n$ *onto* $\mathbb{R}^n$. Amer. Math. Monthly 122 (2015), 268–271.

[Gr12] J. Groves, *An elementary counterexample in the compact-open topology.* Amer. Math. Monthly 119 (2012), 693–694.

[GS79] J. Ginsburg and B. Sands, *Minimal Infinite Topological Spaces.* Amer. Math. Monthly 86 (1979), 574–576.

[Ha78] D. Hartig, *Classroom Notes: An Important Functor in Analysis and Topology.* Amer. Math. Monthly 85 (1978), 41–43.

[Ha11] D. Hathaway, *Using Continuity Induction.* College Math. Journal 42 (2011), 229–231.

[HC] P.L. Clark, *Honors Calculus.* `http://math.uga.edu/~pete/2400full.pdf`

[He46a] E. Hewitt, *On two problems of Urysohn.* Ann. of Math. (2) 47 (1946), 503–509.

[He46b] E. Hewitt, *A remark on density characters.* Bull. Amer. Math. Soc. 52 (1946), 641–643.

[Hi83] D. Higgs, *Iterating the derived set function.* Amer. Math. Monthly 90 (1983), 693–697.

[HM09] C. Hecker and R.P. Millspaugh, *Limit sets and closed sets in separable metric spaces.* Missouri J. Math. Sci. 21 (2009), 78–82.

[Hm77] R.C. Haworth and R.A. McCoy, *Baire spaces.* Dissertationes Math. (Rozprawy Mat.) 141 (1977), 73 pp.

[Hu81] H. Hueber, *On Uniform Continuity and Compactness in Metric Spaces.* Amer. Math. Monthly 88 (1981), 204–205.

[In65] A.J. Insel, *A Note on the Hausdorff Separation Property in First Countable Spaces.* Amer. Math. Monthly 72 (1965), 289–290.

[JGR91] D. Janković, M. Ganster and I. Reilly, *A characterization of Baire spaces.* Amer. Math. Monthly 98 (1991), 624–625.

[JH90] D. Janković and T.R. Hamlett, *New topologies from old via ideals.* Amer. Math. Monthly 97 (1990), 295–310.

[Jo99] H. Johnston, *Transversality for homology manifolds.* Topology 38 (1999), 673–697.

[Ka] I. Kaplansky, *Set theory and metric spaces.* Allyn and Bacon Series in Advanced Mathematics. Allyn and Bacon, Inc., Boston, Mass., 1972.

[Ka07] I. Kalantari, *Induction over the continuum.* (English summary) Induction, algorithmic learning theory, and philosophy, 145–154, Log. Epistemol. Unity Sci., 9, Springer, Dordrecht, 2007.

[Ke50] J.L. Kelley, *Convergence in topology.* Duke Math. J. 17, (1950). 277–283.

[Ke] J.L. Kelley, *General topology.* D. Van Nostrand Company, Inc., Toronto-New York-London, 1955.

[Kh23] A. Khinchin, Fund. Math. 4 (1923), 164–166.

[Kh49] A. Khinchin, *The simplest linear continum.* Uspehi Matem. Nauk (N.S.) 4, (1949). no. 2 (30), 180–197.

[Ki99] S.S. Kim, *A Characterization of the Set of Points of Continuity of a Real Function.* Amer. Math. Monthly 106 (1999), 258–259.

[Ko88] R. Kopperman, *All topologies come from generalized metrics.* Amer. Math. Monthly 95 (1988), 89–97.

[Ko09] J.J. Koliha, *Mean, meaner, and the meanest mean value theorem.* Amer. Math. Monthly 116 (2009), 356–361.

[Kr74] M.R. Krom, *Cartesian products of metric Baire spaces.* Proc. Amer. Math. Soc. 42 (1974), 588–594.

[KS72] A.P. Kirman and D. Sondermann. *Arrow's theorem, many agents, and invisible dictators.* J. Econom. Theory 5 (1972), 267–277.

[KT08] P. Komjáth and V. Totik, *Ultrafilters.* Amer. Math. Monthly 115 (2008), 33–44.

[Le60] N. Levine, *Remarks on Uniform Continuity in Metric Spaces.* Amer. Math. Monthly 67 (1960), 562–563.

[Le65] N. Levine, *Spaces in Which Compact Sets are Never Closed.* Amer. Math. Monthly 72 (1965), 635–638.

[Le82] H. Leinfelder, *A unifying principle in real analysis.* Real Anal. Exchange 8 (1982/83), no. 2, 511–518.

[Ma47] E. Marczewski, *Séparabilité et multiplication cartésienne des espaces topologiques.* Fund. Math. 34 (1947), 127–143.

[Ma88] K.D. Magill, Jr., *Universal topological spaces.* Amer. Math. Monthly 95 (1988), 942–946.

[Ma93] M. Mandelkern, *A short proof of the Tietze-Urysohn extension theorem.* Arch. Math. (Basel) 60 (1993), 364–366.

[Ma20] É. Matheron, *Three proofs of Tychonoff's theorem.* Amer. Math. Monthly 127 (2020), 437–443.

[Ma47] E. Marczewski, *Séparabilité et multiplication cartésienne des espaces topologiques.* Fund. Math. 34 (1947), 127–143.

[Mc78] R.A. McCoy, *Second Countable and Separable Function Spaces.* Amer. Math. Monthly 85 (1978), 487–489.

[Mi70] J.V. Michalowicz, *On Characterization of Uniformizable Spaces.* Amer. Math. Monthly 77 (1970), 1095.

[MO] `https://mathoverflow.net/questions/181226`

[Mo10] E.H. Moore, *On a form of general analysis*, Yale Colloquium Lectures, 1910.

[Mo15] E.H. Moore, *Definition of limit in general integral analysis*, Proc. Nat. Acad. Sci. vol 1 (1915), p. 628.

[MR68] R.M.F. Moss and G.T. Roberts, *A creeping lemma.* Amer. Math. Monthly 75 (1968), 649–652.

[MS22] E.H. Moore and H.L. Smith, *A general theory of limits*, Am. Jour. 44 (1922), 102-121.

[Mu] J.R. Munkres, *Topology: a first course.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

[Ne86] A. Neubrunnová, *On a unifying principle in real analysis.* Acta Math. Univ. Comenian. 48/49 (1986), 123–126 (1987).

[Na] L. Nachbin, *Topology and order.* Van Nostrand Mathematical Studies, No. 4 D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London 1965.

[No52] M.J. Norris, *Integrability of continuous functions.* Amer. Math. Monthly 59 (1952), 244–245.

[NP88] D.J. Newman and T.D. Parsons, *On Monotone Subsequences.* Amer. Math. Monthly 95 (1988), 44–45.

[NZ07] S.B. Nadler, Jr. and D.M. Zitney, *Pointwise products of uniformly continuous functions on sets in the real line.* Amer. Math. Monthly 114 (2007), 160–163.

[Os14] M.S. Osborne, *Hausdorffization and Such.* Amer. Math. Monthly 121 (2014), 727–733.

[Ox] J.C. Oxtoby, *Measure and category. A survey of the analogies between topological and measure spaces.* Second edition. Graduate Texts in Mathematics, 2. Springer-Verlag, New York-Berlin, 1980.

[Ox61] J.C. Oxtoby, *Cartesian products of Baire spaces.* Fund. Math. 49 (1960/61), 157–166.

[Pe26] O. Perron, Jber. Deutsch. Math. Verein. 35 (1926), 194–203.

[Pe90] M. Petkovšek, *Ambiguous numbers are dense.* Amer. Math. Monthly 97 (1990), 408–411.

[Po44] E.S. Pondiczery, *Power problems in abstract spaces.* Duke Math. J. 11 (1944), 835–837.

[Ro76] I. Rosenholtz, *Another Proof that Any Compact Metric Space is the Continuous Image of the Cantor Set.* Amer. Math. Monthly 83 (1976), 646–647.

[RS64] K.A. Ross and A.H. Stone, *Products of Separable Spaces.* Amer. Math. Montly 71 (1964), 398–403.

[RS15] M. Raman-Sundström, *A pedagogical history of compactness.* Amer. Math. Monthly 122 (2015), 619–635.

[Rud] W. Rudin, *Principles of mathematical analysis.* Third edition. International Series in Pure and Applied Mathematics. McGraw-Hill, New York-Auckland-Düsseldorf, 1976.

[Run] V. Runde, Runde, *A taste of topology.* Universitext. Springer, New York, 2005.

[Sa84] T. Šalát, *Remarks on unifying principles in real analysis.* Real Anal. Exchange 10 (1984/85), no. 2, 343–348.

[Sh72] P. Shanahan, *A unified proof of several basic theorems of real analysis.* Amer. Math. Monthly 79 (1972), 895–898.

[Sh74] P. Shanahan, *Addendum to: "A unified proof of several basic theorems of real analysis" (Amer. Math. Monthly 79 (1972), 895–898).* Amer. Math. Monthly 81 (1974), 890–891.

[Si08] T.Q. Sibley, *Sublimital Analysis.* Mathematics Magazine, 81 (2008), 369–373.

[Sm57] M.F. Smiley, *Filters and equivalent nets*, Am. Math. Monthly, Vol. 64, No. 5 (May, 1957), 336-338.

[S] M. Spivak, *Calculus.* Fourth edition.

[SS] L.A. Steen and J.A. Seebach, Jr., *Counterexamples in topology.* Second edition. Springer-Verlag, New York-Heidelberg, 1978.

[St70] L.A. Steen, *A direct proof that a linearly ordered space is hereditarily collectionwise normal.* Proc. Amer. Math. Soc. 24 (1970), 727–728.

[T] A.D. Taylor, *Social choice and the mathematics of manipulation.* Outlooks. Cambridge University Press, Cambridge; Mathematical Association of America, Washington, DC, 2005.

[Tr20] T. Trujillo, *Condorcet's paradox and ultrafilters.* Amer. Math. Monthly 127 (2020), 169–174.

[Tu97] T.W. Tucker, *Rethinking rigor in calculus: the role of the mean value theorem.* Amer. Math. Monthly 104 (1997), 231–240.

[Ul03] D.C. Ullrich, *The Ascoli-Arzelà theorem via Tychonoff's theorem.* Amer. Math. Monthly 110 (2003), 939–940.

[vRS] A.C.M. van Rooij and W.H. Schikhof, *A second course on real functions.* Cambridge University Press, Cambridge-New York, 1982.

[Ve97] D.J. Velleman, *Characterizing continuity.* Amer. Math. Monthly 104 (1997), 318–322.

[W] R.C. Walker, *The Stone-Čech compactification.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 83. Springer-Verlag, New York-Berlin, 1974.

[We23] N. Weiner, *Differential space.* J. Math.Phys.58 (1923), 131–174.

[Wh44] H. Whitney, *The self-intersections of a smooth n-manifold in $2n$-space.* Ann. of Math. (2) 45 (1944), 220–246.

[Wi] S. Willard, *General topology.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont. 1970.