# A FIRST COURSE IN TOPOLOGY: EXPLAINING CONTINUITY (FOR MATH 112, SPRING 2005)

PAUL BANKSTON

CONTENTS

## 1. Epsilons and Deltas

In this course we take the overarching view that the mathematical study called topology grew out of an attempt to make precise the notion of *continuous function* in mathematics. This is one of the most difficult concepts to get across to beginning calculus students, not least because it took centuries for mathematicians themselves to get it right. The intuitive idea is natural enough, and has been around for at least four hundred years. The mathematically precise formulation dates back only to the ninteenth century, however. This is the one involving those pesky epsilons and deltas, the one that leaves most newcomers completely baffled. Why, many ask, do we even bother with this confusing definition, when there is the original intuitive one that makes perfectly good sense? In this introductory section I hope to give a believable answer to this quite natural question.

Let us begin with the classical intuitive definition of what it means for a real-valued function of a real variable to be continuous at a point. By way of notation, we write $f : X \to Y$ to denote such a function. (Often the word *map* is used instead of *function*; the two words are regarded as synonymous in these Notes.) $X$ and $Y$ are sets, in this case $Y$ is the real line $\mathbb{R}$ and $X$ is a subset of $\mathbb{R}$, $X$ is called the **domain** of the function, $Y$ is called the **range** of the function, and $f$ assigns exactly one point $y \in Y$ to each point $x \in X$. We usually write $y = f(x)$ to say this. (This is the familiar way; actually, to be precise, $f$ is a set of ordered pairs $\langle x, y\rangle$ from the cartesian product $X \times Y$ such that, for any $x \in X$ there is a unique $y \in Y$ such that $\langle x, y\rangle \in f$.)

**Definition 1.1** (Continuity at a Point: Intuitive)**.** *Let $X$ be a set of real numbers, $x_0 \in X$, and $f : X \to \mathbb{R}$ a function. $f$ is* ***continuous at*** *$x_0$ if, whenever $x \in X$ is close to $x_0$, then $f(x)$ is close to $f(x_0)$. $f$ is* ***discontinuous at*** *$x_0$ if $f$ is not continuous at $x_0$.*

At first glance, this is quite appealing. However, upon a second look, the definition begs the question of what we mean by the expression *close to*. Without a precise explication of when one point is close to another, this definition of continuity is a castle in the air, not very useful for teasing out the finer points of continuous behavior. The amazing fact is that it actually served for centuries as the main working explication of continuity. Generations of mathematicians made use of it; the good ones were able to discover important results anyway, the mediocre ones made lots of mistakes. The most accurate conception of a function that was continuous at each point of a real interval appears to have lain in the view that the graph of the function had no "gaps," it was "all of one piece." Nevertheless there seemed to be fairly wide-spread confusion between continuity and differentiability, well into the 1800s.

This is where epsilons and deltas come in. By the early 1800s developing standards of mathematical rigor made it imperative to get a firmer handle on continuity. The big breakthrough came with the realization, largely attributed to Augustin-Louis Cauchy (1789–1857) in the 1820s, that one didn't need an *absolute* notion of closeness, only a *relative* one. So, by making $x$ close enough to $x_0$ in response to a given standard of closeness (to $f(x_0)$), one could guarantee that $f(x)$ would fall within that standard of closeness to $f(x_0)$.

The idea can be likened to a problem in archery. Suppose you're trying to aim an arrow to hit a circular target that's four feet across and one hundred feet down range. Assuming the existence of an ideal aim that will send the arrow to the exact center of the target, how much could you be off that aim and still send the arrow somewhere onto the target?

Where the epsilon $\epsilon$ comes in is the standard of closeness you have to meet (getting the arrow within two feet of dead center), where the delta $\delta$ comes in is as a measure of how far off perfect aim you can be and still meet the $\epsilon$-standard. So, given that you're within $\delta$ of perfect aim, your result will be within $\epsilon$ of dead center. Now let's try this idea in mathematical language.

**Definition 1.2** (Continuity at a Point: Precise)**.** *Let $X$ be a set of real numbers, $x_0 \in X$, and $f : X \to \mathbb{R}$ a function. $f$ is **continuous at** $x_0$ if, whenever $\epsilon$ is a positive real number, there exists a positive real number $\delta$ (depending, possibly, on both $\epsilon$ and the point $x_0$) such that: if $x \in X$ and $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. $f$ is **discontinuous at** $x_0$ if $f$ is not continuous at $x_0$. $f$ is **continuous on** $X$ if $f$ is continuous at each point of $X$.*

But wait! Isn't this another castle in the air? Don't we need to define precisely what all the symbols mean? Well, yes indeed. But this is not too difficult, given a calculus student's basic understanding of the real line. Let's work an example.

**Example 1.3** (Continuity of Affine Functions)**.** *Let $m$ and $b$ be fixed real numbers, and define $f(x) := mx + b$. ($f$ is called an **affine** map.) We show $f$ is continuous at any $x_0$ as follows: Given $\epsilon > 0$, we wish to find $\delta > 0$ such that, if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| = |(mx + b) - (mx_0 + b)| = |m(x - x_0)| = |m||x - x_0| < \epsilon$. Clearly, if $m \neq 0$, then setting $\delta = \frac{\epsilon}{|m|}$ (or anything smaller) will work; otherwise $f$ is constantly $b$, and we may choose $\delta$ to be anything we like. (Note that, in the case $m \neq 0$, $\delta$ depends on $\epsilon$, but is independent of $x_0$. In the case $m = 0$, $\delta$ is independent even of $\epsilon$.)*

Of course we are glossing over some important details, namely the precise set-theoretic construction of the real line as a mathematical entity. Such a construction can be done, and it is highly instructive for a student to see it, but we don't need to go into the details here. (The interested reader may consult any of the many textbooks on beginning real analysis for an exhaustive treatment.) What is important for our purposes is that the usual algebraic and order-theoretic structure of the real line makes it into a *complete Archimedean ordered field*. Let's consider briefly what these words mean before going further. First, the usual addition and multiplication operations on $\mathbb{R}$ satisfy the **field** axioms:

- (F1) The *associative* laws $(x + y) + z = x + (y + z)$ and $(xy)z = x(yz)$ hold.
- (F2) The *commutative* laws $x + y = y + x$ and $xy = yx$ hold.
- (F3) The *distributive* law, $x(y + z) = xy + xz$, of multiplication over addition holds.
- (F4) 0 and 1 are, respectively, the *additive* and *multiplicative identity elements*; i.e., the laws $x + 0 = x = x1$ hold.

(F5) Every real number $x$ has an *additive inverse* $-x$; i.e., the law $x + (-x) = 0$ holds.
(F6) Every nonzero real number $x$ has an *multiplicative inverse* $x^{-1}$; i.e., the law $xx^{-1} = 1$ holds.

Next, there are the **order** axioms, most conveniently given by saying there is a designated subset $P$ of $\mathbb{R}$, the set of *positive* elements of the field, satisfying:

(O1) $1 \in P$ and $0 \notin P$.
(O2) If $x$ and $y$ are in $P$, then so are $x + y$ and $xy$.
(O3) If $x \neq 0$, then either $x \in P$ or $-x \in P$.

We write $x > 0$ as an abbreviation for $x \in P$, and $x < y$ as an abbreviation for $y - x > 0$. Finally we have the two axioms for ordered fields that, in fact, characterize the real line:

(O4) *(Archimedean Property)* Given any real number $x$, there is a natural number $n$ such that $x \leq n$. Equivalently, given any $x > 0$, there is a natural number $n > 0$ such that $\frac{1}{n} < x$. (In the context of abstract ordered fields, it is more proper to say that $x \leq (1 + \cdots + 1)$, where we take the sum of $n$ copies of the multiplicative identity element.)
(O5) *(Least Upper Bound (Completeness) Property)* Given any nonempty subset $X \subseteq \mathbb{R}$, if $X$ is *bounded above*; i.e., if there is some *upper bound* for $X$, an element $b$ such that $x \leq b$ for all $x \in X$, then there is a *least* upper bound $b_0$ for $X$ (i.e., $b_0$ is an upper bound for $X$, but no smaller real number is an upper bound for $X$).

**Remark 1.4.** *(For the reader with some algebraic background.) The real line is unique as a complete Archimedean ordered field. That is, if $F$ is any ordered field with addition $+_F$, multiplication $*_F$, identity elements $0_F$ and $1_F$, and set $P_F$ of positive elements, then there is a one-one map $\varphi$ from $\mathbb{R}$ onto $F$ such that both $\varphi$ and its function inverse $\varphi^{-1}$ are* ordered field isomorphisms. *I.e., $\varphi(0) = 0_F$, $\varphi(1) = 1_F$, $\varphi(x+y) = \varphi(x) +_F \varphi(y)$, $\varphi(xy) = \varphi(x) *_F \varphi(y)$, and $\varphi(x) \in P_F$ for all $x > 0$.*

The main point of difference between Definitions 1.1 and 1.2 is that there is a satisfactory explication of all the symbols and words used in defining "continuous at $x_0$" in the latter case, but not in the former. The notion "$x$ is close to $x_0$" is merely subjective; there are no truly *infinitesimal* real numbers, other than zero itself. This forces us to conclude that "close" must mean "equal," that Definition 1.1 merely restates that $f$ is a function, and absolutely nothing new gets defined.

Here's a slightly more sophisticated argument for why there can be no resurrecting Definition 1.1. Suppose there were some way of defining *close to* on the real line so that the same functions would come up continuous, regardless of which definition you used. For simplicity, we write $x \sim y$ to mean $x$ is "close" to $y$, however we may have come up with such a notion. By the very intended nature of a close-to relation, it is entirely reasonable to insist that $x \sim x$ always holds (the **reflexivity** condition), and that $y \sim x$ holds whenever $x \sim y$ holds (the **symmetry** condition). That said, pick an arbitrary $x_0 \in \mathbb{R}$, and suppose there is some $x_1 \neq x_0$ such that $x_1 \sim x_0$. (Either there is such an $x_1$ or there isn't; say for this particular $x_0$ there is.) With these two points chosen, let $y_1$ be any real number, and define $f(x)$ to be the affine map whose graph is the straight line going through the points $\langle x_0, x_0 \rangle$ and $\langle x_1, y_1 \rangle$. This line has slope $m = \dfrac{y_1 - x_0}{x_1 - x_0}$, well defined because the denominator is nonzero. $f$ is continuous at $x_0$, in the sense of Definition 1.2, as was demonstrated in Example 1.3. Assuming $f$ is also continuous at $x_0$ in the sense of Definition 1.1, and since $x_1 \sim x_0$, we have that $f(x_1) \sim f(x_0)$ too. But $f(x_1) = y_1$ and $f(x_0) = x_0$, so this says that $y_1 \sim x_0$. Since $y_1$ may be chosen arbitrarily, once we have $x_0$ and $x_1$, we infer the following from this argument: *If there are two distinct points that are close to one another, then every point is close to every other point.*

So there are only two possibilities for the closeness relation: either all points are close to all other points, or no point is close to any point other than itself. In either case, *every* function is continuous at each point in its domain, in the sense of Definition 1.1. Since there are discontinuous functions in the sense of Definition 1.2 (see Exercise 1.6 (4) below), we have obtained a contradiction. Hence: *There can be no relation of closeness on the real line that captures the notion of continuity put forward in Definition 1.2.*

We conclude this introductory section with a promise. One of the goals of this course is to prove the following two supremely important theorems of elementary analysis, underpinning all the major theoretical results of first-semester calculus. These theorems, though easy to state and intuitively appealing, are notoriously difficult to prove from first principles; and are therefore taken on faith, even in reasonably rigorous calculus courses. By the time we develop the necessary machinery, though, the theorems will fall out as easy corollaries.

**Theorem 1.5** (Twin Pillars of Single-Variable Calculus)**.** *Suppose $X = [a, b]$; i.e., $X$ is the bounded closed interval consisting of real numbers $x$ such that $a \leq x \leq b$. Suppose $f : X \to \mathbb{R}$ is continuous on $X$.*

(i) (Extreme Value Theorem)*: There exist $c, d \in [a, b]$ such that, for all $x \in [a, b]$, $f(c) \leq f(x) \leq f(d)$. (I.e., $f$ achieves an absolute maximum and an absolute minimum on $[a, b]$.)*

(ii) (Intermediate Value Theorem)*: If $d$ lies between $f(a)$ and $f(b)$, then there exists some $c \in [a, b]$ such that $f(c) = d$. (I.e., if $f$ takes on two given values, then it takes on all values in between.)*

**Exercises 1.6.** *(Throughout these Notes, the more challenging exercises are preceded with an asterisk.)*

(1) *Show that $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) := x^2$, is continuous at $x_0 = 1$.*

(2) ** Show that the squaring map $f$ in Exercise 1.6 (1) is continuous on $\mathbb{R}$ (Note how your choice of $\delta$ depends on* both *$x_0$ and $\epsilon$.)*

(3) *Let $X$ be a subset of $\mathbb{R}$. A point $x_0 \in X$ is an **isolatated point** of $X$ if there is some $\epsilon > 0$ such that no point of $X$ lies within $\epsilon$ of $x_0$ (i.e., if*

*$|x-x_0| < \epsilon$ and $x \in X$, then $x = x_0$). Let $f : X \to \mathbb{R}$ be any function, with $x_0$ an isolated point of $X$. Show $f$ is continuous at $x_0$.*

(4) *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by saying $f(x) = 0$ if $x < 0$, and $f(x) = 1$ otherwise. Show that $f$ is* not *continuous on $\mathbb{R}$.*

(5) *Suppose $X = (a, b)$; i.e., $X$ is the open interval consisting of real numbers $x$ such that $a < x < b$. Suppose $x_0 \in X$ and $f : X \to \mathbb{R}$ is such that the derivative $f'(x_0) := \lim_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0}$ exists at $x_0$. Show that $f$ is continuous at $x_0$.*

(6) *Given $x_0 \in \mathbb{R}$, find a function $f : \mathbb{R} \to \mathbb{R}$ that is discontinuous at $x_0$* only.

(7) *Review the proof of the mean value theorem from calculus, and check where Theorem 1.5 is used. After that, look how it enters into the development of the Riemann integral for continuous functions. Even such a basic fact as the one saying a function with zero derivative on an open interval must be constant relies on Theorem 1.5.*

(8) *Consider the set $\mathbb{Q}$ of* rational *numbers, those real numbers that can be represented as fractions of integers. Then, by restricting addition, multiplication, and order to $\mathbb{Q}$, we get an ordered field. Show that the ordered field of rational numbers is Archimedean, but not complete.*

(9) ** Consider the set $\mathbb{C}$ of* complex *numbers. This consists of all formal sums $a + ib$ (or,equivalently, ordered pairs $\langle a, b\rangle$ where $a$ and $b$ are real numbers), with addition and multiplication defined by the laws $(a + ib) + (c + id) := (a + c) + i(b + d)$ and $(a + ib)(c + id) := (ac - bd) + i(ad + bc)$. (Think of $i = \langle 0, 1\rangle$ as the "imaginary unit," chosen to satisfy $i^2 = -1$. Then it makes sense to multiply two complex numbers as if they were "binomials," and collect terms. The operations of addition and multiplication may be formally defined in the plane without reference to imaginary numbers, but it's less well motivated that way.) Show that the field axioms are satisfied, but that it is impossible to specify a set $P$ of "positive" complex numbers, in such a way that $\mathbb{C}$ becomes an ordered field.*

(10) *In this, as well as the following two exercises, refer to the least upper bound property (O5) above. Formulate a "greatest lower bound" property. Show that, given $\mathbb{R}$ has the least upper bound property, $\mathbb{R}$ also has the greatest lower bound property. [Hint: If $A \subseteq \mathbb{R}$ is bounded below, let $L_A$ be the set of lower bounds of $A$.]*

(11) *Show that least upper bounds are unique when they exist. I.e., show a set cannot have two least upper bounds.*

(12) *If $a < b$ in $\mathbb{R}$ and $A$ is either the closed interval $[a, b]$ or the open interval $(a, b)$, show that the least upper bound of $A$ is $b$.*

## 2. Distance Functions in Euclidean Space

What makes Definition 1.2 work is that the real line is equipped with a well-defined notion of *distance* between points $x$ and $y$, namely $|x - y|$. This is called the **euclidean distance** on the real line, is based on the algebraic structure of the reals, and can be readily extended to higher-dimensional euclidean space. The best way to introduce the euclidean distance on $n$-space $\mathbb{R}^n$ is to use some elementary linear algebra and first define the **dot product** between vectors $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ and $\mathbf{y} = \langle y_1, \ldots, y_n \rangle$ in $\mathbb{R}^n$ via the formula

$$\mathbf{x} \cdot \mathbf{y} := \sum_{i=1}^{n} x_i y_i$$

From this notion, one defines the **euclidean norm** $|\mathbf{x}|$ of a vector $\mathbf{x}$ to be $\sqrt{\mathbf{x} \cdot \mathbf{x}}$, and the **euclidean distance** between vectors $\mathbf{x}$ and $\mathbf{y}$ to be $|\mathbf{x} - \mathbf{y}|$. This is the straight-line distance between the two points, as can be justified using the Pythagorean theorem of classical geometry. Note that in the case $n = 1$, the dot product is just the usual product of real numbers and the euclidean norm is the absolute value. In the case $n = 2$, when complex numbers are identified with points in the plane, the euclidean norm amounts to the complex modulus.

Technically, each distinct $n$ gives rise to a different operation of vector sum, dot product, etc. However, since corresponding operations are so closely related as we pass from one euclidean space to another, we use the same notation throughout for the sake of simplicity. You should bear in mind nevertheless that when we write a relation such as $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$, the vertical bars on the left apply to real numbers, those on the right to vectors.

With this machinery now in place, the corresponding notion of continuity of functions taking points in one euclidean space to points in another is almost a *verbatim* restatement of Definition 1.2.

**Definition 2.1** (Continuity at a Point in Euclidean Space)**.** *Let $X$ be a subset of real $m$-space $\mathbb{R}^m$, $\mathbf{x}_0 \in X$, and $f : X \to \mathbb{R}^n$ a function. $f$ is* ***continuous at*** $\mathbf{x}_0$ *if, whenever $\epsilon$ is a positive real number, there exists a positive real number $\delta$ such that: if $\mathbf{x} \in X$ and $|\mathbf{x} - \mathbf{x}_0| < \delta$, then $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$.*

Note that, in this definition, the function $f$ takes points in euclidean $m$-space to points in euclidean $n$-space, and $m$ need not equal $n$. Thus we are dealing with two distinct (but closely related) distance functions. We take this moment to isolate the salient properties, the *metric* properties, of the euclidean distance.

**Theorem 2.2** (Metric Properties of Euclidean Distance)**.** *For any $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ in $\mathbb{R}^n$ we have:*

(i) (Positivity) $|\mathbf{x} - \mathbf{y}| \geq 0$; $|\mathbf{x} - \mathbf{y}| = 0$ *if and only if* $\mathbf{x} = \mathbf{y}$.
(ii) (Symmetry) $|\mathbf{x} - \mathbf{y}| = |\mathbf{y} - \mathbf{x}|$.
(iii) (Triangle Inequality) $|\mathbf{x} - \mathbf{y}| \leq |\mathbf{x} - \mathbf{z}| + |\mathbf{z} - \mathbf{y}|$.

*Proof.* Positivity is just a consequence of the facts that a sum of squares of real numbers is always nonnegative, and that if a sum of squares of real numbers is zero, then all those real numbers must be zero too. Symmetry is even easier to prove. As for the triangle inequality, which says, roughly, that the shortest path between two points is a straight line, a small amount of work is required.

Recall from beginning linear algebra that if $\mathbf{x}$ and $\mathbf{y}$ are two nonzero vectors in $\mathbb{R}^n$, and if $0 \leq \theta \leq \pi$ is the angle at which they meet, then $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos \theta$. (This follows from the parallelogram law of vector addition, plus the law of cosines from classic trigonometry.) In particular, because the cosine function is bounded above by 1, we infer that $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$. This inequality (called the Cauchy-Schwarz inequality, after Augustin-Louis Cauchy and, later, Herman Schwarz (1843–1921)) even holds if one or more of the vectors is zero because both sides then evaluate to zero.

Now, given two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, we have: $|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = (\mathbf{x} \cdot \mathbf{x}) + 2(\mathbf{x} \cdot \mathbf{y}) + (\mathbf{y} \cdot \mathbf{y}) \leq |\mathbf{x}|^2 + 2|\mathbf{x} \cdot \mathbf{y}| + |\mathbf{y}|^2 \leq |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2$. Observing that if $a$ and $b$ are nonnegative real numbers, then $a^2 \leq b^2$ if and only if $a \leq b$, we conclude that $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$ always holds.

Finally, suppose $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in $\mathbb{R}^n$ are given. Then, from the previous paragraph, we may write: $|\mathbf{x} - \mathbf{y}| = |(\mathbf{x} - \mathbf{z}) + (\mathbf{z} - \mathbf{y})| \leq |\mathbf{x} - \mathbf{z}| + |\mathbf{z} - \mathbf{y}|$, and the proof is complete.

□

The importance of Theorem 2.2 is that it acts as a bridge to more abstract settings: it isolates the salient features of euclidean distance that provide a reasonable explication of continuity. It so happens there are lots of other "distance functions" that also satisfy the conclusions of Theorem 2.2. Let's first consider two classic examples before crossing the bridge in the next section.

**Example 2.3** (The Taxicab Distance Function)**.** *In $\mathbb{R}^2$, let's define the **taxicab norm** $|\mathbf{x}|_t$ for $\mathbf{x} = \langle x_1, x_2 \rangle$ to be $|x_1| + |x_2|$. Then, following the definition of the euclidean distance function from the euclidian norm, we define the **taxicab distance** between $\mathbf{x}$ and $\mathbf{y}$ to be $|\mathbf{x} - \mathbf{y}|_t$. (Imagine that the plane is laid out in a grid, with "streets" that go only north-south and east-west. To get from one point to another, you have to follow a "street route.") It is a complete triviality to show that the taxicab distance satisfies the positivity and symmetry conditions of Theorem 2.2; it is only slightly less trivial to show that $|\mathbf{x} + \mathbf{y}|_t \leq |\mathbf{x}|_t + |\mathbf{y}|_t$ (see Exercise 2.5 (1) below), from which the triangle inequality instantly follows, exactly as in the proof of Theorem 2.2.*

**Example 2.4** (The Discrete Distance Function)**.** *Let's stick with $\mathbb{R}^2$, but now define the **discrete norm** $|\mathbf{x}|_d$ for $\mathbf{x}$ to be $0$ if $\mathbf{x} = \mathbf{0}$, and to be $1$ otherwise. I.e., $|\mathbf{x}|_d$ is either $0$ or $1$, depending upon whether or not $\mathbf{x}$ is the zero vector. Then, following the definition of the euclidean (resp., taxicab) distance function from the euclidian (taxicab) norm, we define the **discrete distance** between $\mathbf{x}$ and $\mathbf{y}$ to be $|\mathbf{x} - \mathbf{y}|_d$. Positivity and symmetry are obvious, as before, and it is an easy exercise to show (see Exercise 2.5 (7) below) that $|\mathbf{x} + \mathbf{y}|_d \leq |\mathbf{x}|_d + |\mathbf{y}|_d$. Again, the triangle inequality instantly follows.*

**Exercises 2.5.** (1) *Prove the following assertion about the taxicab norm:* $|\mathbf{x}+\mathbf{y}|_t \leq |\mathbf{x}|_t + |\mathbf{y}|_t$.

(2) *Define the **max norm** $|\mathbf{x}|_m$ of a point $\mathbf{x} = \langle x_1, x_2 \rangle$ in $\mathbb{R}^2$ to be max$\{|x_1|, |x_2|\}$. Then define the **max distance**, after the fashion of the euclidean and taxicab distances, and show that the conclusion of Theorem 2.2 holds for this new distance function.*

(3) *For any $\mathbf{x} \in \mathbb{R}^2$, prove that $|\mathbf{x}|_m \leq |\mathbf{x}| \leq |\mathbf{x}|_t \leq 2|\mathbf{x}|_m$.*

(4) *In $\mathbb{R}^2$, graph the solution sets of the inequalities $|\mathbf{x}| < 1$, $|\mathbf{x}|_t < 1$, and $|\mathbf{x}|_m < 1$.*

(5) *Define the **square norm** $|\mathbf{x}|_2$ of a point $\mathbf{x} = \langle x_1, x_2 \rangle$ in $\mathbb{R}^2$ to be $x_1^2 + y_1^2$. Then define the **square distance**, after the fashion of the euclidean and taxicab distances, and show that the conclusion of Theorem 2.2 does* not *necessarily hold for this new distance function.*

(6) *In $\mathbb{R}^2$, graph the solution sets of the inequalities $|\mathbf{x}|_d < 1$ and $|\mathbf{x}|_d < 1.000001$.*

(7) *Prove the following assertion about the discrete norm: $|\mathbf{x}+\mathbf{y}|_d \leq |\mathbf{x}|_d + |\mathbf{y}|_d$.*

## 3. Metrics and Metric Spaces

As mentioned in the previous section, Theorem 2.2 provides us with a bridge to a more abstract (and therefore more widely applicable) setting in which to talk about continuity. That theorem may be viewed as a litmus test whose only criterion for the suitability of any given distance function is whether or not it satisfies the metric conditions of positivity, symmetry, and the triangle inequality. All the examples considered in Section 2 ultimately arise from the very rich algebraic and order-theoretic structure of the real line. But all the extra structure that happens to support the definition of a given distance function is irrelevant for the purposes of applying the litmus test. At the end of the day, what matters is whether the three metric conditions have been met.

The study of topology is very firmly rooted in set theory. For this reason it is time to pause for a little explanation of set-theoretic notions we use in this course. (We don't bring them in all at once, only just as we need them.) The notion of *set* is taken to be primitive. Intuitively it is a family (or collection, or aggregate) of *elements*; the statement that $x$ is an element of the set $X$ is denoted $x \in X$. Two sets are defined to be equal if and only if they have the same elements. (So the usual way to show, say, $X = Y$ is: first, pick $x \in X$ arbitrarily and show $x \in Y$; then pick $y \in Y$ arbitrarily and show $y \in X$.) If every element of $X$ is also an element of $Y$, we say $X$ is a **subset** of $Y$ (in symbols, $X \subseteq Y$). If a set $X$ has just $x_1, \dots, x_n$ for elements, we use curly bracket notation and write $X = \{x_1, \dots, x_n\}$. If $X$ consists of all elements $y \in Y$ such that some condition $\dots y \dots$ holds, then we write $X = \{y \in Y : \dots y \dots\}$. For example, given that the set of natural numbers is $\mathbb{N} := \{0, 1, 2, \dots\}$, the set of even natural numbers is $\{x \in \mathbb{N} : x = 2y, \text{for some } y \in \mathbb{N}\}$. If the universal set $Y$ is clear from the context, we sometimes suppress its mention in the notation. (So, with some care, we may write the set of even natural numbers in the simplified notation $\{x : x = 2y, \text{for some } y \in \mathbb{N}\}$.

The **cartesian product** $X \times Y$ of two sets $X$ and $Y$ is defined to be the set of all ordered pairs $\langle x, y\rangle$, where $x \in X$ and $y \in Y$. This, of course, begs the question of just what exactly is an ordered pair. $\langle x, y\rangle$ should itself be a set that depends on $x$ and $y$ only, and *in that order*. You want to be able to say that pairs $\langle x, y\rangle$ and $\langle u, v\rangle$ are equal just in case $x = u$ and $y = v$. One way to do this is to define $\langle x, y\rangle := \{\{x\}, \{x, y\}\}$ (see Exercise 3.9 (1) below).

This leads us to our first truly abstract definition, partially weaning us from the real numbers and euclidean space.

**Definition 3.1** (Metric Space)**.** *Let $X$ be a set. By a **metric** on $X$, we mean a two-place real-valued function $d : X \times X \to \mathbb{R}$ satisfying, for all $x, y, z \in X$:*

(i) (Positivity) $d(x, y) \geq 0$*;* $d(x, y) = 0$ *if and only if* $x = y$.
(ii) (Symmetry) $d(x, y) = d(y, x)$.
(iii) (Triangle Inequality) $d(x, y) \leq d(x, z) + d(z, y)$.

*A **metric space** is a pair $\langle X, d\rangle$, where $X$ is a set (the underlying set of **points**) and $d$ is a metric on $X$.*

**Examples 3.2.** (i) *The pairs $\langle \mathbb{R}^n, d\rangle$ are metric spaces, where $n \geq 1$ and $d$ is the corresponding euclidean distance function (Theorem 2.2).*

(ii) *The pairs* $\langle \mathbb{R}^2, d\rangle$ *are metric spaces, where* $d$ *is the taxicab (resp., max, discrete) distance function (Example 2.3 and Exercise 2.5 (2)).*
(iii) *The pair* $\langle \mathbb{R}^2, d\rangle$*, where* $d$ *is the square distance function of Exercise 2.5 (5), is* not *a metric space, because the triangle inequality is not satisfied.*
(iv) *The pair* $\langle X, d\rangle$*, where* $d(x, y)$ *is* $0$ *if* $x = y$*, and* $1$ *otherwise, is a metric space (see Exercise 3.9 (6) below).* $d$ *is called the* ***discrete metric*** *on* $X$*.*

**Definition 3.3** (The Subspace Metric)**.** *Let* $\langle X, d\rangle$ *be a metric space,* $A$ *a subset of* $X$*. Then it makes sense to restrict* $d : X \times X \to \mathbb{R}$ *to* $A \times A \subseteq X \times X$*. This restricted distance function, defined for pairs of points from* $A$ *and denoted* $d|A$ *(i.e., for* $\langle x, y\rangle \in A \times A$*,* $(d|A)(x, y) := d(x, y)$*), still satisfies the metric conditions of Definition 3.1.* $d|A$ *is called the* ***metric on*** $A \subseteq X$ ***induced by*** $d$*, more succinctly the* ***subspace metric*** *on* $A$*. And the metric space* $\langle A, d|A\rangle$ *is called a* ***metric subspace*** *of* $\langle X, d\rangle$*.*

**Example 3.4** (Great Circle Distance)**.** *Consider the earth as a perfect spherical ball, sitting in euclidean* 3*-space. The shortest* practical *distance between two points on the surface of the earth is the so-called* great circle *distance; i.e., the length of the shortest arc of a circle lying on the sphere, of maximal radius and containing the two points. This is never the distance induced by the euclidean metric in* $\mathbb{R}^3$*. In order to achieve a shortest-distance path between two points on the surface of the earth, one would need to burrow underground. The* nautical mile *is based on the lengths of arcs of great circles. The great circle arc length subtended by an angle of one minute (i.e., one-sixtieth of a degree) is the definition of a nautical mile, and measures out to just over* 6000 *feet.*

The reason for introducing metric subspaces at this point is to remove a bit of complexity in Definition 2.1, where the domain of the map in question is a subset $X$ of $\mathbb{R}^m$, equipped with the euclidean metric. In terms of Definition 3.3, $X$ is a metric subspace of $\mathbb{R}^m$. Once we know what the metric on $X$ actually is, there is no need to consider $\mathbb{R}^m$ at all. The following is practically a *verbatim* restatement of Definition 2.1, but is incredibly more wide reaching. It is of fundamental importance to note that Definitions 1.2, 2.1, and 3.5 are saying essentially the same thing, only in increasingly broader contexts.

**Definition 3.5** (Continuity at a Point in Metric Space)**.** *Let* $\langle X, d\rangle$ *and* $\langle Y, e\rangle$ *be metric spaces,* $x_0 \in X$*, and* $f : X \to Y$ *a function.* $f$ *is* ***continuous at*** $x_0$ *if, whenever* $\epsilon$ *is a positive real number, there exists a positive real number* $\delta$ *such that: if* $x \in X$ *and* $d(x, x_0) < \delta$*, then* $e(f(x), f(x_0)) < \epsilon$*.* $f$ *is* ***continuous on*** $X$ *if* $f$ *is continuous at each point of* $X$*.*

By way of a simple exercise in the use of Definition 3.5, let's prove the following.

**Proposition 3.6.** *Suppose* $\langle X, d\rangle$ *and* $\langle Y, e\rangle$ *are metric spaces, and* $f : X \to Y$ *is a function for which there is a constant* $C > 0$ *such that* $e(f(x), f(y)) \leq Cd(x, y)$

*for all $x, y \in X$. Then $f$ is continuous.*

*Proof.* Pick $x_0 \in X$, and let $\epsilon > 0$ be given. Since we want $e(f(x), f(x_0))$ to be $< \epsilon$, and we already know that $e(f(x), f(x_0)) \leq Cd(x, x_0)$, it suffices to make sure $Cd(x, x_0) < \epsilon$. So, if we make $\delta$ any positive number not exceeding $\frac{\epsilon}{C}$ (well definied because $C$ is assumed to be positive), and if $d(x, x_0) < \delta$, then $Cd(x, x_0) < C\delta \leq C\frac{\epsilon}{C} = \epsilon$, as desired. (Note that, just as in the argument in Example 1.3, the choice of $\delta$ does not depend on the particular point $x_0$. Note also that this proposition includes constant functions, for then $e(f(x), f(y))$ is always zero.)

$\square$

In the study of continuity, the particular *values* that a metric takes on are far less important than how distances *relate* to one another. As an illustration of this slightly vague statement (which will be made precise in Section 6), let us show that continuity of real-valued functions defined on the plane is unaffected as we move between the euclidean and the taxicab metrics.

**Example 3.7** (Change of Metric on the Plane)**.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function, and suppose we fix the euclidean metric on $\mathbb{R}$. Then $f$ is continuous at a point with respect to the euclidean metric if and only if $f$ is continuous at that point with respect to the taxicab metric of Example 2.3. To see this, pick $\mathbf{x}_0 \in \mathbb{R}^2$, and $\epsilon > 0$. Assuming that $f$ is continuous at $\mathbf{x}_0$ with respect to the euclidean metric, we know that there is a $\delta > 0$ such that, if $|\mathbf{x} - \mathbf{x}_0| < \delta$, then $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$. Now, by Exercise 2.5 (3), we know that $|\mathbf{y}| \leq |\mathbf{y}|_t \leq 2|\mathbf{y}|$ always holds. Thus, if $|\mathbf{x} - \mathbf{x}_0|_t < \delta$, then $|\mathbf{x} - \mathbf{x}_0| < \delta$ as well, and therefore $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$, as desired. This tells us that $f$ is continuous at $\mathbf{x}_0$ with respect to the taxicab metric. For the other direction, suppose there is a $\delta > 0$ such that $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$ whenever $|\mathbf{x} - \mathbf{x}_0|_t < \delta$. Thus if $|\mathbf{x} - \mathbf{x}_0| < \frac{\delta}{2}$, then $|\mathbf{x} - \mathbf{x}_0|_t < 2\frac{\delta}{2} = \delta$, and $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$. This tells us that $f$ is continuous at $\mathbf{x}_0$ with respect to the euclidean metric.*

We end this section with an important metric example from real analysis. Certain well-known theorems from calculus are assumed without proof; we promise to prove them later in the course.

**Example 3.8** (Function Space)**.** *We denote by $[0, 1]$ the closed unit interval in the real line; i.e., $[0, 1] := \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. We then denote by $C([0, 1])$ the set of all functions $f : [0, 1] \to \mathbb{R}$ such that $f$ is continuous with respect to the euclidean metric in both domain and range. (So elements are continuous real-valued functions.) Now, given $f \in C([0, 1])$, define the (generalized)* ***taxicab norm*** *$|f|_t$ as a Riemann integral:*

$$|f|_t := \int_0^1 |f(x)|\, dx$$

*Two comments are in order:*

(i) *$|f|_t$ is always well defined because of the fact, based on the extreme value theorem (Theorem 1.5 (i)), that says continuous real-valued functions on bounded closed intervals are Riemann integrable.*

(ii) *The taxicab norm makes sense in the setting of any euclidean space, just let $|\mathbf{x}|_t$ equal the sum of the absolute values of the coordinates of* $\mathbf{x}$. *Since* $\mathbf{x} \in \mathbb{R}^n$ *may be thought of as a function from the set* $\{1, 2, \ldots, n\}$ *to* $\mathbb{R}$, *and since the Riemann integral is a "continuous version of summing," it is reasonable to apply the modifier* taxicab *in defining* $|f|_t$, *for* $f \in C([0, 1])$.

*So, as in Example 2.3, let's define the* ***taxicab distance*** *between* $f$ *and* $g$ *in* $C([0, 1])$ *to be* $|f - g|_t$ *(where, as in calculus,* $(f - g)(x) := f(x) - g(x)$*). Is this a* bona fide *metric on* $C([0, 1])$*? Let's check symmetry first:* $|f - g|_t = \int_0^1 |f(x) - g(x)|\ dx = \int_0^1 |g(x) - f(x)|\ dx = |g - f|_t$. *That was easy; now let's see about the triangle inequality. This follows immediately if we can show that* $|f + g|_t \leq |f|_t + |g|_t$ *always holds. But the left-hand side is* $\int_0^1 |f(x) + g(x)|\ dx$. *Now the integrand* $|f(x) + g(x)|$ *is always* $\leq$ *the integrand* $|f(x)| + |g(x)|$*; so, by familiar properties of the Riemann integral, we have:* $|f + g|_t = \int_0^1 |f(x) + g(x)|\ dx \leq \int_0^1 (|f(x)| + |g(x)|)\ dx = \int_0^1 |f(x)|\ dx + \int_0^1 |g(x)|\ dx = |f|_t + |g|_t$. *Finally, having saved the best for last, let's check positivity. Interestingly enough, this turns out to be the most difficult condition to verify. Of course* $|f - g|_t$ *is always* $\geq 0$, *but what if* $|f - g|_t = 0$. *Does that imply that* $f$ *and* $g$ *are the same function? We can easily answer this question in the affirmative if we can show the following: If* $f : [0, 1] \to \mathbb{R}$ *is a continuous function such that* $f(x) \geq 0$ *for all* $x \in [0, 1]$, *and if* $f(x_0) > 0$ *for some* $x_0 \in [0, 1]$, *then* $\int_0^1 f(x)\ dx > 0$. *This fact depends crucially on the continuity of* $f$ *at* $x_0$. *Suppose* $f(x_0) = \epsilon > 0$. *Then there is a* $\delta > 0$ *such that, if* $x \in [0, 1]$ *and* $|x - x_0| < \delta$, *then* $|f(x) - f(x_0)| = |f(x) - \epsilon| < \epsilon$. *That implies that, for all* $x \in [0, 1]$ *within* $\delta$ *of* $x_0$, *we have* $f(x) > \frac{\epsilon}{2}$. *Thus the graph of* $f$ *lies above the horizontal line* $y = \epsilon$ *for* $x$ *lying in an interval, of positive width, containing* $x_0$. *Since* $f(x) \geq 0$ *for all* $x \in [0, 1]$, *we infer that* $\int_0^1 f(x)\ dx$ *is at least as large as the area of a rectangle of positive width and height.*

**Exercises 3.9.** (1) *Defining the ordered pair* $\langle x, y\rangle$ *to be the set* $\{\{x\}, \{x, y\}\}$, *show that, for any* $u, v, s, t$, $\langle u, v\rangle = \langle s, t\rangle$ *if and only if* $u = s$ *and* $v = t$.

(2) *Refer to Example 3.7, and show that the same conclusion obtains if we replace the taxicab metric with the max metric of Exercise 2.5 (2).*

(3) *Refer to Example 3.8, and find the generalized taxicab distance between* $f$ *and* $g$, *where* $f(x) := x^2$ *and* $g(x) := \sin(\frac{\pi}{2}x)$.

(4) ** Refer to Example 3.8, and define the (generalized)* ***max norm*** $|f|_m$ *of* $f \in C([0, 1])$ *to be the absolute maximum value that* $|f(x)|$ *takes. (In symbols,* $|f|_m := max\{|f(x)| : x \in [0, 1]\}$.*) Justify this definition, show the corresponding distance function to be a metric, and verify that* $|f|_t$ *is always* $\leq |f|_m$. *Show by example that for any* $\epsilon > 0$ *there exists some* $f \in C([0, 1])$ *such that* $|f|_m = 1$ *and* $|f|_t < \epsilon$. *(Compare with the statement of Exercise 2.5 (3).)*

(5) *Refer to Example 3.8, and define the (generalized)* ***dot product*** $f \cdot g$ *of* $f, g \in C([0,1])$ *to be the integral* $\int_0^1 f(x)g(x)\,dx$*, and then define the (generalized)* ***euclidean norm*** $|f|$ *(also called the* $L^2$-norm*) to be* $\sqrt{f \cdot f}$*. Use the (generalized) Cauchy-Schwarz inequality* $|f \cdot g| \leq |f||g|$ *to prove that the corresponding distance function is a metric. How do* $|f|$ *and* $|f|_m$ *(see Exercise 3.9 (4)) generally compare?*

(6) *Refer to Example 3.2 (iv), and show that the discrete metric really is a metric (compare with Example 2.4). If* $\langle X, d\rangle$ *and* $\langle Y, e\rangle$ *are metric spaces, and* $d$ *is the discrete metric, what does it mean for a function* $f : X \to Y$ *to be continuous?*

(7) ** Let* $\langle X, d\rangle$ *and* $\langle Y, e\rangle$ *be metric spaces, with* $f : X \to Y$ *a function. Let* $A \subseteq X$ *be equipped with the induced (subspace) metric, and consider the restriction* $f|A$ *of* $f$ *to* $A$ *(i.e.,* $f|A : A \to Y$*, and* $(f|A)(x) := f(x)$*, for* $x \in A$*. Show that if* $x_0 \in A$ *and* $f$ *is continuous at* $x_0$*, then* $f|A$ *is continuous at* $x_0$*. Is the converse always true? More precisely, if* $x_0 \in A$ *and* $f$ *is discontinuous at* $x_0$*, does it necessarily follow that* $f|A$ *is discontinuous at* $x_0$*?*

(8) *Let* $\langle X, d\rangle$ *be a metric space, with* $x_0 \in X$ *fixed. Define* $f : X \to \mathbb{R}$ *by the rule* $f(x) := d(x, x_0)$*, and show* $f$ *is continuous as a function from* $\langle X, d\rangle$ *to* $\mathbb{R}$ *(equipped with the euclidean metric).*

## 4. Some Topology of Metric Spaces

One of the popular definitions of *topology* is "rubber sheet geometry," and suggests the operations you could perform on a plane geometric figure, a triangle, say, (viewed as a union of three line segments, not as the enclosed area) and keep the "essential" qualities of that figure intact. In ordinary plane geometry, these operations are quite restricted; namely you're allowed translations and rotations only. A sequence of such operations is commonly called a **rigid motion**; the result of a rigid motion on a triangle is precisely another triangle of the same shape and size (but possibly moved over and upside down). In this situation the two triangles are said to be **congruent**. In "rubber sheet" geometry, the class of admissible operations is much broader than just the rigid motions. As the word *rubber* suggests, you're allowed lots of stretching and bending (but no tearing). While preserving all the euclidean features of a plane figure is summed up in the word *congruent*, the preservation of all the topological features of that figure is summed up in the word *homeomorphic*. In this course we will see that three-corneredness is not a topological property of a triangle because a square, even a circle, is homeomorphic to a triangle. What *is* a topological property is the fact that a triangle is *connected*; i.e., comes in "one piece." Even after cutting the triangle by removing a point, you still have something that is connected. However, upon the removal of any two points, the result is two connected pieces. This tells us that a triangle, a line segment, and a figure-eight are all in distinct homeomorphism classes; i.e., no two of these three plane figures are topologically "the same."

With the help of topology, we can make all these vague intuitive ideas crystal clear. As a first step, we introduce the basic topological features of metric spaces.

**Definition 4.1** (Neighborhood of a Point)**.** *suppose $\langle X, d\rangle$ is a metric space, $x_0$ is a point of $X$, and $\epsilon > 0$. We denote by $B_d(x_0, \epsilon)$ the set $\{x : d(x, x_0) < \epsilon\}$, the* ***$d$-ball neighborhood with center*** *$x_0$* ***and radius*** *$\epsilon$. (When the metric $d$ is understood, we often drop the subscript and simply write $B(x_0, \epsilon)$.) If $x_0 \in X$ and $A \subseteq X$ are given, we say $A$ is a* ***$d$-neighborhood*** *of $x_0$ if there is some $\epsilon > 0$ such that $B_d(x_0, \epsilon) \subseteq A$. (When confusion is not likely to arise, we suppress mention of the metric; also we frequently abbreviate* neighborhood *with* nbd.*)*

**Example 4.2.** *In the plane $\mathbb{R}^2$, let $e$ and $t$ be the euclidean and taxicab metrics, respectively. Then $B_e(\mathbf{0}, 1) = \{\langle x, y\rangle : \sqrt{x^2 + y^2} < 1\}$, i.e., the open disk, of radius 1, centered at the origin. On the other hand, $B_t(\mathbf{0}, 1) = \{\langle x, y\rangle : |x| + |y| < 1\}$. The easiest way to graph this is to consider each quadrant separately and graph the corresponding equality. In the first quadrant, for example, the bounding line segment is given by $x + y = 1$; in the second quadrant the relevant equality is $-x + y = 1$. The resulting ball then has the shape of a diamond (not including the boundary segments).*

Let us now recast the definition of continuity at a point in the language of neighborhoods. To begin, we introduce some notation regarding functions. Given a function $f : X \to Y$ between sets and a subset $A$ of $X$, the **image of $A$ under $f$** is the set $f[A] := \{f(x) : x \in A\}$. If $B$ is now a subset of $Y$, the **inverse image of $B$ under $f$** is the set $f^{-1}[B] := \{x : f(x) \in B\}$. The following is a simple, but very important, rewording of Definition 3.5.

**Theorem 4.3** (Neighborhood Characterization of Continuity at a Point)**.** *Let $\langle X, d\rangle$ and $\langle Y, e\rangle$ be metric spaces, $x_0 \in X$, and $f : X \to Y$ a function. Then $f$ is continuous at $x_0$ if and only if whenever $E$ is a nbd of $f(x_0)$ in $Y$, there is a nbd $D$ of $x_0$ in $X$ such that $f[D] \subseteq E$.*

*Proof.* Suppose first that $f$ is continuous at $x_0$, and that $E$ is any given nbd of $f(x_0)$. Then there is some $\epsilon > 0$ such that $B_e(f(x_0), \epsilon) \subseteq E$. By continuity, we know there exists a $\delta > 0$ such that whenever $x \in B_d(x_0, \delta)$, we have that $f(x) \in B_e(f(x_0), \epsilon)$. That is, $f[B_d(x_0, \delta)] \subseteq B_e(f(x_0), \epsilon) \subseteq E$. $D := B_d(x_0, \delta)$ is the nbd of $x_0$ we want. For the converse, let $\epsilon > 0$ be given, and set $E := B_e(f(x_0), \epsilon)$. Then there is a nbd $D$ of $x_0$ such that $f[D] \subseteq E$. Let $\delta > 0$ be such that $B_d(x_0, \delta) \subseteq D$. Then $f[B_d(x_0, \delta)] \subseteq B_e(f(x_0), \epsilon)$.

□

**Definition 4.4** (Open Sets and Closed Sets in Metric Spaces)**.** *Let $\langle X, d\rangle$ be a metric space, with $U \subseteq X$. $U$ is a $d$-**open** set (or, simply, an **open** set) if $U$ is a neighborhood of each of its elements. A set $C \subseteq X$ is a **closed set** if the complement of $C$ (i.e., $X \setminus C := \{x \in X : x \notin C\}$) is open.*

So, to check whether $U \subseteq X$ is an open set, one generally picks arbitrary $x_0 \in U$ and tries to find some $\epsilon > 0$ such that $B_d(x_0, \epsilon) \subseteq U$. The usual procedure for checking whether $C \subseteq X$ is closed is to pick $x_0 \in X$ outside $C$ (but arbitrary otherwise) and to try to find $\epsilon > 0$ such that $B_d(x_0, \epsilon)$ misses $C$ altogether. (When sets $A$ and $B$ fail to share any points, they are said to be **disjoint**. This says their intersection is empty; in symbols, $A \cap B = \emptyset$.)

**Theorem 4.5** (Topological Properties of Metrics)**.** *Let $\langle X, d\rangle$ be a fixed metric space.*

(i) *The empty set $\emptyset$ and the universal set $X$ are both open and closed.*
(ii) *If $A$ and $B$ are both open (resp., closed) sets, then $A \cap B$ (resp., $A \cup B$) is also open (resp., closed).*
(iii) *If $\{A_i : i \in I\}$ is any family of open (resp., closed) sets, then the union $\bigcup_{i\in I} A_i$ (resp., the intersection $\bigcap_{i\in I} A_i$) is also open (resp., closed).*
(iv) *For any $x_0 \in X$ and $\epsilon > 0$, $B(x_0, \epsilon)$ is an open set.*
(v) *For any $x_0 \in X$ and $\epsilon > 0$, $B[x_0, \epsilon] := \{x : d(x, x_0) \leq \epsilon\}$ is a closed set containing $B(x_0, \epsilon)$.*
(vi) *Every finite subset of $X$ is closed.*

*Proof. Ad (i)*: See Exercise 4.11 (1) below.

*Ad (ii)*: Assume $A$ and $B$ are both open sets. We need to show that $A \cap B$ is a nbd of each of its points. Given $x_0 \in A \cap B$, we have, since $A$ is a nbd of $x_0$, some $\epsilon_A > 0$ such that $B(x_0, \epsilon_A) \subseteq A$. Similarly, there is some $\epsilon_B > 0$ such that $B(x_0, \epsilon_B) \subseteq B$. These two ball neighborhoods are concentric. Hence, if we let $\epsilon$ be the minimum of $\epsilon_A$ and $\epsilon_B$, we find that $B(x_0, \epsilon)$ is contained in both $A$ and $B$. Therefore $B(x_0, \epsilon) \subseteq A \cap B$.

Now suppose $A$ and $B$ are both closed sets. Then, by definition, $X \setminus A$ and $X \setminus B$ are both open sets; so, by the paragraph above, $(X \setminus A) \cap (X \setminus B)$ is also open. By the classic DeMorgan laws of basic set theory (see Exercise 4.11 (7) below), this intersection is $X \setminus (A \cup B)$. This tells us $A \cup B$ is closed.

*Ad (iii)*: Assume, for each $i$ in our (possibly infinite) index set $I$, $A_i$ is an open set. If $x_0 \in \bigcup_{i \in I} A_i$, then, by definition of set union, there is some $i_0 \in I$ such that $x_0 \in A_{i_0}$. Since $A_{i_0}$ is a nbd of $x_0$, there is some $\epsilon > 0$ with $B(x_0, \epsilon) \subseteq A_{i_0}$. But $A_{i_0}$ is contained in the big union; hence $B(x_0, \epsilon) \subseteq \bigcup_{i \in I} A_i$.

If each $A_i$ is now assumed to be closed, we use the DeMorgan laws again; i.e., $X \setminus (\bigcap_{i \in I} A_i) = \bigcup_{i \in I}(X \setminus A_i)$.

*Ad (iv)*: Suppose $x \in B(x_0, \epsilon)$. We must find a $\delta > 0$ such that $B(x, \delta) \subseteq B(x_0, \epsilon)$. Let $d(x, x_0) = \eta$. Then $0 < \eta < \epsilon$, so $\delta := \epsilon - \eta > 0$. If $y \in B(x, \delta)$, then $d(y, x) < \delta$. So, by the triangle inequality, $d(y, x_0) \leq d(y, x) + d(x, x_0) < \delta + \eta = \epsilon$; so $y \in B(x_0, \epsilon)$.

*Ad (v)*: Containment is obvious; what is less clear is the assertion of closedness. If $x \notin B[x_0, \epsilon]$, then we have $\eta := d(x, x_0) > \epsilon$. If we let $\delta := \eta - \epsilon > 0$, then we claim that $B(x, \delta) \cap B[x_0, \epsilon] = \emptyset$. Indeed, suppose not. Then we have some $y$ in the intersection; so $d(y, x_0) \leq \epsilon$ and $d(y, x) < \delta$. But then, using the triangle inequality again, we have $\eta = d(x, x_0) \leq d(x, y) + d(y, x_0) < \delta + \epsilon = \eta$. A real number cannot be strictly less than itself, so we have a contradiction.

*Ad (vi)*: If $\{x_1, \ldots, x_n\}$ is a finite subset of $X$ and $x_0$ is not in this set, let $\epsilon > 0$ be the minimum of all the finitely many positive distances from $x_0$ to each $x_i$, $1 \leq i \leq n$. Then $B(x_0, \epsilon)$ is disjoint from $\{x_1, \ldots, x_n\}$.

$\square$

**Example 4.6** (Infinite Intersections of Open Sets)**.** *Although finite intersections of open sets are open, infinite intersections hardly ever are. Consider a metric space $\langle X, d \rangle$, with $x_0 \in X$, and consider the infinite family $B(x_0, \frac{1}{k})$, $k = 1, 2, \ldots$. Because of the Archimedean property of the reals, there is no positive real number that is less than $\frac{1}{k}$ for* every *positive natural number $k$; hence we know that $\bigcap_{k \geq 1} B(x_0, \frac{1}{k}) = \{x_0\}$. While it is possible for a singleton subset to be open, it is usually not the case. For example, this never happens in $\mathbb{R}^n$ with the euclidean metric.*

**Definition 4.7** (Limit/Isolated Point of a Set)**.** *Let $\langle X, d \rangle$ be a metric space, $A$ a subset of $X$, and $x_0$ a point of $X$ (not necessarily of $A$). $x_0$ is called a **limit point** of $A$ if every nbd of $x_0$ contains points of $A$ other than $x_0$ itself. If $x_0 \in A$ but $x_0$ is not a limit point of $A$, we call $x_0$ an **isolated point** of $A$.*

The use of the word *limit* in Definition 4.7 may ring some bells: does this have anything to do with limits of sequences? The answer is yes, and we will have more to say on the subject when we talk about convergence in Section 10.

**Examples 4.8** (Sets in the Real Line)**.** *Let us consider various examples of sets in the real line $\mathbb{R}$, equipped with the usual (euclidean) metric.*

(i) *The open intervals $(a, b)$ are open sets because they are ball neighborhoods of their midpoints (see Theorem 4.5 (iv)). Unbounded open intervals, like $(a, \infty)$, are open sets as well because they are unions of bounded ones. I.e., $(a, \infty) = \bigcup_{n \in \mathbb{N}} (a, a + n)$. Every point of $[a, b]$ is a limit point of $(a, b)$. No point of an interval of positive length is an isolated point of that interval.*

(ii) *Intervals that include their endpoints are closed, but not open. For example, $[a, b]$ is closed because its complement in $\mathbb{R}$ is $(-\infty, a) \cup (b, \infty)$, an open set. $[a, b]$ is not open because it is not a nbd of either of its endpoints. ($[a, b]$ is a nbd of each of its other points, though.)*

(iii) *The set $\mathbb{Q}$ of* rational *numbers (i.e., those real numbers that are representable as fractions of integers) is neither open nor closed. This follows from the facts that:*
  (a) *between any two real numbers lies a rational number; and*
  (b) *between any two real numbers lies an irrational number.*

  *For the same reason, every real number is a limit point of $\mathbb{Q}$.*

(iv) *The set $\mathbb{Z}$ if integers is closed, but not open. Furthermore, each integer point is an isolated point of $\mathbb{Z}$. (See Exercise 4.11 (9) below.)*

We now return to our main theme, continuity. The following characterization of continuity at a point most closely resembles our original intuitive treatment in Definition 1.1.

**Theorem 4.9** (Limit Point Characterization of Continuity at a Point)**.** *Let $\langle X, d\rangle$ and $\langle Y, e\rangle$ be metric spaces, $x_0 \in X$, and $f : X \to Y$ a function. Then $f$ is continuous at $x_0$ if and only if whenever $A$ is a subset of $X$ such that $x_0$ is a limit point of $A$, we have that $f(x_0)$ is either a limit point of $f[A]$ or a member of $f[A]$.*

*Proof.* Assuming $f$ is continuous at $x_0$, suppose $A \subseteq X$ has $x_0$ as a limit point. If $f(x_0)$ is neither a limit point of $f[A]$ nor a member of $f[A]$, then there is some nbd $E$ of $f(x_0)$ that is disjoint from $f[A]$. Using Theorem 4.3, and by continuity of $f$ at $x_0$, there is some nbd $D$ of $x_0$ such that $f[D] \subseteq E$. Since $x_0$ is a limit point of $A$, there must be some $x \in D \cap A$, hence $f(x) \in f[D \cap A]$. But (see Exercise 4.11 (5) below) $f[D \cap A] \subseteq f[D] \cap f[A]$, which, in turn, is contained in $E \cap f[A] = \emptyset$. This is a contradiction.

For the converse, suppose $x_0 \in X$ and $E$ is a nbd of $f(x_0)$. We need to find a nbd $D$ of $x_0$ such that $f[D] \subseteq E$. Suppose no such $D$ exists (heaven forbid). Then, for each natural number $k \geq 1$, we have $f[B_d(x_0, \frac{1}{k})]$ not contained in $E$. For each $k \geq 1$, then, let $y_k \in f[B_d(x_0, \frac{1}{k})] \setminus E$ witness this noncontainment. And, once we've done this, pick $x_k \in B_d(x_0, \frac{1}{k})$ such that $y_k = f(x_k)$. No $x_k$ is $x_0$ because the points $f(x_k)$ lie outside of $E$, for $k \geq 1$. Since every nbd of $x_0$ contains $B_d(x_0, \frac{1}{k})$ for some suitably large $k$ (the Archimedean property again), we infer that $A := \{x_1, x_2, \dots\}$ has $x_0$ for a limit point. But then it is not the case that $f(x_0) \in f[A]$; nor is it the case that $f(x_0)$ is a limit point of $f[A]$. This gives us a contradiction.

□

We are now in a position to characterize continuity in terms of several different topological notions.

**Theorem 4.10** (Characterizations of Continuity)**.** *Let $\langle X, d\rangle$ and $\langle Y, e\rangle$ be metric spaces, with $f : X \to Y$ a function. The following are equivalent:*

(a) *$f$ is continuous on $X$.*
(b) *For every $U \subseteq Y$ that is e-open, $f^{-1}[U] \subseteq X$ is d-open.*
(c) *For every $C \subseteq Y$ that is e-closed, $f^{-1}[C] \subseteq X$ is d-closed.*

*Proof. Ad ((a) $\Longrightarrow$ (b))*: Assume that $U \subseteq Y$ is $e$-open. It suffices to prove that $f^{-1}[U]$ is a nbd of each of its points. So let $x_0$ be arbitrily chosen from $f^{-1}[U]$. Then, by definition, $f(x_0) \in U$. Now $U$ is a nbd of $f(x_0)$; so, by the assumption (a), in the form of Theorem 4.3, there is a nbd $D$ of $x_0$ such that $f[D] \subseteq U$. But then $D \subseteq f^{-1}[U]$. Any superset of a nbd of a point is also a nbd of the point. Thus $f^{-1}[U]$ is a nbd of $x_0$.

*Ad ((b) $\Longrightarrow$ (c))*: Supposed $C \subseteq Y$ is $e$-closed. Then $Y \setminus C$ is $e$-open. By assumption (b), $f^{-1}[Y \setminus C]$ is $d$-open. By elementary Boolean properties of inverse images of functions, this set is also $X \setminus f^{-1}[C]$. Thus $f^{-1}[C]$ is $d$-closed.

*Ad ((c) $\Longrightarrow$ (b))*: This is almost exactly the same as the argument in the last paragraph.

*Ad ((b) $\Longrightarrow$ (a))*: It suffices to show that $f$ is continuous at each point $x_0$ of $X$. So let $\epsilon > 0$ be given. Then, by Theorem 4.5 (iv), $B_e(f(x_0), \epsilon)$ is an $e$-open set. By assumption (b), $f^{-1}[B_e(f(x_0), \epsilon)]$ is a $d$-open set containing $x_0$. Hence there is a $\delta > 0$ such that $B_d(x_0, \delta) \subseteq f^{-1}[B_e(f(x_0), \epsilon)]$. This now implies that $f[B_d(x_0, \delta)] \subseteq B_e(f(x_0), \epsilon)$, completing the proof.

□

**Exercises 4.11.** (1) *Show that the empty set and the universal set are always open subsets of any metric space.*

(2) *Suppose that $\langle X, d\rangle$ is a metric space, where $d$ is the discrete metric. What exactly are the open sets in this case? What are the closed sets? What does it mean for a point to be a limit point of a set in the discrete context?*

(3) *Let $e$ and $t$ be the euclidean and the taxicab metrics on the plane $\mathbb{R}^2$. Show that every e-open set is a t-open set, and* vice versa. *If $d$ is now the discrete metric, show every e-open set is a d-open set, but* not vice versa.

(4) *Let $\langle X, d\rangle$ be a metric space, with $x$ and $y$ distinct points of $X$. Show there are open sets $U$ containing $x$ and $V$ containing $y$, such that $U \cap V = \emptyset$. This is called the **Hausdorff property** of metric spaces (after Felix Hausdorff (1868–1942)).*

(5) ** Suppose $f : X \to Y$ is a function between sets, and that $\{A_i : i \in I\}$ is a family of subsets of $X$.*
   (a) *Show $f[\bigcup_{i\in I} A_i] = \bigcup_{i\in I} f[A_i]$.*
   (b) *Show $f[\bigcap_{i\in I} A_i] \subseteq \bigcap_{i\in I} f[A_i]$, but that equality need not hold, even in the case $I$ has just two indices.*
   (c) *If $A \subseteq X$, show by a single example that $f[X \setminus A]$ and $Y \setminus f[A]$ need not be set-theoretically related (i.e., neither need be a subset of the other).*

(6) *Suppose $f : X \to Y$ is a function between sets, and that $\{B_i : i \in I\}$ is a family of subsets of $Y$.*
   (a) *Show $f^{-1}[\bigcup_{i\in I} B_i] = \bigcup_{i\in I} f^{-1}[B_i]$.*
   (b) *Show $f^{-1}[\bigcap_{i\in I} B_i] = \bigcap_{i\in I} f^{-1}[B_i]$.*
   (c) *If $B \subseteq Y$, show that $f^{-1}[Y \setminus B] = X \setminus f^{-1}[B]$.*

(7) *Let $X$ be a set, with $\{A_i : i \in I\}$ a family of subsets of $Y$. Prove the following two* DeMorgan laws*:*
   (a) $X \setminus (\bigcup_{i\in I} A_i) = \bigcap_{i\in I}(X \setminus A_i)$.
   (b) $X \setminus (\bigcap_{i\in I} A_i) = \bigcup_{i\in I}(X \setminus A_i)$.

(8) *Refer to Theorem 4.9 and discuss the effect of replacing "either a limit point of $f[A]$ or a member of $f[A]$" with the shorter "a limit point of $f[A]$."*

(9) *Show that the set $\mathbb{Z}$ if integers is closed, but not open, as a subset of $\mathbb{R}$, equipped with the usual metric. Show further that each integer point is an isolated point of $\mathbb{Z}$.*

## 5. Topologies and Topological Spaces

Given any metric space $\langle X, d\rangle$, there is the associated family of $d$-open sets, as discussed in the last section. All the related notions of *closed set*, *neighborhood of a point*, *limit point of a set*, and so on, may be defined in terms of the open sets alone. (For example, $x \in X$ is a limit point of $A \subseteq X$ if and only if every open set containing $x$ also contains points of $A$ other than $x$.) Just as Theorem 2.2 singled out the *metric* properties of distance functions needed to free up the definition of continuity from the euclidean context, Theorem 4.5 (i, ii, iii) showed us the purely topological features of metrics. Theorems 4.3, 4.9, and 4.10 then served to show us how to to wean continuity notions completely from all dependence on the structure of the real line.

Our first definition parallels Definition 3.1, in that it uses a theorem as its motivation. Essential features of euclidean distance were abstracted to constitute the definition of metric and of metric space. Now we abstract essential features of metrics to constitute the definition of *topology* and *topological space*.

**Definition 5.1** (Topological Space)**.** *Let $X$ be a set. By a **topology on** $X$, we mean a family $\mathcal{T}$ of subsets of $X$ (the $\mathcal{T}$**-open sets**) satisfying the following:*

(T1) *The empty set $\emptyset$ and the universal set $X$ are members of $\mathcal{T}$.*
(T2) *If $A$ and $B$ are both members of $\mathcal{T}$, then so is the intersection $A \cap B$.*
(T3) *If $\{A_i : i \in I\}$ is any family of members of $\mathcal{T}$, then so is the union $\bigcup_{i\in I} A_i$.*

*A **topological space** is a pair $\langle X, \mathcal{T}\rangle$, where $X$ is a set (the underlying set of **points**) and $\mathcal{T}$ is a topology on $X$.*

**Remark 5.2** (On Terminology)**.** *The use of the word* topology *to denote both a mathematical subject area* and *an object of mathematical study may seem a bit peculiar. It seems to be a historical accident that someone coined the word "topology" to apply to the collection of open sets; maybe it would have been better called a "topo," so that the study of such things would be properly called topology (or maybe "topo space theory"). Go figure!*

**Examples 5.3.** (i) *Suppose $\langle X, d\rangle$ is a metric space. Then the collection $\mathcal{T}_d$ of $d$-open sets forms a topology on $X$, called the **metric topology induced by** $d$. If $\langle X, \mathcal{T}\rangle$ is such that $\mathcal{T} = \mathcal{T}_d$ for some metric $d$ on $X$, we say that $\langle X, \mathcal{T}\rangle$ is a **metrizable** topological space. (It often happens that a topology on a set is defined using no mention of a metric, yet is indeed metrizable. A fundamental problem in general topology is to provide purely topological criteria on a topological space that ensure metrizability.)*

(ii) *Let $X$ be any set, and define $\mathcal{T}$ to be the two-set family $\{\emptyset, X\}$. Then this family satisfies the conditions for being a topology; it's called the **trivial** topology, and includes only those sets it's required to. At the opposite extreme is to define $\mathcal{T}$ to be the family of all subsets of $X$. This is called the **discrete** topology:* every *set is open in this topology. The discrete topology is precisely $\mathcal{T}_d$, where $d$ is the discrete metric. (See Exercise 5.12 (1) below.)*

(iii) *Let $X$ be any set, and define a subset $A$ of $X$ to be **cofinite** in $X$ if $X \setminus A$ is finite. We now define $\mathcal{T}$ to be the family consisting of the empty set,* *plus all the sets cofinite in $X$. This is called the* ***cofinite*** *(or sometimes the* ***finite-complement****) topology. Of course, if $X$ is itself finite, then the cofinite topology is just the discrete topology since every subset of a finite set is cofinite in the set. Checking that this is indeed a topology in all cases is straightforward (see Exercise 5.12 (6) below); in the case $X$ is an infinite set, the cofinite topology is not metrizable. To see this, recall Exercise 4.11 (4). Every metrizable topology satisfies the Hausdorff property; i.e., when $x$ and $y$ are two distinct points, there exist open sets $U$ and $V$ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$. However, if $X$ is infinite, and $U$ and $V$ are disjoint nonempty open (i.e., cofinite) sets, then $V$ is a subset of $X \setminus U$, which is finite. That is a contradiction; hence* all nonempty open sets in this topology must overlap.

(iv) *Suppose we start with a topological space $\langle X, \mathcal{T} \rangle$, with $Y$ a fixed subset of $X$. We now form the family $\mathcal{T}_Y$, consisting of sets $U \cup A$, where $U \in \mathcal{T}$ and $A \subseteq Y$. This is a new topology on $X$. Every $\mathcal{T}$-open set $U$ is also $\mathcal{T}_Y$-open because $U = U \cup \emptyset$ and $\emptyset \subseteq Y$. Let's check the Boolean closure conditions for being a topology. If $U \cup A$ and $V \cup B$ are arbitrary members of $\mathcal{T}_Y$, then so is $(U \cup A) \cap (V \cup B) = (U \cap V) \cup ((U \cap B) \cup (A \cap V) \cup (A \cap B))$. Similarly, if $\{U_i \cup A_i : i \in I\}$ is a family of members of $\mathcal{T}_Y$, then so is the union $\bigcup_{i \in I}(U_i \cup A_i) = (\bigcup_{i \in I} U_i) \cup (\bigcup_{i \in I} A_i)$. If $Y = \emptyset$, we get $\mathcal{T}_Y = \mathcal{T}$; if $Y = X$, we get the discrete topology on $X$. The prototypical example of this phenomenon is popularly called the* ***Michael line*** *(after Ernest A. Michael): $\langle X, \mathcal{T} \rangle$ is the usual (i.e., euclidean-metric-induced) topology on the real line, and $Y$ is the set of irrational numbers. (In particular, all irrational singleton sets are open, but no rational singletons are.) This topological space is an important source of examples in more advanced courses.*

**Definition 5.4** (The Subspace Topology)**.** *Let $\langle X, \mathcal{T} \rangle$ be a topological space, $A$ a subset of $X$. Then it makes sense to intersect each $U \in \mathcal{T}$ with $A$. The resulting collection, denoted $\mathcal{T}|A := \{U \cap A : U \in \mathcal{T}\}$ is easily seen to satisfy the conditions for being a topology on $A$, and is called the* ***topology on*** *$A$* ***induced by*** *$\mathcal{T}$, or the* ***subspace topology*** *on $A$. The pair $\langle A, \mathcal{T}|A \rangle$ is called a* ***topological subspace*** *of $\langle X, \mathcal{T} \rangle$; a set of the form $U \cap A$, where $U \in \mathcal{T}$, is frequently called* ***open in*** *$A$ (or* ***open relative to*** *$A$).*

Just in case you were wondering, the subspace metric gives rise to the subspace topology. We take this up in Section 6.

Exactly as with the topological study of metric spaces, we have the same derived notions of neighborhood, closed set, etc. For future reference, we collect them here.

**Definition 5.5** (Derived Topological Notions)**.** *Let $\langle X, \mathcal{T} \rangle$ be a fixed topological space.*

(i) (Neighborhood of a Point) *If $x \in X$ and $A \subseteq X$, we say $A$ is a* ***neighborhood*** *(abbreviated nbd) of $x$ if there is some $U \in \mathcal{T}$, with $x \in U \subseteq A$. (So the open sets are precisely the nbds of each of their points.)*

(ii) (Closed Set) *Let $A \subseteq X$. $A$ is a* ***closed set*** *if $X \setminus A \in \mathcal{T}$.*

(iii) (Limit Point of a Set) *If* $x \in X$ *and* $A \subseteq X$*, we say* $x$ *is a* ***limit point*** *of* $A$ *if every nbd of* $x$ *has points, other than* $x$*, in common with* $A$.
(iv) (Isolated Point of a Set) *if* $x \in X$ *and* $A \subseteq X$*, we say* $x$ *is an* ***isolated point*** *of* $A$ *if* $U \cap A = \{x\}$*, for some nbd* $U$ *of* $x$.

We now have what amounts to a reiteration of Theorem 4.5; its proof is a straightforward use of the DeMorgan laws.

**Theorem 5.6** (Boolean Properties of the Collection of Closed Sets)**.** *Let* $\langle X, \mathcal{T}\rangle$ *be a fixed topological space.*
(i) *Both* $\emptyset$ *and* $X$ *are closed.*
(ii) *If* $A$ *and* $B$ *are both closed sets, then so is the union* $A \cup B$.
(iii) *If* $\{A_i : i \in I\}$ *is any family of closed sets, then so is the intersection* $\bigcap_{i\in I} A_i$.

In addition to the notions in Definition 5.5, essentially repetitions of what was defined in Section 4, we introduce the three new, but highly important ideas of closure, interior, and boundary. These owe their definitions to Theorem 5.6.

**Definition 5.7** (Closure, Interior, and Boundary of a Set)**.** *Let* $\langle X, \mathcal{T}\rangle$ *be a fixed topological space,* $A$ *a subset of* $X$.
(i) (Closure) *The intersection of all closed sets containing* $A$ *is a closed set, called the* ***closure*** *of* $A$ *and denoted* $Cl_{\mathcal{T}}(A)$*; it is the smallest (in the sense of set inclusion) closed set containing* $A$*. A point* $x$ *is in* $Cl_{\mathcal{T}}(A)$ *just in case either* $x \in A$ *or* $x$ *is a limit point of* $A$.
(ii) (Interior) *The union of all open sets contained in* $A$ *is an open set, called the* ***interior*** *of* $A$ *and denoted* $Int_{\mathcal{T}}(A)$*; it is the largest (in the sense of set inclusion) open set contained in* $A$*. A point* $x$ *is in* $Int_{\mathcal{T}}(A)$ *just in case some nbd of* $x$ *is a subset of* $A$.
(iii) (Boundary) $Cl_{\mathcal{T}}(A) \cap Cl_{\mathcal{T}}(X \setminus A)$ *is a closed set, and is called the* ***boundary*** *of* $A$*. This set is denoted* $Bd_{\mathcal{T}}(A)$*. A point* $x$ *is in* $Bd_{\mathcal{T}}(A)$ *just in case every nbd of* $x$ *intersects both* $A$ *and* $X \setminus A$.

(When confusion is unlikely to arise, we drop the subscripts that specify which topology the closures, etc. are relative to.) The Boolean properties of the closure and interior operators behave rather predictably; there is less that can be said about the boundary operator.

**Theorem 5.8** (Boolean Properties of Closure, Interior, and Boundary)**.** *Let* $\langle X, \mathcal{T}\rangle$ *be a fixed topological space, with* $\{A_i : i \in I\}$ *a fixed family of subsets of* $X$.
(i) $Cl(A) = Int(A) = A$ *and* $Bd(A) = \emptyset$ *whenever* $A$ *is either* $\emptyset$ *or* $X$.
(ii) $X \setminus Cl(A) = Int(X \setminus A)$ *and* $X \setminus Int(A) = Cl(X \setminus A)$*, for any* $A \subseteq X$.
(iii) $\bigcup_{i\in I} Cl(A_i) \subseteq Cl(\bigcup_{i\in I} A_i)$*. Equality holds if* $I$ *is finite; equality needn't hold otherwise.*
(iv) $\bigcap_{i\in I} Cl(A_i) \supseteq Cl(\bigcap_{i\in I} A_i)$*. Equality needn't hold, even if* $I$ *is finite.*

(v) $\bigcap_{i\in I} Int(A_i) \supseteq Int(\bigcap_{i\in I} A_i)$. *Equality holds if $I$ is finite; equality needn't hold otherwise.*
(vi) $\bigcup_{i\in I} Int(A_i) \subseteq Int(\bigcup_{i\in I} A_i)$. *Equality needn't hold, even if $I$ is finite.*
(vii) $Bd(A) = Bd(X \setminus A)$ *for any* $A \subseteq X$.
(viii) $Bd(A \cup B) \subseteq (Bd(A) \cap Cl(X \setminus B)) \cup (Bd(B) \cap Cl(X \setminus A))$
(ix) $Bd(A \cap B) \subseteq (Bd(A) \cap Cl(B)) \cup (Bd(B) \cap Cl(A))$

*Proof. Ad (i)*: This is immediate, because both the empty set and the universal set are open and closed relative to any topology.

*Ad (ii)*: $X \setminus \mathrm{Cl}(A)$ is an open set contained in $X \setminus A$, and is hence contained in $\mathrm{Int}(X \setminus A)$ (= the largest open set contained in $X \setminus A$). On the other hand, a point $x$ in $\mathrm{Int}(X \setminus A)$ has a nbd that is contained in $X \setminus A$; hence $x$ has a nbd that is disjoint from $A$. For $x$ to be in $\mathrm{Cl}(A)$, it is necessary for every nbd of $x$ to intersect $A$. Thus $x \notin \mathrm{Cl}(A)$; i.e., $x \in X \setminus \mathrm{Cl}(A)$. The second equality is proved similarly.

*Ad (iii)*: Suppose $x \in \bigcup_{i\in I} \mathrm{Cl}(A_i)$. Then $x \in \mathrm{Cl}(A_j)$ for some $j \in I$. So if $U$ is a nbd of $x$, then $U$ must intersect $A_j$. Thus $U$ must intersect $\bigcup_{i\in I} A_i$. This tells us that $x \in \mathrm{Cl}(\bigcup_{i\in I} A_i)$.

To see that the containment can be proper (i.e., that equality needn't hold always), let our topological space be the real line with the usual topology. Let $I := \{2, 3 \ldots\}$, and set $A_i := [\frac{1}{i}, 1]$, for $i \in I$. Then $\bigcup_{i\in I} \mathrm{Cl}(A_i) = (0, 1]$ $(:= \{x \in \mathbb{R} : 0 < x \leq 1\})$; however $\mathrm{Cl}(\bigcup_{i\in I} A_i) = [0, 1]$. (See Exercise 5.12 (2) below.)

To see that equality does hold when $I$ is finite, let's assume $I := \{1, \ldots, n\}$. We've already shown that $\mathrm{Cl}(A_1) \cup \cdots \cup \mathrm{Cl}(A_n) \subseteq \mathrm{Cl}(A_1 \cup \cdots \cup A_n)$, so let's show the reverse inclusion. But each $\mathrm{Cl}(A_i)$ is a closed set containing $A_i$, and a finite union of closed sets is a closed set. Thus $\mathrm{Cl}(A_1) \cup \cdots \cup \mathrm{Cl}(A_n)$ is a closed set containing $A_1 \cup \cdots \cup A_n$. Since $\mathrm{Cl}(A_1 \cup \cdots \cup A_n)$ is the *smallest* closed set containing $A_1 \cup \cdots \cup A_n$, we infer that $\mathrm{Cl}(A_1) \cup \cdots \cup \mathrm{Cl}(A_n) \supseteq \mathrm{Cl}(A_1 \cup \cdots \cup A_n)$, and so equality holds.

*Ad (iv)*: Each $\mathrm{Cl}(A_j)$ is a closed set containing $\bigcap_{i\in I} A_i$; hence $\bigcap_{i\in I} \mathrm{Cl}(A_i)$ is a closed set containing $\bigcap_{i\in I} A_i$. It therefore contains the closure of the intersection.

To see that equality needn't hold, even when the index set is finite, let our topological space again be the real line with the usual topology, and put $A := [0, 1)$ and $B := (1, 2]$ (usual half-open interval notation). Then $A \cap B = \emptyset$, so $\mathrm{Cl}(A \cap B) = \emptyset$ too. On the other hand, $\mathrm{Cl}(A) \cap \mathrm{Cl}(B) = [0, 1] \cap [1, 2] = \{1\}$.

*Ad (v)*: This is similar to the proof of (iii) above.

*Ad (vi)*: This is similar to the proof of (iv) above.

*Ad (vii)*: This is immediate from the definition of the boundary operator.

*Ad (viii)*: $\mathrm{Bd}(A \cup B) = \mathrm{Cl}(A \cup B) \cap \mathrm{Cl}(X \setminus (A \cup B)) = (\mathrm{Cl}(A) \cup \mathrm{Cl}(B)) \cap \mathrm{Cl}((X \setminus A) \cap (X \setminus B))$, by (iii) above, plus an application of the DeMorgan laws. Now, by (iv) above, the right-hand side is contained in $(\mathrm{Cl}(A) \cup \mathrm{Cl}(B)) \cap \mathrm{Cl}(X \setminus A) \cap \mathrm{Cl}(X \setminus B)$. Using the Boolean distributive law of intersection over union, and using the definition of the boundary operator, the right-hand side is now equal to $(\mathrm{Bd}(A) \cap \mathrm{Cl}(X \setminus B)) \cup (\mathrm{Bd}(B) \cap \mathrm{Cl}(X \setminus A))$.

*Ad (ix)*: This is similar to (viii) above.

□

**Examples 5.9.** (i) *If $\mathcal{T}$ is the discrete topology on $X$, then $Cl(A) = Int(A) = A$ and $Bd(A) = \emptyset$, for each $A \in X$.*

(ii) *If $\mathcal{T}$ is the usual topology on $\mathbb{R}$, then $Cl((a,b]) = [a,b]$, $Int((a,b]) = (a,b)$, and $Bd((a,b]) = \{a,b\}$. Furthermore, $Cl(\mathbb{Q}) = Bd(\mathbb{Q}) = \mathbb{R}$ and $Int(\mathbb{Q}) = \emptyset$.*

(iii) *If $\langle X, d\rangle$ is a metric space, it is tempting to conjecture that $Cl(B_d(x_0,\epsilon))$ is always equal to $B_d[x_0,\epsilon]$ (recall Theorem 4.5 (v)). While the former is always contained in the latter, and equality does indeed hold when $d$ is the euclidean metric in $\mathbb{R}^n$, strict inequality can occur. For example, let $d$ be a discrete metric and let $\epsilon = 1$ (See Exercise 5.12 (7) below).*

Armed with Theorem 4.3, we are now in a position to give a definition of continuity in the general topological context, and to characterize continuity in various topological terms.

**Definition 5.10** (Continuity at a Point in Topological Space)**.** *Let $\langle X, \mathcal{T}\rangle$ and $\langle Y, \mathcal{U}\rangle$ be topological spaces, $x_0 \in X$, and $f : X \to Y$ a function. $f$ is **continuous at** $x_0$ if, whenever $E$ is a nbd of $f(x_0)$ in $Y$, there is a nbd $D$ of $x_0$ in $X$ such that $f[D] \subseteq E$. $f$ is **continuous on** $X$ if $f$ is continuous at each point of $X$.*

The following improves on Theorem 4.10; its proof is almost identical.

**Theorem 5.11** (Characterizations of Continuity)**.** *Let $\langle X, \mathcal{T}\rangle$ and $\langle Y, \mathcal{U}\rangle$ be topological spaces, with $f : X \to Y$ a map. The following are equivalent:*

(a) *$f$ is continuous on $X$.*
(b) *For every $U \subseteq Y$ that is $\mathcal{U}$-open, $f^{-1}[U] \subseteq X$ is $\mathcal{T}$-open.*
(c) *For every $C \subseteq Y$ that is $\mathcal{U}$-closed, $f^{-1}[C] \subseteq X$ is $\mathcal{T}$-closed.*
(d) *For every $A \subseteq X$, $f[Cl_{\mathcal{T}}(A)] \subseteq Cl_{\mathcal{U}}(f[A])$.*

*Proof.* The equivalence of (a), (b), and (c) is an almost word-for-word repeat of the proof of Theorem 4.10. Let's now prove the equivalence of (a) and (d).

*Ad ((a) $\Longrightarrow$ (d))*: Let $A \subseteq X$ be given, and suppose $y_0 \in f[\mathrm{Cl}_{\mathcal{T}}(A)]$. Then we may choose $x_0 \in \mathrm{Cl}_{\mathcal{T}}(A)$ such that $y_0 = f(x_0)$. Let's assume, for the sake of contradiction, that $y_0 \notin \mathrm{Cl}_{\mathcal{U}}(f[A])$. Then there is a nbd $E$ of $y_0$ in $Y$ such that $E \cap f[A] = \emptyset$. By continuity at $x_0$, there is a nbd $D$ of $x_0$ in $X$ such that $f[D] \subseteq E$. This is a contradiction, since $x_0 \in \mathrm{Cl}_{\mathcal{T}}(A)$, and so $D$ must intersect $A$. This proves that $f[\mathrm{Cl}_{\mathcal{T}}(A)] \subseteq \mathrm{Cl}_{\mathcal{U}}(f[A])$.

*Ad ((d) $\Longrightarrow$ (a))*: Let $x_0 \in X$ be given, with $E$ a nbd of $f(x_0)$ in $Y$. We need to find a nbd $D$ of $x_0$ with $f[D] \subseteq E$. Assuming this is not the case, we have, for each nbd $D$ of $x_0$, a point $x_D \in D$ such that $f(x_D) \notin E$. Let $A := \{x_D :$

$D$ is a nbd of $x_0\}$. Then every nbd $D$ of $x_0$ contains $x_D$ (at least), so must intersect $A$. Thus $x_0 \in \mathrm{Cl}_{\mathcal{T}}(A)$. By our assumption (d), though, we have $f(x_0) \in \mathrm{Cl}_{\mathcal{U}}(f[A])$. But $E$ is a nbd of $f(x_0)$ that is disjoint from $f[A]$, by the definition of the points $x_D$ that constitute $A$. This is a contradiction. Hence there must be some nbd $D$ of $x_0$ such that $f[D] \subseteq E$.

□

**Exercises 5.12.** (1) *Show that the discrete metric on a set $X$ (as defined in Example 3.2 (iv)) gives rise to the discrete topology on $X$ (as defined in Example 5.3 (ii)). Is there another metric that gives rise to the discrete topology?*

(2) *Give $\mathbb{R}$ the usual topology, and show that $Cl(\bigcup_{n\geq 2}(\frac{1}{n}, 1]) = [0, 1]$.*

(3) ** Show by example that the equalities $Bd(A \cup B) = Bd(A) \cup Bd(B)$ and $Bd(A \cap B) = Bd(A) \cap Bd(B)$ need not always hold.*

(4) *An open set in a topological space is called **regular-open** if it is the interior of its own closure. Give an example of an open set in the real line that is regular-open, and an example of an open set in the real line that is not. How are $U$ and $Int(Cl(U))$ generally related when $U$ is open?*

(5) *A subset $D$ of a topological space $X$ is **dense in** $X$ if every nonempty open set intersects $D$. Show that $\mathbb{Q}$ is a dense subset of $\mathbb{R}$, in the usual topology, and that the intersection of two dense open sets is generally a dense open set. Show further that the intersection of two dense sets can be empty.*

(6) *Refer to Example 5.3 (iii) and show that, for any set $X$, the collection of cofinite subsets, together with the empty set, constitute a topology on $X$.*

(7) *Refer to Example 5.9 (iii) and show that the equality $Cl_d(B_d(x_0, \epsilon)) = B_d[x_0, \epsilon]$ need not always hold for metric spaces. How are the two sets related in general?*

(8) *Let $X$ be a two-element set. How many families of subsets are there? Which ones are topologies? Now repeat the exercise when $X$ is a three-element set.*

(9) *Let $\langle X, \mathcal{T}\rangle$ be a topological space, with $A \subseteq X$. Show all $\mathcal{T}|A$-closed subsets of $A$ to be of the form $C \cap A$, where $C$ is a $\mathcal{T}$-closed subset of $X$. Furthermore, if $B \subseteq A$, show $Cl_{\mathcal{T}|A}(B) = Cl_{\mathcal{T}}(B) \cap A$.*

(10) *Let $\langle X, \mathcal{T}\rangle$ be a topological space, with $A \subseteq X$. Show that if $A$ is $\mathcal{T}$-open (resp., $\mathcal{T}$-closed) and $B \subseteq A$ is $\mathcal{T}|A$-open (resp., $\mathcal{T}|A$-closed), then $B$ is $\mathcal{T}$-open (resp., $\mathcal{T}$-closed).*

## 6. Comparison of Topologies on a Set

As we have seen, a topology on a set is a particularly well-behaved collection of subsets of that set. So far we have seen topologies arising from metrics; also, with the definition of the cofinite topology on an infinite set, we have examples of topologies that don't arise in this way. In later sections we will explore other important ways of devising topologies; before we do, however, we must take the time to consider the very important issue of how two topologies on the same set may be compared.

Given a set $X$, denote by $\wp(X)$ the collection of all subsets of $X$, the **power set of** $X$. We have seen this collection before, under a different name: the discrete topology on $X$. The discrete topology on $X$ is the *largest* topology on $X$, in the sense that every topology on $X$ (indeed, every collection of subsets of $X$, topology or not) is a subset of $\wp(X)$. Another word to be used in this context beside *largest* is *finest*. Here is the official definition.

**Definition 6.1** (Comparison of Topologies)**.** *If $\mathcal{T}$ and $\mathcal{U}$ are two topologies on a set $X$, we say $\mathcal{T}$ **is coarser than** $\mathcal{U}$ (equivalently, $\mathcal{U}$ **is finer than** $\mathcal{T}$) if $\mathcal{T} \subseteq \mathcal{U}$.*

When $\mathcal{T}$ is coarser than $\mathcal{U}$, every $\mathcal{T}$-open set is a $\mathcal{U}$-open set, but not necessarily *vice versa*. One way to detect this is to proceed as follows: First arbitrarily pick a point $x \in X$ and a $\mathcal{T}$-open nbd $V$ of $x$. Next, try to find a $\mathcal{U}$-open set $U$ with $x \in U \subseteq V$. If this is always possible, we have each $\mathcal{T}$-open set represented as a union of $\mathcal{U}$-open sets, and hence itself a $\mathcal{U}$-open set.

**Examples 6.2.** (i) *The trivial topology $\{\emptyset, \{X\}\}$ is coarser than every topology on $X$, the discrete topology $\wp(X)$ finer.*

(ii) *Let $d$ be any metric on $X$, with $\mathcal{T}$ the cofinite topology on $X$. Then $\mathcal{T}$ is coarser than $\mathcal{T}_d$. Indeed, if $x \in V \in \mathcal{T}$, then there is some finite set $F := \{x_1, \ldots, x_n\}$ such that $V = X \setminus F$. Let $\epsilon := min\{d(x, x_i) : 1 \leq i \leq n\}$. Then $\epsilon > 0$ and we have $x \in B_d(x, \epsilon) \subseteq V$. If $X$ is infinite, then (see Example 5.3 (iii)) any two nonempty $\mathcal{T}$-open sets must overlap; so $\mathcal{T}$ is not metrizable, since it does not satisfy the Hausdorff property (see Exercise 4.11 (4)). In particular, $\mathcal{T}$ is* strictly *coarser than $\mathcal{T}_d$.*

**Definition 6.3** (Comparison of Metrics)**.** *Let $d$ and $e$ be two metrics on a set $X$. We say $d$ **is coarser** (resp., **finer**) **than** $e$ if $\mathcal{T}_d$ is coarser (resp., finer) than $\mathcal{T}_e$. We say $d$ and $e$ are **equivalent metrics** if each is coarser than the other.*

**Example 6.4** (Equivalence of Metrics in the Plane)**.** *We recall Example 3.7. There we showed that a real-valued function defined on $\mathbb{R}^2$ is continuous with respect to the euclidean metric $e$ if and only if it is continuous with respect to the taxicab metric $t$. Here we show that that is true precisely because the two metrics are equivalent, a fact that also depends on the result of Exercise 2.5 (3), namely that the inequalities $|\mathbf{x}| \leq |\mathbf{x}|_t \leq 2|\mathbf{x}|$ always hold.*

*Given the e-open set $U$ and $\mathbf{x} \in U$, first pick $\epsilon > 0$ such that $B_e(\mathbf{x}, \epsilon) \subseteq U$. Then $B_t(\mathbf{x}, \epsilon) \subseteq B_e(\mathbf{x}, \epsilon)$. Since (Theorem 4.5 (iv)) t-ball nbds are t-open sets, we have* *shown that $\mathcal{T}_t$ is finer than $\mathcal{T}_e$. The opposite direction is shown similarly (here we use $B_e(\mathbf{x}, \frac{\epsilon}{2}) \subseteq B_t(\mathbf{x}, \epsilon)$), so we infer that $\mathcal{T}_t = \mathcal{T}_e$.*

Given a metric space $\langle X, d\rangle$ and a subset $A$ of $X$, there are two clear ways of obtaining a "subspace topology" on $A$. First we could restrict $d$ to $A$ and then take the topology from the restricted metric; second we could take the metric topology associated with $d$ and then restrict that topology to $A$. The happy news is that we get the same topology on $A$ either way.

**Theorem 6.5** (Subspace Metric vs Subspace Topology)**.** *Let $\langle X, d\rangle$ be a metric space, $A$ a subset of $X$. Then the metric topology on $A$ induced by $d|A$ coincides with the subspace topology on $A$ induced by the metric topology on $X$. In symbols, $\mathcal{T}_{d|A} = \mathcal{T}_d|A$.*

*Proof.* By the definition of restricted metric, if $x \in A$ and $\epsilon > 0$, then $B_{d|A}(x, \epsilon) = B_d(x, \epsilon) \cap A$. So suppose $U \in \mathcal{T}_{d|A}$, with $x \in U$. Then there is some $\epsilon > 0$ with $B_d(x, \epsilon) \cap A \subseteq U$. Since $B_d(x, \epsilon) \cap A$ is $\mathcal{T}_d|A$-open, we infer that $\mathcal{T}_{d|A} \subseteq \mathcal{T}_d|A$. The reverse inclusion is just as easy.

□

In analysis the concept of boundedness is of central importance; its definition is most easily couched in the language of metric spaces.

**Definition 6.6** (Bounded Sets and Bounded Metrics)**.** *Let $\langle X, d\rangle$ be a metric space, $A$ a subset of $X$. Then $A$ is $d$-**bounded** if $A$ is contained in some $d$-ball neighborhood. $d$ is a **bounded** metric if $d[X \times X]$ is a set of real numbers that is bounded with respect to the euclidean metric on $\mathbb{R}$.*

**Examples 6.7.**

(i) *While each two points are a finite distance apart in $\mathbb{R}^n$ with the euclidean metric $e$, there is no upper bound to how far apart points can be. Thus $e$ is an unbounded metric; $\mathbb{R}^n$ itself is an unbounded set in this metric.*

(ii) *The discrete metric $d$ on a set $X$ is a bounded metric; $B_d(x, \epsilon) = X$ just in case $\epsilon > 1$ (or $X$ consists of a single point and $\epsilon$ is arbitrary).*

The message of the next result is that boundedness is not "topological."

**Theorem 6.8** (Existence of Bounded Metrics)**.** *Let $\langle X, d\rangle$ be a metric space. then there is a bounded metric $\overline{d}$ on $X$ that is equivalent to $d$; i.e., such that $\mathcal{T}_{\overline{d}} = \mathcal{T}_d$.*

*Proof.* Given $d$, we simply define $\overline{d}$ as a "truncated" version of $d$: if $d(x, y) < 1$, $\overline{d}(x, y) := d(x, y)$; otherwise, $\overline{d}(x, y) := 1$.

Let's first show $\overline{d}$ is a metric. Clearly positivity and symmetry give us no problem, but the triangle inequality needs some scrutiny. Suppose we're given $x, y, z \in X$; let's show $\overline{d}(x, y)) > \overline{d}(x, z) + \overline{d}(z, y)$ can never hold. Indeed, the only time $\overline{d}(x,y)$ differs from $d(x,y)$ is when $d(x,y) \geq 1$ and $\overline{d}(x,y) = 1$. Since the triangle inequality holds for $d$, the only way the contrary inequality above can occur is for the left-hand side to be 1 and the right-hand side to be $< 1$. But then the right-hand side equals $d(x,z) + d(z,y)$, which is $\geq d(x,y) = \overline{d}(x,y)$. This is a contradiction, so the triangle inequality does indeed hold for $\overline{d}$.

Now suppose $U$ is a $d$-open set, with $x \in U$. Find $\epsilon > 0$ such that $B_d(x,\epsilon) \subseteq U$. We lose no generality in assuming $\epsilon \leq 1$, so $B_d(x,\epsilon) = B_{\overline{d}}(x,\epsilon)$. This shows $\overline{d}$ is finer than $d$. The argument showing $d$ to be finer than $\overline{d}$ is almost identical; hence $d$ and $\overline{d}$ are equivalent metrics.

□

We end this section with a metric construction that has proven to be invaluable in the pursuit of general theorems that say when a given topological space is metrizable.

**Example 6.9** (Another Function Space)**.** *For any nonempty set $I$ we denote by $\mathbb{R}^I$ the set of all functions $f : I \to \mathbb{R}$. (This notation is consistent with the notation $\mathbb{R}^n$ for euclidean space because an $n$-tuple is nothing more than a function from $\{1,\ldots,n\}$ to $\mathbb{R}$.) Letting $\overline{e}$ be the truncated euclidean metric on $\mathbb{R}$, we have the set $\{\overline{e}(f(i),0) : i \in I\}$ for each $f \in \mathbb{R}^I$, a nonempty set of reals that is bounded above by 1. By the completeness property, this set has a least upper bound, which we denote by $|f|_s$. This is often called the* ***supremum norm****, not to be confused with the generalized max norm of Exercise 3.9 (4) to which it is related. (Even though a supremum exists, a maximum may not; viz., $f(x) := \frac{2}{\pi}\arctan(x)$ in $\mathbb{R}^{\mathbb{R}}$.) When $I$ is finite, say with $n$ elements, we obtain euclidean $n$-space with the truncated max norm. This gives rise to the usual (euclidean) topology on $\mathbb{R}^n$ (see Exercise 6.10 (4) below).*

**Exercises 6.10.** (1) *Let $\langle X, \mathcal{T}\rangle$ be a metrizable topological space, with $A$ a subset of $X$. Show that $\mathcal{T}|A$ is a metrizable topology on $A$.*

(2) *Let $X$ be a set, with $\iota_X : X \to X$ denoting the identity map; i.e., $\iota_X(x) := x$ for $x \in X$. Let $\mathcal{T}$ and $\mathcal{U}$ be two topologies on $X$. Show that $\mathcal{T} \subseteq \mathcal{U}$ if and only if $\iota_X$ is continuous as a function from $\langle X, \mathcal{U}\rangle$ to $\langle X, \mathcal{T}\rangle$.*

(3) ** Refer to Example 3.8 and to Exercise 3.9 (4). Show that the taxicab metric is* strictly *finer than the max metric on $C([0,1])$. [Hint: Use Exercise 3.9 (4) to show that $B_t(0,\epsilon)$ is not contained in $B_m(0,\frac{1}{2})$ for any $\epsilon > 0$.]*

(4) ** Refer to Example 6.9 and show that when $I = \{1,\ldots,n\}$, the metric arising from the supremum norm induces the usual topology on euclidean $n$-space.*

(5) *Let $X$ be a set with exactly two elements. List all topologies on $X$, indicating how any two of them relate on the coarse-fine scale. Repeat this exercise for a set with exactly three elements.*

(6) *Let $X$ be a set, with $\mathcal{T}$ and $\mathcal{U}$ two topologies on $X$. Show that $\mathcal{T}$ is coarser than $\mathcal{U}$ if and only if: for every function $f$ with domain $X$, if $f$ is continuous with respect to $\mathcal{T}$, then $f$ is continuous with respect to $\mathcal{U}$.*

(7) *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $A$ a subset of $X$, and let $\iota_X|A : A \to X$ denote the inclusion map; i.e., the identity map $\iota_X$ on $X$, restricted to $A$. Let $\mathcal{U}$ be a topology on the set $A$. Show that $\mathcal{U}$ is finer than $\mathcal{T}|A$ if and only if $\iota_X|A$ is continuous with respect to $\mathcal{U}$.*

(8) *Let $\{\mathcal{T}_i : i \in I\}$ be a family of topologies on a set $X$. Show that $\bigcap_{i\in I} \mathcal{T}_i$ is also a topology on $X$, coarser than each $\mathcal{T}_i$.*

(9) *Show by example that the union of two topologies on a set is* not *necessarily a topology on the set. Use Exercise 6.10 (8) above to show that if $\{\mathcal{T}_i : i \in I\}$ is any family of topologies on a set $X$, then there is a smallest topology $\mathcal{T}$ that is finer than each $\mathcal{T}_i$.*

## 7. Bases and Subbases for Topologies

Recall that when we defined the topology $\mathcal{T}_d$ from the metric $d$ on $X$, we used the $d$-ball neighborhoods $B_d(x,\epsilon)$ in a key way: $U \in \mathcal{T}_d$ just in case, for each $x \in U$, there is an $\epsilon > 0$ such that $B_d(x,\epsilon) \subseteq U$. This says that the $\mathcal{T}_d$-open sets are precisely the unions of $d$-ball neighborhoods. This example gives a powerful impetus to isolate the role that is played by these special open sets.

**Definition 7.1** (Base for a Topology)**.** *Let $\mathcal{T}$ be a topology on a set $X$, with $\mathcal{B}$ a subfamily of $\mathcal{T}$. $\mathcal{B}$ is called a* ***base*** *for $\mathcal{T}$ (sometimes the word* basis *is used instead of* base*) if each $\mathcal{T}$-open set is a union of sets in $\mathcal{B}$; equivalently, if for each $U \in \mathcal{T}$ and each $x \in U$, there is a $B \in \mathcal{B}$ with $x \in B \subseteq U$.*

**Examples 7.2.**
(i) *Let $\langle X, d\rangle$ be a metric space, and define $\mathcal{B}$ to consist of all $d$-ball neighborhoods of the form $B(x,\frac{1}{n})$, where $x \in X$ and $n \in \{1,2,\dots\}$. Then $\mathcal{B}$ is a base for $\mathcal{T}_d$. Indeed, given $U \in \mathcal{T}_d$ and $x \in U$, we first find $\epsilon > 0$ such that $B(x,\epsilon) \subseteq U$. Next, using the Archimedean property of the reals, we find a positive natural number $n$ large enough so that $\frac{1}{n} < \epsilon$. then we have $B(x,\frac{1}{n}) \subseteq B(x,\epsilon) \subseteq U$, as desired.*
(ii) *Let $\mathcal{T}$ be the discrete topology on $X$. Then $\mathcal{B} := \{\{x\} : x \in X\}$ is a base for $\mathcal{T}$.*

The question naturally arises as to the specification of conditions on a family $\mathcal{B}$ of subsets of $X$ that characterize when that family is a base for *some* topology on $X$. Is every such family a topological base? By taking a hint from the examples in 7.2, we know there must be some restrictions, but they're not necessarily those that define a topology. After all, the collection of ball neighborhoods for the euclidean metric on the plane (i.e., the usual open disks) satisfies *none* of the defining conditions for a topology: all such nbds are nonempty; they're all bounded, so $\mathbb{R}^2$ is not a ball nbd; and neither the intersection nor the union of two disjoint open disks is necessarily another open disk. Surprisingly, there are very simple criteria that characterize topological bases.

**Theorem 7.3** (Characterizing Bases)**.** *Let $\mathcal{B}$ be a family of subsets of a set $X$. Then $\mathcal{B}$ is a base for some topology on $X$ if and only if $\mathcal{B}$ satisfies the following two conditions:*
(B1) *For each $x \in X$, there is some $B \in \mathcal{B}$ such that $x \in B$.*
(B2) *If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then there is some $B \in \mathcal{B}$ with $x \in B \subseteq B_1 \cap B_2$.*

*Proof.* Let's first suppose that $\mathcal{B}$ is a base for topology $\mathcal{T}$. if $x \in X$, then, since $X \in \mathcal{T}$, we have some $B \in \mathcal{B}$ such that $x \in B$. Thus (B1) holds. If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then, since $B_1$ and $B_2$ are also in $\mathcal{T}$, their intersection is $\mathcal{T}$-open. Hence there is some $B \in \mathcal{B}$ with $x \in B \subseteq B_1 \cap B_2$, showing (B2) holds.

For the converse, assume $\mathcal{B}$ satisfies both (B1) and (B2). We show that $\mathcal{T} := \{\bigcup \mathcal{C} : \mathcal{C} \subseteq \mathcal{B}\}$ (i.e., all unions of subfamilies of $\mathcal{B}$) is a topology for which $\mathcal{B}$ is a base. (Obviously, by the definition of $\mathcal{T}$, $\mathcal{B}$ is a base for $\mathcal{T}$ once we show $\mathcal{T}$ is actually a topology.)

*Ad (T1)*: Let $\mathcal{C}$ be the empty subfamily of $\mathcal{B}$. Then the union of all sets in that family is empty; hence we have $\emptyset \in \mathcal{T}$. If $\mathcal{C} = \mathcal{B}$, then $\bigcup \mathcal{C} = X$ because of (B1). Thus $X \in \mathcal{T}$.

*Ad (T3)*: Let $\{C_i : i \in I\}$ be a family of members of $\mathcal{T}$. We need to show $\bigcup_{i\in I} C_i \in \mathcal{T}$. By the definition of $\mathcal{T}$, there is, for each $i \in I$, a subfamily $\mathcal{C}_i$ of $\mathcal{B}$, the union of whose members is $C_i$. Let $\mathcal{C}$ be the union $\bigcup_{i\in I} \mathcal{C}_i$, a subfamily of $\mathcal{B}$. Then $\bigcup_{i\in I} C_i = \bigcup \mathcal{C}$, another member of $\mathcal{T}$.

*Ad (T2)*: Let $C, D \in \mathcal{T}$. We need to show $C \cap D \in \mathcal{T}$. By the definition of $\mathcal{T}$, we pick $\{C_i : i \in I\}$ and $\{D_j : j \in J\}$, subfamilies of $\mathcal{B}$, such that $C = \bigcup_{i\in I} C_i$ and $D = \bigcup_{j\in J} D_j$. By the Boolean distributivity of finite intersection over arbitrary union, we have $C \cap D = \bigcup_{\langle i,j\rangle \in I\times J}(C_i \cap D_j)$. Now, by (B2), each $C_i \cap D_j$ is a union of members of $\mathcal{B}$, and is hence in $\mathcal{T}$. Hence $C \cap D$ is a union of members of $\mathcal{T}$, and is therefore a member of $\mathcal{T}$ by (T3) just proved.

$\square$

The next simple result shows how we may restrict bases to subsets and obtain subspace topologies.

**Theorem 7.4** (Restricting Bases). *Let $\mathcal{B}$ be a base for the topology $\mathcal{T}$ on $X$, and suppose $A$ is a subset of $X$. Then the collection $\mathcal{B}|A := \{B \cap A : B \in \mathcal{B}\}$, the **restriction of** $\mathcal{B}$ **to** $A$, is a base for the subspace topology $\mathcal{T}|A$.*

*Proof.* If $U \cap A$ is a typical relativized open set in $\mathcal{T}|A$, then, since $\mathcal{B}$ is a base for $\mathcal{T}$, there is some family $\{B_i : i \in I\} \subseteq \mathcal{B}$ with $U = \bigcup_{i\in I} B_i$. But then $U \cap A = (\bigcup_{i\in I} B_i) \cap A = \bigcup_{i\in I}(B_i \cap A)$. This shows that $\mathcal{B}|A$ is a base for $\mathcal{T}|A$.

$\square$

Besides metrics, there are other sources of topologies; and it is nice to be able to specify those topologies via easily-defined bases. One of the most fruitful of these sources is linear orderings.

**Definition 7.5** (Linear Orderings). *Let $X$ be any set. A **linear ordering** on $X$ is a binary relation $R$ on $X$ (i.e., $R$ consists of pairs of elements of $X$, $R \subseteq X\times X$) that satisfies the following three conditions. (We write $xRy$ as shorthand for $\langle x, y\rangle \in R$.)*

(L1) (Irreflexivity) *$xRx$ never holds.*
(L2) (Transitivity) *If $x, y, z \in X$, $xRy$, and $yRz$, then $xRz$.*
(L3) (Trichotomy) *For any $x, y \in X$, either $xRy$, or $x = y$, or $yRx$.*

*A **linearly ordered set** is a pair $\langle X, R\rangle$, where $R$ is a linear ordering on $X$. When confusion is unlikely to occur, we abuse notation and write $<$ (or $<_X$) for $R$ when $R$ is a linear ordering.*

**Examples 7.6.** (i) *The usual ordering on the real line $\mathbb{R}$ is a linear ordering, as are the restricted orderings on the rational numbers $\mathbb{Q}$, the integers $\mathbb{Z}$, and the natural numbers $\mathbb{N}$.*

(ii) *Review the definition of* ordered field *in Section 1. The ordering is disguised as a subset $P$ of "positive" elements of the field, and we write $x < y$ to mean that $y - x \in P$. Let's check that the conditions for being a linear ordering are satisfied:*

Ad (L1)*: If $x < x$ were true, then $0 = x - x$ would be in $P$. This contradicts axiom (O1) for ordered fields.*

Ad (L2)*: Suppose $x < y$ and $y < z$. Then $y - x \in P$ and $z - y \in P$; so, by ordered field axiom (O2) $(y - x) + (z - y) = z - x \in P$. Hence $x < z$.*

Ad (L3)*: Suppose $x$ and $y$ are distinct elements of the field. Then $x - y \neq 0$. By axiom (O3), then, either $x - y \in P$ (in which case $y < x$) or $y - x = -(x - y) \in P$ (in which case $x < y$).*

(iii) *In the plane $\mathbb{R}^2$, define $\langle x, y\rangle < \langle u, v\rangle$ to mean $x < y$ and $u < v$. This attempt at a linear ordering on $\mathbb{R}^2$ satisfies (L1) and (L2), but not (L3): $\langle 0, 1\rangle$ and $\langle 1, 0\rangle$ are not comparable in this ordering, for example. (Relations that satisfy (L1) and (L2) are called **partial orderings**. One of the most important sources of partial orderings is the power set of a set, with the ordering being inclusion.)*

(iv) *Back to the plane, we now define $\langle x, y\rangle << \langle u, v\rangle$ just in case: either $x < u$, or $x = u$ and $y < v$. This is called the **lexicographic ordering** on the plane. Let's check that the axioms hold for $<<$. We leave (L1) and (L2) to the reader; as for (L3), suppose $\langle x, y\rangle$ and $\langle u, v\rangle$ are given. If they're unequal, we could have $x < u$; then $\langle x, y\rangle << \langle u, v\rangle$. If $u < x$, the opposite inequality holds for $<<$. Suppose they're unequal and $x = u$. Then $y \neq v$; so either $y < v$ or $v < y$. In either case the two pairs are $<<$-comparable.*

Now we connect linear orderings with topology. The construction of a topology from a linear ordering mostly mimics how we define the usual topology on the real line via bounded open intervals. However, due to the very non-typical nature of the real line as a linearly ordered set (i.e., because we do not get to assume order completeness), we need to sharpen our terminology.

**Definition 7.7** (Convex Sets and Intervals)**.** *Given a linearly ordered set $\langle X, <\rangle$, a subset $C$ of $X$ is called **convex** if, whenever $x < y < z$ is true and both $x$ and $z$ are in $C$, then $y \in C$ too. (So the convex subsets of a linearly ordered set are the subsets that are "closed under betweenness.") A convex set $C$ is **bounded below** (resp., **bounded above**) if there is an element $x \in X$ satisfying $x \leq y$ (resp., satisfying $x \geq y$) for each $y \in C$. $C$ is **bounded** if it is bounded both below and above.*

*A convex set $I$ is called an **interval** if it is either empty; or, if not and bounded below (resp., above), then it has a greatest lower (resp., least upper) bound. Because of the trichotomy property of linear orders, a greatest lower (resp., least upper) bound is unique, when it exists, and is called the **left** (resp., **right**) **end point** of $I$. An interval may be bounded, in which case it takes one of the four forms: $(a, b) := \{x \in X : a < x < b\}$ (bounded open); $[a, b] := \{x \in X : a \leq x \leq b\}$ (bounded closed); $[a, b) := \{x \in X : a \leq x < b\}$ (bounded half-open-right); or $(a, b] := \{x \in X : a < x \leq b\}$ (bounded half-open-left). An interval is called a **right-looking ray** if it is either of the form $(a, \infty) := \{x \in X : a < x\}$ (open) or of the form $[a, \infty) := \{x \in X : a \leq x\}$ (closed). (The notion **left-looking ray*** *is defined similarly.) A **ray**, then is a set that is either a right-looking ray or a left-looking ray. (Note: rays may or may not be bounded sets.)*

**Remarks 7.8.** (i) *The set $(\sqrt{2}, \sqrt{3})$ is a bounded open interval in $\mathbb{R}$; its end points are the irrational numbers $\sqrt{2}$ and $\sqrt{3}$. The set $(\sqrt{2}, \sqrt{3}) \cap \mathbb{Q}$ is, therefore, a bounded convex subset of $\mathbb{Q}$ that fails to be an interval. It is, nevertheless, an open set in the order topology on $\mathbb{Q}$.*

(ii) *In the case of the real line, the distinction between intervals and convex sets melts away because of the least upper bound property (alias completeness). In this situation a convex set is either:*
- *all of $\mathbb{R}$;*
- *a right-looking ray;*
- *a left-looking ray; or*
- *a bounded interval.*

**Theorem 7.9** (Order Topologies)**.** *Let $\langle X, < \rangle$ be a linearly ordered set, and let $\mathcal{B}$ consist of all bounded open intervals, all open rays, and $X$ itself. Then $\mathcal{B}$ is a base for a topology on $X$, the **order topology $\mathcal{T}_<$ induced by** $<$.*

*Proof.* All we need to do is verify (B1) and (B2) for $\mathcal{B}$. (B1) is trivial, since we're including $X$ for good measure: every member of $X$ is contained in a member of $\mathcal{B}$. As for (B2), we can actually show $\mathcal{B}$ is closed under finite intersections: the intersection of two bounded open intervals is a (possibly empty) bounded open interval; the intersection of two open rays looking in opposite directions is a bounded open interval; the intersection of two open rays looking in the same direction is also an open ray looking in that direction.

□

**Remarks 7.10.** *Referring to Theorem 7.9 above:*

(i) *If $X$ consists of a single point, then all open rays and bounded open intervals are empty. Hence we must include $X$ in $\mathcal{B}$ to cover this one degenerate case.*

(ii) *If $X$ has more than one point, but has, say, a maximal (or minimal) element $m$, then no bounded open interval will be a nbd of $m$. This requires the inclusion of open rays in $\mathcal{B}$ to cover the case where a nondegenerate linearly ordered set has at least one end point.*

(iii) *If $X$ has neither a maximal nor a minimal element, then we may take $\mathcal{B}$ to consist of just the bounded open intervals.*

**Definition 7.11** (Orderable and Suborderable Spaces)**.** *A topological space $\langle X, \mathcal{T} \rangle$ is **orderable** if there is a linear ordering $<$ on $X$ such that the collection consisting of bounded open intervals, open rays, plus $X$ itself, forms a base for $\mathcal{T}$. A space is **suborderable** if it can be viewed as a subspace of an orderable space.*

**Remarks 7.12.** (i) *The definition of* suborderable *in Definition 7.11 is slightly vague; we will correct this in Section 8 when we properly introduce the idea of homeomorphism.*

(ii) *Although most of the familiar examples of orderable spaces have metrics associated with them that give rise to the same topology, there are examples of orderable spaces that are not metrizable, and* vice versa. *Also it turns out that while topological subspaces of metrizable spaces are metrizable (see Exercise 6.10 (1) above), the corresponding statement does not hold for orderability: topological subspaces of orderable spaces may not themselves be orderable. We will see examples supporting these claims in the sequel, once we have developed the necessary machinery.*

As we have seen, both metrics and linear orderings give rise to topologies. We saw in Theorem 6.5 that the result of restricting a metric to a subset and then taking the associated metric topology is the same as that of first taking the metric topology on the big set and then restricting that topology to the subset. Can the same be said for linear orderings? The answer turns out to be no, but the two subset topologies are comparable, nonetheless. By way of notation, if $\langle X, <\rangle$ is a linearly ordered set and $A \subseteq X$, $< |A$ is the relation $<$ restricted to pairs of points from $A$. Clearly $< |A$ is a linear ordering on $A$.

**Theorem 7.13** (Subset Topologies in the Order Context)**.** *Let $\langle X, <\rangle$ be a linearly ordered set, $A$ a subset of $X$. Then the order topology on $A$ induced by $< |A$ is coarser than the subspace topology on $A$ induced by the order topology on $X$. In symbols, $\mathcal{T}_{<|A} \subseteq \mathcal{T}_{<}|A$. These two topologies can differ in general.*

*Proof.* Let $\mathcal{B}$ be the standard base for $\mathcal{T}_{<}$, consisting of all bounded open intervals and open rays on $\langle X, <\rangle$. Then, by Theorem 7.4, $\mathcal{B}|A$ is a base for $\mathcal{T}_{<}|A$. A base for $\mathcal{T}_{<|A}$ consists of sets of the form $I \cap A$, where $I$ is a bounded open interval or open ray *with endpoints in* $A$. Thus we have a base for $\mathcal{T}_{<|A}$ that is a subfamily of $\mathcal{B}|A$; and so $\mathcal{T}_{<|A} \subseteq \mathcal{T}_{<}|A$.

To show strict inequality can occur, Consider $A := [0, 1) \cup \{2\}$ as a subset of $\mathbb{R}$ with the usual ordering. Then, from the perspective of $\mathcal{T}_{<|A}$, $A$ has no isolated points: each nbd of 2 must contain elements of $[0, 1)$. On the other hand, from the perspective of $\mathcal{T}_{<}|A$, the point 2 is indeed isolated.

□

As we have seen, the use of bases to specify topologies is a very powerful tool; it is usually much easier to recognize a basic open set than it is to recognize an open set in general. The only drawback to defining a topology via a distinguished base is the need to verify that what you're putting forward as a base actually satisfies the conditions (B1) and (B2) of Theorem 7.3. There is a more general approach to specifying a topology, one in which there is no need to verify any conditions beforehand. It's founded on the following simple result.

**Theorem 7.14.** *Let $X$ be a set, and suppose $\mathcal{S} \subseteq \wp(X)$. Then the collection $\mathcal{B}_{\mathcal{S}} := \{\bigcap \mathcal{F} : \mathcal{F}$ is a finite subset of $\mathcal{S}\}$; i.e., the collection of all finite intersections of sets in $\mathcal{S}$, is a topological base.*

*Proof.* Let's verify the conditions (B1) and (B2).

*Ad (B1)*: If $\mathcal{F}$ is the empty subcollection of $\mathcal{S}$, then $\bigcap \mathcal{F}$ consists of all points of $X$ that lie in *every* member of $\mathcal{F}$. Since there are no members of $\mathcal{F}$ to worry about, *every* point of $X$ must lie in $\bigcap \mathcal{F}$; i.e., $\bigcap \mathcal{F} = X$. Since we now have $X \in \mathcal{B}_{\mathcal{S}}$, (B1) is automatically satisfied.

*Ad (B2)*: Any family of sets that is already closed under finite intersections clearly satisfies (B2). So let $F_1, \ldots, F_n \in \mathcal{B}_{\mathcal{S}}$, say we have finite subcollections $\mathcal{F}_1, \ldots, \mathcal{F}_n$ of $\mathcal{S}$ such that $F_i = \bigcap \mathcal{F}_i$, $1 \leq i \leq n$. Then one easily checks that $F_1 \cap \cdots \cap F_n = \bigcap(\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n)$. Since a finite union of finite subcollections is still a finite subcollection, we infer that $F_1 \cap \cdots \cap F_n \in \mathcal{B}_{\mathcal{S}}$.

□

Theorem 7.14 then motivates the next definition.

**Definition 7.15** (Subbase for a Topology)**.** *Let $\mathcal{T}$ be a topology on a set $X$, with $\mathcal{S}$ a subfamily of $\mathcal{T}$. $\mathcal{S}$ is called a* ***subbase*** *for $\mathcal{T}$ (sometimes the word* subbasis *is used instead of* subbase*) if each $\mathcal{T}$-open set is a union of finite intersections of members of $\mathcal{S}$; equivalently, if for each $U \in \mathcal{T}$ and each $x \in U$, there is a finite subcollection $\mathcal{F} \subseteq \mathcal{S}$ with $x \in \bigcap \mathcal{F} \subseteq U$.*

**Theorem 7.16** (Simple Properties of Subbases)**.** *Let $X$ be a set, $\mathcal{S} \subseteq \wp(X)$.*

(i) *$\mathcal{S}$ is a subbase for a unique topology $\mathcal{T}_{\mathcal{S}}$ on $X$; $\mathcal{B}_{\mathcal{S}} :=$ $\{\bigcap \mathcal{F} : \mathcal{F}$ is a finite subset of $\mathcal{S}\}$ is a base for that topology.*

(ii) *$\mathcal{T}_{\mathcal{S}}$ is coarser than any topology on $X$ that contains the family $\mathcal{S}$.*

(iii) *If $\mathcal{S}$ is a topological base, then it is a base for $\mathcal{T}_{\mathcal{S}}$.*

*Proof. Ad (i)*: This is essentially the content of Theorem 7.14.

*Ad (ii)*: If $\mathcal{S} \subseteq \mathcal{T}$, where $\mathcal{T}$ is a topology, then every union of finite intersections of members of $\mathcal{T}$ (hence of members of $\mathcal{S}$) is a member of $\mathcal{T}$. Thus $\mathcal{T}_{\mathcal{S}} \subseteq \mathcal{T}$.

*Ad (iii)*: Suppose $\mathcal{S}$ is a topological base; i.e., it satisfies conditions (B1) and (B2) of Theorem 7.3. Let $\mathcal{T}$ be the topology basically generated by $\mathcal{S}$; i.e., $\mathcal{T}$ consists of all unions of members of $\mathcal{S}$. Then $\mathcal{S} \subseteq \mathcal{T}$; so, by (ii) above, $\mathcal{T}_{\mathcal{S}} \subseteq \mathcal{T}$. But every union of members of $\mathcal{S}$ is trivially a union of finite intersections of members of $\mathcal{S}$, so $\mathcal{T} \subseteq \mathcal{T}_{\mathcal{S}}$, and equality holds.

□

**Examples 7.17** (Simple Subbases, Complicated Topologies)**.**

(i) *If $\langle X, < \rangle$ is a linealy ordered set, then the family of open rays forms a subbase for the order topology on $X$.*

(ii) *Refer to Example 6.9. For each $i \in I$ and $U$ open in $\mathbb{R}$, define $[i \to U]$ to be the set $\{f \in \mathbb{R}^I : f(i) \in U\}$. This collection of subsets of $\wp(\mathbb{R}^I)$ forms a subbase for an important topology on $\mathbb{R}^I$, the* ***topology of pointwise convergence****. Other well-studied topologies on function spaces are defined in a similar manner. While it is relatively easy to deal with individual* *subbasic sets in these topologies, it is quite difficult to manipulate even finite intersections of such sets.*

(iii) *Start with any topological space $\langle X, \mathcal{T}\rangle$, and let $F(X) := \{C \in \wp(X) : C \text{ is closed and nonempty}\}$. $F(X)$ now forms the set of points of a new topological space, a* ***hyperspace over*** *$X$. The* ***Vietoris topology*** *(after Leopold Vietoris (1891–2002!)) on $F(X)$ is subbasically generated by two kinds of sets: For each $U \in \mathcal{T}$, $[U]_* := \{C \in F(X) : C \subseteq U\}$ and $[U]^* := \{C \in F(X) : C \cap U \neq \emptyset\}$.*

**Exercises 7.18.** (1) *Let $\langle X, \mathcal{T}\rangle$ be any topological space, and suppose $x_0$ is an isolated point of $X$. If $\mathcal{B}$ is any base for $\mathcal{T}$, then show $\{x_0\} \in \mathcal{B}$.*

(2) *Let $\mathcal{B}$ consist of all bounded open intervals $(a, b)$ in $\mathbb{R}$, where both $a$ and $b$ are rational numbers. Show $\mathcal{B}$ is a base for the usual topology on $\mathbb{R}$.*

(3) *Let $\mathcal{B}$ consist of all open rectangles $(a, b) \times (c, d)$, where $a, b, c, d$ are all rational numbers. Show $\mathcal{B}$ is a base for the euclidean topology on $\mathbb{R}^2$.*

(4) *Let $\mathcal{B}$ consist of all bounded half-open-right intervals $[a, b)$, where $a$ and $b$ are real numbers. Show that $\mathcal{B}$ is a base for a topology, the so-called* ***lower limit*** *topology, that is strictly finer than the usual topology. ($\mathbb{R}$ with this topology is also frequently referred to as the* ***Sorgenfrey line*** *$\mathbb{L}$ (after R. H. Sorgenfrey).)*

(5) ** Referring to the construction in Exercise 7.18 (4) above, if we define $\mathcal{B}$ to consist of just the bounded half-open-right intervals with rational endpoints, show that we obtain a base for a topology that is strictly coarser than the lower limit topology.*

(6) *Refer to the lexicographic order on the plane (Example 7.6 (iv)), and sketch the bounded open interval $(\langle 0, 1\rangle, \langle 1, 0\rangle)$.*

(7) *Let $X$ be a set, and define $\mathcal{S} := \{X \setminus \{x\} : x \in X\}$. What is the topology subbasically generated by $\mathcal{S}$?*

(8) *Refer to Theorem 7.13, and show equality holds when $X = \mathbb{R}$ and $A$ is either $\mathbb{Q}$ or $\mathbb{Z}$.*

(9) *Give the closed unit interval $[0, 1]$ its usual order topology. Show that bounded open intervals $(a, b)$, for $0 \leq a < b \leq 1$, are insufficient for basically generating this topology.*

(10) *Refer to Example 7.17 (ii), and define $[F \to U]$ to consist of all $f \in \mathbb{R}^I$ such that $f[F] \subseteq U$, where $F \subseteq I$ is finite and $U \subseteq \mathbb{R}$ is usual-open. Show that the collection of such sets subbasically generates the same topology as does the smaller collection of sets $[i \to U]$.*

(11) ** Refer to Example 7.17 (iii). For each finite collection* $\{U_1, \ldots, U_n\}$ *of open sets of* $\langle X, \mathcal{T} \rangle$*, define* $[U_1, \ldots, U_n]$ *to consist of all* $C \in F(X)$ *such that* $C \subseteq U_1 \cup \cdots \cup U_n$ *and* $C \cap U_i \neq \emptyset$ *for each* $1 \leq i \leq n$*. Show that the collection of all such sets* $[U_n, \ldots, U_n]$ *forms a base for the Vietoris topology on* $F(X)$*.*

(12) ** Show that* $\mathbb{R}^2$*, with the lexicographic ordering (Example 7.6 (iv)) does not satisfy the least upper bound property. Give an example of a bounded open convex subset of this ordering that is not an interval.*

(13) *Suppose* $\mathcal{T}$ *and* $\mathcal{U}$ *are topologies on a set* $X$*, with bases* $\mathcal{B}$ *and* $\mathcal{C}$*, respectively. Show that* $\mathcal{T}$ *is coarser than* $\mathcal{U}$ *if and only if whenever* $x \in X$ *and* $B \in \mathcal{B}$ *contains* $x$*, there is a* $C \in \mathcal{C}$ *such that* $x \in C \subseteq B$*.*

## 8. Homeomorphisms and Topological Properties

At the beginning of Section 4, we talked about "rubber sheet geometry" as being the popular—if slightly whimsical—way of explaining that you could deform, distort, and otherwise bend or mutilate, any topological object as long as the original object and the transformed object were "homeomorphic." In this section, we make clear exactly what that means.

Recall that a function $f : X \to Y$ between sets is **one-one** (or **injective**, or an **injection**) if $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$ (i.e., $f$ sends distinct points to distinct points). $f$ is **onto** (or **surjective**, or a **surjection**) if every $y \in Y$ is $f(x)$ for some $x \in X$ (i.e., $f$ hits everything in $Y$). $f$ is **bijective** (or a **bijection**) if $f$ is both one-one and onto (i.e., $f$ is a perfect "code" between elements of $X$ and elements of $Y$). In that case, we have the function inverse $f^{-1} : Y \to X$; $f^{-1}(y)$ is the unique $x \in X$ such that $f(x) = y$ (i.e., $f^{-1}$ is Captain Midnight's intergalactically-patented decoder ring).

**Definition 8.1** (Homeomorphisms and Homeomorphic Spaces)**.** *A **homeomorphism** from space $\langle X, \mathcal{T}\rangle$ to space $\langle Y, \mathcal{U}\rangle$ is a bijection $f : X \to Y$ such that both $f$ and $f^{-1}$ are continuous on their respective domains. We write $\langle X, \mathcal{T}\rangle \cong \langle Y, \mathcal{U}\rangle$ to indicate that the two spaces are **homeomorphic**; i.e., that there is a homeomorphism from one to the other. (The inverse of a homeomorphism is a homeomorphism too.)*

**Remark 8.2.** *Referring back to Remark 1.4, we see that the notion of* homeomorphism *is very much akin to that of* isomorphism *in algebra. In both cases we are trying to say that two sets-with-structure are "the same" (only one is painted green, an irrelevancy). Points corresponding to one another in a homeomorphism behave the same* topologically *within their respective spaces. In the same way, two elements corresponding to one another in an isomorphism behave the same* algebraically *within their respective algebraic structures (e.g., groups, fields, etc.) There is a similar situation in biology: the skeleton of a human may be put in one-one correspondence with that of a chimpanzee, or even a bat. (This is mostly true; biologists may want to take issue here; but I'm playing fast and loose with the facts to make a real point.) One way to do this is simply to put all the bones of a (long deceased) chimpanzee in one bag, all the bones of a (similarly deceased) bat in another, and match the bones randomly. While this says that the two sets of bones have the same* cardinality*, it tells us nothing else of comparative anatomical interest. A much more informative correspondence pairs up bones that occupy the same relative structural position or evolutionary development: the left femur of Mr. Chimp is matched up with the left femur of Mr. Bat, and so on. (Such corresponding body parts are called* homologous *by zoologists.) This is significantly more like what we mean by homeomorphism or isomorphism in mathematics.*

*An important point to make here is that homeomorphisms and isomorphisms may ignore details that are deemed irrelevant for the purposes of highlighting similarities. Back to our skeletal example, the correspondence wilfully ignores the anatomical* role *of corresponding bones and bone groups: the joints of the foot of a chimpanzee allow for grasping, the joints of the hand of a bat allow for flying. None of this is true for human feet and hands.*

*In similar fashion, certain differing geometric details (e.g., angles, distances between points) pertaining to two homeomorphic figures are ignored as being irrelevant for the purposes of topological study. Those same geometric details are considered relevant, though, if the study is of sufficiently high resolution. But even then there are details left out. For example, two triangles may be considered congruent, even though one of them is upside down relative to the other. This difference may be considered negligible for the purposes of studying plane geometry, but not for those of studying visual perception: a picture of a familiar human face, for example, may be quite unrecognizable if viewed upside down.*

**Examples 8.3.** (i) *The identity map $\iota_X$ from a space $\langle X, \mathcal{T}\rangle$ to itself is a homeomorphism (the **equality** homeomorphism).*

(ii) *If $\langle X, \mathcal{T}\rangle \cong \langle Y, \mathcal{U}\rangle$ are homeomorphic topological spaces, then $\mathcal{T}$ is the discrete topology on $X$ if and only if each singleton set $\{x\} \subseteq X$ is $\mathcal{T}$-open, if and only if each singleton set $\{y\} \subseteq Y$ is $\mathcal{U}$-open, if and only if $\mathcal{U}$ is the discrete topology on $Y$. Thus a bijection $f : X \to Y$ may be viewed as a homeomorphism between discrete spaces. In this case, we say that $X$ and $Y$ have the same **cardinality**.*

(iii) *Consider the function $f(x) := \frac{2}{\pi}\arctan(x)$, taking the usual real line $\mathbb{R}$ bijectively onto the bounded open interval $(-1, 1) \subseteq \mathbb{R}$. Both $f$ and its inverse $f^{-1}(x) := \tan(\frac{\pi}{2}x)$, $x \in (-1, 1)$, are continuous; hence we have established a homeomorphism between $\mathbb{R}$ and the bounded open interval $(-1, 1)$. (Homeomorphisms do indeed ignore distances between points.)*

(iv) *If $\langle X, \mathcal{T}\rangle \cong \langle Y, \mathcal{U}\rangle$ are homeomorphic topological spaces, then $\mathcal{T}$ is the cofinite topology on $X$ if and only if $\mathcal{U}$ is the cofinite topology on $Y$ (see Exercise 8.15 (1) below).*

(v) *The function $f(t) := \langle\cos(t), \sin(t)\rangle$ maps the bounded half-open-right interval $[0, 2\pi)$ bijectively onto the standard unit circle $S^1 := \{\langle x, y\rangle \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. This function is indeed continuous, but its inverse is not: If $A := S^1 \cap \{\langle x, y\rangle : x > 0$ and $y < 0\}$, then $\langle 1, 0\rangle$ is a limit point of $A$. But $f^{-1}(\langle 1, 0\rangle) = 0$ and $f^{-1}[A] = (\frac{3}{2}\pi, 2\pi)$. Hence the inverse image of the point $\langle 1, 0\rangle$ is not in the closure of the inverse image of $A$. [Note that what we just did was to show a particular bijection fails to be a homeomorphism; we did* not *show that $S^1$ and $[0, 2\pi)$ fail to be homeomorphic. While this is indeed true, we have to develop some more technology to prove it. (See Example 11.16 (iii) and Exercise 13.16 (13).)]*

The idea of two spaces being homeomorphic is how we say that they are "topologically equal." We can also use homeomorphisms to say that one space is "topologically a subspace" of another. The following definition sharpens the fairly loose language, "can be viewed as a subspace," found in Definition 7.11.

**Definition 8.4** (Embeddings and Embeddability)**.** *An **embedding** from space $\langle X, \mathcal{T}\rangle$ to space $\langle Y, \mathcal{U}\rangle$ is a function $f : X \to Y$ such that $f$ is a homeomorphism from $X$ to the image $f[X]$ (with the subspace topology). We say $\langle X, \mathcal{T}\rangle$ is **embeddable in** $\langle Y, \mathcal{U}\rangle$ if there exists an embedding from the first space to the second.*

So, in Definition 7.11, we should properly say that a space is **suborderable** if it is embeddable in an orderable space.

**Examples 8.5.** *(All euclidean spaces below are equipped with the euclidean topology.)*

(i) *The inclusion map $\iota_X|A$ from a subspace $\langle A, \mathcal{T}|A\rangle$ to $\langle X, \mathcal{T}\rangle$ is an embedding (the* ***inclusion*** *embedding).*
(ii) *Each $\mathbb{R}^m$ is embeddable in $\mathbb{R}^n$, as long as $m \leq n$. One way to do this is to define $f(x_1, \dots, x_m) := \langle x_1, \dots, x_m, 0, \dots, 0\rangle$ (see Exercise 8.15 (2) below).*
(iii) *The unit circle $S^1$ (see Example 8.3 (v)) is not embeddable in $\mathbb{R}$, and is hence not homeomorphic to $[0, 2\pi)$. (We'll be able to prove this, once we have studied connectedness in Section 11.)*
(iv) *Let $X$ be the parabolic arc $\{\langle x, y\rangle \in \mathbb{R}^2 : 0 \leq x \leq 1 \text{ and } y = x^2\}$. Then $X$ is embeddable in $\mathbb{R}$; indeed, the map $f : X \to \mathbb{R}$, given by $f(x, y) = x$, is an embedding (see Exercise 8.15 (2) below).*

Our next task is to introduce some function notions related to continuity. Recall from Theorem 5.11 that a function is continuous if and only if inverse images of open (resp., closed) sets in the range are open (resp., closed) in the domain. What about images of open/closed sets in the domain?

**Definition 8.6** (Open/Closed Maps)**.** *Let $\langle X, \mathcal{T}\rangle$ and $\langle Y, \mathcal{U}\rangle$ be topological spaces, with $f : X \to Y$ a function.*

(i) *$f$ is an* ***open*** *map if $f[U]$ is $\mathcal{U}$-open whenever $U \subseteq X$ is $\mathcal{T}$-open.*
(ii) *$f$ is a* ***closed*** *map if $f[C]$ is $\mathcal{U}$-closed whenever $C \subseteq X$ is $\mathcal{T}$-closed.*

Note that in the definition above, we make no mention of whether $f$ is continuous. It turns out the three functional properties are completely independent: a function may possess one or more of them, and none of the others. Before going further into this claim, we establish a simple, but important, connection with homeomorphisms and embeddings.

**Theorem 8.7** (Characterizing Homeomorphisms)**.** *Let $f$ be a function from space $\langle X, \mathcal{T}\rangle$ to space $\langle Y, \mathcal{U}\rangle$. The following are equivalent:*

(a) *$f$ is a homeomorphism.*
(b) *$f$ is a continuous open bijection.*
(c) *$f$ is a continuous closed bijection.*

*Proof. Ad ((a) $\Longrightarrow$ (b))*: If $f$ is a homeomorphism, then $f$ is a bijection, and both $f$ and $f^{-1}$ are continuous. To show $f$ is open, let $U \subseteq X$ be $\mathcal{T}$-open. Then, because $f$ is a bijection, $f[U] = ((f^{-1})^{-1})[U]$, a $\mathcal{U}$-open set. Thus $f$ is an open map.

*Ad ((b) $\Longrightarrow$ (c))*: If $f$ is a continuous open bijection, and $C \subseteq X$ is $\mathcal{T}$-closed, then, because $f$ is a bijection, $f[C] = f[X \setminus (X \setminus C)] = Y \setminus f[X \setminus C]$, a $\mathcal{U}$-closed set. Thus $f$ is continuous and closed.

*Ad ((c) $\Longrightarrow$ (a))*: If $f$ is a continuous closed bijection, we need to show $f^{-1}$ is continuous. But, for any $\mathcal{T}$-closed $C$, $((f^{-1})^{-1})[C] = f[C]$, a $\mathcal{U}$-closed set. By Theorem 5.11, this shows $f^{-1}$ to be continuous. Therefore $f$ is a homeomorphism. □

**Theorem 8.8** (Characterizing Embeddings)**.** *Let $f$ be a function from space $\langle X,\mathcal{T}\rangle$ to $\langle Y,\mathcal{U}\rangle$. The following are equivalent:*

(a) *$f$ is an embedding.*
(b) *$f$ is a continuous injection that is* ***relatively open*** *(i.e., $f$ is an open map onto its image in $Y$).*
(c) *$f$ is a continuous injection that is* ***relatively closed*** *(i.e., $f$ is a closed map onto its image in $Y$).*

*Proof. Ad ((a) $\Longrightarrow$ (b))*: The fact that $f$ is a homeomorphism onto its image immediately tells us that $f$ is a continuous injection. Let $U \subseteq X$ be a $\mathcal{T}$-open set, and let $g : f[X] \to X$ be the inverse of $f$ (here viewed as a bijection from $X$ to $f[X]$). Then $f[U] = g^{-1}[U]$. Since $g$ is continuous and $f[X]$ has the subspace topology induced by $\mathcal{U}$, there is a $\mathcal{U}$-open $V \subseteq Y$ such that $V \cap f[X] = g^{-1}[U] = f[U]$. This says that $f$ is a relatively open map.

*Ad ((b) $\Longrightarrow$ (c))*: Assume that $f$ is a continuous injection that is relatively open, and suppose $C \subseteq X$ is $\mathcal{T}$-closed. Then $U := X \setminus C$ is $\mathcal{T}$-open, so $f[U] = V \cap f[X]$, for some $\mathcal{U}$-open set $V \subseteq Y$. Because $f$ is injective, we then have $f[C] = f[X \setminus U] = (Y \setminus f[U]) \cap f[X] = (Y \setminus (V \cap f[X])) \cap f[X] = (Y \setminus V) \cap f[X]$, showing that $f$ is a relatively closed map.

*Ad ((c) $\Longrightarrow$ (a))*: Assume that $f$ is a continuous injection that is relatively closed, and let $g : f[X] \to X$ be the inverse of $f$ (again viewed as a bijection from $X$ to $f[X]$). To show $f$ is an embedding, it suffices to show that $g$ is continuous. Indeed, if $C \subseteq X$ is $\mathcal{T}$-closed, then $g^{-1}[C] = f[C] = D \cap f[X]$ for some $\mathcal{U}$-closed set $D \subseteq Y$. By Theorem 5.11 and Exercise 5.12 (9), $g$ is indeed continuous. □

**Examples 8.9.** (i) *Give $\mathbb{R}$ the usual topology, but let $[0,1]$ be given a topology that is strictly coarser than the subspace topology. Then the inclusion map $\iota_{\mathbb{R}}|[0,1] : [0,1] \to \mathbb{R}$ is a closed map that is neither open nor continuous (see Exercise 8.15 (3) below).*

(ii) *As in (i) above, but with $\mathbb{Q}$ given the subspace topology. Then $\iota_{\mathbb{R}}|\mathbb{Q}$ is an embedding that is neither open nor closed because $\mathbb{Q}$ is neither open nor closed as a subset of $\mathbb{R}$.*

(iii) *Let $\pi_i : \mathbb{R}^2 \to \mathbb{R}$, $i = 1,2$, be the two standard* ***projection*** *maps; i.e, for $\langle x_1, x_2\rangle \in \mathbb{R}^2$, $\pi_i(x_1,x_2) := x_i$. Given that the line and plane have their euclidean topologies, we see that the projection maps are each continuous and open, but not closed:*

Ad (continuous)*: Given $U \subseteq \mathbb{R}$, $\pi_1^{-1}[U] = U \times \mathbb{R}$ and $\pi_2^{-1}[U] = \mathbb{R} \times U$. Open rectangles are euclidean-open sets (see Exercise 7.18 (3)), hence the projection maps are each continuous.*

Ad (open)*: If $(a,b) \times (c,d)$ is an open rectangle with real endpoints, then its image under $\pi_1$ (resp., $\pi_2$) is $(a,b)$ (resp., $(c,d)$). By Exercise*

*7.18 (3), the open rectangles form a base for the euclidean topology on the plane; by Exercise 4.11 (5), the images of open sets under the projection maps are open.*

Ad (not closed)*: Let* $C := \{\langle x, y\rangle \in \mathbb{R}^2 : x > 0,\ y > 0,\ \text{and } xy = 1\}$*. Then* $C$ *is euclidean-closed (see Exercise 8.15 (4) below), but the image of* $C$ *under either projection map is* $(0, \infty)$*, a set that is not euclidean-closed. (Note that each projection map, restricted to* $C$*, is an embedding into* $\mathbb{R}$*.)*

(iv) *(This is a deep theorem, the* invariance of domain theorem*, due to L. E. J. Brouwer (1881–1966). While easy to state, its proof is beyond the scope of this course.) Let* $U \subseteq \mathbb{R}^n$ *be an open subspace of euclidean* $n$*-space, with* $f : U \to \mathbb{R}^n$ *an embedding. Then* $f$ *is an open map.*

(v) *(This is also a deep theorem, from complex analysis, and is called the* open mapping theorem*. Its proof too is beyond the scope of this course.) Give the complex plane* $\mathbb{C}$ *the euclidean topology, and suppose* $U \subseteq \mathbb{C}$ *is open, but not the union of two nonempty disjoint open sets. Let* $f : U \to \mathbb{C}$ *be analytic (i.e., the complex derivative* $f'(z)$ *exists for all* $z \in U$*) and non-constant. Then* $f$ *is an open map.*

The topic of concern to us next is that of *topological property*. And to do it justice, we need to talk about composition of functions. This is the same composition of functions that you have seen since calculus (remember the chain rule?), and has a very simple definition.

**Definition 8.10** (Function Composition)**.** *Let* $X$*,* $Y$*, and* $Z$ *be sets, with* $f : X \to Y$ *and* $g : Y \to Z$ *functions. Then the* ***composition*** *of* $f$ *and* $g$*, written* $g \circ f$*, is a function from* $X$ *to* $Z$*, defined by the rule* $(g \circ f)(x) := g(f(x))$*.*

The following is easy to prove, but extremely important.

**Theorem 8.11** (Closure under Composition)**.** *Let* $\langle X, \mathcal{T}\rangle$*,* $\langle Y, \mathcal{U}\rangle$*, and* $\langle Z, \mathcal{V}\rangle$ *be spaces, with* $f : X \to Y$ *and* $g : Y \to Z$ *functions. If both* $f$ *and* $g$ *are continuous maps (resp., open maps, closed maps, embeddings, homeomorphisms), then so is the composition* $g \circ f$*.*

*Proof.* See Exercise 8.15 (5) below. □

As an immediate application of Theorem 8.11, we may consider all real bounded open intervals and open rays to be topologically equivalent.

**Theorem 8.12** (Open Intervals/Rays in the Real Line)**.** *Let the real line* $\mathbb{R}$ *be given its usual topology. Then all bounded open intervals and open rays in* $\mathbb{R}$ *are homeomorphic to* $\mathbb{R}$*, and hence are homeomorphic to each other.*

*Proof.* By Example 8.3 (iii), we know that $\mathbb{R} \cong (-1, 1)$. If $a < b$ are real numbers, then the affine map taking $-1$ to $a$ and $1$ to $b$ yields $(-1, 1) \cong (a, b)$. Thus all bounded open intervals are homeomorphic to $\mathbb{R}$; and, by Theorem 8.11, homeomorphic to each other. Now the exponential function $f(x) := \exp(x) = e^x$ gives rise to a homeomorphism from $\mathbb{R}$ to $(0, \infty)$. Compose that with a suitable affine map, and we get $\mathbb{R} \cong (a, \infty)$ and $\mathbb{R} \cong (-\infty, a)$ for any $a \in \mathbb{R}$.

□

**Remarks 8.13.** (i) *Theorem 8.11 tells us that the relation of being homeomorphic gives us an* equivalence *relation between topological spaces. The signal conditions that are satisfied by* $\cong$ *are:*

(E1) (Reflexivity) $\langle X, \mathcal{T} \rangle \cong \langle X, \mathcal{T} \rangle$.

(E2) (Symmetry) *If* $\langle X, \mathcal{T} \rangle \cong \langle Y, \mathcal{U} \rangle$*, then* $\langle Y, \mathcal{U} \rangle \cong \langle X, \mathcal{T} \rangle$*. (This is closely related to the idea of symmetry found in Theorem 2.2 (ii) and Definition 3.1 (ii). Also the postulated "close to" relation* $\sim$ *discussed in Section 1 was both reflexive and symmetric.)*

(E3) (Transitivity) *If* $\langle X, \mathcal{T} \rangle \cong \langle Y, \mathcal{U} \rangle$ *and* $\langle Y, \mathcal{U} \rangle \cong \langle Z, \mathcal{V} \rangle$*, then* $\langle X, \mathcal{T} \rangle \cong \langle Z, \mathcal{V} \rangle$*. (This is the same idea of transitivity as condition (L2) in Definition 7.5.)*

*Two spaces that are homeomorphic are said to be in the same* ***homeomorphism class****. We will see other equivalence relations later on when we study quotient maps, the topologist's version of "cut-and-paste."*

(ii) *Once we develop the topological theory of connectedness in Section 11, we will have an extremely powerful tool to show that bounded open intervals, bounded half-open intervals, and bounded closed intervals are in separate homeomorphism classes.*

Now back to the notion of *topological property*. Somewhat vaguely—because it is beyond the scope of this course to get into the formal distinctions between sets and proper classes—a property (or predicate, or class) of topological spaces is **topological** if two homeomorphic spaces either both share the property or both fail to share it. (A bit more abstractly, a topological property is a union of homeomorphism classes: if one member of a homeomorphsm class satisfies the property, then they all do.) We will be spending much of our time in the sequel identifying and studying some of the more key topological properties; here is a small list of examples and non-examples for starters.

**Examples 8.14** (Some Topological Properties)**.** (i) $\langle$`discrete`$\rangle$*: A space* $\langle X, \mathcal{T} \rangle$ *satisfies this property just in case* $\mathcal{T}$ *is the discrete topology on the set* $X$*. This property is topological because of Example 8.3 (ii).*

(ii) $\langle$`metrizable`$\rangle$*: If* $\langle X, \mathcal{T} \rangle$ *is a metrizable topological space,* $d$ *is a metric on* $X$ *such that* $\mathcal{T} = \mathcal{T}_d$*, and* $h$ *is a homeomorphism from* $\langle X, \mathcal{T} \rangle$ *to space* $\langle Y, \mathcal{U} \rangle$*, then (see Exercise 8.15 (6) below) the map* $e : Y \times Y \to \mathbb{R}$*, given by* $e(y_1, y_2) := d(f^{-1}(y_1), f^{-1}(y_2))$*, is a metric on* $Y$ *whose induced topology is* $\mathcal{U}$*. This makes* $\langle Y, \mathcal{U} \rangle$ *metrizable as well.*

(iii) $\langle$`Hausdorff`$\rangle$*: A space satisfying this property (see Exercise 4.11 (4)) is one where two distinct points have disjoint neighborhoods. (See Exercise 8.15 (6) below.)*

(iv) $\langle$`orderable`$\rangle$*: A homeomorphism from one orderable space to a second topological space will induce a linear ordering on the second space that, in turn, will induce the given topology on the second space. (See Exercise 8.15 (6) below.)*

(v) $\langle$`closed subset of the reals`$\rangle$*: This is not a topological property. The closed subset* $[0, \infty)$ *is homeomorphic to the non-closed subset* $[0, 1)$*. (See Exercise 8.15 (7) below.)*

**Exercises 8.15.** (1) *Refer to Example 8.3 (iv) and prove that if two spaces are homeomorphic and one of them has the cofinite topology, then so does the other. (So having the cofinite topology is a topological property.)*

(2) *Show the functions defined in Example 8.5 (ii, iv) are embeddings.*

(3) *Prove the assertion in Example 8.9 (i).*

(4) *Prove the assertion in Example 8.9 (iii) that* $C$ *is a closed subset of* $\mathbb{R}^2$*.*

(5) *Prove Theorem 8.11.*

(6) *Prove the assertions in Example 8.14 (ii, iii, iv).*

(7) *Prove the assertion in Example 8.14 (v).*

(8) *Imagine the letters of the alphabet,*

A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z

*as geometrical figures in the plane, and group them according to homeomorphism class.*

(9) ** A continuous map* $f$ *from space* $\langle X, \mathcal{T}\rangle$ *to space* $\langle Y, \mathcal{U}\rangle$ *is called a* ***co-retraction*** *if there is a continuous map* $g : Y \to X$ *such that* $g \circ f = \iota_X$*. Show that a co-retraction is always an embedding.*

(10) *Give an example of a map that is both open and closed, but continuous at no point of its domain.*

(11) ** Suppose* $f : \mathbb{R} \to \mathbb{R}$ *is a polynomial function; i.e., there are real numbers* $a_0, a_1, \ldots, a_n$*, such that* $f(x) := a_0 + a_1x^1 + \cdots + a_nx^n$*. Give conditions that characterize when* $f$ *is a homeomorphism.*

## 9. The Basic Lower-Level Separation Axioms

Topological properties come in several "flavors," each dictated by a certain vague intuition that manifests itself into one or more precise formulations. One flavor is "separating stuff;" another is "being in one piece;" still another is "being small and tidy." It is the first of these flavors we take up in this section, manifesting itself in the so-called *separation axioms*. (The second flavor, by the way, is manifested in various formulations of connectedness, the third in formulations of compactness.)

The separation axioms take on many forms, according to how we want to separate things. One way is via open sets, another is via continuous real-valued functions. We concentrate on open-set separation in this section. We refer to the separation axioms in this section as "lower-level" because they deal with separating one point from another. The "upper-level" axioms, dealing with the separation of sets, will be considered later on. By way of notation in the sequel, we will often suppress mention of the topology on a space $\langle X, \mathcal{T} \rangle$, especially when only one topology on a set is being considered. Thus we may talk of a *space* $X$, or an *open set* in $X$, or a *continuous map* $f : X \to Y$, when the meaning is clear as to which topology is being considered. Similarly, we speak of $A \subseteq X$ as a *subspace* of $X$, with the intention that the usual subspace topology is being considered on $A$.

**Definition 9.1** (The $\mathsf{T}_0$, or Kolmogorov, Axiom)**.** *A topological space $X$ is called a $\mathsf{T}_0$ **space** (or a **Kolmogorov space**, after Andrei N. Kolmogorov (1903–1987)) if, for each two points $x \neq y$ in $X$, at least one of them has an open nbd not containing the other. That is, there is an open set $U \subseteq X$ such that: either $x \in U$ and $y \notin U$, or $y \in U$ and $x \notin U$.*

As a notational aside, the usage $\mathsf{T}_0$ suggests the lowest rung on a ladder of axioms designated $\mathsf{T}_n$, for various numbers $n$. (The $\mathsf{T}$ is for *Trennung*, German for *separation*.) This is the weakest separation axiom usually studied; its main application in recent years being in the formal semantics of computer programming languages. We consider two useful examples.

**Examples 9.2** (Some $\mathsf{T}_0$ Spaces *et al*)**.**

(i) *Any metric space is clearly a $\mathsf{T}_0$ space.*

(ii) *Since any topology on a single-point set is the discrete topology, the simplest example of a $\mathsf{T}_0$ space that doesn't satisfy any stronger separation conditions (to be taken up shortly) is **Sierpiński space** $\mathbb{S}$ (after Wacław Sierpiński (1882–1969)). The underlying set consists of two points, only one of which is isolated. Clearly, any two such spaces are homeomorphic.*

(iii) *Given the real line $\mathbb{R}$, consider the set $\mathcal{B}$ consisting of all open right-looking rays $(a, \infty)$. Then $\mathcal{B}$ is easily seen to be a topological base; indeed it may be made into a topology merely by adding the empty and universal sets. Given any two points $x \neq y$, say $x < y$, and pick $a$ strictly between $x$ and $y$. Then $(a, \infty)$ is an open nbd of $y$ not containing $x$. If we call this space the **right-looking real line**, we may also define the **left-looking real line** in a similar fashion. Of course the right-looking real line and the left-looking real line are homeomorphic spaces (see Exercise 9.11 (1) below).*

(iv) *The trivial topology on any set with more than one point fails to be a* $\mathsf{T}_0$ *topology, since no nonempty open set may exclude any point.*

We now introduce the next rung of the ladder.

**Definition 9.3** (The $\mathsf{T}_1$, or Fréchet, Axiom)**.** *A topological space $X$ is called a* $\mathsf{T}_1$ ***space*** *(or a* ***Fréchet space****, after Maurice R. Fréchet (1878–1973)) if, for each two points $x \neq y$ in $X$, each one of them has an open nbd not containing the other. That is, there is an open nbd $U$ of $x$ and an open nbd $V$ of $y$ such that $y \notin U$ and $x \notin V$.*

Clearly every $\mathsf{T}_1$ space is a $\mathsf{T}_0$ space; the following is the basic result about $\mathsf{T}_1$ spaces.

**Theorem 9.4** (The Finite-Set Criterion)**.** *$X$ is a $\mathsf{T}_1$-space if and only if all finite subsets of $X$ are closed. (Consequently, any topology on a set is a $\mathsf{T}_1$ topology if and only if that topology is finer than the cofinite topology on the set.)*

*Proof.* Suppose $X$ is $\mathsf{T}_1$, and let $F := \{x_1, \ldots, x_n\}$ be a finite subset of $X$. It suffices to show that each point outside $F$ has a nbd that misses $F$. Indeed, let $x \notin F$. For each $1 \leq i \leq n$, there is an open nbd $U_i$ of $x$ that does not contain $x_i$. Thus $U_1 \cap \cdots \cap U_n$ is an open nbd of $x$ that contains no element of $F$.

Now suppose each finite subset of $X$ is closed, and let $x \neq y$ be two points of $X$. Then, since singleton sets are closed, $X \setminus \{y\}$ is an open nbd of $x$ not containing $y$, and $X \setminus \{x\}$ is an open nbd of $y$ not containing $x$. Thus $X$ is a $\mathsf{T}_1$ space.

□

**Examples 9.5** (Some $\mathsf{T}_1$ Spaces *et al*)**.**

(i) *Any metric space is clearly a* $\mathsf{T}_1$ *space.*

(ii) *The Sierpiński two-point space $\mathbb{S}$ of Example 9.2 (ii) is a* $\mathsf{T}_0$ *space without being a* $\mathsf{T}_1$ *space.*

(iii) *The right-looking real line of Example 9.2 (iii) is a* $\mathsf{T}_0$ *space without being a* $\mathsf{T}_1$ *space.*

(iv) *Let $\langle X, \mathcal{T} \rangle$ be a* $\mathsf{T}_1$ *space, and let $p$ be a point not in $X$. Consider the set $X \cup \{p\}$, and consider the family $\mathcal{B} := \mathcal{T} \cup \{(C \cup \{p\}) : C \subseteq X \text{ cofinite}\}$ of subsets of $X \cup \{p\}$. Then (see Exercise 9.11 (2) below) $\mathcal{B}$ is a base for a* $\mathsf{T}_1$ *topology on $X \cup \{p\}$, whose restriction to $X$ is the original topology $\mathcal{T}$. Such a construction, called a* ***one-point extension****, is a well-known method for building new spaces from old ones.*

The third rung of the ladder is arguably the most useful and well-studied of the three we consider in this section. We saw it first in Exercise 4.11 (4); now we make it official.

**Definition 9.6** (The $\mathsf{T}_2$, or Hausdorff, Axiom)**.** *A topological space $X$ is called a* $\mathsf{T}_2$ ***space*** *(or a* ***Hausdorff space****, after Felix Hausdorff (1868–1942)) if, for each* *two points $x \neq y$ in $X$, there are disjoint open sets $U$ and $V$ with $x \in U$ and $y \in V$.*

**Examples 9.7** (Some $\mathsf{T}_2$ Spaces *et al*)**.**
(i) *Any metric space is a $\mathsf{T}_2$ space (see Exercise 4.11 (4)).*
(ii) *The cofinite topology on any infinite set is a $\mathsf{T}_1$ topology that is not a $\mathsf{T}_2$ topology (see Example 5.3 (iii)).*
(iii) *Referring to Example 9.5 (iv), the one-point extension of a discrete space is a $\mathsf{T}_2$ space (that is not discrete if $X$ is infinite). Clearly any two points in $X$ are isolated, so there is no problem separating them. If $x \in X$, set $U := \{x\}$ and $V := (X \cup \{p\}) \setminus \{x\}$. Then both $U$ and $V$ are open sets in the extension topology, $x \in U$, $p \in V$, and $U \cap V = \emptyset$. (The discreteness condition on $X$ is actually necessary for $X \cup \{p\}$ to be Hausdorff (see Exercise 9.11 (3) below).)*

We say a topological property is **hereditary** if whenever $X$ has the property and $A \subseteq X$, then $A$ has the property in its subspace topology. Theorem 6.5 tells us, then, that metrizability is a hereditary topological property; the following says the same for the three separation properties introduced in this section.

**Theorem 9.8.** *$\mathsf{T}_n$ is a hereditary property, for $n = 0, 1, 2$.*

*Proof.* Let's show the assertion for $T_2$; the others are similar. If $X$ is Hausdorff and $A$ is a subspace of $X$, pick two points $x \neq y$ in $A$. Since the big space is Hausdorff, there are open sets $U$ and $V$ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$. Then $U \cap A$ and $V \cap A$ are disjoint relatively open nbds of $x$ and $y$, respectively, showing that $A$ is Hausdorff too.

□

We end this section with two nontrivial applications of the Hausdorff axiom. The first shows us how to build discrete subsets inductively; the second makes an important statement about mappings.

Given a space $X$ and $A \subseteq X$, we say $A$ is a **discrete subset of** $X$ if its subspace topology is discrete. This is equivalent to saying that for each $a \in A$, there is an open nbd $U$ of $a$ such that $U \cap A = \{a\}$.

**Theorem 9.9.** *An infinite Hausdorff space contains an infinite discrete subset.*

*Proof.* Suppose $X$ is infinite and Hausdorff. If each point of $X$ is isolated, then $X$ is an infinite discrete subset of itself. So assume $X$ has a nonisolated point, say $x$. Then $x$ can have no finite nbds (see Exercise 9.11 (12) below). We build our infinite discrete subset using induction on the natural numbers. First let $x_1$ be any point different from $x$. By Hausdorffness, we may choose disjoint open nbds $U_1$ of $x_1$ and $V_1$ of $x$. In the next step of the induction, we work in the open subspace $V_1$, also a Hausdorff space, by Theorem 9.8. $V_1$ is infinite, so we may pick $x_2 \in V_1$ different from $x$. Again using Hausdorffness, we pick disjoint open nbds $U_2$ of $x_2$, $V_2$ of $x$, both contained in $V_1$. We now repeat the procedure in $V_2$, continuing like this forever. More properly, we assume at the $n$th stage of the induction, $n > 1$, that we have points $x_1, \dots, x_n$ in $X$, open sets $U_1, \dots, U_n$, and open sets $V_1, \dots, V_n$ such that the following conditions hold:

- $x_i \in U_i$, $1 \leq i \leq n$;
- $x \in V_i$, $1 \leq i \leq n$;
- $U_i \cap V_i = \emptyset$, $1 \leq i \leq n$; and
- $U_{i+1} \cup V_{i+1} \subseteq V_i$, $1 \leq i \leq n-1$.

The key resides in the facts that the sets $U_i$ are disjoint from each other and that $U_1 \cup \cdots \cup U_n$ is disjoint from $V_n$. Since $V_n$ is infinite, we may choose $x_{n+1} \in V_n$ different from $x$ and use Hausdorffness to find disjoint open sets $U_{n+1}$ and $V_{n+1}$, contained in $V_n$, such that $x_{n+1} \in U_{n+1}$ and $x \in V_{n+1}$. In this way, the conditions above still hold; only now $n$ may be replaced by $n+1$.

Let $A := \{x_1, x_2, \dots\}$. Then $A$ is infinite because the inductive process above ensures $x_{n+1} \notin \{x_1, \dots, x_n\}$. $A$ is discrete because each $U_i$ contains no $x_j$ when $j \neq i$. Thus $A$ is an infinite discrete subset of $X$.

□

Recall (Exercise 8.15 (9)) that a continuous map $f : X \to Y$ is a **co-retraction** if there is a continuous map $g : Y \to X$ such that $g \circ f = \iota_X$. (The map $g$ is called a **retraction**.)

**Theorem 9.10.** *Co-retractions with Hausdorff ranges are closed embeddings.*

*Proof.* Let $f : X \to Y$ be a co-retraction, with "left-inverse" $g : Y \to X$ (i.e., $g$ is continuous and $g \circ f = \iota_X$). In Exercise 8.15 (9), you are asked to show $f$ is an embedding; so we assume that as already established. What's left is to show $f$ is a closed map when $Y$ is assumed to be Hausdorff. Since embeddings are relatively closed maps (Theorem 8.8), all we need to show (see Exercise 5.12 (10)) is that $f[X]$ is a closed subset of $Y$.

For convenience, let's write $A := f[X]$ and define $r : Y \to A$ by $r(y) := f(g(y))$. Then $r$ is a continuous map from $Y$ onto the subspace $A$, and $r \circ r = f \circ (g \circ f) \circ g = f \circ g = r$; i.e., $r(a) = a$ for each $a \in A$.

Now suppose $A$ is not closed in $Y$. Then there is a limit point $y$ of $A$ such that $y \notin A$. Since $y \notin A$, we know that $y \neq r(y)$. Here is where we use Hausdorffness: Pick open sets $U$, $V$ such that $y \in U$, $r(y) \in V$, and $U \cap V = \emptyset$. But $r$ is continuous. Therefore there is an open set $W$ containing $y$ such that $r[W] \subseteq V$. Let $N := U \cap W$. Then $N$ is an open nbd of $y$ that is disjoint from $V$, and such that $r[N] \subseteq V$.

As a consequence of this, we have $N \subseteq r^{-1}[V]$; in particular $N \cap A \subseteq r^{-1}[V]$. But $y$ is a limit point of $A$, so $N \cap A \neq \emptyset$. Finally, since $r$ is a retraction; i.e., $r \circ r = r$, $N \cap A \subseteq r^{-1}[N]$. We thus have a nonempty set, namely $N \cap A$, as a subset of $r^{-1}[N] \cap r^{-1}[V] = r^{-1}[N \cap V] = r^{-1}[\emptyset] = \emptyset$, a contradiction. Thus $A$ must have contained all its limit points all along, so must be closed after all. This completes the proof.

□

**Exercises 9.11.** (1) *Referring to Example 9.2 (iii), show that the collection of open right-looking rays forms a base for a topology on $\mathbb{R}$, and show that* *the right-looking real line and the left-looking real line are homeomorphic.*

(2) *Prove the assertion in Example 9.5 (iv).*

(3) ** Referring to Example 9.7 (iii), show that if the one-point extension* $X \cup \{p\}$ *is Hausdorff, then* $X$ *must be discrete.*

(4) *Suppose* $\mathcal{T}$ *and* $\mathcal{U}$ *are two topologies on* $X$*, with* $\mathcal{T} \subseteq \mathcal{U}$*, and assume that* $\mathcal{T}$ *is a* $\mathsf{T}_n$ *topology for some fixed* $n = 0, 1, 2$*. Then* $\mathcal{U}$ *is a* $\mathsf{T}_n$ *topology also.*

(5) *Suppose* $\mathcal{T}$ *and* $\mathcal{U}$ *are two* $\mathsf{T}_1$ *topologies on* $X$*. Show that* $\mathcal{T} \cap \mathcal{U}$ *is a* $\mathsf{T}_1$ *topology also. Show that the same statement fails for* $\mathsf{T}_0$*. (It also fails for* $\mathsf{T}_2$*, but is harder to show.)*

(6) *Are any of the properties* $\mathsf{T}_n$*,* $n = 0, 1, 2$ *preserved by continuous bijections? I.e., if* $f : X \to Y$ *is a continuous bijection, and* $X$ *is a* $\mathsf{T}_n$ *space, can the same be said for* $Y$*?*

(7) *Show that every suborderable space is Hausdorff.*

(8) *A space* $X$ *is* ***functionally Hausdorff*** *if, for every two points* $x \neq y$ *in* $X$*, there is a continuous function* $f : X \to [0, 1]$ *(usual topology) such that* $f(x) = 0$ *and* $f(y) = 1$*. Show that a functionally Hausdorff space is Hausdorff.*

(9) ** Consider all polynomial functions* $f : \mathbb{R}^2 \to \mathbb{R}$*, e.g.,* $f(x, y) := x^2 + y^2 - 1$*, and define the* ***zero set*** $Z_f$ *of such a polynomial* $f$ *to be the set of all pairs* $\langle x, y\rangle$ *such that* $f(x, y) = 0$*. (In our example,* $Z_f$ *is the unit-radius circle in the plane, centered at the origin.) Now let* $\mathcal{Z} := \{Z_f : f \text{ is a polynomial function from } \mathbb{R}^2 \text{ to } \mathbb{R}\}$*, and define the* ***Zariski topology*** *(after Oscar Zariski (1899–1986)) to be the smallest topology having each member of* $\mathcal{Z}$ *as a* closed *set. Show this topology satisfies the* $\mathsf{T}_1$ *axiom. (The Zariski topology, though non-Hausdorff, is an important tool in the study of algebraic geometry.)*

(10) *Characterize the discrete subsets of a space equipped with the cofinite topology.*

(11) *Show that any finite* $\mathsf{T}_1$ *topological space must be discrete.*

(12) *Show that if a point in a* $\mathsf{T}_1$ *space has a finite nbd, then that point is an isolated point.*

(13) ** Show that the infinite discrete set* $A$ *constructed in the proof of Theorem 9.9 has at most one limit point.*

(14) ** Let* $f : X \to Y$ *and* $g : X \to Y$ *be continuous functions, where* $Y$ *is a Hausdorff space, and suppose that there is a dense (in the sense of Exercise*

*5.12 (5)) subset* $D$ *of* $X$ *such that* $f(x) = g(x)$ *for all* $x \in D$*. Show that* $f(x) = g(x)$ *for all* $x \in X$*. [Hint: First show that* $\{x \in X : f(x) = g(x)\}$ *is always a closed subset of* $X$*.]*

## 10. Convergence

The idea of convergence of numerical sequences should ring some familiar bells from when you took calculus. In particular, sequences can be used to characterize continuity when we're dealing with real-valued functions of a real variable. Sequences play a pivotal role in advanced calculus and analysis courses; we present sequences in a topology course because they still have an important bearing on continuity, even in very general contexts. And, at this point in the course, we are able to give a quite elegant and clear (I hope) account of at least some of the role convergence plays in mathematics. The following is a litany of basic definitions we will need in this section.

**Definition 10.1** (Sequences, Accumulation, Convergence)**.** *Let $X$ be a topological space. By a* ***sequence*** *in $X$ we mean a function $s : D \to X$, where $D \subseteq \mathbb{N}$ is infinite. (Note that the infinite subsets of $\mathbb{N}$ are precisely the nonempty subsets with no last element.) If $n \in D$, we frequently write $s_n$ instead of $s(n)$; the* ***image*** *(or, sometimes,* ***trace****) of the sequence $s$ is just the set $s[D]$, as usual. (So sequences, even ones with finite images, are infinite objects because they are functions with infinite domains.) A* ***subsequence*** *of sequence $s : D \to X$ is a restriction $s|E$, where $E$ is an infinite subset of $D$. The subsequence $s|E$ is called a* ***tail*** *of $s$ if the set $E$ is actually cofinite in $D$. (Note that the cofinite subsets of $D$ are those subsets that contain all members of $D$ beyond a certain point; i.e., for some $n_0 \in D$, we have $n \in E$ whenever $n \in D$ and $n \geq n_0$.) The sequence $s$* ***accumulates at*** *(resp.,* ***converges to****) the point $x \in X$ if each nbd of $x$ contains the image of a subsequence (resp., tail) of $s$. We write $s \vdash x$ (resp., $s \to x$) to indicate that $s$ accumulates at (resp., conveges to) $x$. If $s \vdash x$ (resp., $s \to x$), we call $x$ an* ***accumulation point*** *(resp., a* ***limit point****) of $s$. Frequently we denote $s : D \to X$ by $\langle s_n \rangle_{n \in D}$, or, when the meaning is clear, simply by $\langle s_n \rangle$.*

**Examples 10.2.**
(i) *The sequence $s : \mathbb{N} \to \mathbb{R}$, given by $s_n := \frac{1}{n+1}$, has $0$ as its only accumulation or limit point.*
(ii) *The sequence $s : \mathbb{N} \to \mathbb{R}$, given by $s_n := (-1)^n$, has the points $\pm 1$ as its two accumulation points. It has no limit point.*
(iii) *The sequence $s : \mathbb{N} \to \mathbb{R}$, given by $s_n := n$, has neither accumulation points nor limit points.*

**Theorem 10.3** (Basic Properties)**.** *Let $X$ be a topological space.*
(i) *If the sequence $s : D \to X$ is* ***eventually constant*** *at $x$; i.e., if some tail $s|E$ of $s$ takes on the constant value $x$, then $s \to x$.*
(ii) *If some subsequence of $s$ converges to $x \in X$, then $s \vdash x$.*
(iii) *If $s \vdash x$ in $X$, then $x$ is in the closure of the trace $s[D]$.*
(iv) *If $s \to x$ and $s[D]$ is an infinite set, then $x$ is a limit point of $s[D]$.*
(v) *If $s \to x$, then $s|E \to x$ for every subsequence $s|E$ of $s$.*

*Proof. Ad (i)*: Suppose $s_n = x$ for all $n$ in some cofinite subset $E$ of $D$. Then clearly every nbd of $x$ contains $s[E]$. Consequently, $s \to x$.

*Ad (ii)*: If $s|E \to x$ for some infinite subset $E \subseteq D$, then every nbd of $x$ contains the image of a tail of $s|E$. Since a tail of a subsequence of $s$ is a subsequence of $s$, we infer that $s \vdash x$.

*Ad (iii)*: If $s \vdash x$, then, for each nbd $U$ of $x$, there is a subsequence $s|E$ of $s$ whose image $s[E]$ is a subset of $U$. Since each such $s[E]$ is nonempty, we infer that $U$ intersects $s[D]$.

*Ad (iv)*: Suppose $s \to x$ and $s[D]$ is infinite. Let $U$ be a nbd of $x$. Then there is a tail $s|E$ of $s$ such that $s[E] \subseteq U$. Since $E$ is cofinite in $D$, $D \setminus E$ must be finite. Since $s[D] = s[E] \cup s[D \setminus E]$, we infer that $s[E]$ must be infinite. Thus each nbd of $x$ contains infinitely many elements of $s[D]$. This says that $x$ is a limit point of $s[D]$.

*Ad (v)*: Suppose $s \to x$, and let $s|E$ be a fixed subsequence of $s$. If $U$ is any nbd of $x$, let $C \subseteq D$ be cofinite in $D$, such that $s[C] \subseteq U$. Since $D \setminus C$ is finite, we must have $E \setminus C$ finite as well. Thus $C \cap E$ is a cofinite subset of $E$. Now $s[C \cap E] \subseteq s[C] \subseteq U$, telling us that every nbd of $x$ contains the image of a tail of $s|E$. Hence $s|E \to x$.

□

In Example 10.2 (ii), we see that it is possible for a sequence of real numbers to have more than one accumulation point. Can there be more than one limit point? The answer is no, if the space in question is Hausdorff

**Theorem 10.4** (Uniqueness of Limit)**.** *If a sequence in a Hausdorff space has a limit point, then that limit point is the only accumulation point of the sequence. In particular, no sequence in a Hausdorff space can have more than one limit point.*

*Proof.* Suppose $s : D \to X$ is a sequence, where $X$ is a Hausdorff space, and suppose $s \to x$. In the interests of a contradiction, suppose further that $y \neq x$ is an accumulation point of $s$. By the Hausdorff property, we may pick disjoint open sets $U$ and $V$, with $x \in U$ and $y \in V$. By the definitions of convergence and accumulation, there is a cofinite subset $C \subseteq D$ and an infinite subset $E \subseteq D$ such that $s[C] \subseteq U$ and $s[E] \subseteq V$. As in the proof of Theorem 10.3 (v), $C \cap E$ is a cofinite subset of $E$; in particular, $C \cap E$ is nonempty, and we have $\emptyset \neq s[C \cap E] \subseteq U \cap V$, contradicting the disjointness of $U$ and $V$. As an immediate consequence of this, no sequence in a Hausdorff space can converge to more than one point of the space. □

Theorem 10.4 justifies the alternative notation $\lim_{n\to\infty} s_n = x$ when $s \to x$. It is, however, a misleading notation unless the sequence takes values in a Hausdorff space. The following example illustrates that, even in a $\mathsf{T}_1$ environment, a sequence may have many limits.

**Example 10.5** (Nonuniquenes of Limit)**.** *Consider the $\mathsf{T}_1$ space $\langle X, \mathcal{T} \rangle$, where $X$ is an infinite set and $\mathcal{T}$ is the cofinite topology on $X$. Let $s : D \to X$ be any sequence. There are three cases to consider:*

- *$s$ hits each point of $X$ only finitely often. If $x \in X$ is arbitrary and $X \setminus F$ is a typical $\mathcal{T}$-nbd of $x$ (i.e., $F$ is finite and $x \notin F$), there is a largest $n \in D$ such that $s(n) \in F$. Hence $X \setminus F$ contains the image of a tail of $s$; consequently* every *point of $X$ is a limit point of $s$.*
- *There is a unique $x_0 \in X$ such that $s(n) = x_0$ for infinitely many $n \in D$. If $x \neq x_0$, then $X \setminus \{x_0\}$ is a $\mathcal{T}$-nbd of $x$ that fails to contain the image of a tail of $s$, so $s$ cannot converge to any point different from $x_0$. On the other hand, if $U$ is a typical $\mathcal{T}$-open nbd of $x_0$, then there is a finite set $F \subseteq X$, with $x_0 \notin F$, such that $U = X \setminus F$. $s$ hits each member of $F$ only finitely often, and $F$ is finite; hence there is a greatest $n \in D$ such that $s(n) \in F$. Consequently, $U$ contains the image of a tail of $s$, and $s$ therefore converges to a unique point of $X$.*
- *There are two (or more) points $x_0 \neq x_1$ in $X$ such that $s$ hits both $x_0$ and $x_1$ infinitely often. Suppose $E_i \subseteq D$ is an infinite set such that $s(n) = x_i$ for all $n \in E_i$, $i = 0, 1$. Then, for $x \notin \{x_0, x_1\}$, $X \setminus \{x_0, x_1\}$ is a $\mathcal{T}$-nbd of $x$ that fails to contain any tail of $s$; for each $i \in \{0, 1\}$, $X \setminus \{x_{1-i}\}$ is a $\mathcal{T}$-nbd of $x_i$ that fails to contain any tail of $s$. Consequently, $s$ has no limit points at all.*

You may recall that there is a sequential characterization of continuity in calculus. That is, if $f : X \to \mathbb{R}$, where $X \subseteq \mathbb{R}$, and if $x_0 \in X$, then $f$ is continuous at $x_0$ if and only if every sequence converging to $x_0$ in $X$ gets sent, via $f$, to a sequence that converges to $f(x_0)$ in $\mathbb{R}$. While the "if" half of the characterization fails to hold in the general topological context, the "only if" half goes through quite easily. (We'll consider what to do about the failed "if" half shortly.)

**Theorem 10.6** (Continuity Implies Convergence Preservation)**.** *Let $f : X \to Y$ be a map from one topological space to another, and suppose $x_0 \in X$. If $f$ is continuous at $x_0$ and $s : D \to X$ is a sequence converging to $x_0$, then $f \circ s : D \to Y$ is a sequence converging to $f(x_0)$.*

*Proof.* Suppose that $f$ is continuous at $x_0$, and let $s : D \to X$ converge to $x_0$. We need to show $f \circ s \to f(x_0)$. If $V$ is any open nbd of $f(x_0)$, use continuity to find an open nbd $U$ of $x_0$ such that $f[U] \subseteq V$. Since $s \to x_0$, there is a tail $s|C$ of $s$ such that $s[C] \subseteq U$. Therefore there is a tail $(f \circ s)|C$ of $f \circ s$ such that $(f \circ s)[C] = f[s[C]] \subseteq f[U] \subseteq V$.

□

For the converse of Theorem 10.6, we need some more control on how sequential convergence connects with the notion of set closure.

**Theorem 10.7** (Sequences and Closure)**.** *Let $X$ be a metrizable space, $A \subseteq X$, and $x \in X$. The following are equivalent:*

(a) *There is a sequence $s : D \to X$, with $s[D] \subseteq A$ and $s \to x$.*
(b) *$x \in Cl(A)$.*

*Proof. Ad ((a) $\Longrightarrow$ (b)):* If there is a sequence $s : D \to X$ converging to $x$, then, by Theorem 10.3 (ii, iii), we have $x \in \mathrm{Cl}(s[D])$. Since $s[D] \subseteq A$, we have $x \in \mathrm{Cl}(A)$.

(Note that this part of the proof does not require any restrictions on the space $X$.)

*Ad ((b) $\Longrightarrow$ (a))*: Suppose $x \in \mathrm{Cl}(A)$. We need to find a sequence $s : D \to X$, whose image lies in $A$, and which converges to $x$. If $x \in A$, we may take our sequence to be constantly $x$ and be done. Otherwise $x$ is a limit point of $A$, not in $A$. Pick a metric $d$ on $X$ that induces the given metrizable topology on $X$. For each $n \geq 1$, pick $s_n \in A \cap B(x, \frac{1}{n})$. Then the sequence $\langle s_1, s_2, \dots \rangle$ does what we want: For if $U$ is any open nbd of $x$, then we pick $n_0$ large enough so that $B(x, \frac{1}{n_0}) \subseteq U$. Since $B(x, \frac{1}{k+1}) \subseteq B(x, \frac{1}{k})$ for each $k \geq 1$, we infer that $s_n \in U$ for all $n \geq n_0$.

□

**Theorem 10.8** (Continuity Equals Convergence Preservation for Metrizable Domains)**.** *Let $f : X \to Y$ be a map from one topological space to another, where $X$ is metrizable, and suppose $x_0 \in X$. The following are equivalent:*

(a) *$f$ is continuous at $x_0$.*
(b) *Whenever $s : D \to X$ is a sequence converging to $x_0$, then $f \circ s : D \to Y$ is a sequence converging to $f(x_0)$.*

*Proof. Ad ((a) $\Longrightarrow$ (b))*: This is Theorem 10.6.

*Ad ((b) $\Longrightarrow$ (a))*: Assume (b) holds. By Theorems 4.9 and 5.11 (d), it suffices to show that whenever $A \subseteq X$ has $x_0$ in its closure, then $f[A]$ has $f(x_0)$ in its closure. So suppose $x_0 \in \mathrm{Cl}(A)$. Since $X$ is metrizable, we infer from Theorem 10.7 that there is a sequence $s : D \to X$ with $s[D] \subseteq A$ and $s \to x_0$. From (b), then, we have $f \circ s \to f(x_0)$. Since $(f \circ s)[D] \subseteq f[A]$, we know that $f(x_0) \in \mathrm{Cl}(f[A])$. (Note that we don't need any additional assumptions for the range space.)

□

In Theorem 10.3, basic properties (ii), (iii), and (iv) beg for converses. Theorem 10.7 gets us a partial converse for (iii) (i.e., we need an added metrizability assumption). Moreover, if we modify its proof, we additionally get a partial converse for (iv) (see Exercise 10.12 (1) below). As for (ii), there is still only a metrizability-conditioned converse.

**Theorem 10.9** (Subsequences)**.** *Suppose $X$ is a metrizable space and $s : D \to X$ is a sequence that accumulates to $x \in X$. Then some subsequence $s|E$ converges to $x$.*

*Proof.* Let $s \vdash x$, and fix a metric $d$ that induces the given topology on $X$. For each $n \geq 1$, set $U_n := B(x, \frac{1}{n})$. Then for $n = 1, 2, \dots$, there is an infinite set $E_n \subseteq D$ such that $s[E_n] \subseteq U_n$. Let $e_1$ now be the least element of $E_1$, let $e_2$ be the least element of $E_2$ greater than $e_1$, etc. This is actually a very simple induction; at the $n$th stage of the construction, we have $e_1 \in E_1, \dots, e_n \in E_n$; we choose $e_{n+1} \in E_{n+1}$ to be the least element of $E_{n+1}$ greater than each $e_i$, $1 \leq i \leq n$. (This is possible because each $E_k$ is infinite.) Now let $E := \{e_1, e_2, \dots\}$. If $U$ is any nbd of $x$, then there is some $n_0 \geq 1$ such that $U_{n_0} \subseteq U$. Let $C := \{e \in E : e \geq n_0\}$. Then $C$ is cofinite in $E$, and $s[C] \subseteq U_{n_0}$ because $U_k \supseteq U_{k+1}$ for each $k \geq 1$. Thus the image of a tail of $s|E$ lies in $U$; hence $s|E \to x$.

□

**Remarks 10.10.** (i) *You may already have noticed that, in the proofs of Theorems 10.7, 10.8, and 10.9, all that we required of the metrizability assumption was the existence of the ball neighborhoods $B(x, \frac{1}{n})$, for $n \geq 1$. The two key properties of this family of nbds are: (i) $B(x, \frac{1}{k}) \supseteq B(x, \frac{1}{k+1})$ for all $k \geq 1$ (i.e., the family forms a **nested sequence of sets**; and (ii) if $U$ is any nbd of $x$, then $B(x, \frac{1}{n}) \subseteq U$ for some (large enough) $n$. A topological space $X$ is called **first countable** if, for each $x \in X$ there is a nested sequence $B_1 \supseteq B_2 \supseteq \ldots$ of open nbds of $x$ such that every nbd of $x$ contains some $B_n$ in the family. The family $\{B_1, B_2, \ldots\}$ is called a **countable nested neighborhood base at** $x$. Metrizable spaces clearly are first countable, but first countable spaces need not be metrizable (see Exercise 10.12 (2) below). Because the proofs of Theorems 10.7–10.9 use no more than first countability, we may replace metrizability by first countability throughout in their statements. We will see more "countability properties" in later sections.*

(ii) *This is as good a point as any to expand on the use of the word* countable *in (i) above. A set is said to be **countable** if it is either finite or can be put in one-one correspondence with the natural numbers $\mathbb{N}$. George Cantor (1845–1918) is credited with initiating in the late 1800s the development of mathematics based on set-theoretic ideas. Of particular importance was the idea of comparing two sets by asking whether they could be put in one-one correspondence with each other; i.e., whether they were of the same cardinality (see Example 8.3). Two basic theorems that he proved, quite counterintuitive at the time, were: (1) that the set of rational numbers is countable, even though it is apparently much "bigger" than the set of natural numbers; while (2) the set of real numbers in uncountable. (It seems Cantor started out believing that both $\mathbb{Q}$ and $\mathbb{R}$ are uncountable, but when he subsequently proved $\mathbb{Q}$ to be countable, he started to believe $\mathbb{R}$ is too. But then came his second surprise: $\mathbb{R}$ is uncountable after all.)*

*So a topological space is first countable just in case each of its points has a nested neighborhood base that also is countable as a family of sets.*

The metrizability (or first countability) assumption in some of the theorems of this section is a *necessary assumption*, in the sense that the wholesale deletion of such an assumption would lead to a false assertion. Take, for example, Theorem 10.9. If we were to delete *metrizable* from the statement, then we would have:

*Suppose $X$ is a space and $s : D \to X$ is a sequence that accumulates to $x \in X$. Then some subsequence $s|E$ converges to $x$.*

How do we go about showing such a statement to be false, if it indeed is false? One way is to try as hard as you can to show it to be true, to look deeply into the proof you have and see whether you truly need the assumption. Sometimes the assumption is reasonably easy to get around in the proof at hand; other times you need to look for a radically new idea. (Quite frankly, this can keep you up nights.)

Well, maybe there are no new ideas lurking about because there can't be. After you've exhausted all the possibilities you can think of, you may start gaining some insight into just *how* crucial the assumption really is in the proof you have, enabling you to build a **counterexample** (i.e., an example that shows a general assertion to be wrong). For instance, the geometrical statement, "All triangles have at least one right angle" is belied by the existence of equilateral triangles, with all angles measuring 60 degrees: the equilateral triangle is then a counterexample to the given general statement. In Theorem 10.4, the Hausdorff assumption is necessary because Example 10.5 affords a $(\mathsf{T}_1)$ counterexample to the version with *Hausdorff* deleted.

Back to our present situation: in order to show the beefed-up version of Theorem 10.9 to be false, you need to find a topological space $X$ and a sequence $s : D \to X$ which accumulates to a point $x \in X$, but such that no subsequence $s|E$ converges to $x$. It turns out such a space does exist, if we believe a certain set-theoretic axiom called the *axiom of choice* (see more advanced texts ). The nice thing about this example is that it also puts to the lie the correspondingly beefed-up versions of Theorems 10.7 and 10.8.

**Example 10.11.** *Our example is a one-point extension, like Example 9.5 (iv), but allows a lot more open nbds of the new point. Start with the natural numbers $\mathbb{N}$, with the discrete topology, and let $p$ be a brand new point. $X := \mathbb{N} \cup \{p\}$ comprises the points of our space; natural numbers $n$ are all isolated points, the point $p$ is intended to be the only nonisolated point. If we were following the prescription of Example 9.5 (iv), we would take as base for our topology, the family $\mathcal{B} := \wp(\mathbb{N}) \cup \{C \cup \{p\} : C \in \mathcal{C}\}$, where $\mathcal{C}$ is the family of cofinite subsets of $\mathbb{N}$. The reason this space is not quite what we want is that it is first countable. (Indeed, see Exercise 10.12 (3) below, it is homeomorphic to the subspace $\{0\} \cup \{\frac{1}{n} : n = 1, 2, \dots\}$ of $\mathbb{R}$, and is hence metrizable.) So what we actually do is add a bunch of new nbds of $p$; and to accomplish this we need the axiom of choice (in the form of* Zorn's lemma, *after M. Zorn (1906–1993)). The upshot is that there exists a family $\mathcal{D} \subseteq \wp(\mathbb{N})$ satisfying:*

(i) $\emptyset \notin \mathcal{D}$;
(ii) $\mathbb{N} \in \mathcal{D}$;
(iii) *if $J \in \mathcal{D}$ and $J \subseteq K \subseteq \mathbb{N}$, then $K \in \mathcal{D}$ (closure under superset);*
(iv) *if $J, K \in \mathcal{D}$, then $J \cap K \in \mathcal{D}$ (closure under finite intersections);*
(v) *every member of $\mathcal{D}$ is infinite (infinity condition) ; and*
(vi) *if $I \cup J \in \mathcal{D}$, then either $I \in \mathcal{D}$ or $J \in \mathcal{D}$ (primality).*

*Given that such a family $\mathcal{D}$ can be found—which it can, with the aid of Zorn's lemma—we let our new base be $\mathcal{B}_{\mathcal{D}} := \wp(\mathbb{N}) \cup \{I \cup \{p\} : I \in \mathcal{D}\}$.*

*We first note that the family $\mathcal{C}$ of cofinite subsets of $\mathbb{N}$ satisfies conditions (i)–(v) above. What it doesn't satisfy is condition (vi): any infinite subset with infinite complement is not in $\mathcal{C}$, even though the union of the two sets is $\mathbb{N} \in \mathcal{C}$. Next we note that, by (v) and (vi), it is immediate that $\mathcal{C} \subseteq \mathcal{D}$. Thus $\mathcal{B}_{\mathcal{D}} \supseteq \mathcal{B}$, as claimed: our topology on $X$ is finer than the one specified in Example 9.5; by Example 9.7 (iii) and Exercise 9.11 (4), then, this space is Hausdorff.*

*Now the space is specified, we need to demonstrate that it does what we want. Let's show that, while $p$ is a limit point of $\mathbb{N}$, there is no sequence $s : D \to X$, with values in $\mathbb{N}$, such that $s \to p$.*

Ad $(p \in \mathrm{Cl}(\mathbb{N}))$*: If $I \in \mathcal{D}$, then $I \neq \emptyset$; so $I \cup \{p\}$ intersects $\mathbb{N}$.*

Ad (No $s \to p$ can have values only in $\mathbb{N}$)*: Suppose, to the contrary, one such $s$ can be found. If $s[D] \notin \mathcal{D}$, then, by (vi), $I := \mathbb{N} \setminus s[D] \in \mathcal{D}$; so we* *immediately get a nbd* $I \cup \{p\}$ *of* $p$ *that cannot contain the image of a tail of* $s$*. Thus it must be the case that* $s[D] \in \mathcal{D}$*; in particular,* $s[D]$ *is an infinite set, by (v). For convenience, write* $s := \langle s_0, s_1, \dots \rangle$*, and set* $t_0 := s_0$*. Using a simple induction, we let* $t_1$ *be the first* $s_n > s_0$*, let* $t_2$ *be the first* $s_n > t_1$*, etc. Thus we have replaced* $s$ *by a strictly increasing subsequence* $t$*; moreover if* $s \to p$ *then* $t \to p$ *as well. As with* $s$*, we then argue that* $t[D] \in \mathcal{D}$*. (The argument just given, obtaining* $t$ *from* $s$*, might have been omitted, being replaced by the more enigmatic: "... without loss of generality, we may assume* $s$ *is strictly increasing ... ") Now let* $t[D] := I \cup J$*, where* $I$ *and* $J$ *are infinite and disjoint. Again using (vi), one of* $I$*,*$J$ *must be in* $\mathcal{D}$*; say it's* $I$*. Then* $I \cup \{p\}$ *is a nbd of* $p$ *that fails to intersect the image of a tail of* $t$*. Consequently,* $t$ *cannot converge to* $p$*.*

*So far we have a counterexample to the version of Theorem 10.7 that omits the metrizability assumption (i.e., the (b)* $\Longrightarrow$ *(a) direction). But now we can see quickly that* $X$ *is also a counterexample to the corresponding version of Theorem 10.9. For let* $s_n := n$*,* $n = 0, 1, \dots$*. Then clearly every nbd* $I \cup \{p\}$*, for* $I \in \mathcal{D}$*, contains the image of a subsequence of* $s$*, namely* $s|I$*. Thus* $p$ *is an accumulation point of* $s$*. However, we have shown that no sequence with values in* $\mathbb{N}$ *can converge to* $p$*. So, in particular, no subsequence of* $s$ *can converge to* $p$*.*

*Finally, to address Theorem 10.8, let* $Y$ *be* $\langle X, \mathcal{T} \rangle$*, where* $\mathcal{T}$ *is the discrete topology. Then the identity map* $\iota_X$ *fails to be continuous at* $p$*. However, if* $s : D \to X$ *is a sequence converging to* $p$ *in the original topology, then (see Exercise 10.12 (4) below)* $s$ *is eventually constant at* $p$*. Consequently,* $\iota_X \circ s = s$ *is eventually constant at* $p$*, and hence converges to* $p$ *in the discrete topology.*

**Exercises 10.12.** (1) ** Show that if* $X$ *is a first countable* $\mathsf{T}_1$ *space and* $x$ *is a limit point of* $A \subseteq X$*, then there is a sequence* $s : D \to X$ *with* $s \to x$ *and* $s[D] \subseteq A$ *infinite. [Hint: Modify the proof of Theorem 10.7. If* $s_1 \in B_1$ *differs from* $x$*, find* $s_2 \in B_n$*, where* $n$ *is least such that* $s_1 \notin B_n$*. Use induction.]*

(2) *Find a first countable* $\mathsf{T}_1$ *topological space that is not metrizable. [Hint: The cofinite topology works. There are more sophisticated examples; the Sorgenfrey line* $\mathbb{L}$ *is first countable* $\mathsf{T}_2$ *and nonmetrizable. (We won't have the technology to demonstrate this, however, until Section 17.)]*

(3) *Prove that the one-point extension of the discrete natural numbers, formed by letting cofinite subsets serve as a nbd base for the new point, is homeomorphic to the subset of* $\mathbb{R}$ *consisting of* $0$*, plus points* $\frac{1}{n}$*,* $n \geq 1$*.*

(4) *Refer to Example 10.11, and show that if* $s : D \to X$ *is a convergent sequence in* $X$*, then* $s$ *is eventually constant.*

(5) *Let* $\mathbb{L}$ *be the Sorgenfrey line (see Exercise 7.18 (4)), and let* $s : \mathbb{N} \to \mathbb{L}$ *be a strictly increasing sequence in* $X$ *(i.e.,* $s_0 < s_1 < \dots$*). Show* $s$ *has no accumulation points.*

(6) *Refer to Example 6.9, where* $\mathbb{R}^{[0,1]}$ *has the supremum norm. Suppose* $s : D \to \mathbb{R}^{[0,1]}$ *is a convergent sequence, and that each* $s_n$ *is a continuous function from* $[0,1]$ *to* $\mathbb{R}$*. Show that the limit of this sequence is also a continuous function.*

(7) *Refer to Example 7.17 (ii), and let* $\mathbb{R}^{[0,1]}$ *have the topology of pointwise convergence. For each* $n \geq 1$*, let* $s_n$ *be the function taking* $x \in [0,1]$ *to* $x^n$*. Does the sequence* $\langle s_1, s_2, \dots \rangle$ *have a limit in* $\mathbb{R}^{[0,1]}$*? What does this say about limits of sequences of continuous functions relative to the topology of pointwise convergence?*

(8) *Prove the* ***bounded convergence theorem****: If* $s : D \to \mathbb{R}$ *is a bounded weakly increasing sequence of real numbers (i.e., there is some* $M \in \mathbb{R}$ *such that* $s_n \leq M$ *for all* $n$*, and* $s_0 \leq s_1 \leq \dots$*), then* $s$ *is convergent.*

(9) ** Consider the basic* ***Fibonacci sequence*** *(after Leonardo Pisano Fibonacci (1170–1250))* $\langle 1, 1, 2, 3, 5, 8, 13, \dots \rangle$ *(where* $s_{n+2} = s_{n+1} + s_n$*, in general). Without peeking on the web, show that the sequence of ratios* $r_n := \dfrac{s_{n+1}}{s_n}$ *converges; and find exactly what the limit of this sequence is.*

(10) ** Define a sequence* $s : D \to \mathbb{R}$ *to be* ***busy*** *if, for each rational number* $r \in \mathbb{Q}$*,* $s^{-1}[r] := \{n \in D : s_n = r\}$ *is infinite. First show that busy sequences exist; and, once you've done that, find the set of all accumulation points of a busy sequence.*

(11) *Let* $X$ *be a topological space. Show that* $X$ *is first countable if and only if each point has a family of open nbds* $\{B_1, B_2, \dots\}$ *such that every nbd of* $x$ *contains some* $B_n$ *in the family (no nestedness assumption).*

## 11. Connectedness

At the beginning of Section 9 we talked about topological properties having certain "flavors." The "separating stuff" flavor gives us the separation axioms $\mathsf{T}_n$, $n = 0, 1, 2$ (plus quite a few more, as we shall see in Section ); in this section we consider the flavor, "coming in one piece." Intuitively, a space $X$ "comes in one piece" if it cannot be *disconnected*; i.e., if it doesn't come in "two separated pieces."

**Definition 11.1** (Connectedness)**.** *Let $X$ be a topological space. A* ***disconnection of*** *$X$ is a two-element family $\{H, K\}$ of disjoint nonempty open subsets of $X$, whose union is $X$. $X$ is* ***disconnected*** *if there exists a disconnection of $X$;* ***connected*** *otherwise. A subset $A$ of $X$ is a* ***connected subset*** *if it is connected as a subspace of $X$. This is equivalent to saying that there are no open subsets $U$ and $V$ of $X$ such that: $U \cap A \neq \emptyset \neq V \cap A$; $U \cap V \cap A = \emptyset$; and $A \subseteq U \cup V$.*

**Remarks 11.2.**
(i) *A family of subsets of $X$ whose union is all of $X$ is called a* ***cover*** *of $X$. If the sets all happen to be open, the cover is called an* ***open*** *cover of $X$. (Similarly, define* ***closed*** *cover, etc.) In the language of open covers, then, a disconnection of $X$ is an open cover of $X$ consisting of two nonempty disjoint sets. They are the two "pieces" mentioned earlier. Of course the two sets should be disjoint; more strongly they should be "separated:" neither set should contain a limit point of the other. Notice that, since $H$ and $K$ are disjoint complementary open sets, they are also both closed sets; so this idea of "separatedness" seems to be reasonably explicated in topological language. Sets that are both closed and open (e.g., the empty set and the whole space) are (rather frivolously) termed* ***clopen****.*
(ii) *Refer to Example 8.9 (v). The domain of the analytic function in the statement of the open mapping theorem is assumed to be an open set that is* connected *(frequently called a* region *in the parlance of complex analysis).*

Here is a result that rephrases connectedness in other terms.

**Theorem 11.3** (Characterizations of Connectedness)**.** *The following are equivalent for any topological space $X$:*
(a) *$X$ is connected.*
(b) *There is no continuous map from $X$ onto the two-point discrete space.*
(c) *The only clopen subsets of $X$ are $\emptyset$ and $X$.*
(d) *Whenever $\mathcal{U}$ is an open cover of $X$ and $x, y$ are points of $X$, there is a finite sequence $\langle U_1, \ldots, U_n \rangle$ of members of $\mathcal{U}$ such that $x \in U_1$, $y \in U_n$, and $U_i \cap U_{i+1} \neq \emptyset$ for $1 \leq i \leq n-1$.*

*Proof. Ad ((a) $\Longrightarrow$ (b))*: Suppose (b) is false. If $f : X \to \{a, b\}$ is a continuous surjection, where both $a$ and $b$ are isolated points, then the set $\{f^{-1}[\{a\}], f^{-1}[\{b\}]\}$ forms a disconnection of $X$. Thus (a) is false.

*Ad ((b) $\Longrightarrow$ (c))*: Suppose (c) is false. If there is a nonempty clopen set $U \subseteq X$ whose complement is also nonempty, then we may construct $f : X \to \{a, b\}$ by assigning points of $U$ to $a$ and points of $X \setminus U$ to $b$. This gives us a continuous surjection, no matter what topology is assigned to $\{a, b\}$; hence (b) is false.

*Ad ((c) $\Longrightarrow$ (d))*: Suppose $\mathcal{U}$ is an open cover of $X$. For the purposes of this proof, let's say points $x$ and $y$ are **$\mathcal{U}$-linkable** if there is a sequence $\langle U_1, \dots, U_n \rangle$, as indicated above. Now fix $x \in X$ and define $U$ to be the set of points $y \in X$ such that $x$ and $y$ are $\mathcal{U}$-linkable. Since $\mathcal{U}$ is an open cover—so every member of $X$ is contained in a member of $\mathcal{U}$, hence $\mathcal{U}$-linkable with itself—we have $x \in U$. The plan is to show that $U$ is clopen. For then, assuming (c), it must be all of $X$; hence (d) holds.

If $y \in U$ is arbitrary, with $\langle U_1, \dots, U_n \rangle$ witnessing the $\mathcal{U}$-joinability of $x$ to $y$, then each $z \in U_n$ is also in $U$, as witnessed by the very same finite sequence of sets. This says that $U$ is an open subset of $X$. If $y \in \mathrm{Cl}(U)$, pick $V \in \mathcal{U}$ such that $y \in V$. Then there is some $z \in V \cap U$, so there is a finite sequence $\langle U_1, \dots, U_n \rangle$ witnessing the fact; i.e., $x \in U_1$, $z \in U_n$, and each $U_i \cap U_{i+1} \neq \emptyset$ for $1 \leq i \leq n-1$. But then the sequence $\langle U_1, \dots, U_n, V \rangle$ witnesses the fact that $y \in U$. This tells us that $U$ is closed in $X$, and (d) therefore holds.

*Ad ((d) $\Longrightarrow$ (a))*: Suppose (a) is false. Then any disconnection $\mathcal{U} := \{H, K\}$ demonstrates the failure of (d).

$\square$

**Examples 11.4.** (i) *Every space with no more than one point is, trivially, connected.*

(ii) *The Sierpiński space $\mathbb{S}$ of Example 9.2 (ii) is connected.*

(iii) *If $X$ is an infinite space with the cofinite topology, then $X$ is connected because any two nonempty open sets must overlap. There is no possibility for a disconnection if this happens.*

(iv) *Let $X$ be the subset of the usual real line consisting of the union of the two closed intervals $[0, 1]$ and $[2.3]$. Then $X$ is disconnected because $\{[0, 1], [2, 3]\}$ is a disconnection of $X$. (Note that $[0, 1]$ is open in $X$ because $[0, 1] = (-1, \frac{3}{2}) \cap X$.)*

(v) *Let $\mathbb{L}$ be the real line with the Sorgenfrey (lower limit) topology (see Exercise 7.18 (4)). Then $\mathbb{L}$ fails to be connected. Indeed, if $x < y$, then pick $a \in (x, y)$. The family $\{(-\infty, a), [a, \infty)\}$ is a disconnection of $\mathbb{L}$ that "disconnects" $x$ and $y$. Since $x$ and $y$ were arbitrarily chosen, this says that no two distinct points of $\mathbb{L}$ lie in a connected subset of $\mathbb{L}$. Topological spaces with this property are called **totally disconnected**.*

(vi) *The usual closed unit interval $[0, 1]$ is connected. To see this, suppose otherwise, that $\{H, K\}$ is a disconnection of $[0, 1]$. We may as well assume $0 \in H$. Then, because $H$ is open in $[0, 1]$, there is some $\epsilon > 0$ such that $[0, \epsilon) \subseteq H$. Thus for every $0 \leq t < \epsilon$, $t$ is a lower bound of $K$. Using the order completeness of the real line, let $k$ be the greatest lower bound of $K$. Then $k \geq \epsilon > 0$. Can $k = 1$? If so, then $K$ has no choice but to be $\{1\}$. But then $K$ is not open in $[0, 1]$, a contradiction. From this we conclude that $0 < k < 1$. Now let $U$ be any open nbd of $k$; without loss of generality, we may assume that $U = (k - \delta, k + \delta)$ for suitably small $\delta > 0$. Now, because $k$ is a lower bound of $K$, it must be the case that $(k - \delta, k) \cap K = \emptyset$. Thus $(k - \delta, k) \subseteq H$; so $U \cap H \neq \emptyset$. This says that*

*$k \in Cl(H)$. Since $H$ is already closed in $[0,1]$ we know $k \in H$. On the other hand, if $(k, k+\delta) \cap K = \emptyset$, then there must be lower bounds of $K$ strictly to the right of $k$. But $k$ is the* greatest *lower bound of $K$; hence we know $U \cap K \neq \emptyset$, and that $k \in Cl(K) = K$. Thus $k \in H \cap K = \emptyset$, a contradiction.*

In general terms, it is more difficult to prove connectedness than the contrary. Proving disconnectedness requires an example of a disconnection; proving connectedness requires showing that no such example can exist. Thus proof of connectedness is often proof by contradiction. There are, however some general theorems that help to convert a connectedness result about one space to one about another space. One of the most basic of these involves continuous mappings; it is refreshingly easy to prove.

**Theorem 11.5** (Connectedness and Continuity)**.** *Let $f : X \to Y$ be a continuous surjection, where $X$ is connected. Then $Y$ is also connected.*

*Proof.* Let $\{H, K\}$ be a disconnection of $Y$. Then the inverse images $f^{-1}[H]$ and $f^{-1}[K]$ are disjoint, open in $X$, both nonempty (because $f$ is onto), and together cover $X$.

□

By adjusting Example 8.3 (v) a little, and noticing that Example 11.4 (vi) actually shows that every bounded closed real interval is connected, we immediately obtain the following.

**Corollary 11.6.** *Any **simple closed curve** (i.e., homeomorphic copy of the standard unit circle $S^1$) is connected.*

Another powerful-but-simple tool in the search for connectedness involves connected subsets.

**Theorem 11.7.** *Suppose $C$ is a connected subset of a space $X$. If $\{H, K\}$ is a disconnection of $X$, then either $C \subseteq H$ or $C \subseteq K$. Consequently, if each two-point subset of $X$ lies in a connected subset of $X$, then $X$ is connected.*

*Proof.* Suppose, to the contrary, that $\{H, K\}$ is a disconnection of $X$ such that both $H \cap C$ and $K \cap C$ are nonempty. Then we have a disconnection of $C$ in $\{H \cap C, K \cap C\}$.

Suppose now that there is a disconnection $\{H, K\}$ of $X$. Pick $x \in H$ and $y \in K$. By assumption, there is a connected subset $C \subseteq X$ containing both $x$ and $y$. But then $C$ intersects both $H$ and $K$, a contradiction.

□

We now have enough ammunition to consider connectedness in spaces of higher (euclidean) dimension.

**Corollary 11.8** (Connectedness in Ball Neighborhoods)**.** *Let $\mathbb{R}^n$ be supplied with the euclidean metric. Then the ball neighborhoods $B(\mathbf{x}, \epsilon)$ are connected.*

*Proof.* Given any two distinct points $\mathbf{y}$ and $\mathbf{z}$ in $B := B(\mathbf{x}, \epsilon)$, the parameterized line segment $S := \{t\mathbf{y} + (1-t)\mathbf{z} : t \in [0,1]\}$ lies entirely in $B$, by simple linear algebra. (Check out the string of inequalities: $|(t\mathbf{y} + (1-t)\mathbf{z}) - \mathbf{x}| = |(t\mathbf{y} + (1-t)\mathbf{z}) - (t\mathbf{x} + (1-t)\mathbf{x})| = |t(\mathbf{y}-\mathbf{x}) + (1-t)(\mathbf{z}-\mathbf{x})| \leq t|\mathbf{y}-\mathbf{x}| + (1-t)|\mathbf{z}-\mathbf{x}| < t\epsilon + (1-t)\epsilon = \epsilon$. $B$ is a **convex set**, in the sense of linear algebra.) $S$ is easily seen to be homeomorphic to $[0,1]$, so is connected. By Theorem 11.7, then, $B$ is connected.

□

One goal of this section is a proof of one of the "twin pillars" mentioned in Theorem 1.5, namely the intermediate value theorem. First we need to characterize connectedness in subsets of the usual real line; we do this in three easy steps.

**Theorem 11.9.** *Let $\{A_i : i \in I\}$ be a family of connected subsets of a space $X$. If $A_i \cap A_j \neq \emptyset$ for each $i, j \in I$, then $\bigcup_{i \in I} A_i$ is connected.*

*Proof.* Set $A := \bigcup_{i \in I} A_i$, and suppose $\{H, K\}$ is a family of two disjoint relatively open subsets of $A$ whose union is $A$. By Theorem 11.7, each $A_i$, being a connected subset of $A$, is contained in either $H$ or in $K$. Pick $i_0 \in I$, say $A_{i_0} \subseteq H$. If $j \in I$ is arbitrary, then it is impossible for $A_j$ to be a subset of $K$ because then $\emptyset \neq A_j \cap A_{i_0} \subseteq K \cap H = \emptyset$. Thus $A = H$, and hence $K = \emptyset$. This tells us that $A$ is connected.

□

**Theorem 11.10** (Connected Subsets of $\mathbb{R}$)**.** *Let $\mathbb{R}$ have the usual topology, with $A \subseteq \mathbb{R}$. Then $A$ is connected if and only if $A$ is an interval.*

*Proof.* Suppose $A$ is not an interval. Then $A$ is not convex; i.e., there exist $x < y < z$ in $\mathbb{R}$ such that $x$ and $z$ are in $A$, but $y$ is not. Then $\{(-\infty, y) \cap A, (y, \infty) \cap A\}$ forms a disconnection of $A$; hence $A$ is not a connected set.

For the converse, suppose $A$ is an interval. If $A$ is a single point, we're done; otherwise we may write $A$ as the union of an expanding sequence of bounded closed intervals. (For example, $(2,5) = \bigcup_{n \geq 1}[2 + \frac{1}{n}, 5 - \frac{1}{n}]$, $[2, \infty) = \bigcup_{n \geq 1}[2, n]$, etc.) Since all (nonsingleton) bounded closed intervals in the real line are homeomorphic to $[0,1]$ (using an argument similar to that in Theorem 8.12), and since $[0,1]$ is connected (by Example 11.4 (vi)), so too is every bounded closed interval. We now use Theorem 11.9: all the intervals in the expanding sequence whose union is the interval $A$ are connected; hence so is $A$ itself.

□

And now, finally, we have one of the advertised pillars of Freshman calculus.

**Corollary 11.11** (Intermediate Value Theorem)**.** *Suppose $X \subseteq \mathbb{R}$ is an interval, with $f : X \to \mathbb{R}$ continuous on $X$. If $a$ and $b$ lie in $X$ and $d \in \mathbb{R}$ lies between $f(a)$ and $f(b)$, then there is some $c$ between $a$ and $b$ such that $f(c) = d$.*

*Proof.* For any $a < b$ lying in the interval $X$, we have $[a, b] \subseteq X$. By Theorem 11.10, $[a, b]$ is connected; by Theorem 11.5, so too is $f[[a, b]] \subseteq \mathbb{R}$. Again using 11.10, $f[[a, b]]$ is an interval, one that contains both $f(a)$ and $f(b)$. So if $d$ lies between $f(a)$ and $f(b)$, then $d \in f[[a, b]]$. Thus there is some $c \in [a, b]$ such that $f(c) = d$.

□

From Theorem 11.9 the union of a bunch of connected subsets all having a single point in common must be connected. This prompts the following concept.

**Definition 11.12** (Components)**.** *Let $X$ be a topological space, with $x \in X$. We denote by $C(x)$ the union of all connected subsets $A \subseteq X$ with $x \in A$. $C(x)$ is called the* ***component of*** *$X$* ***containinng*** *$x$; a subset of $X$ is called simply a* ***component of*** *$X$ if it is equal to come $C(x)$.*

Before going on to identify some of the key properties of components, we prove the following useful result. It too is a way of obtaining new connected sets from old ones.

**Theorem 11.13.** *Let $A$ be a connected subset of a space $X$, with $B$ any set such that $A \subseteq B \subseteq Cl(A)$. Then $B$ is connected.*

*Proof.* Suppose we have a disconnection of $B$. That means there are sets $U$ and $V$, open in $X$, such that: $(U \cap B) \cap (V \cap B) = U \cap V \cap B = \emptyset$, $B \subseteq U \cup V$, and both $U \cap B$ and $V \cap B$ are nonempty. Since $B \subseteq \mathrm{Cl}(A)$, every nbd of a point of $B$ must intersect $A$. Hence both $U \cap A$ and $V \cap A$ are nonempty. Since $(U \cap A) \cap (V \cap A) \subseteq (U \cap B) \cap (V \cap B) = \emptyset$ and $A \subseteq B \subseteq U \cup V$, we see that $\{U \cap A, V \cap A\}$ is a disconnection of $A$, contradicting the fact that $A$ is connected.
□

**Theorem 11.14** (Basic Properties of Components)**.** *Let $X$ be a topological space.*

(i) *For any $x \in X$, $C(x)$ is connected; if $D \supseteq C(x)$ is a connected subset of $X$, then $D = C(x)$ (i.e., $C(x)$ is a* ***maximally connected subset of*** *$X$).*
(ii) *$X$ is connected if and only if $C(x) = X$ for every $x \in X$; $X$ is totally disconnected if and only if $C(x) = \{x\}$ for every $x \in X$.*
(iii) *For any $x, y \in X$, either $C(x) = C(y)$ or $C(x) \cap C(y) = \emptyset$.*
(iv) *$C(x)$ is closed in $X$.*

*Proof. Ad (i)*: The collection of connected subsets of $X$ containing the point $x$ is easily seen to satisfy the hypothesis of Theorem 11.9. Hence its union, $C(x)$, must be connected. If $D$ is a connected set containing $C(x)$, then, in particular, $D$ is a connected set containing $x$; hence is one of the sets making up the union $C(x)$. Therefore $D \subseteq C(x)$, and the two sets must be equal.

*Ad (ii)*: If $X$ is connected, then $X$ is one of the connected subsets of $X$ containing arbitrary $x \in X$. Thus $X \subseteq C(x)$; and therefore $C(x) = X$ for every $x \in X$. The converse is just as easy.

*Ad (iii)*: Suppose $x, y \in X$ are given. If $C(x) \cap C(y) \neq \emptyset$, then, by Theorem 11.9 and (i) above, $C(x) \cup C(y)$ is a connected superset of both $C(x)$ and $C(y)$. Thus, again by (i) above, $C(x) = C(x) \cup C(y) = C(y)$.

*Ad (iv)*: This is immediate from Theorem 11.13 and (i) above: $\mathrm{Cl}(C(x))$ is a connected superset of $C(x)$; hence the two sets must be equal. Thus $C(x)$ is a closed subset of $X$.

□

The next idea is of major importance in the study of connectedness. For one thing, it allows us to distinguish many spaces as being topologically different.

**Definition 11.15** (Cut Sets and Cut Points)**.** *Let $X$ be a connected space. A subset $A$ of $X$ is called a **cut set** of $X$ if $X \setminus A$ is disconnected. If $A$ consists of the single point $x$, we say $x$ is a **cut point** of $X$.*

It is easy to show that any homeomorphism between connected spaces preserves the property of being a cut set/point; i.e., if $X$ is connected, $A$ is a (non)cut set of $X$, and $h : X \to Y$ is a homeomorphism, then $Y$ is connected and $h[A]$ is a (non)cut set of $Y$. In particular, if two spaces have a differing number of cut points or of noncut points, then they cannot be homeomorphic. We end this section with a number of examples where cut sets give us important information about a space.

**Examples 11.16.** (i) *Bounded closed, half-open, and open intervals in $\mathbb{R}$ have 2, 1, and 0 noncut points, respectively. Hence they are in separate homeomorphism classes, as promised in Example 8.13 (ii). Also, any unbounded interval in $\mathbb{R}$ can have at most one noncut point; hence cannot be homeomorphic to a bounded closed interval. Of course any two half-open intervals, bounded or not, are homeomorphic to one another, as are any two open intervals.*

(ii) *If $n > 1$, then $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}$. To see this, it is an easy geometrical exercise to show that $\mathbb{R}^n$ has no cut points for $n \geq 1$.*

(iii) *In Example 8.5 (iii), we promised to show that the unit circle $S^1$ is not embeddable in the reals. This is true because: $S^1$ is connected (Corollary 11.6); every connected nonsingleton subset of $\mathbb{R}$, being an interval, has at least one cut point (see the proof of Theorem 11.10); and $S^1$ has no cut points at all. Indeed (see Exercise 11.18 (10) below), the removal of any point from $S^1$ leaves a space homeomorphic to $\mathbb{R}$.*

(iv) *If $A \subseteq \mathbb{R}^n$, $n > 1$, and $A$ is countable, then $A$ is not a cut set for $\mathbb{R}^n$. To see this, take the case $n = 2$ (the argument being essentially the same in higher dimensions). If $\mathbf{x}$ and $\mathbf{y}$ are two distinct points in $\mathbb{R}^n \setminus A$, let $S$ be the straight line segment with endpoints $\mathbf{x}$ and $\mathbf{y}$; then let $L$ be the perpendicular bisector of $S$ (so $L$ is the straight line perpendicular to $S$, going through the midpoint of $S$. Now pick any point $\mathbf{z}$ on $L$, and let $B_{\mathbf{z}}$ be the union of the line segment joining $\mathbf{x}$ to $\mathbf{z}$, with the line segment joining $\mathbf{z}$ to $\mathbf{y}$. Then each $B_{\mathbf{z}}$ is connected and contains both $\mathbf{x}$ and $\mathbf{y}$; moreover, if $\mathbf{z}$ and $\mathbf{w}$ are any two distinct points on the line $L$, then $B_{\mathbf{z}} \cap B_{\mathbf{w}} \cap A = \emptyset$. Now the points on a straight line may easily be put in one-one correspondence with $\mathbb{R}$ (indeed,* *via a homeomorphism). Thus $L$ must be uncountable. Hence, if each $B_{\mathbf{z}}$ were to intersect $A$, that would force $A$ to be uncountable too. The upshot is that there is some $\mathbf{z} \in L$ such that $B_{\mathbf{z}} \subseteq \mathbb{R}^2 \setminus A$. By Theorem 11.7, $\mathbb{R}^2 \setminus A$ is connected.*

(v) *(This is a deep theorem, the* curve theorem, *first claimed in 1887 by Camille Jordan (1838–1922). It has a deceptively simple statement, making it seem almost obvious, but Jordan's original proof was completely wrong. It wasn't until 1905 when the first correct proof appeared, due to Oswald Veblen (1880–1960). Although we now have the mathematical vocabulary to understant the statement of this extremely wide-reaching theorem, its proof is beyond the scope of this course.) If $S$ is a simple closed curve lying in $\mathbb{R}^2$ (i.e., $S$ is the image under an embedding from the standard unit circle into the plane), then $S$ is a cut set of $\mathbb{R}^2$. Moreover, $\mathbb{R}^2 \setminus S$ consists of exactly two components, exactly one of which is bounded relative to the euclidean metric.*

(vi) *One way to distinguish topologically $\mathbb{R}^m$ from $\mathbb{R}^n$ when $m > n > 1$ is via cut sets and the topological theory of dimension. (This turns out to be very hard mathematics.) For example, it is not terribly difficult to show that no simple closed curve in $\mathbb{R}^m$, $m > 2$, is a cut set therein. This, coupled with the Jordan curve theorem stated in (v) above, lets us know that no $\mathbb{R}^m$ is homeomorphic to $\mathbb{R}^2$, unless $m = 2$.*

**Exercises 11.17.** (1) *Show that a connected $\mathsf{T}_1$ space with more than one point has no isolated points, but that it is possible for a connected $\mathsf{T}_0$ space with more than one point to have isolated points.*

(2) *Show that if a topology on a set is coarser than a connected topology on the set, then the coarse topology is connected also.*

(3) *Show that the space $\mathbb{Q}$ of rational numbers, with the usual topology, is totally disconnected (i.e., components are singletons).*

(4) ** Show that $\mathbb{R}^2$ with the topology induced by the lexicographic ordering (see Example 7.6 (iv)) is not connected. What are the components of this space?*

(5) ** Let $X$ be a connected space, with $A$ a connected subset. Show the following:*
  (a) *If $U$ is a clopen subset of $X \setminus A$, then $U \cup A$ is connected.*
  (b) *If $C$ is a component of $X \setminus A$, then $X \setminus C$ is connected.*

(6) *Describe precisely just what are the connected subsets of a space whose topology is the cofinite topology.*

(7) *Let $\{C_n : n \in \mathbb{N}\}$ be a countable family of connected subsets of a space $X$, satisfying $C_n \cap C_{n+1} \neq \emptyset$ for each $n \in \mathbb{N}$. Show that $\bigcup_{n\in\mathbb{N}} C_n$ is connected.*

(8) *For any point $x$ of a space $X$, define the* ***quasicomponent*** $Q(x)$ *of $x$ to be the set of points $y \in X$ such that there is no disconnection $\{H, K\}$ of $X$ with $x \in H$ and $y \in K$. Prove the following basic results about quasicomponents:*
   (a) *$X$ is connected if and only if $Q(x) = X$ for every $x \in X$.*
   (b) *For any $x, y \in X$, either $Q(x) = Q(y)$ or $Q(x) \cap Q(y) = \emptyset$.*
   (c) *Every quasicomponent is a union of components.*
   (d) *For each $x \in X$, $Q(x)$ is the intersection of all clopen sets in $X$ that contain $x$.*

(9) *Define a space $X$ to be* ***zero-dimensional*** *if it has a base for its topology that consists of clopen sets. Show that if $X$ is zero-dimensional and $\mathsf{T}_0$, then $X$ is Hausdorff and $Q(x) = \{x\}$ for every $x \in X$.*

(10) *Let $x$ be a point on the unit circle $S^1$. Show that $S^1 \setminus \{x\}$ is homeomorphic to $\mathbb{R}$.*

(11) *Show that if $X$ has finitely many components, then each component is a clopen set.*

## 12. Path Connectedness

In spite of the amazing power and generality of the notion of connectedness studied in the last section, there is something eerily "negativistic" about its definition: the *existence* of connectednss is conditioned by the *nonexistence* of a disconnection. There is another connectedness notion, one based on the simple idea of "walking from here to there and not wandering off the set." It may be more intuiuively appealing; it's certainly more direct.

**Definition 12.1** (Path Connectedness)**.** *Let $X$ be a topological space, $x$ and $y$ two points of $X$. By a* ***path from*** *$x$* ***to*** *$y$ in $X$ we mean a continuous map $f : [a, b] \to X$ such that $f(a) = x$ and $f(b) = y$, where $[a, b]$ is any bounded closed interval in $\mathbb{R}$. The space $X$ is* ***path connected*** *if, for each two points of $X$, there is a path from one to the other. A subset $A$ of $X$ is a* ***path connected subset*** *if $A$ is path connected as a subspace of $X$.*

**Remark 12.2.** *Just as a sequence should not be confused with its trace (or image) in the range space, so too a path should not be confused with its trace. (The trace of a path (or sequence) is much easier to depict in a drawing than is the path itself.) In particular, the natural order on $[a, b]$ suggests that a path has a "direction." (Recall line integrals in calculus.) Moreover, without more stringent conditions on the path (e.g., being one-one), the trace of a path needn't be "skinny:" it is quite possible for a path to be surjective onto a square, cube, or higher-dimensional object; i.e., to be "space-filling."*

Connectedness and path connectedness are indeed related, but somewhat subtly. Here are some of the less subtle things you can say.

**Theorem 12.3** (Basic Consequences of Path Connectedness)**.**

(i) *A path connected space is connected.*

(ii) *A continuous image of a path connected space is path connected.*

(iii) *Let $\{A_i : i \in I\}$ be a family of path connected subsets of a space $X$. If $A_i \cap A_j \neq \emptyset$ for each $i, j \in I$, then $\bigcup_{i \in I} A_i$ is path connected.*

*Proof. Ad (i)*: Suppose $X$ is path connected. If $x$ and $y$ are any two points of $X$, then there is a path $f : [0, 1] \to X$ with $f(0) = x$ and $f(1) = y$. By Example 11.4 (vi), $[0, 1]$ is connected; by Theorem 11.5, so is $f[[0, 1]]$. The latter is then a connected subset of $X$ containing both $x$ and $y$; hence, by Theorem 11.7, $X$ is connected.

*Ad (ii)*: (This is the path-connectedness version of Theorem 11.5.) Suppose $g : X \to Y$ is a continuous surjection, where $X$ is path connected; and let $y$ and $z$ be two points of $Y$. We need to find a path in $Y$ that joins $y$ and $z$. Indeed, since is onto, there are $u, v \in X$ such that $g(u) = y$ and $g(v) = z$. Since $X$ is path connected, there is a path $f : [0, 1] \to X$ with $f(0) = u$ and $f(1) = v$. Thus $g \circ f$ is a path in $Y$ joining $y$ and $z$.

*Ad (iii)*: (This is the path-connectedness version of Theorem 11.9.) Let $A := \bigcup_{i\in I} A_i$, with $x$ and $y$ two points of $A$. Then thre are indices $i, j \in I$ such that $x \in A_i$ and $y \in A_j$. Let $z \in A_i \cap A_j$. Then there are paths $f : [0,1] \to A_i$ and $g : [0,1] \to A_j$ with $f(0) = x$, $f(1) = g(0) = z$, and $g(1) = y$ (because both $A_i$ and $A_j$ are path connected). We now pull a little trick: define $h : [0,1] \to A$ by letting $h(t) := f(2t)$, $0 \le t \le \frac{1}{2}$, and letting $h(t) := g(2t-1)$, $\frac{1}{2} \le t \le 1$. Note first that, when $t = \frac{1}{2}$, we have $f(2t) = f(1) = z = g(0) = g(2t-1)$; so there is no ambiguity in the definition of $h$ at $t = \frac{1}{2}$ (or anywhere else). It is then an easy exercise (see Exercise 12.6 (1) below) to show that $h$ is continuous. Since $h(0) = f(0) = x$ and $h(1) = g(1) = y$, we infer that $A$ is path connected.

□

We shall see later that connected spaces need not be path connected, so there is no full converse th Theorem 12.3 to be hoped for. However,there is a very nice partial converse.

**Theorem 12.4.** *Any connected open subset of euclidean space is path connected.*

*Proof.* Let $\mathbb{R}^n$ be equipped with the euclidean metric, and suppose $U \subseteq \mathbb{R}^n$ is connected and open. Since the open ball neighborhoods constitute a base for the euclidean topology, we may form an open cover $\mathcal{B}$ of $U$ with these sets. (I.e., for each $x \in U$, there is some $\epsilon_x > 0$ with $B(x, \epsilon_x) \subseteq U$; hence $U = \bigcup_{x\in U} B(x, \epsilon_x)$.) By Corollary 11.8, each open ball nbd is connected. In fact, because one may connect any two points in such a nbd using a straight line segment lying entirely in that nbd, it is immediate that the open ball nbds are path connected as well.

So let $\mathbf{x}$ and $\mathbf{y}$ be any two points in $U$. By Theorem 11.3, there are nbds $B_1, \dots, B_n$ from $\mathcal{B}$ such that $\mathbf{x} \in B_1$, $\mathbf{y} \in B_n$, and $B_i \cap B_{i+1} \ne \emptyset$ for $1 \le i \le n-1$. By Theorem 12.3 (iii), plus an easy induction, the union $B_1 \cup \cdots \cup B_n$ is path connected. The two points $\mathbf{x}$ and $\mathbf{y}$ can now be joined by a path in $B_1 \cup \cdots \cup B_n$, and hence by a path in the larger set $U$.

□

We end this section with a single counterexample to the following two conjectures:

(i) Every connected space is path connected.
(ii) (Path-connectedness version of Theorem 11.13) The closure of a path connected subset of a space is path connected.

**Example 12.5.** *In the euclidean plane $\mathbb{R}^2$, we construct the "topologist's square wave" as follows: First let $V_0$ be the vertical line segment $\{0\} \times [-1,1]$. Second, for each natural number $n \ge 1$, let $V_n$ be the vertical line segment $\{\frac{1}{n}\} \times [-1,1]$. Third, for each natural number $n \ge 1$, let $H_n$ be the horizontal line segment $[\frac{1}{n+1}, \frac{1}{n}] \times \{(-1)^{n+1}\}$. (So $H_1$ joins the top of $V_1$ to the top of $V_2$, $H_2$ joins the bottom of $V_2$ to the bottom of $V_3$, etc.) We now let $S := \bigcup_{n\ge1}(V_n \cup H_n)$. $S$ is a "square wave" that packs an infinite amount of oscillation into a finite amount of space. It is not hard to show that $S$ is homeomorphic to $[0,\infty)$, so is path connected.*

*Define $\overline{S} := V_0 \cup S$. Then $\overline{S} = Cl(S)$, and is hence connected, by Theorems 12.3 (i) and 11.13. We are done once we show $\overline{S}$ is not path connected.*

*Suppose* $f : [a,b] \to \overline{S}$ *is a path joining* $\langle 0,0\rangle \in V_0$ *to* $\langle 1,-1\rangle \in S$. *Since* $V_0$ *is closed in* $\mathbb{R}^2$, $f^{-1}[V_0]$ *is closed in* $[a,b]$*; so this set must have a maximal element, say* $c$. *That is, we have* $a \le c < b$ *such that* $f(c) \in V_0$ *and* $f(t) \in S$ *for* $c < t \le b$. *For convenience, let's set* $c = 0$ *and* $b = 1$. *So then* $f(0) \in V_0$ *and* $f(t) = \langle x(t), y(t)\rangle \in S$ *for* $t \in (0,1]$, $f(1) = \langle 1,-1\rangle$. *Now the coordinate function* $x(t)$ *is a continuous map from* $[0,1]$ *to itself, taking on the values* $0$ *and* $1$. *Hence, by the intermediate value theorem, it must be a surjection. Thus we may pick a sequence* $\langle t_0, t_1, \dots\rangle \to 0$ *in* $(0,1]$ *such that* $y(t_n) = (-1)^n$. *But the continuity of* $f$ *implies that the sequence* $\langle f(t_0), f(t_1), \dots\rangle$ *converges, and that means the sequence* $\langle y(t_0), y(t_1), \dots\rangle$ *also converges (see Theorem 10.6). This is a contradiction, and we conclude that no such path* $f$ *can exist.*

**Exercises 12.6.** (1) *Suppose* $\{A, B\}$ *is a closed cover of* $X$, *with* $f : A \to Y$ *and* $g : B \to Y$ *continuous maps that agree on* $A \cap B$ *(i.e.,* $f(x) = g(x)$ *for all* $x \in A \cap B$*). Then there is a unique continuous* $h : X \to Y$ *such that* $h|A = f$ *and* $h|B = g$.

(2) *Prove that a space* $X$ *is path connected if and only if there is a point* $x_0 \in X$ *such that every point of* $X$ *may be joined to* $x_0$ *by a path in* $X$.

(3) *If* $X$ *is a space and* $x \in X$, *we define the* ***path component*** $P(x)$ *of* $x$ *to be the set of all points* $y \in X$ *such there is a path in* $X$ *that joins* $x$ *and* $y$. *Prove the following basic facts about path components:*
  (a) $X$ *is path connected if and only if* $P(x) = X$ *for avery* $x \in X$.
  (b) *For any* $x, y \in X$, *either* $P(x) = P(y)$ *or* $P(x) \cap P(y) = \emptyset$.
  (c) *Every component is a union of path components.*

(4) *Refer to Exercise (3) above and show that each path component of an open subset of euclidean space is also an open subset.*

(5) *Explain how Example 12.5 disproves the assertion that path components are closed sets (à la Theorem 11.14 (iv)).*

## 13. Compactness

We now come to the flavor referred to at the beginning of Section 9 as "being small and tidy." The most intuitively appealing mathematical explication of this is the one you see in a real analysis course: A subset of $\mathbb{R}^n$ is **small and tidy** just in case it is bounded with respect to the standard euclidean metric (i.e., of finite volume), and closed (i.e., with no "loose ends"). While this notion makes perfectly good sense mathematically, its dependence on boundedness for its definition makes it nontopological: switch from the euclidean metric to a topologically equivalent bounded metric, and *all* the closed sets now get to be small and tidy.

In this section we investigate three genuinely topological notions that all turn out to be equivalent in the metrizable realm; and indeed turn out to be equivalent to small and tidy when we're in euclidean space with the euclidean metric. These three notions are variations on what we term *compactness* in topology.

**Definition 13.1** (Three Notions of Compactness)**.** *Let $X$ be a topological space. $X$ is said to be:*

(i) ***compact*** *if every open cover of $X$ has a finite subcover (i.e., whenever $\mathcal{U}$ is an open cover of $X$, there is a finite subfamily $\mathcal{U}_0 \subseteq \mathcal{U}$ that is also a cover of $X$);*
(ii) ***countably compact*** *if every* countable *open cover of $X$ has a finite subcover; and*
(iii) ***limit point compact*** *if every infinite subset of $X$ has a limit point in $X$.*

*A subset $A$ of $X$ is a* ***(countably/limit point) compact subset*** *if it is (countably/limit point) compact as a subspace of $X$.*

Of the three notions, compactness is the strongest; immediately followed by countable compactness.

**Theorem 13.2.** *Every compact space is countably compact; every countably compact space is limit point compact.*

*Proof.* Clearly every compact space is countably compact; suppose $X$ is a space that is not limit point compact. Then there is an infinite subset $A \subseteq X$ such that $A$ has no limit point in $X$. We may as well assume $A$ is countably infinite, because no subset of $A$ can have a limit point if $A$ doesn't. Write $A = \{a_1, a_2, \dots\}$. Then, since no point of $A$ is a limit point of $A$, we know that there are open sets $U_n$, $n = 1, 2, \dots$, such that $U_n \cap A = \{a_n\}$. Also, since no point of $X \setminus A$ is a limit point of $A$, we know that $A$ must be closed. So let $\mathcal{U}$ be the family $\{X \setminus A\} \cup \{U_1, U_2, \dots\}$. Clearly $\mathcal{U}$ is a countable open cover with no finite subcover.

□

So to show a space is compact in one form or another, the best result is that you show it is compact in the unmodified sense. Conversely, if you want to show a space fails to have one of the three compactness properties, the best result is that it fails to be limit point compact.

**Examples 13.3.** (i) *If a discrete space is finite, then it is compact; if it is infinite, then it fails to be limit point compact.*

(ii) *Euclidean space fails to be limit point compact: the set of integer points on any one of the coordinate axes is infinite, discrete, and closed.*

(iii) *The usual closed unit interval* $[0, 1]$ *is limit point compact. To see this, suppose* $A$ *is an infinite subset if* $[0, 1]$*. Then either* $[0, \frac{1}{2}] \cap A$ *or* $[\frac{1}{2}, 1] \cap A$ *is infinite (maybe both are, but at least one has to be). Let* $I_1$ *be one of these two subintervals, chosen to have infinite intersection with* $A$*. We now repeat the process, choosing the left half of* $I_1$ *or the right half, depending upon its having infinite intersection with* $A$*. (Try to rephrase this more formally in terms of induction.) Then we have a nested sequence* $I_1 \supseteq I_2 \supseteq \dots$ *of closed subintervals of* $[0, 1]$*, each one of which having infinite intersection with* $A$*. Let* $I_n = [a_n, b_n]$*. Then, we have* $a_1 \leq a_2 \leq \dots$ *and* $b_1 \geq b_2 \geq \dots$ *In addition, we have* $b_n - a_n = \frac{1}{2^n}$*, for all* $n = 1, 2, \dots$ *Since each* $b_k$ *is an upper bound of* $\{a_1, a_2, \dots\}$*, the least upper bound* $a$ *of* $\{a_1, a_2, \dots\}$ *is a lower bound of* $\{b_1, b_2, \dots\}$*. Thus the greatest lower bound* $b$ *of* $\{b_1, b_2, \dots\}$ *is* $\geq a$*. Since* $[a, b] \subseteq [a_n, b_n]$ *for each* $n = 1, 2, \dots,$ *we see that* $b - a \leq \lim_{n\to\infty}(b_n - a_n) = 0$*; so* $a = b$*. The claim is that* $a$ *is a limit point of* $A$*. Indeed because the* $a_n$ *converge to* $a$ *from below and the* $b_n$ *converge to* $a$ *from above, every* $\epsilon$*-nbd of* $a$ *must contain some* $I_n$*, and hence must contain infinitely many elements of* $A$*.*

(iv) *The argument in (iii) above can be extended to the unit square* $[0, 1]^2$ *in* $\mathbb{R}^2$*. The only substantial difference is that now, instead of subdividing into two congruent subintervals, we subdivide into four congruent subsquares. Likewise, if we want to extend the argument to the unit cube* $[0, 1]^3$*, the subdivision involves eight congruent subcubes. In this way we can show that all finite-dimensional hypercubes (i.e., cartesian powers of* $[0, 1]$ *sitting in euclidean space) are limit point compact. (see Exercise 13.16 (9) below).*

Like connectedness and path connectedness, all three compactness properties are preserved under continuous images.

**Theorem 13.4** (Compactness and Continuity)**.** *Let* $f : X \to Y$ *be a continuous surjection, where* $X$ *is (countably/limit point) compact. Then* $Y$ *is also (countably/limit point) compact.*

*Proof.* We prove the statement for compactness; you are asked in Exercise 13.16 (2) to show the corresponding statements for countable compactness and limit point compactness.

Suppose $\mathcal{V}$ is an open cover of $Y$. Then $\mathcal{U} := \{f^{-1}[V] : V \in \mathcal{V}\}$ is an open cover of $X$ because $f$ is continuous. By compactness of $X$, there is a finite subcover; i.e., there are $V_1, \dots, V_n \in \mathcal{V}$ such that $\{f^{-1}[V_1], \dots f^{-1}[V_n]\}$ is a finite subcover of $\mathcal{U}$. Then we claim $\{V_1, \dots, V_n\}$ is a finite subcover of $\mathcal{V}$. To see this, suppose $y \in Y$ is given. Then $y = f(x)$ for some $x \in X$, since $f$ is a surjection. Since $\{f^{-1}[V_1], \dots, f^{-1}[V_n]\}$ is an open cover of $X$, there is some $i$, $1 \leq i \leq n$, such that $x \in f^{-1}[V_i]$. Thus $y = f(x) \in V_i$.

□

In Theorem 9.8 we showed that the separation axioms $\mathsf{T}_n$, $n = 0, 1, 2$, are hereditary; i.e., subsets of $\mathsf{T}$ spaces are $\mathsf{T}_n$. If we want to make a similar statement about our three compactness properties, we need to restrict the subsets we look at. Define a topological property to be **closed-hereditary** if whenever $X$ is a space with the property and $A$ is a closed subset of $X$, then $A$ has the property too, in its subspace topology.

**Theorem 13.5.** *Compactness, countable compactness, and limit point compactness are closed-hereditary properties.*

*Proof.* We prove the statement for compactness; you are asked in Exercise 13.16 (3) to show the corresponding statements for countable compactness and limit point compactness.

Suppose $X$ is compact and that $A \subseteq X$ is closed. Let $\mathcal{U}$ be an open cover of $A$; i.e., $\mathcal{U}$ is a family of open subsets of $X$ whose union contains $A$. We need to find a finite subfamily of $\mathcal{U}$ that covers $A$; but before we can use our compactness hypothesis, we need to get an open cover of $X$. Since $A$ is closed, though, we can get an open cover of $X$ by adding the set $X \setminus A$ to $\mathcal{U}$. That is, $\mathcal{V} := \{X \setminus A\} \cup \mathcal{U}$ is an open cover of $X$. Since $X$ is compact, there are $U_1, \dots, U_n \in \mathcal{U}$ such that $\{X \setminus A, U_1, \dots, U_n\}$ is a finite subcover of $\mathcal{V}$. Since $X \setminus A$ misses $A$ altogether, it must be the case that $\{U_1, \dots, U_n\}$ covers $A$. This shows $A$ is compact.

□

Compactness and the Hausdorff separation axiom work very well in combination, each balancing the other.

**Theorem 13.6.** *The compact subsets of a Hausdorff space are closed.*

*Proof.* Suppose $A$ is a compact subset of the Hausdorff space $X$. Given $x \in X \setminus A$, we need to find an open nbd $V$ of $x$ such that $V \cap A = \emptyset$. For each $a \in A$, find open nbds $U_a$ of $a$ and $V_a$ of $x$ such that $U_a \cap V_a = \emptyset$ (because $X$ is Hausdorff). Next, because $A$ is compact, there are finitely many sets $U_{a_1}, \dots, U_{a_n}$ that cover $A$. Let $V := \bigcap_{i=1}^n V_i$. Then $V$ is an open nbd of $x$ that misses $\bigcup_{i=1}^n U_i \supseteq A$. Thus $V \cap A = \emptyset$, as desired; and $A$ is therefore closed in $X$.

□

**Theorem 13.7.** *The compact subsets of euclidean space are closed and bounded (in the usual metric).*

*Proof.* In view of Theorem 13.6, all we need show is boundedness. Suppose $A \subseteq R^n$ is unbounded. Then, for $n = 1, 2, \dots$, $A \setminus B(\mathbf{0}, n) \neq \emptyset$. Thus $\{B(\mathbf{0}, n) : n = 1, 2 \dots\}$ is an open cover of $A$ such that no finite subfamily covers $A$. This open cover witnesses the noncompactness of $A$.

□

**Theorem 13.8.** *Let $f : X \to Y$ be a continuous surjection, where $X$ is compact and $Y$ is Hausdorff. Then $f$ is a closed map. If $f$ is also an injection, then it is a* *homeomorphism.*

*Proof.* First we check is that $f$ is a closed map. Once that is done, we may cite Theorem 8.7 to do the rest. Indeed, suppose $A \subseteq X$ is closed. By Theorem 13.5, $A$ is compact; and by Theorem 13.4, $f[A]$ is a compact subset of the Hausdorff space $Y$. $f[A]$ is therefore closed in $Y$, by Theorem 13.6. □

Our next goal in this course is to provide a proof for the second "pillar" mentioned in Theorem 1.5, namely the extreme value theorem. Before we can do that, however, we need to be able to strengthen Example 13.3 (iii) to say that bounded closed intervals in the real line are compact (not merely limit point compact).

**Theorem 13.9.** *For $\mathsf{T}_1$ spaces, limit point compactness and countable compactness are equivalent.*

*Proof.* By Theorem 13.2, all we need show is that limit point compact $\mathsf{T}_1$ spaces are countably compact. We prove the contrapositive. Suppose $X$ has a countable open cover $\{U_1, U_2, \dots\}$ such that no finite subfamily covers $X$. For each $n = 1, 2, \dots,$ let $V_n := \bigcup_{i=1}^{n} U_i$. We then have a new open cover $\{V_1, V_2, \dots\}$ of $X$ such that $V_1 \subseteq V_2 \subseteq \dots$ and such that each $V_n$ is a proper subset of $X$. Because of this, there is a subsequence $n_1 < n_2 < \dots$ such that $V_{n_i}$ is a proper subset of $V_{n_{i+1}}$ for each $i = 1, 2 \dots$; hence we may assume, without loss of generality, that each inclusion $V_i \subseteq V_{i+1}$ is proper.

We now form $A := \{a_1, a_2, \dots\}$, where $a_1 \in V_1$ and $a_{n+1} \in V_{n+1} \backslash V_n$, $n = 1, 2, \dots$ $A$ is clearly an infinite set; we show it can have no limit point in $X$. Indeed, fix $x \in X$, and let $m \geq 1$ be chosen such that $x \in V_m$. Then $V_m$ is a nbd of $x$ that misses all of $A$, except possibly for $\{x_1, \dots, x_m\}$. Using the $\mathsf{T}_1$ axiom, there is a nbd $W$ of $x$ that misses the finite closed set $\{x_1, \dots, x_m\} \setminus \{x\}$. Then $W \cap V_m$ is a nbd of $x$ whose intersection with $A$ includes no point of $A$, except possibly for $x$ itself. This tells us $A$ has no limit point in $X$; hence $X$ fails to be limit point compact. □

Next we need to identify an arena where countable compactness implies compactness. A relatively hard theorem, beyond the scope if this course, is that the two compactness notions are equivalent for metrizable spaces. The good news is that we don't actually need that result to accomplish the goals of this section; the bad news is that we need to make a small digression.

**Definition 13.10** (Second Countability)**.** *A topological space $X$ is* ***second countable*** *if the topology of $X$ has a base consisting of countably many sets.*

**Examples 13.11.** (i) *If $X$ is a countably infinite set with the cofinite topology, then (by a standard counting argument, similar to how you show the set of rational numbers is countable) $X$ has only countably many open sets. This makes $X$ second countable.*

(ii) *By Exercise 7.18 (2), the set of bounded open intervals with rational end points in the real line forms a countable base for the usual topology on* $\mathbb{R}$*. More generally (see Exercise 13.16 (9) below), each euclidean space* $\mathbb{R}^n$ *is second countable.*
(iii) *Second countable spaces are first countable: form a nbd base at a point by considering only those members of the countable base that contain the point.*
(iv) *Any uncountable space with the discrete topology is metrizable (hence first countable), but not second countable (see Exercise 13.16 (10)).*

The following is a technical result that greatly simplifies many compactness arguments.

**Theorem 13.12.** *Let* $X$ *be a topological space, with base* $\mathcal{B}$ *for the open sets of* $X$*. If every open cover of* $X$ *by members of* $\mathcal{B}$ *has a finite subcover, then* $X$ *is compact.*

*Proof.* Start with an open cover $\mathcal{U}$ of $X$. For each $x \in X$, pick $U_x \in \mathcal{U}$ such that $x \in U_x$. Then, because $\mathcal{B}$ is a base for the open sets, pick $B_x \in \mathcal{B}$ such that $x \in B_x \subseteq U_x$. Then, by our hypothesis, there are finitely many points $x_n, \dots, x_n \in X$ such that $X = \bigcup_{i=1}^n B_{x_i}$. But then $X = \bigcup_{i=1}^n U_{x_i}$ too; hence $\mathcal{U}$ has a finite subcover.

□

As an immediate consequence of Theorems 13.2, 13.9, and 13.12, we may now conclude the following.

**Corollary 13.13.** *For second countable spaces, countable compactness and compactness are equivalent. Hence, for second countable* $\mathsf{T}_1$ *spaces, limit point compactness, countable compactness, and compactness are all equivalent.*

And now the advertised "second pillar:"

**Corollary 13.14** (Extreme Value Theorem)**.** *Suppose* $f : [a, b] \to \mathbb{R}$ *is a continuous function. Then there exist* $c, d \in [a, b]$ *such that, for all* $x \in [a, b]$*,* $f(c) \leq f(x) \leq f(d)$*.*

*Proof.* By Example 13.3 (iii), $[a, b]$ is limit point compact; by Corollary 13.13, it is compact. Thus, by Theorem 13.4, $f[[a, b]]$ is a compact subset of $\mathbb{R}$. By Theorem 13.7, $f[[a, b]]$ is both closed and bounded. Thus there are $m, M \in f[[a, b]]$ such that, for all $x \in [a, b]$, $m \leq f(x) \leq M$. Let $c, d \in [a, b]$ be such that $f(c) = m$ and $f(d) = M$.

□

The next result, often presented in undergraduate analysis courses, justifies the metric-dependent definition of "small and tidy" given at the beginning of this section. It is due to Eduard Heine (1821–1881) and Emile Borel (1871–1956).

**Theorem 13.15** (Heine-Borel)**.** *A subset of euclidean space is compact if and only if it is closed and bounded (in the usual metric)*

*Proof.* In view of Theorem 13.7, we need only show that the closed bounded subsets of $\mathbb{R}^n$ are compact. Indeed, if $A$ is bounded, then there is a cube $[a,b]^n$ that contains it. By Example 13.3 (iv), the cube is limit point compact; by Example 13.11 (ii) and Corollary 13.13, the cube is compact as well. (Second countability is a hereditary property, see Exercise 13.16 (11) below.) Finally, if $A$ is also closed, we conclude that $A$ is compact, by Theorem 13.5. $\square$

**Exercises 13.16.** (1) *Define a space $X$ to be **sequentially compact** if every sequence in $X$ has a convergent subsequence. Show that a metrizable space is sequentially compact if and only if it is limit point compact. (What happens to this statement if you replace* metrizable *with* first countable and $\mathsf{T}_1$*? Can you prove either the "if" part or the "only if" part after removing first countability or the $\mathsf{T}_1$ condition?)*

(2) *Show that continuous surjections preserve the properties of countable and limit point compactness.*

(3) *Show that countable compactness and limit point compactness are closed-hereditary properties.*

(4) *Let $\mathcal{T}$ and $\mathcal{U}$ be two topologies on the point set $X$, with $\mathcal{T} \subseteq \mathcal{U}$. Show that if $\mathcal{U}$ is a compact topology, then so is $\mathcal{T}$.*

(5) *Let $\langle X, \mathcal{T} \rangle$ be a compact Hausdorff space. Prove the following:*
  (a) *If $\mathcal{U}$ is a topology on $X$ that is strictly coarser than $\mathcal{T}$, then $\mathcal{U}$ is compact but not Hausdorff.*
  (b) *If $\mathcal{U}$ is a topology on $X$ that is strictly finer than $\mathcal{T}$, then $\mathcal{U}$ is Hausdorff but not compact.*

(6) *A collection $\mathcal{F}$ of subsets of a set $X$ has the **finite intersection property** if $\bigcap \mathcal{F}_0 \neq \emptyset$ for every finite $\mathcal{F}_0 \subseteq \mathcal{F}$. Show that a topological space is compact if and only if every collection of closed subsets with the finite intersection property has nonempty intersection.*

(7) *Show that the conclusion of Theorem 13.6 fails for compact $\mathsf{T}_1$ spaces. [Hint: try an infinite set with the cofinite topology.]*

(8) *Using the argument of Example 13.3 (iii), fill in the details of the statements made in Example 13.3 (iv).*

(9) *Show that euclidean space is second countable. [Hint: Pick ball nbds of the form $B(\mathbf{x}, \frac{1}{n})$, where each coordinate of $\mathbf{x}$ is rational and $n$ is a positive natural number.]*

(10) *Suppose $X$ is an uncountable discrete topological space. Show $X$ is first countable, but not second countable.*

(11) *Prove Cantor's intersection theorem: If $C_1 \supseteq C_2 \supseteq \dots$ is a nested sequence of nonempty closed bounded subsets of $\mathbb{R}$, then $\bigcap_{n\geq 1} C_n \neq \emptyset$.*

(12) *Prove that second countability is a hereditary property.*

(13) *Using an argument centering around compactness, show that $S^1$ and $[0, 2\pi)$ are nonhomeomorphic.*

## 14. Product Spaces

In Section 3 we gave the official definition of cartesian product of two sets, in order to put the definition of metric on a firm footing. It is nearly impossible to overemphasize the importance of this simple construction in mathematics; in this section we use the cartesian product to create new topological spaces from old ones.

**Definition 14.1** (The Product Topology)**.** *Let $X = \langle X, \mathcal{T}\rangle$ and $Y = \langle Y, \mathcal{U}\rangle$ be two topological spaces. We define the **product topology** on the cartesian product $X \times Y$ of the underlying point sets by taking sets of the form $U \times V$, $U \in \mathcal{T}$, $V \in \mathcal{U}$, as a subbase. When we write $X \times Y$, where $X$ and $Y$ are topological spaces, it is the product topology that we are intentionally placing on the cartesian product, unless we explicitly state otherwise.*

**Remark 14.2.** *The definition above extends, by an easy induction, to arbitrary finite products $X_1 \times \cdots \times X_n$; however it does not tell us how to define the product topology on infinite products. The topic of infinite products is quite extensive and deep, and is covered in more advanced courses.*

The first result is quite simple, but extremely helpful

**Theorem 14.3.** *Let $X$ and $Y$ be spaces, with $\mathcal{B}$ and $\mathcal{C}$ bases for the topologies of $X$ and $Y$, respectively. Then the collection $\{B \times C : \langle B, C\rangle \in \mathcal{B} \times \mathcal{C}\}$ is a base for the product topology.*

*Proof.* We first note that if $U_1 \times V_1$ and $U_2 \times V_2$ are typical subbasic open sets in the product topology, then their intersection is $(U_1 \cap U_2) \times (V_1 \cap V_2)$. Since topologies are closed under finite intersections, we infer that our collection of subbasic open sets is, in fact, a base for the product topology. So suppose $W$ is an open set in the product topology, with $\langle x, y\rangle \in W$. Then we can find $X$-open set $U$ and $Y$-open set $V$ with $\langle x, y\rangle \in U \times V \subseteq W$. Because $\mathcal{B}$ (resp., $\mathcal{C}$) is a base for the topology on $X$ (resp., $Y$), there is some $B \in \mathcal{B}$ and some $C \in \mathcal{C}$ such that $x \in B \subseteq U$ and $y \in C \subseteq V$. Thus $\langle x, y\rangle \in B \times C \subseteq U \times V \subseteq W$.

$\square$

Given spaces $X$ and $Y$, there are canonical maps $p : X \times Y \to X$ and $q : X \times Y \to Y$, defined by $p(x, y) := x$ and $q(x, y) := y$. These are what are called the **projection maps** (first encountered in Example 8.9 (iii), in connection with the euclidean plane), and are continuous and open (see Exercise 14.10 (1) below). Thus each factor (i.e., $X$ and $Y$) is a continuous image of the product. Moreover, each factor embeds in the product. Indeed, given arbitrary $y_0 \in Y$, the subset $X \times \{y_0\}$ of $X \times Y$ is a homeomorphic copy of $X$ (embedded as a closed subset when $Y$ is a $\mathsf{T}_1$ space, see Exercise 14.10 (2) below). Likewise, for each $x_0 \in X$, $\{x_0\} \times Y$ is a homeomorphic copy of $Y$ (closed in $X \times Y$ when $X$ is a $\mathsf{T}_1$ space). These observations, plus what we know about the various topological properties we have studied so far, allow us to make the following assertion.

**Theorem 14.4** (Big Product Preservation Result)**.** *Let $\mathcal{P}$ be any of the properties: $\mathsf{T}_n$ ($n = 0, 1, 2$), connectedness, path connectedness, compactness. Then, for any spaces $X$ and $Y$: $X \times Y$ has property $\mathcal{P}$ if and only if both $X$ and $Y$ have property $\mathcal{P}$.*

*Proof.* Since both $X$ and $Y$ embed in $X \times Y$, they share any hereditary property enjoyed by the product. This includes the separation properties $\mathsf{T}_n$, for $n = 0, 1, 2$. If $X \times Y$ is connected (path connected, compact, countably compact, limit point compact), then so are $X$ and $Y$ because they are continuous images of the product (see Theorems 11.5, 12.3 (ii), and 13.4).

Now assume both $X$ and $Y$ have property $\mathcal{P}$; we show $X \times Y$ does as well.

*Ad* $\mathsf{T}_0$: Suppose $\langle x, y\rangle$ and $\langle u, v\rangle$ are two distinct points in $X \times Y$. Then either $x \neq u$ or $y \neq v$; say the former is true. If there is an $X$-open nbd $U$ of $x$ that misses $u$, then $U \times Y$ is an $X \times Y$-open nbd of $\langle x, y\rangle$ that misses $\langle u, v\rangle$. The other cases are treated similarly, and we infer that $X \times Y$ is a $\mathsf{T}_0$ space. (Note that we need for $Y$ to be $\mathsf{T}_0$ if it so happens that $x = u$ and $y \neq v$.)

*Ad* $\mathsf{T}_n$*,* $n = 1, 2$: See Exercise 14.10 (3) below.

*Ad connectedness*: Assume both $X$ and $Y$ are connected. We show the product to be connected by showing every pair of points to be contained in a connected subset, and then citing Theorem 11.7.

Pick two points $\langle x, y\rangle$ and $\langle u, v\rangle$ in $X \times Y$. Let $A := \{x\} \times Y$ and $B := X \times \{v\}$. Then $A$, being homeomorphic to $Y$, is connected; $B$ is connected for the obvious corresponding reason. $A$ contains $\langle x, y\rangle$ and $B$ contains $\langle u, v\rangle$, so $A \cup B$ contains them both. Since $A \cap B = \{\langle x, v\rangle\} \neq \emptyset$, we know $A \cup B$ is connected (Theorem 11.9). This proves our assertion.

*Ad path connectedness*: See Exercise 14.10 (4) below.

*Ad compactness*: Suppose both $X$ and $Y$ are compact, and let $\mathcal{U}$ be an open cover of $X \times Y$. By Theorem 13.12, we may take $\mathcal{U}$ to consist of sets of the form $U \times V$, where $U$ and $V$ are open in $X$ and $Y$, respectively. For each $x \in X$, let $\mathcal{U}_x := \{U \times V \in \mathcal{U} : x \in U\}$. Then each $\mathcal{U}_x$ is an open cover of $\{x\} \times Y$, which is compact because it is homeomorphic to $Y$ (by Exercise 14.10 (2) below). So, for each $x \in X$ let $\mathcal{U}_x^0$ be a finite subfamily of $\mathcal{U}_x$ that covers $\{x\} \times Y$. Now, for each $x \in X$, let $U_x$ be the finite intersection of all sets $U$ such that $U \times V \in \mathcal{U}_x^0$, for some $V$. Then $\mathcal{U}_x^0$ is an open cover of $U_x \times Y$. Now we use the compactness of $X$ to find finitely many points $x_1, \dots, x_n$ such that the sets $U_{x_1}, \dots, U_{x_n}$ cover $X$. Then $\mathcal{U}_{x_1} \cup \dots \cup \mathcal{U}_{x_n}$ is a finite subcover of $\mathcal{U}$.

□

**Remark 14.5.** *One of the reasons that the every-open-cover-has-a-finite-subcover version of compactness won prominence is that it is preserved by finite (and even infinite) products. The same cannot be said for countable compactness or limit point compactness, however. In a 1953 paper, Jiři Novák ["On the Cartesian product of two compact spaces,"* Fundamenta Mathematicae*, vol. 40, 106–112] gives a sophisticated construction of two countably compact Hausdorff spaces $X$ and $Y$ such that $X \times Y$ fails to be countably compact. Since countable compactness and limit point compactness are equivalent in the Hausdorff context (see Theorem 13.9), this* *example serves to show that limit point compactness also fails to be preserved by finite products. In Exercise 14.10 (5) below, you are asked to show that sequential compactness is indeed so preserved.*

In many areas of mathematics, the cartesian product of sets serves as the natural carrier of extra structure. We have talked about topological structure in this section, and have alluded in previous sections to ways in which one may induce metric and order structures on a product. In algebra, group theory, for example, one easily constructs a group operation on the cartesian product, given such operations on the factors. Even in measure theory, one has a natural definition of product measure, enabling a more abstract version of the theorem of Fubini that students see in multivariable calculus (i.e., double integrals as iterated single integrals). We conclude this section with a brief investigation of how metric and order structure, two ways to get topologies, are consistent with topological structure when we take products. Here is a more precise statement of the problem.

**Definition 14.6** (Product Metrics)**.** *Suppose $\langle X, d\rangle$ and $\langle Y, e\rangle$ are two metric spaces. Then we may define a distance function $d \times e$ on $X \times Y$ by setting $(d \times e)(\langle x, y\rangle, \langle u, v\rangle) := \sqrt{d(x, u)^2 + e(y, v)^2}$.*

**Theorem 14.7.** *Suppose $\langle X, d\rangle$ and $\langle Y, e\rangle$ are two metric spaces. Then the product metric $d \times e$ on $X \times Y$ gives rise to the product topology resulting from the metric topologies $\mathcal{T}_d$ and $\mathcal{T}_e$.*

*Proof.* This is quite straightforward, and is left as an exercise. (See Exercise 14.10 (7) below.)

□

**Definition 14.8** (Product Orderings)**.** *Suppose $\langle X, R\rangle$ and $\langle Y, S\rangle$ are two linear orderings. Then we may define a binary relation $R \times S$ on $X \times Y$ by setting $\langle x, y\rangle R \times S\langle u, v\rangle$ just in case either $xRu$, or $x = u$ and $ySv$. This is called the **lexicographic ordering** on the the product.*

**Theorem 14.9.** *Suppose $\langle X, R\rangle$ and $\langle Y, S\rangle$ are two linear orderings. Then the product ordering $R \times S$ gives rise to a topology that is generally strictly finer than the product topology resulting from the order topologies $\mathcal{T}_R$ and $\mathcal{T}_S$.*

*Proof.* □

For simplicity, let's assume there are no end points. (The reader may easily provide the extra cases to cover their existence.) Then a base for the product topology consists of open rectangles $(a, b) \times (c, d)$, where $(a, b)$ is an open $R$-interval and $(c, d)$ is an open $S$-interval. Let $\langle x, y\rangle \in (a, b) \times (c, d)$. Then $\{x\} \times (c, d)$ is an open interval in the lexicographic ordering $R \times S$, which contains the point $\langle x, y\rangle$, and which is itself contained in the rectangle $(a, b) \times (c, d)$. This shows the lexicographic ordering gives rise to a topology finer than the product topology. It is strictly finer as long as there is a nonisolated point in $X$ (see Exercise 14.10 (8) below).

**Exercises 14.10.** (1) *Show that the canonical projection maps from $X \times Y$ to $X$ and to $Y$ are continuous and open. (Not necessarily closed, though; see Example 8.9 (iii).)*

(2) *Consider the product space $X \times Y$. For each $x_0 \in X$, show that $\{x_0\} \times Y$ is a homeomorphic copy of $Y$ embedded—as a closed subset, in case $X$ is a $\mathsf{T}_1$ space—in $X \times Y$.*

(3) *Prove that $X \times Y$ is a $\mathsf{T}_n$ space if both $X$ and $Y$ are $\mathsf{T}_n$ spaces, $n = 1, 2$.*

(4) *Show that the product of two path connected spaces is path connected.*

(5) *Refer to Exercise 13.16 (1) for the definition of* sequentially compact*, and show that the product of two sequentially compact spaces (or any finite number, for that matter) is sequentially compact.*

(6) *Let $\mathbb{L}$ be the real line with the Sorgenfrey (lower limit) topology (see Exercise 7.18 (4)). Show that the subspace $\Delta := \{\langle x, x\rangle : x \in \mathbb{L}\}$ of $\mathbb{L}^2$ is homeomorphic to $\mathbb{L}$. What about the subspace $\Delta' := \{\langle x, -x\rangle : x \in \mathbb{L}\}$?*

(7) *Prove Theorem 14.7. (See Exercise 7.18 (13).)*

(8) *Complete the proof of Theorem 14.9.*

(9) **Let $X$ and $Y$ be topological spaces, with $Y$ compact. Show that the canonical projection onto $X$ is a closed map.*

## 15. Quotient Spaces

Besides the topological product construction, the other principal way of making new spaces from old is to form quotients. Readers who have seen some abstract algebra will recall how one produces quotient groups by "modding out" a normal subgroup: two elements are deemed equivalent if the product of one with the inverse of the other is in the normal subgroup; this partitions the original group into equivalence classes (called cosets), which become the elements of the quotient group. There is a similar story in the theory of rings (instead of normal subgroups we have ideals), as well as other algebraic systems. In each case there is a special way to make two points "equivalent," and then to take equivalence classes to be new points. The trick is to define the algebraic operations on these sets of equivalence classes.

The same thing happens in the topological context, only here we can pick any equivalence relation we like on the original space. The tricky part (only not very tricky) is to define a suitable topology on the set of equivalence classes. The following preliminary definition should be familiar to anyone who has seen some abstract algebra; it pervades much of mathematics.

**Definition 15.1** (Equivalence Relations)**.** *Let $X$ be a set. A binary relation $R \subseteq X \times X$ is called an* ***equivalence relation*** *if it satisfies the following three conditions.*

(E1) (Reflexivity) *$xRx$ always holds.*
(E2) (Symmetry) *$yRx$ holds whenever $xRy$ holds.*
(E3) (Transitivity) *If $x, y, z \in X$, $xRy$, and $yRz$, then $xRz$.*

*Given $x \in X$, the* ***equivalence class*** *of $x$, denoted $[x]_R$ (or, more simply, $[x]$ if there's no likely confusion) is the set $\{y \in X : xRy\}$. Because of the conditions $E_1 - -E_3$, any two distinct equivalence classes must be disjoint. Hence the $R$-equivalence classes $[x]$, for $x \in X$, form a* ***partition*** *of the set $X$; i.e., a cover of $X$ by pairwise disjoint subsets. $x$ is called a* ***representative*** *of the equivalence class; two elements $x$ and $y$ are both representatives of the same equivalence class just in case $xRy$. The set of $R$-equivalence classes is denoted $X/R$, the $R$-****quotient*** *set.*

**Examples 15.2.** (i) *Consider the underlying set to be the set $\mathbb{Z}$ of integers, and define $mRn$ just in case $m - n$ is a multiple of 3. (In number-theoretic terms $m$ is congruent to $n$ modulo 3.) There are three equivalence classes, depending on whether the remainder is 0,1, or 2 after an integer is divided by 3. That is, the equivalence classes are $\{\dots, -6, -3, 0, 3, 6, \dots\}$, $\{\dots, -5, -2, 1, 4, 7, \dots\}$, and $\{\dots, -4, -1, 2, 5, \dots\}$. So when each equivalence class is collapsed to a point, the quotient set $\mathbb{Z}/R = \{[0], [1], [2]\}$ has three elements in it. It is the "mod 3 integers," better known as $\mathbb{Z}/3\mathbb{Z}$, or simply $\mathbb{Z}_3$.*

(ii) *(For those with the group-theoretic background.) Given a group $G$ and a normal subgroup $N$, we declare $xR_N y$ just in case $xy^{-1} \in N$. The equivalence class represented by $x \in G$ is the coset $xN$ of multiples $xn$, for $n \in N$. The quotient set $G/R_N$, also known as $G/N$, can be made into a group by defining the product $(xN)(yN)$ to be the coset $(xy)N$. The fact that this all* *works is covered in any first course in abstract algebra.*

(iii) *Consider $X$ to be the closed unit interval $[0,1]$, and define $xRy$ just in case either $x = y$ or $|x - y| = 1$. Thus the equivalence classes are just the singleton sets $\{x\}$, for $0 < x < 1$, as well as the doubleton set $\{0,1\}$. (So the end points have been identified to a single point, forming a "circle.")*

(iv) *Consider $X$ to be the closed unit square $[0,1]^2$, and define $\langle x,y\rangle R\langle u,v\rangle$ just in case either $\langle x,y\rangle = \langle u,v\rangle$ or $x = u$ and $!y - v! = 1$. Thus the equivalence classes are the singleton sets $\{\langle x,y\rangle\}$, for $0 \leq x \leq 1$ and $0 < y < 1$, as well as the doubleton sets $\{\langle x,0\rangle,\langle x,1\rangle\}$, for $0 \leq x \leq 1$. (So the top and bottom edges have been "glued together" to form a "tube.")*

(v) *Let $f : X \to Y$ be a function between sets, and define the* ***kernel*** *$ker(f)$ of $f$ to be the set $\{\langle x,y\rangle \in X^2 : f(x) = f(y)\}$. Then $ker(f)$ is easily seen to be an equivalence reltion on $X$.*

**Definition 15.3** (The Quotient Topology)**.** *Let $X = \langle X,\mathcal{T}\rangle$ be a topological space, with $R$ an equivalence relation on $X$. We define the* ***quotient topology*** *on the set $X/R$ of $R$-equivalence classes by declaring the set $\mathcal{V} \subseteq X/R$ open just in case the union of all the equivalence classes in the family $\mathcal{V}$ is open in $X$. We denote the quotient topology $\mathcal{T}/R$; so, in symbols, we have $\mathcal{T}/R := \{\mathcal{V} \subseteq X/R : \bigcup\mathcal{V} \in \mathcal{T}\}$.*

Given any set $X$ and equivalence relation $R$ on $X$, there is the natural quotient map $q_R : X \to X/R$, given by the assignment $q_R(x) := [x]_R$. Using this function, the definition of the quotient topology may be rephrased by declaring $\mathcal{V}$ to be $\mathcal{T}/R$-open just in case $q_R^{-1}[\mathcal{V}]$ is $\mathcal{T}$-open.

**Proposition 15.4.** *Let $X$ be a topological space, $R$ an equivalence relation on $X$. Then the quotient topology on $X/R$ is the smallest topology on the quotient set such that the quotient map $q_R$ is continuous.*

*Proof.* See Exercise 15.11 (1) below.

□

**Definition 15.5.** *An* ***identification map*** *is a surjection $f : X \to Y$ such that, for any subset $V$ of $Y$, $V$ is open in $Y$ if and only if $f^{-1}[V]$ is open in $X$.*

So clearly, if $q_R : X \to X/R$ is the natural quotient map, then $q_R$ is an identification map. The following two results constitute the topological analogue of what is generally known as the "fundamental homomorphism theorem" in abstract algebra.

**Theorem 15.6** (Transgression)**.** *Let $f : X \to Y$ be a continuous map, and suppose $R$ is any equivalence relation on $X$ such that $R \subseteq ker(f)$. Then there is a unique continuous $g : X/R \to Y$ such that $g \circ q_R = f$.*

*Proof.* Since $g \circ q_R$ needs to agree with $f$, there is no choice but to define $g([x])$ to be $f(x)$. That $g$ is well defined follows from the assumption that if $xRy$ in $X$, then $f(x) = f(y)$. Where well definedness can go wrong is when we define a function on an equivalence class, but that function depends on the choice of representative for that equivalence class. But in this case, if $[x] = [y]$, then $xRy$. Since $f(x) = f(y)$ then, it doesn't matter which representative we choose.

To show $g$ is continuous, suppose $V$ is open in $Y$. Then $g^{-1}[V]$ is open in $X/R$ just in case $q_R^{-1}[g^{-1}[V]]$ is open in $X$. But, because $g \circ q_R = f$, this set is just $f^{-1}[V]$, open in $X$ because $f$ is continuous.

□

**Corollary 15.7.** *Let $f : X \to Y$ be an identification map, and suppose $R = ker(f)$. Then the map $g$ defined in Theorem 15.6 is a homeomorphism between $X/R$ and $Y$.*

*Proof.* First, if $y \in Y$, then $y = f(x)$ for some $x \in X$; hence $y = g([x])$. Thus $g$ is a continuous surjection. Now suppose $g([x]) = g([y])$. Then $f(x) = f(y)$; and, since $R$ is the kernel of $f$, we know $xRy$. Thus $[x] = [y]$; this tells us $g$ is a continuous bijection. (This is the best we can hope for if $f$ is assumed to be merely a continuous surjection, not necessarily an identification map. See Exercise 15.11 (2) below.)

Finally, suppose $V \subseteq Y$ is not open in $Y$. Then, because $f$ is an identification map, $f^{-1}[V]$ is not open in $X$. But $f^{-1}[V] = q_R^{-1}[g^{-1}[V]]$. And, since $q_R$ is also an identification map, $g^{-1}[V]$ can't be open in $X/R$. This is the contrapositive to the statement that $g$ is an open continuous bijection; hence a homeomorphism.

□

Next we relate identification maps to other kinds of map.

**Theorem 15.8.** *Let $f : X \to Y$ be a continuous surjection. If $f$ is either an open map or a closed map, then $f$ is an identification map.*

*Proof.* Suppose first that $f$ is an open map. If $V$ is an open subset of $Y$, then $f^{-1}[V]$ is open in $X$ because $f$ is continuous. If $f^{-1}[V]$ is open in $X$, then $V = f[f^{-1}[V]]$ is open in $Y$ because $f$ is an open map.

Now suppose $f$ is a closed map. Then the sentence immediately preceding may be phrased: If $f^{-1}[V]$ is open in $X$, then $X \setminus f^{-1}[V] = f^{-1}[Y \setminus V]$ is closed in $X$. Thus $Y \setminus V = f[f^{-1}[Y \setminus V]]$ is closed in $Y$ because $f$ is a closed map. Therefore $V$ is open in $Y$.

□

**Corollary 15.9.** *Let $f : X \to Y$ be a continuous surjection, where $X$ is compact and $Y$ is Hausdorff. Then $f$ is an identification map.*

*Proof.* This is immediate from Theorems 13.8 and 15.8.

□

We end this brief introduction to identification maps and quotients with the following simple application of the quotient method.

**Example 15.10** (Example 15.2 (iii) Revisited)**.** *Consider $X$ to be the closed unit interval $[0,1]$, and define $xRy$ just in case either $x = y$ or $|x - y| = 1$. Thus the equivalence classes are just the singleton sets $\{x\}$, for $0 < x < 1$, as well as the doubleton set $\{0,1\}$. We give $X/R$ the quotient topology, and show that it is homeomorphic to the standard unit circle $S^1 := \{\langle x, y\rangle \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. (See Example 8.3 (v).) The function $f : [0,1] \to S^1$, defined by $f(t) := \langle \cos(2\pi t), \sin(2\pi t)\rangle$ is a continuous surjection whose kernel is equal to $R$. $f$ is an identification map by Corollary 15.3; hence the derived map $g : X/R \to S^1$ is a homeomorphism by Corollary 15.7.*

**Exercises 15.11.**

(1) *Prove Proposition 15.4.*

(2) *Show how Corollary 15.7 can go wrong if we assume that $f$ is just a continuous surjection. [Hint: think discrete topology for $X$.]*

(3) *Let $X$ be a space, with $A$ a subset of $X$. We define the equivalence relation $R_A$ by stipulating that $[x] = A$ for $x \in A$, and $[x] = \{x\}$ otherwise. We denote $X/R_A$ simply as $X/A$. Show that if $X$ is a $\mathsf{T}_1$ space, then $X/A$ is a $\mathsf{T}_1$ space if and only if $A$ is closed in $X$. (What would be the analogous statement if you wanted to replace $\mathsf{T}_1$ with $\mathsf{T}_2$?)*

(4) *(See notation in Exercise 15.11 (3) above.) Show that $\mathbb{R}/[0,1]$ is homeomorphic to $\mathbb{R}$.*

(5) *(See notation in Exercise 15.11 (3) above.) Given a topological space $X$, define the* ***cone*** *$\mathsf{C}X$ over $X$ to be the quotient space $(X \times [0,1])/(X \times \{1\}$. (So the top edge of the "rectangle" $X \times [0,1]$ is identified to a point.) What familiar spaces are formed by taking the cone over: (i) the two-point discrete space; (ii) the closed unit interval; and (iii) the unit circle?*

(6) *Start with the closed unit square $[0,1]^2$, and define the equivalence relation $R$ by taking equivalence classes to be the singletons $\{\langle x, y\rangle\}$ for $0 \le x \le 1$, $0 < y < 1$, as well as the doubletons $\{\langle x, 0\rangle, \langle 1-x, 1\rangle\}$, $0 \le x \le 1$. How does $[0,1]^2/R$ differ topologically from the "tube" in Example 15.2 (iv)? (You might enjoy making paper models of the two spaces and experimenting.)*

(7) *The cone over a three-point discrete space is called a* ***triod****. How does the triod differ topologically from the cone over a two-point discrete space?*

## 16. The Basic Upper-Level Separation Axioms

In Section 9, we introduced the three "lower" separation axioms $\mathsf{T}_n$, $n = 0, 1, 2$, based on separating points from one another using open sets. In this section we continue in the vein of separation by open sets, adding the two "upper" axioms $\mathsf{T}_3$ and $\mathsf{T}_4$. The area of set-theoretic topology connected with these axioms has a rich history, with many deep results. Due to the introductory nature of this course, however, we can offor only a small taste of the subject.

In the first new axiom, we separate points and closed sets.

**Definition 16.1** (The $\mathsf{T}_3$, or Vietoris, Axiom). *A topological space $X$ is called a $\mathsf{T}_3$ **space** (or a **regular space**, or sometimes a **Vietoris space**, after Leopold Vietoris (1891–2002)) if $X$ is a $\mathsf{T}_1$ space with the property that for each closed set $A \subseteq X$ and each point $x \in X \setminus A$, there are disjoint open sets $U$ and $V$ with $A \subseteq U$ and $x \in V$.*

Clearly regular spaces are Hausdorff; the next example shows that regularity is a more restrictive property.

**Example 16.2** (A $\mathsf{T}_2$ Space that is not $\mathsf{T}_3$). *Let the real line $\mathbb{R}$ be given the topology $\mathcal{T}$, with subbase consisting of: (i) usual open intervals of the form $(a, b)$, for $a < b$, ; and (ii) the set $\mathbb{Q}$ of rational numbers. (So a typical basic open set may be either a usual bounded open interval or a bounded open interval intersected with $\mathbb{Q}$.) Then $\mathcal{T}$ is finer than the usual topology; hence, by Exercise 9.11 (4), it is a Hausdorff topology. We show it is not regular. Indeed, pick any rational number, say 0. The set of rationals is a $\mathcal{T}$-open set, so the set $A := \mathbb{R} \setminus \mathbb{Q}$ is $\mathcal{T}$-closed and does not contain the point 0. Assume $U$ and $V$ are disjoint $\mathcal{T}$-open sets, where $0 \in U$ and $A \subseteq V$. Then we may take $U$ to be of the form $(-r, r) \cap \mathbb{Q}$, for some real $r > 0$. Suppose $x$ is an irrational number in $(-r, r)$. Then $V$ contains all rational numbers sufficiently close to $x$; hence $V$ must intersect $U$. This is a contradiction, so we may infer that $\mathcal{T}$ is a Hausdorff topology that is not regular.*

We now show that Theorem 9.8 applies for the $\mathsf{T}_3$ axiom.

**Theorem 16.3.** *Regularity is a hereditary property.*

*Proof.* Suppose $X$ is a regular space, with $Y$ a subspace of $X$. Then a typical $Y$-closed set is of the form $A \cap Y$, where $A$ is $X$-closed. Let $y \in Y \setminus A$. Since $X$ is regular, there are disjoint $X$-open sets $U$ containing $y$ and $V$ containing $A$. Then $U \cap Y$ and $V \cap Y$ are disjoint $Y$-open sets containing $y$ and $A \cap Y$, respectively. This shows $Y$ is regular in its subspace topology.

□

The following is an easy, but very useful, paraphrase of regularity.

**Proposition 16.4.** *Let $X$ be a $\mathsf{T}_1$ space. Then $X$ is regular if and only if, whenever $x \in X$ and $U$ is an open neighborhood of $x$, then there exists an open set $V$*

*with $x \in V \subseteq CL(V) \subseteq U$.*

*Proof.* Suppose $X$ is regular, and pick $x \in U$, where $U$ is open in $X$. Then $A := X \setminus U$ is $X$-closed and doesn't contain $x$. So pick open $V$ containing $x$ and open $W$ containing $A$ such that $V \cap W = \emptyset$. Then $X \setminus W$ is an $X$-closed set that contains $V$; hence $\mathrm{Cl}(V) \cap W = \emptyset$. This tells us that $\mathrm{Cl}(V) \subseteq X \setminus A = U$.

The converse is Exercise 16.16 (1) below.

□

The next result extends Theorem 14.4.

**Theorem 16.5.** *Let $X$ and $Y$ be topological spaces. Then $X \times Y$ is regular if and only if both $X$ and $Y$ are regular.*

*Proof.* Recalling that each factor embeds as a subspace of the product (Exercise 14.10 (2)), and that regularity is a hereditary property (Theorem 16.3), we conclude that each factor is regular whenever the product is.

For the converse, suppose both $X$ and $Y$ are regular, and use the paraphrasing in Proposition 16.4. Pick $W$ open in $X \times Y$, with $p := \langle x, y\rangle \in W$. Then we can find an $X$-open neighborhood $U$ of $x$ and a $Y$-open neighborhood $V$ of $y$ such that $U \times V \subseteq W$. Using the regularity of the factor spaces (along with Proposition 16.4), we obtain an $X$-open set $U'$ and a $Y$-open set $V'$ such that $x \in U' \subseteq \mathrm{Cl}(U') \subseteq U$ and $y \in V' \subseteq \mathrm{Cl}(V') \subseteq V$. Then $p \in U' \times V' \subseteq \mathrm{Cl}(U' \times V') \subseteq \mathrm{Cl}(U') \times \mathrm{Cl}(V') \subseteq U \times V \subseteq W$. This shows the product is regular.

□

So far we have no examples of regular spaces; the following theorem is meant to redress that.

**Theorem 16.6.** *Let $X$ be a topological space. If $X$ is either compact Hausdorff, metrizable, or suborderable, then $X$ is regular.*

*Proof.* Assume first that $X$ is compact Hausdorff, and let $A$ be a closed subset of $X$. Fix $x \in X \setminus A$. By the Hausdorff assumption, for each $y \in A$ there are disjoint open neighborhoods $U_y$ of $y$ and $V_y$ of $x$. Let $\mathcal{U} := \{U_y : y \in A\}$. Then $\mathcal{U}$ is an open cover of $A$. Since $A$ is closed in $X$, Theorem 13.5 tells us that $A$ is compact. Thus there is a finite subfamily $\{U_{y_1}, \ldots, U_{y_n}\}$ of $\mathcal{U}$ that covers $A$. Let $U := U_{y_1} \cup \cdots \cup U_{y_n}$ and $V := V_{y_1} \cap \cdots \cap V_{y_n}$. Then $U$ is an open neighborhood of $A$, $V$ is an open neighborhood of $x$ (being a *finite* intersection of open neighborhoods of $x$), and $U \cap V = \emptyset$. Thus $X$ is regular.

Now assume that $X$ is metrizable, and let $d$ be a metric that gives rise to the topology on $X$. In anticipation of using Proposition 16.4, suppose $U$ is an open neighborhood of $x$ in $X$. Then there is an open $d$-ball neighborhood $B_d(x, \epsilon) \subseteq U$. Let $V := B_d(x, \frac{\epsilon}{2})$. Then (see Theorem 4.5 (v)) $x \in V \subseteq \mathrm{Cl}(V) \subseteq B_d[x, \frac{\epsilon}{2}] \subseteq B_d(x, \epsilon) \subseteq U$.

Finally suppose $X$ is suborderable. By Theorem 16.3, we may as well assume that $X$ is actually orderable; say, $X = \langle X, <\rangle$, where we take the standard open base for the topology on $X$ to consist of bounded open intervals, open rays, and the set $X$ itself. Let $U$ be an open neighborhood of $x$ in $X$, and suppose $x$ is not an end point. (The case where $x$ is an end point is part of Exercise 16.16 (3) below.) Then we have $x \in (a,b) \subseteq U$, for some $a < x < b$. Suppose it is the case that there is some $a < c < x$, but nothing strictly between $x$ and $b$. Then $(a,b) = (a,x]$. Set $V := (c,b)$. Then $x \in V = (c,x] \subseteq \mathrm{Cl}(V) = [c,x] \subseteq U$. The other cases are handled similarly.

□

The next (and strongest) separation axiom is a natural progression from regularity. But, aside from the surface similarity in formulation, it is amazingly different in its behavior.

**Definition 16.7** (The $\mathsf{T}_4$, or Tietze, Axiom)**.** *A topological space $X$ is called a $\mathsf{T}_4$ **space** (or a **normal space**, or sometimes a **Tietze space**, after Heinrich F. Tietze (1880–1964)) if $X$ is a $\mathsf{T}_1$ space with the property that for each pair of disjoint closed sets $A, B \subseteq X$, there are disjoint open sets $U$ and $V$ with $A \subseteq U$ and $B \subseteq V$.*

**Remark 16.8.** *Some authors make a distinction between regularity and the $\mathsf{T}_3$ axiom, reserving the $\mathsf{T}_1$ assumption only for the latter. (Likewise, they distinguish normality from the $\mathsf{T}_4$ axiom.) Because of the elementary nature of this course, we conflate the two notions.*

Clearly normal spaces are regular, and in Example 16.11 below we show that regular spaces needn't be normal. This, however, turns out not to be an elementary result. The main tools in the theory of normal spaces are known as Urysohn's lemma (after Pavel S. Urysohn (1898–1924)) and Tietze's extension theorem; we state these important results without proof.

**Theorem 16.9** (Main Characterizations of Normality)**.**

(i) (Urysohn's Lemma) *Let $X$ be a $\mathsf{T}_1$ space. Then $X$ is normal if and only if, whenever $A$ and $B$ are disjoint closed subsets of $X$, then there exists a continuous $f : X \to [0,1]$ such that $f[A] = \{0\}$ and $f[B] = \{1\}$.*

(ii) (Tietze's Extension Theorem) *Let $X$ be a $\mathsf{T}_1$ space. Then $X$ is normal if and only if, whenever $A$ is a closed subset of $X$ and $f : A \to \mathbb{R}$ is continuous, there exists a continuous $F : X \to \mathbb{R}$ such that $F(x) = f(x)$ whenever $x \in A$ (i.e., $F|A = f$; $F$* extends *$f$). Moreover, if $f[X] \subseteq [a,b] \subseteq \mathbb{R}$, then we may find $F$ so that $F[X] \subseteq [a,b]$ as well.*

**Remarks 16.10.**

(i) *Referring to Exercise 9.11 (5), the condition given in Urysohn's lemma could well be labeled "functionally normal." But while functionally Hausdorff and Hausdorff are distinct properties, functionally normal and normal are not.*

(ii) *If, in Urysohn's lemma, you assume one of the closed subsets $A$ or $B$ to be a single point, then you get a separation property called* complete regularity *(not functional regularity, as you might expect). It turns out that* *complete regularity is a highly-studied property, also known as the Tychonoff property, after Andrei N. Tychonoff (1906–1993). Among the main results conserning this property are: (i) products and subspaces of completely regular spaces are completely regular; and (ii) a space is completely regular if and only if it can be embedded as a subspace of a compact Hausdorff space.*

**Example 16.11** (A $\mathsf{T}_3$ Space that is not $\mathsf{T}_4$)**.** *See Exercise 14.10 (6). Given the Sorgenfrey line $\mathbb{L}$; i.e., the real line equipped with topology basically generated by half-open intervals $[a, b)$, the topological product $\mathbb{L}^2$ is commonly referred to as the Sorgenfrey* plane. *This is a very popular counterexample to many reasonable topological conjectures; in particular, it is a regular space that is not normal.*

Ad (regular)*: We first show that $\mathbb{L}$ is regular, and then apply Theorem 16.5 to infer that $\mathbb{L}^2$ is regular too. So pick a point $x$ and an open $\mathbb{L}$-neighborhood $U$ of $x$. Then there is a basic $\mathbb{L}$-open neighborhood $[a, b)$, containing $x$ and contained within $U$. In the usual order on the real line, then, we have $a \leq x < b$; so we may pick a real number $c$ with $x < c < b$. Then we have $x \in [a, c) \subseteq Cl([a, c)) = [a, c] \subseteq [a, b) \subseteq U$; hence $\mathbb{L}$ is regular, by Proposition 16.4. This completes the proof that the Sorgenfrey plane is a $\mathsf{T}_3$ space.*

Ad (not normal)*: In Exercise 14.10 (6), you are asked to determine whether $\Delta' := \{\langle x, -x\rangle : x \in \mathbb{L}\}$ is homeomorphic to $\mathbb{L}$. It actually is not; in fact it is a discrete subspace of the Sorgenfrey plane. (To see this, consider the relatively open set $\{\langle x, -x\rangle\} = \{\langle x, -x\rangle\} \cap [x, x + 1) \times [-x, -x + 1)$.) Not only is it discrete, it is also closed. (Given $\langle a, b\rangle \notin \Delta'$, it is easy to find an open set $[a, a+\epsilon)\times[b, b+\epsilon)$ that misses $\Delta'$.) Now let $A$ be any subset of $\Delta'$. Since $A$ is closed in $\Delta'$ and $\Delta'$ is closed in $\mathbb{L}^2$, $A$ is closed in $\mathbb{L}^2$. By the same token, $\Delta' \setminus A$ is also closed in $\mathbb{L}^2$. So let $f_A : \Delta' \to [0, 1]$ take $A$ to $\{0\}$ and $\Delta' \setminus A$ to $\{1\}$. Then each such $f_A$ is continuous; moreover, if $A$ and $B$ are distinct subsets of $\Delta'$, then $f_A \neq f_B$. So assume $\mathbb{L}$ to be normal. Then, by The Tietze extension theorem (Theorem 16.9 (ii)), each $f_A$ extends to a continuous map $F_A : \mathbb{L}^2 \to [0, 1]$. Moreover, if $A$ and $B$ are distinct subsets of $\Delta'$, then $F_A \neq F_B$. This says that there are at least as many continuous $[0, 1]$-valued functions on the Sorgenfrey plane as there are functions from the real line to the real line. For those with the set-theoretical background, this cardinal is strictly bigger than the cardinality of the real line. On the other hand, the set $\mathbb{Q}^2$, consisting of points with rational coordinates, is dense in the Sorgenfrey plane. But $\mathbb{Q}$ is countable, and the number of $[0, 1]$-valued functions on $\mathbb{Q}$ is exactly the cardinality of the real line. Since any continuous function from a space to $[0, 1]$ is determined by its values on a dense subset of that space (see Exercise 9.11 (14)), there can be no more continuous $[0, 1]$-valued functions defined on $\mathbb{L}^2$ than there are real numbers. This contradiction shows that the Sorgenfrey plane cannot be a normal space.*

There is a weak analogue of Theorem 16.3 for normality; the full analogue is false.

**Theorem 16.12.** *Normality is a closed-hereditary property.*

*Proof.* Suppose $X$ is a normal space, with $Y$ a closed subspace of $X$. Let $A$ and $B$ be disjoint $Y$-closed sets. Since $Y$ is $X$-closed, it follows (see Exercise 5.12 (10)) that both $A$ and $B$ are $X$-closed. By the normality assumption on $X$, there are disjoint $X$-open sets $U \subseteq A$ and $V \supseteq B$. Then $U \cap Y$ and $V \cap Y$ are disjoint $Y$-open sets separating $A$ and $B$.

□

To address the question of normality and the product construction, only half of Theorem 16.5 remains when regularity is strengthened to normality.

**Theorem 16.13.** *Let $X$ and $Y$ be topological spaces. If $X \times Y$ is normal, then both $X$ and $Y$ are normal.*

*Proof.* We again recall that each factor embeds as a subspace of the product, but add that the embedded subspaces may be taken to be closed (Exercise 14.10 (2)) when the factor spaces are $\mathsf{T}_1$ spaces. By Theorem 16.12, then, we infer that the factor spaces are normal whenever the product is.

□

**Remarks 16.14.** (i) *Many quite innocent-looking problems concerning normality are surprisingly difficult. For example, one way to show that normality is not closed under the taking of products is to use the Sorgenfrey plane. All you need add is a proof that the Sorgenfrey line is normal (see Example 17.7 below) and use Example 16.11. It was a long-unsolved problem, due to Clifford H. Dowker (1912–1982), whether the product of a normal space with the closed unit interval is necessarily normal. The negative answer was provided by Mary Ellen Rudin (1924–) in the early 1970s.*

(ii) *One may easily build on the proof that compact Hausdorff spaces are regular (Theorem 16.6) to show that they are, in fact, normal (see Exercise 16.16 (4) below). Then, since the Sorgenfrey line is normal, it is completely regular as well (see Remark 16.10 (ii)). Since complete regularity is preserved under the taking of products, we infer that the Sorgenfrey plane is also completely regular. Now it is precisely the completely regular spaces that may be embedded as subspaces of compact Hausdorff spaces. Thus there is a normal space that contains a homeomorphic copy of the nonnormal space $\mathbb{L}^2$.*

The analogue of Theorem 16.6 for normality is still true, and we end this section with a proof for the metrizable case. (The suborderable case requires a much more sophisticated argument, and will not be proved here.)

**Theorem 16.15.** *Every metrizable space is normal.*

*Proof.* Let $d$ be a metric on $X$ that gives rise to its metrizable topology, and suppose $A$ and $B$ are disjoint closed subsets of $X$. For any $a \in A$, $a$ is *not* in $B$ and $B$ is closed. Therefore there is a real $\epsilon_a > 0$ such that the open $d$-ball $B_d(a, \epsilon_a)$ misses $B$. By the same token, for each $b \in B$, there is some real $\epsilon_b > 0$ such that the open $d$-ball $B_d(b, \epsilon_b)$ misses $A$.

We now define $U := \bigcup\{B(a, \frac{\epsilon_a}{2}) : a \in A\}$ and $V := \bigcup\{B(b, \frac{\epsilon_b}{2}) : b \in B\}$. Then clearly $U$ is an open neighborhood of $A$ that misses $B$ and $V$ is an open neighborhood of $B$ that misses $A$; we need to show that $U$ and $V$ also miss each other. Indeed, suppose there is some $x \in U \cap V$. Then for some $a \in A$ and some $b \in B$, we have $x \in B_d(a, \frac{\epsilon_a}{2}) \cap B_d(b, \frac{\epsilon_b}{2})$. Suppose first that $\epsilon_a \leq \epsilon_b$. Then $d(a,b) \leq d(a,x) + d(x,b) < \frac{\epsilon_a}{2} + \frac{\epsilon_b}{2} \leq \frac{\epsilon_b}{2} + \frac{\epsilon_b}{2} = \epsilon_b$. But this says that $a \in B_d(b, \epsilon_b)$, contradicting the fact that $B_d(b, \epsilon_b)$ misses $A$. If it turns out that $\epsilon_a \geq \epsilon_b$, then we infer that $b \in B_d(a, \epsilon_a)$, with another contradiction.

□

**Exercises 16.16.** (1) *Complete the proof of Proposition 16.4.*

(2) *Devise (and prove) an analogous version of Proposition 16.4 that characterizes normality.*

(3) *Complete the proof of Theorem 16.6 (i.e., the orderable case where the point $x$ is an end point).*

(4) *Show that compact Hausdorff spaces are normal [Hint: build on the proof of regularity in Theorem 16.6.]*

(5) *Show that if $X$ is a connected compact Hausdorff space with more than one point, then there is a continuous map from $X$ onto $[0,1]$. Conclude that $X$ must then have at least as many points as there are real numbers.*

(6) *A topological space $X$ is* ***hereditarily normal*** *if each subspace of $X$ is normal. Show that metrizable spaces are hereditarily normal.*

(7) *(Refer to Exercise 11.17 (9).) Show that a zero-dimensional $\mathsf{T}_0$ space is regular. (Try showing it's completely regular to boot.)*

(8) *A space $X$ is called* ***locally compact*** *if every point has a neighborhood base consisting of open sets whose closures are compact. Show that a locally compact Hausdorff space is regular. (It's actually completely regular.)*

(9) *If $X$ is a finite set, how many topologies on $X$ are normal?*

(10) ** (See Exercise 11.17 (8) above.) Show that, in a compact Hausdorff space, the component of a point and the quasicomponent of that point are the same set. [Hint: Use a contradiction argument, employing Exercise 16.16 (4) above, to show $Q(x)$ is connected when the space is compact Hausdorff.]*

## 17. Some Countability Conditions

Recall (Remark 10.10 (i)) that we defined a topological space $X$ to be first countable if, for each $x \in X$ there is a nested sequence $B_1 \supseteq B_2 \supseteq \dots$ of open nbds of $x$ such that every nbd of $x$ contains some $B_n$ in the family. We saw that metrizable spaces clearly are first countable; we will see in this section (Example 17.5) that the Sorgenfrey line $\mathbb{L}$ is first countable and regular (indeed, normal) but nonmetrizable.

A property stronger than first countability is second countability, introduced in Definition 13.10. We showed in Corollary 13.13 that, for second countable spaces, ones having countable bases, the notions of compactness and of countable compacness are equivalent. In this section we explore further how second countability plays a significant role, especially in the theory of metrizable spaces.

**Theorem 17.1.** *Let $X$ be a second countable topological space. Then:*

(i) *$X$ is **separable**; i.e., $X$ has a countable dense subset.*
(ii) *$X$ has the **Lindelöf Property** (after Ernst L. Lindelöf, 1870–1946); i.e., every open cover of $X$ has a countable subcover.*

*Proof. Ad (i)*: Let $\mathcal{B} := \{B_1, B_2, \dots\}$ be a countable base of nonempty open sets for $X$. For each $n = 1, 2, \dots$, pick $x_n \in B_n$. Then $D := \{x_1, x_2, \dots\}$ is a countable subset of $X$, which we now show to be dense. Indeed, if $U$ is a nonempty $X$-open set and $x \in U$, then there exists some $n$ such that $x \in B_n \subseteq U$. Hence $x_n \in U$, so $U \cap D \neq \emptyset$.

*Ad (ii)*: Let $\mathcal{B}$ be as above, and suppose $\mathcal{U}$ is an open cover of $X$. For each $x \in X$, pick $U_x \in \mathcal{U}$ such that $x \in U_x$. Next, pick whole number $n_x$ such that $x \in B_{n_x} \subseteq U_x$. Then $\{B_{n_x} : x \in X\}$ is a countable open cover of $X$, each of whose members is contained in some member of $\mathcal{U}$. So, for each $x \in X$, let $U_{n_x} \in \mathcal{U}$ be picked so that $B_{n_x} \subseteq U_{n_x}$. Then $\{U_{n_x} : x \in X\}$ is a countable subfamily of $\mathcal{U}$ that covers $X$.

□

**Remarks 17.2.** (i) *Because euclidean $n$-space $\mathbb{R}^n$ (usual topology) is second countable, it is separable. Also, because of the rational numbers, the Sorgenfrey line $\mathbb{L}$ is also separable. (We'll see later that $\mathbb{L}$ is* not *second countable, however.)*

(ii) *Note how nicely the Lindelöf property fits in with countable compactness: countably compact + Lindelöf = compact.*

(iii) *Note that second countability is a hereditary property. It turns out that while neither separability nor the Lindelöf property is hereditary (hard), separability is open-hereditary and the Lindelöf property is closed-hereditary. (See Exercises 17.9 below.) Because second countability is hereditary, and because it implies both separability and the Lindelöf property, we infer that second countable spaces are both* hereditarily *separable and* hereditarily *Lindelöf.*

(iv) *Theorem 13.12 tells us that we may check compactness using members of an open base for the topology. The same proof (practically* verbatim*) works* *for checking the Lindelöf property.*

The following is a result that proves very useful if you want to study a space with a particularly nice base for its open sets.

**Theorem 17.3.** *Suppose $X$ is second countable, and $\mathcal{C}$ is a (not necessarily countable) base for the open sets of $X$. Then there is a countable base for $X$ consisting of members of $\mathcal{C}$.*

*Proof.* Let $\mathcal{B}$ be a countable base for $X$. Then $X$ is hereditarily Lindelöf, so each member of $\mathcal{B}$ may be written as a countable union of members of the open base $\mathcal{C}$. For each $B \in \mathcal{B}$, let $\mathcal{C}_B$ be a countable subfamily of $\mathcal{C}$ such that $B = \bigcup \mathcal{C}_B$. Then $\bigcup\{\mathcal{C}_B : B \in \mathcal{B}\}$ is an open base for $X$, consisting of members of $\mathcal{C}$.

□

The following is the central result of this section, and underscores the importance of metrizability in topology.

**Theorem 17.4.** *For metrizable spaces, the notions of second countability, separability, and the Lindelöf property are all equivalent.*

*Proof.* Suppose $X$ is a metrizable space, say with compatible metric $d$. Our plan is to show that second countability follows from either of the two other properties; in view of Theorem 17.1, this will be enough.

*Ad (separable) $\Longrightarrow$ (second countable)*: Let $A$ be a countable dense subset of $X$, and let $\mathcal{B} := \{B_d(a, \rho) : a \in A, \rho \in \mathbb{Q}^+\}$. Then $\mathbb{B}$ is a countable family of $X$-open sets; it remains to show $\mathcal{B}$ is an open base. Indeed, suppose $U$ is an $X$-open set, with $x \in U$. Pick $\epsilon > 0$ such that $B_d(x, \epsilon) \subseteq U$. Since $A$ is dense in $X$, there is some $a \in A \cap B_d(x, \frac{\epsilon}{2})$. Then $d(a, x) < \frac{\epsilon}{2}$, so $\frac{\epsilon}{2} < \epsilon - d(a, x)$. Let $\rho \in \mathbb{Q}^+$ be such that $\frac{\epsilon}{2} < \rho < \epsilon - d(a, x)$. Then $d(a, x) < \frac{\epsilon}{2} < \rho$, so $x \in B_d(a, \rho)$. On the other hand, if $y \in B_d(a, \rho)$, then $d(x, y) \leq d(a, x)+d(a, y) < d(a, x)+\rho < d(a, x)+(\epsilon-d(a, x)) = \epsilon$. Thus $x \in B_d(a, \rho) \subseteq B_d(x, \epsilon) \subseteq U$.

*Ad (Lindelöf) $\Longrightarrow$ (second countable)*: For each positive integer $n$, let $\mathcal{U}_n$ be the open cover $\{B_d(a, \frac{1}{n}) : x \in X\}$. Then there is a countable subcover; i.e., a countable subset $A_n$ of $X$ such that $\mathcal{U}_n := \{B_d(a, \frac{1}{n}) : a \in A_n\}$ covers $X$. Let $\mathcal{B} := \bigcup_{n=1}^{\infty} \mathcal{U}_n$. Then $\mathcal{B}$ is a countable union of countable collections, and is hence countable. We claim it is an open base. As before, suppose $U$ is an $X$-open set, with $x \in U$. Pick $\epsilon > 0$ such that $B_d(x, \epsilon) \subseteq U$. Let $n$ be a positive whole number large enough so that $\frac{1}{n} < \epsilon$. Then $\mathcal{U}_{2n}$ covers $X$, so there is some $a \in A_{2n}$ such that $x \in B_d(a, \frac{1}{2n})$. Suppose $y \in B_d(a, \frac{1}{2n})$. Then $d(x, y) \leq d(a, x) + d(a, y) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n} < \epsilon$. Thus $x \in B_d(a, \frac{1}{2n}) \subseteq B_d(x, \epsilon) \subseteq U$.

□

In the hint to Exercise 10.12 (2), we mention that the Sorgenfrey line is a first countable Hausdorff space that is nonmetrizable, deferring a proof to this section.

We now are in a position to provide that proof.

**Example 17.5** (A First Countable $\mathsf{T}_3$ Space that is not Metrizable)**.** *Indeed, we show the Sorgenfrey line $\mathbb{L}$ is: (i) first countable, (ii) regular, (iii) separable, (iv) Lindelöf, but (v) not second countable. By Theorem 17.4, it is therefore nonmetrizable.*

Ad (i)*: $\mathbb{L}$ is first countable because, for any $x \in \mathbb{L}$, the family $\{[x, x + \frac{1}{n}) : n = 1, 2, \dots\}$ is a countable neighborhood base for the open neighborhoods of $x$.*

Ad (ii)*: This was shown in Example 16.11. ($\mathbb{L}$ is even normal, but this was stated without proof.)*

Ad (iii)*: The countable set $\mathbb{Q}$ of rational numbers is not only dense in the usual real topology, but dense in the finer Sorgenfrey topology as well.*

Ad (iv)*: Let $\mathcal{U}$ be an open cover of $\mathbb{L}$. By Remark 17.2 (iv), we may assume all sets in $\mathcal{U}$ to be of the form $[a, b)$, with $a < b$. Now let $\mathcal{V} := \{(a, b) : [a, b) \in \mathcal{U}$; i.e., we form $\mathcal{V}$ by removing the left-hand end points from the sets in $\mathcal{U}$. Then $\mathcal{V}$ is a family of $\mathbb{R}$-open sets. Let $V := \bigcup \mathcal{V}$. If $x \in A := \mathbb{L} \setminus V$, then $x$ must be the left-hand end point of some $[x, b_x) \in \mathcal{U}$, but cannot be in any $[a, b) \in \mathcal{U}$, with $a < x$. Suppose $x$ and $y$ are both in $A$, with $x < y$. Then it must be the case that $b_x \leq y$. Hence $[x, b_x)$ and $[y, b_y)$ are disjoint. Since $x < b_x$ for each $x \in A$, there is a rational number $q_x$, with $x < q_x < b_x$. Since $q_x < q_y$ for $x < y$ in $A$, and since there are only countably many rational numbers, we may infer that $A$ is a countable set.*

*Now $\mathbb{R}$ is second countable, so it is hereditarily Lindelöf (Remark 17.2 (iii)). Thus there is a countable subcollection $\{(a_n, b_n) : n = 1, 2, \dots\}$ of $\mathcal{V}$ that covers $V$. If we add on the end points, we have a countable subcollection $\{[a_n, b_n) : n = 1, 2, \dots\}$ of $\mathcal{U}$ that covers $V$; i.e., a countable subcollection of $\mathcal{U}$ that covers all but a countable number of points. For each such point, throw in a member of $\mathcal{U}$ that contains that point. By this process we add only a countable number of extra sets from $\mathcal{U}$; hence $\mathcal{U}$ has a countable subcover.*

Ad (v)*: Suppose, for the sake of a contradiction, that $\mathbb{L}$ is second countable. Then, by Theorem 17.3, there is a countable base consisting of sets of the form $[a, b)$, with $a < b$. Let $\mathcal{B} := \{[a_n, b_n) : n = 1, 2, \dots\}$ be such a base. Since the real line is uncountabl, though, there must be some $x \in \mathbb{L}$ which is not equal to any point $a_n$. Let $U := [x, x + 1)$. Since $U$ is an open neighborhood of $x$, there must be some $n \geq 1$ with $x \in [a_n, b_n) \subseteq U$. But this is clearly impossible; hence $\mathbb{L}$ cannot be second countable.*

One of the results of Theorem 14.4 is that the product of two compact spaces is compact. Considering how similar compactness and the Lindelöf property are on the surface, it is tempting to conjecture that the corresponding assertion holds for the latter property as well as the former. It doesn't. In Example 17.5 above, we showed that the Sorgenfrey line is Lindelöf; we will finish this section by showing that the Sorgenfrey plane, the product of the Sorgenfrey line with itself, is not. We will use the fact (see Example 16.11) that $\mathbb{L}^2$ is not normal, but first we need a result that connects normality with the Lindelöf property.

**Theorem 17.6.** *Every Lindelöf regular space is normal.*

*Proof.* Let $A$ and $B$ be disjoint closed subsets of $X$. By the regularity assumption, we may choose, for each $a \in A$, an open neighborhood $U_a$ of $a$ such that $\mathrm{Cl}(U_a) \cap B = \emptyset$. Likewise, for each $b \in B$, we may choose an open neighborhood $V_b$ of $b$ such that $\mathrm{Cl}(V_b) \cap A = \emptyset$.

By the Lindelöf assumption, since $A$ and $B$ are both closed (see Exercise 17.9 (3)), we may find a countable subcollection $\{U_1, U_2, \dots\}$ of $\{U_a : a \in A\}$ that covers $A$. Likewise, we may find a countable subcollection $\{V_1, V_2, \dots\}$ of $\{V_b : b \in B\}$ that covers $B$.

For each $n = 1, 2, \dots$, define the open sets $U'_n := U_n \setminus (\bigcup_{i=1}^n \mathrm{Cl}(V_i))$ and $V'_n := V_n \setminus (\bigcup_{i=1}^n \mathrm{Cl}(U_i))$. If $a \in A$, say $a \in U_n$, then $a$ is not in the closure of any $V_k$, so $a \in U'_n$. Thus $U := \bigcup_{n=1}^\infty U'_n$ is an open set containing $A$. Likewise $V := \bigcup_{n=1}^\infty V'_n$ is an open set containing $B$.

It remains to show $U$ and $V$ are disjoint. Indeed, suppose not. If $x \in U \cap V$, then there exist $m, n \geq 1$ such that $x \in U'_m \cap V'_n$. Suppose $m \leq n$. Then we have $x \in U'_m \subseteq U_m \subseteq \mathrm{Cl}(U_m)$. On the other hand, $x \in V'_n$, so $x \notin \bigcup_{i=1}^n \mathrm{Cl}(U_i)$; in particular, $x \notin \mathrm{Cl}(U_m)$. If $m \geq n$, a similar contradiction occurs; hence $U$ and $V$ must be disjoint. This shows $X$ is normal.

□

**Example 17.7** (A Lindelöf Space whose Square is not Lindelöf)**.** *Our example is the Sorgenfrey line* $\mathbb{L}$*. We showed it to be regular in Example 16.11 and to be Lindelöf in Example 17.5. (So, by Theorem 17.6, it is actually normal.) We also showed in Example 16.11 that the Sorgenfray plane* $\mathbb{L}^2$ *is regular, but not normal. Therefore it cannot be Lindelöf, again by Theorem 17.6.*

**Remark 17.8.** *Because* $\mathbb{Q}^2$ *is a countable dense subset of* $\mathbb{L}^2$*, we see that the Sorgenfrey plane is also an example of a separable space that is not Lindelöf.*

**Exercises 17.9.**

(1) *Show that separability is an open-hereditary property.*

(2) *Show that the Lindelöf property is a closed-hereditary property.*

(3) *Adapt the argument in Example 17.5 to show that the Sorgenfrey line is* hereditarily *Lindelöf.*

(4) *Let $X$ be a topological space. A point $x \in X$ is called a **P-point** if, whenever $U_1, U_2, \dots$ are countably many open neighborhoods of $x$, there is an open set $U$ with $x \in U \subseteq \bigcap_{n=1}^\infty U_n$. Show that, if every point of $X$ is a P-point, then the intersection of any countable family of open subsets of $X$ is open in $X$. ($X$ is called a **P-space** when this happens.)*

(5) **Suppose $X$ is a regular P-space (see Exercise 17.9 (5)). Show that $X$ is zero-dimensional (see Exercise 16.16 (7)).*

(6) *Show that any second countable P-space must be discrete.*

(7) *A topological space $X$ is said to* ***satisfy the countable chain condition*** *if there is no collection of uncountably many pairwise disjoint open subsets of $X$. Show that every separable space satisfies the countable chain condition.*

(8) *Show that any set with the cofinite topology is a separable space.*

(9) *A topological space $X$ is said to be a* ***Baire*** *space, after René-Louis Baire (1874–1932), if the intersection of countably many dense open subsets of $X$ is dense in $X$ (see Exercise 5.12 (5)). Show that the rational line $\mathbb{Q}$ is* not *a Baire space. (Any euclidean space $\mathbb{R}^n$ is a Baire space, as is any compact Hausdorff space. This arises from what is known as the Baire category theorem.)*

(10) **Let $f : X \to Y$ be a continuous surjection, where $X$ is compact Hausdorff, and $Y$ is Hausdorff. Show that if $X$ is second countable, then $Y$ is too.*

(11) *Let $X$ be a second countable space, and suppose $\mathcal{B}$ is any (possibly uncountable) base for the topology on $X$. Then there is a countable base for $X$ that is contained in $\mathcal{B}$.*

## 18. Further Reading

Topology is a huge subject, with many branches. We were able to just barely scratch the surface in this course. The following four texts are some of the best I have seen; there are many others. (Just query The Marquette Library webpage under the LC call number QA 611.) All four are on reserve at Memorial Library.

Fred H. Croom: Principles of Topology, 1989.

George F. Simmons: Introduction to Topology and Modern Analysis, 1963.

James R. Munkres: Topology, 2000.

Lynn A. Steen and J. Arthur Seebach: Counterexamples in Topology, 1978.

Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI 53201-1881

*E-mail address*: `paulb@mscs.mu.edu`