Math 344-1: Introduction to Topology

Northwestern University, Lecture Notes

Written by Santiago Cañez

These are notes which provide a basic summary of each lecture for Math 344-1, the first quarter of "Introduction to Topology", taught by the author at Northwestern University. The book used as a reference is the 2nd edition of *Topology* by Munkres. Watch out for typos! Comments and suggestions are welcome.

Contents

Lecture 1: Topological Spaces

Why topology? Topology provides the most general setting in which we can talk about continuity, which is good because continuous functions are amazing things to have available. Topology does this by providing a general setting in which we can talk about the notion of "near" or "close", and it is this perspective which I hope to make more precise as we go on. In particular, at some point we'll come back and discuss why the properties which open sets are required to have in the definition of a topology are the right ones to have if you want to capture the idea that open sets should give a way to measure "nearness".

For now we point out that "nearness" in this sense cannot depend on distance nor length, since such things are not "topological" concepts. Imagine taking a sphere and stretching it out in one direction to make it thinner and thinner; this procedure does not change the "topology" of the sphere (whatever that means), but it does affect distance. Similarly, we would say that the surface of a donut and a coffee mug are the same topologically since we can (continuously!) deform one into the other, but such deformations will certainly affect distance and length. The need to define "near" without making use of distance is what, in my mind, serves as a guide to the modern definition of a topology.

Open in $\mathbb{R}^2$. open in $\mathbb{R}^2$, unions and intersections

Definition. defn of topology

Examples. $\mathbb{R}^2$, $\mathbb{R}$, $\mathbb{R}^n$ standard
discrete, trivial
cofinite

Line with two origins. Take two "points" p and q and consider the set

$$(\mathbb{R} - \{0\}) \cup \{p\} \cup \{q\}.$$

(The idea is that we replace the origin 0 in $\mathbb{R}$ with *two* new points.) The *line with two origins* is this set equipped with the following topology. First, any ordinary open set in $\mathbb{R}$ which does not contain 0 remains open in the line with two origins. For open sets U in $\mathbb{R}$ which *do* contain 0, we introduce two copies of U, each containing one of the two new "origins" p and q; to be clear, for U open in $\mathbb{R}$ with $0 \in U$, we take

$$(U - \{0\}) \cup \{p\} \text{ and } (U - \{0\}) \cup \{q\}$$

to be open sets in the line with two origins.

Picture this space as an ordinary line, only, as the name suggests, with two origins, usually drawn with one on top of the other. These two origins in a sense share the same open sets. To get another visualization, imagine taking two copies of $\mathbb{R}$ (so, two lines) and gluing each point in the first to the corresponding point in the second *except* for the two origins; the space resulting from this gluing procedure is the line with two origins. We'll talk about such gluing constructions later when we discuss *quotient* topologies.

Relation between "open' and "near". To start to give some intuition as to why open sets are good things to consider, recall the definition of what it means for a subset of $\mathbb{R}^2$ to be open: $U \subseteq \mathbb{R}^2$ is open if for any $p \in U$, there exists $r > 0$ such that $B_r(p) \subseteq U$. This is saying that U is open if any for any point p inside of it, points *close enough* to p (as measure by r) are still within

U. The standard drawing of a subset of $\mathbb{R}^2$ which is not open since it contains part of its boundary does not have this property, since we can have points get arbitrarily close to this boundary without belonging to the set in question itself.

But to highlight that this notion of "near" does not depend on distance, consider two distinct points in $\mathbb{R}^2$. Certainly if these points are drawn far enough apart we can easily surround each by open disks which do not intersect each other. The point is that no matter how visually close these points appear to be to one another (say the distance between them is the size of an electron), this is still true: the open disks we need might be incredibly small, but they still exist. Thus such points can still be "separated" in a topological sense, and so really aren't that "near" each after all. However, in the line with two origins something new happens: the two origins themselves cannot be separated in this way. To be precise, the claim is that there do not exist open sets containing the two origins which are disjoint, which is true since any open set containing one origin has to intersect an open set containing the other by the way in which defined open sets in that topology. Intuitively this says that the two origins are "arbitrarily close" to one another, even though there is no notion of "distance" defined a priori in this space.

Lecture 2: More on Topologies

Warm-Up 1. For a set X, the *cofinite* topology (also called the *finite complement* topology) on X is the one where we take as open sets $\emptyset$ and complements of finite sets. (Equivalently, we take as closed sets X itself and finite sets.) We showed that this indeed gives a topology on X, but we'll omit the details here since this can be found in the book.

Closed sets. We introduced the term "closed set" earlier than the book does, so we record it here. A subset of a topological space is *closed* if its complement is open. Note that the properties that open sets have in the definition of a topology on X then give the corresponding properties of closed sets: $\emptyset$ and X are closed, the intersection of arbitrarily many closed sets is closed, and the union of finitely many closed sets is closed.

Warm-Up 2. The Zariski topology on $\mathbb{R}^2$ is the one whose closed sets are common zero sets of polynomials in two variables. A problem on the first homework asks to show that this indeed gives a topology on $\mathbb{R}^2$, and here we verify just two special cases: show that if f and g are each polynomials in two variables, then $V(f) \cap V(g)$ and $V(f) \cup V(g)$ are closed, where $V(h)$ denotes the set of zeroes of h.

Indeed, a point in $V(f) \cap V(g)$ is one which is a zero of f and g simultaneously, meaning that it is a common zero of the polynomials in the set $\{f, g\}$. Hence

$$V(f) \cap V(g) = V(\{f, g\}),$$

so $V(f) \cap V(g)$ is closed. Now, a point in $V(f) \cup V(g)$ is one which is a zero either of f or g. But to say that $f(x, y) = 0$ or $g(x, y) = 0$ is the same as saying that $f(x, y)g(x, y) = 0$ since a product is zero when one factor is zero. Hence

$$V(f) \cup V(g) = V(fg),$$

so $V(f) \cup V(g)$ is closed. (Think about why this fails if we try to do the same for the union of infinitely many closed sets.)

Here is one last observation. Recall that the Zariski topology on $\mathbb{R}$ is defined in an analogous way, where we take zero sets of polynomials in one variable. In this case, since a nonzero polynomial in one variable can only have finitely many roots (it has no more than the degree of the polynomial), we see that any closed set in the Zariski topology on $\mathbb{R}$ is either $\mathbb{R}$ itself or consists of finitely many points. But this is precisely the characterization of the cofinite topology on $\mathbb{R}$, so we conclude that the cofinite and Zariski topologies on $\mathbb{R}$ are one and the same. This is not true in $\mathbb{R}^n$ for $n > 1$, where these two topologies are different.

Zariski vs Euclidean. As we've seen, the standard parabola $y = x^2$ defines a closed subset of $\mathbb{R}^2$ in the Zariski topology since it is the zero set of the polynomial $y - x^2$. Now, this set is also closed in the standard Euclidean topology on $\mathbb{R}^2$, which we can see either by convincing ourselves that its complement is open (in the "drawing small open disks" sense) or by recalling some facts from analysis, namely that the set of zeroes of any continuous function always defines a closed set. Moreover, it is true that the set of common zeroes of any collection of polynomials is closed in $\mathbb{R}^2$ in the standard topology for a similar reason. This implies in fact that any set which is open (respectively closed) in the Zarisksi topology on $\mathbb{R}^2$ is also open (respectively closed) in the standard topology, so the Zariski topology is *coarser* than the standard topology.

However, it is not true that any set which is closed in the standard topology is also closed in the Zariski topology. For instance, the graph of $y = e^x$ is closed in the standard topology and yet we claim that it is not closed in the Zariski topology. Now, $y - e^x$ is certainly not a polynomial in two variables (infinite polynomials don't count!), but this alone does not guarantee that its zero set is not open in the Zariski topology since there *could* be a polynomial in two variables which had the same zero set as $y - e^x$; there isn't, but this is somewhat difficult to prove, so take my word for it. Thus, the Zariski topology is actually *strictly coarser* than the standard topology on $\mathbb{R}^2$.

Coarse/fine topologies. We will often resort to defining topologies by specifying that they should be the *coarsest* ones in which some stated property should be true. To be precise, to say that $\mathcal{T}$ is the coarsest topology satisfying some property means that if $\mathcal{T}'$ is any other topology satisfying that same property, we should have $\mathcal{T} \subseteq \mathcal{T}'$. In practice this means that we allow as open sets whatever we need in order to guarantee that the stated property holds, and then we also take as open sets anything else we need to include to ensure we get a topology, but no more. (So, the coarsest topology in which a property holds is the one which has the fewest open sets needed to ensure that property holds.) This should become clearer as we actually start using this terminology. For now, notice that in the cofinite topology on a set, single points are always closed, and indeed we can characterize the cofinite topology on a set as the coarsest one in which this is true.

Lecture 3: Bases

Warm-Up. We describe the coarsest topology on $X = \{a, b, c, d, e\}$ in which $\{a, b\}$ and $\{b, d\}$ are closed.***FINISH***

Motivation for bases. We motivated the definition of a *basis* for a topology on a set by considering the case of open disks in $\mathbb{R}^2$. The point is the following: say we define $U \subseteq \mathbb{R}^2$ to be open if for any $p \in U$ there exists a disk $B_r(q)$ such that $p \in B_r(q) \subseteq U$. The subtlety is that now we are no longer requiring that the disk be centered at p itself; this is important, since the notion of "centered at" has no meaning in a general topological setting since there is no such thing as "distance" in general. The question is: if we use definition of open, how do we show that the intersection of two open sets is still open? If you work through the details, this boils down to showing that if p is in the intersection of two open disks

$$p \in B_r(q) \cap B_s(m),$$

there exists exists a third open disk $B_t(n)$ containing p and contained in this intersection:

$$p \in B_t(n) \subseteq B_r(q) \cap B_s(m).$$

This is precisely the second condition needed in the definition of a basis, and the point is that it is essential in showing that the intersection of open things is still open.

Definition. defn of basis

Do we get a topology? actually get a topology

Examples of bases. Open disks form a basis for the standard topology on $\mathbb{R}^2$. Note that this statement actually says two things: first that open disks form a basis for *a* topology, and second that the topology they generated *is* the standard topology. Such considerations are important to distinguish when we talk about a given collection of open sets forming a basis for a topology we already have in mind. The fact that the topology generated by open disks is the standard topology just comes from the fact that we defined "open" in the standard topology in terms of open disks.

But bases aren't unique! For instance, the collection of all open squares (regions enclosed by squares but excluding the boundary) also form a basis for the standard topology on $\mathbb{R}^2$, as does the collection of all open diamonds. We'll be able to see this more simply next time by noting that these bases arise from certain *metrics*.

Examples on $\mathbb{R}$. $\mathbb{R}_\ell$, $\mathbb{R}_K$, compare

Lecture 4: Metric Spaces

Open in topology generated by a basis. By definition, an open set in a topology generated by a basis is one which can be written as a union of basis elements. To make this condition simpler to work with, here is an equivalent formulation: U is open in the topology generated by a basis $\mathcal{B}$ if and only if for each $p \in U$, there exists $B \in \mathcal{B}$ such that $p \in B \subseteq U$. Indeed, picking such a basis element B_p for each $p \in U$ allows us to express U as

$$U = \bigcup_{p \in U} B_p,$$

which shows that U is open in the topology generated by $\mathcal{B}$. This is meant to be the analog of how we originally defined open sets in $\mathbb{R}^2$ using open disks.

Warm-Up. If $\mathcal{T}$ and $\mathcal{T}'$ are two topologies on a set generated by bases $\mathcal{B}$ and $\mathcal{B}'$ respectively, we showed that $\mathcal{T}'$ is finer than $\mathcal{T}$ if and only if for each $B \in \mathcal{B}$ and $p \in B$, there exists $B' \in \mathcal{B}'$ such that $p \in B' \subseteq B$. This allows us to characterize fineness/coarseness in terms of a basis. This is proved in the book, so we omit the proof here.

Metric spaces. We introduced metric spaces earlier than the book does in order to have a large class of examples of topological spaces. Indeed, metric spaces are the most intuitive topological spaces we have available, and understanding their properties goes a long way towards making sense of general topological notions. You can read about metric spaces and metric topologies in Section 20 of the book. That open balls with respect to a metric always form a basis for a topology (the metric topology) is left to the homework, but it is also in the book.

The one thing to keep in mind, however, is that metric spaces are very "nice" topological spaces, and won't illustrate on their own all the things that can happen in general. So, while they serve to give good intuition, you should avoid getting the habit of thinking of all topological spaces as if they were metric spaces.

Metrics on $\mathbb{R}^n$. Here are three metrics on $\mathbb{R}^n$, the so-called *Euclidean* metric d_E, the *box* metric d_{box}, and the *taxicab* metric d_{taxi}:

$$\begin{aligned} d_E((x_1,\ldots,x_n),(y_1,\ldots,y_n)) &= \sqrt{(x_1-y_1)^2+\cdots+(x_n-y_n)^2} \\ d_{box}((x_1,\ldots,x_n),(y_1,\ldots,y_n)) &= \max\{|x_1-y_1|,\ldots,|x_n-y_n|\} \\ d_{taxi}(x_1,\ldots,x_n),(y_1,\ldots,y_n)) &= |x_1-y_1|+\cdots+|x_n-y_n|. \end{aligned}$$

In $\mathbb{R}^2$, open balls with respect to the Euclidean metric are disks, open balls with respect to the box metric are squares, and open balls with respect to the taxicab metric are diamonds.

The point is that, even though these metrics are different, they all generated the *same* topology on $\mathbb{R}^n$, which is the standard topology. As a consequence of the Warm-Up, this can be shown by showing that an open ball with respect to the one metric always contains an open ball with respect to any of the other metrics; this will be left to a discussion problem.

Uniform topology on $\mathbb{R}^\omega$.

Metrizability. include discrete metric

Lecture 5: Product Topology

Warm-Up 1. open balls do give basis

Hausdorff spaces. For the sake of the Warm-Up today, we defined the notion of a Hausdorff space, which the book does soon enough. We say that a topological space X is *Hausdorff* if for any distinct $p,q \in X$, there exist open sets U containing p and V containing q such that $U \cap V = \emptyset$. (We say that p and q can be *separated* by open sets. We might also say that distinct points are "topologically distinguishable".)

Warm-Up 2. We show that metric spaces are always Hausdorff. Suppose X is a metric space with metric d and that $p,q \in X$ are distinct. Then $d(p,q) > 0$. We claim that $B_{d(p,q)/2}(p)$ and $B_{d(p,q)/2}(q)$ are then disjoint open sets containing p and q respectively. Indeed, if there exists $x \in B_{d(p,q)/2}(p) \cap B_{d(p,q)/2}(q)$, then

$$d(x,p) < \frac{d(p,q)}{2} \text{ and } d(x,q) < \frac{d(p,q)}{2},$$

so the triangle inequality gives

$$d(p,q) \le d(p,x) + d(x,q) < \frac{d(p,q)}{2} + \frac{d(p,q)}{2} = d(p,q).$$

This is not possible, so there is not such x and hence $B_{d(p,q)/2}(p)$ and $B_{d(p,q)/2}(q)$ are disjoint as claimed. Hence these are open sets separating p and q, so X is Hausdorff.

Non-metric spaces. If the topology on a topological space arises from a metric, we say that that space is *metrizable*. The Warm-Up says that any metrizable space must be Hausdorff, so we can now give examples of topologies which do not arise from metrics. For instance, the cofinite topology on an infinite set is not Hausdorff (any nonempty set open set in such a topology only excludes finitely many points, so any two such open sets will always have infinitely many points in common and so are not disjoint) and so cannot be given by a metric. The Zariski topology on $\mathbb{R}^n$ is also non-Hausdorff (we'll come back to this later), and so is also not given by a metric.

However, note that we can also have Hausdroff spaces which are not metrizable. For instance, $\mathbb{R}_\ell$ ($\mathbb{R}$ with the lower limit topology) is actually Hausdroff, but it turns out not metrizable. Showing that there is no metric on $\mathbb{R}$ which gives the lower limit topology is not something we can do just yet, but will follow from some other properties of metric spaces we'll look at later. (If you want to hear the buzzwords now, the key fact is that a metric space is "separable" if and only if it is "second countable"; $\mathbb{R}_\ell$ is separable but is not second countable, so it can't be metrizable.)

Finite products. The product topology is introduced in the finite case in Section 15 of the book, and in the infinite case in Section 19. In the finite case the product and box topologies are one and the same, but are crucially different in the infinite case.

Here we single out one aspect of the product topology in the finite case we looked at in class, which is essentially in the book if you read between the lines but is not made explicit. The claim is that the product topology on $X_1 \times \cdots \times X_n$ is the coarsest one in which the preimage of any open set under any project is itself open, i.e. for any $i = 1, \dots, n$

$$pr_i^{-1}(U) \text{ is open in } X_1 \times \cdots \times X_n \text{ whenever } U \text{ is open in } X_i.$$

Here, the i-th projection $pr_i : X_1 \times \cdots X_n \to X_i$ is the function which picks out i-th components:

$$pr_i(x_1, \dots, x_n) = x_i.$$

The condition given above in terms of preimages is (as we'll soon see) precisely what it means to say that each projection is continuous, so the claim is that the product topology is the coarsest one relative to which all projections are continuous.

To prove this, suppose $\mathcal{T}$ is any topology on $X_1 \times \cdots \times X_n$ having the property that the preimage of any open set under any projection is open in $X_1 \times \cdots \times X_n$. We want to show that $\mathcal{T}$ is finer than the product topology. To this end, suppose $U \subseteq X_1 \times \cdots \times X_n$ is open in the product topology. Then U can be written as the union of open sets of the form $U_{1\alpha} \times \cdots \times U_{n\alpha}$:

$$U = \bigcup_{\alpha \in I} (U_{1\alpha} \times \cdots \times U_{n\alpha})$$

for α in some indexing set I and where $U_{i\alpha}$ is open in X_i for each α. The preimage of such a $U_{i\alpha}$ under the projection pr_i is

$$pr_i^{-1}(U_{i\alpha}) = X_1 \times \cdots \times \underbrace{U_{i\alpha}}_{i\text{-th location}} \times \cdots X_n,$$

which we can write using product notation more succinctly as

$$pr_i^{-1}(U_{i\alpha}) = \prod_{j=1}^{n} U_j, \text{ where } U_i = U_{i\alpha} \text{ and } U_j = X_j \text{ for } i \neq j.$$

By the assumption on $\mathcal{T}$ this preimage is open in T. But then the intersection of finitely many such preimages is also open in $\mathcal{T}$, and such an intersection is precisely of the form

$$U_{1\alpha} \times \cdots \times U_{n\alpha} = pr_1^{-1}(U_{1\alpha}) \cap \cdots \cap pr_n^{-1}(U_{n\alpha}).$$

Thus

$$U = \bigcup_{\alpha \in I} (U_{1\alpha} \times \cdots \times U_{n\alpha})$$

is open in $\mathcal{T}$ as well, and hence $\mathcal{T}$ is finer than the product topology as claimed.

What goes wrong in $\mathbb{R}^\omega$? We finished with illustrating why we have to careful when trying to define the "product topology" in the case of infinite products. First, we can attempt to generalize the case we had for finite product as is and declare that the topology we want is the one generated by products of open sets. In the case of $\mathbb{R}^\omega$ (the space of infinite sequences of real numbers), this would say that the topology we want is the one generated by the basis consisting of things of the form

$$U_1 \times U_2 \times U_3 \times \cdots$$

where each U_i is open in $\mathbb{R}$. The topology arising in this way is the *box* topology on $\mathbb{R}^\omega$, which is now distinguished from the *product* topology we'll define next time. For instance, the infinite product

$$(-1, 1) \times \left(-\tfrac{1}{2}, \tfrac{1}{2}\right) \times \left(-\tfrac{1}{3}, \tfrac{1}{3}\right) \times \cdots,$$

where the i-th term is $(-\frac{1}{i}, \frac{1}{i})$, is open in the box topology on $\mathbb{R}^\omega$.

To see why the box topology is in some sense the "wrong" one to consider, take the sequence of elements in $\mathbb{R}^\omega$ given by

$$\left(\tfrac{1}{n}, \tfrac{1}{n}, \tfrac{1}{n}, \ldots\right).$$

To be clear, the first term in this sequence is $(1, 1, 1, \ldots)$, the second term is $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \ldots)$, and so on. (So we are taking a "sequence of sequences".) The question is: does this sequence in $\mathbb{R}^\omega$ converge? We'll define what *convergence* means in an arbitrary topological space next time, but for now we're just thinking about it in an intuitive sense. You would hope that since the sequence $\frac{1}{n}$ in $\mathbb{R}$ converges to 0, the sequence we're looking at in $\mathbb{R}^\omega$ should converge to

$$(0, 0, 0, \ldots) \in \mathbb{R}^\omega.$$

However, this is NOT true in the box topology! In fact, the sequence

$$\left(\tfrac{1}{n}, \tfrac{1}{n}, \tfrac{1}{n}, \ldots\right)$$

in $\mathbb{R}^\omega$ does not converge at all with respect to the box topology, the problem being that in a sense the box topology has "too many" open sets. However, this sequence WILL converge as we expect it to with respect to the product topology. We'll elaborate on all this next time, but is essentially the key distinguishing feature of the product topology vs the box topology.

Lecture 6: More on Products

Warm-Up. Denote $\mathbb{R}^n$ with the Zariski topology by $\mathbb{R}^n_{Zar}$. We will determine the relation between $\mathbb{R}^2_{Zar}$ and the product topology on $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$. (Of course, as sets both of these spaces are just $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.) First, recalling that the Zariski topology on $\mathbb{R}$ is the same as the cofinite topology, we note that closed sets in $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$ (apart from $\mathbb{R} \times \mathbb{R}$ itself) are of the form

$$\{\text{finite set}\} \times \mathbb{R},\ \mathbb{R} \times \{\text{finite set}\},\ \{\text{finite set}\} \times \{\text{finite set}\},$$

or finite unions of such things. (In general, if A is closed in X and B is closed in Y, then $A \times B$ is closed in $X \times Y$ under the product topology, which I encourage you to justify on your own.) Furthermore, these three types of closed subsets are finite unions of closed sets of the form

$$\{\text{point}\} \times \mathbb{R},\ \mathbb{R} \times \{\text{point}\},\ \{\text{point}\} \times \{\text{point}\},$$

so if each of these is open in $\mathbb{R}^2_{Zar}$ we will be able to conclude that anything open in $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$ is open in $\mathbb{R}^2_{Zar}$, meaning that $\mathbb{R}^2_{Zar}$ is finer than $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$. The set

$$\{a\} \times \mathbb{R}$$

is the vertical line $x = a$, which is the zero set of the polynomial $x - a$ and hence is closed in $\mathbb{R}^2_{Zar}$; the set

$$\mathbb{R} \times \{b\}$$

is the horizontal line $y = b$, and hence is closed in $\mathbb{R}^2_{Zar}$ since it is the zero set of $y - b$; and a single point $\{(a, b)\}$ is the common zero set of the collection of polynomials given by $\{x - a, y - b\}$, so is also closed in $\mathbb{R}^2_{Zar}$. Thus $\mathbb{R}^2_{Zar}$ is finer than $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$.

But we claim that the opposite inclusion does not hold: $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$ is not finer than $\mathbb{R}^2_{Zar}$. Indeed, the parabola $y = x^2$ is closed in $\mathbb{R}^2_{Zar}$ since it is the zero set of $y - x^2$, but this is not closed in $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$. Indeed, note that the types of closed sets in $\mathbb{R}_{Zar} \times \mathbb{R}_{Zar}$ mentioned above are all finite or collections of lines, and the parabola $y = x^2$ is none of these forms.

Convergence. We introduced the notion of convergence for sequences in a different spot than when the book does, so we record it here. A sequence (p_n) in a space X *converges* to $p \in X$ if for any open set U containing p, there exists $N \in \mathbb{N}$ such that $p_n \in U$ for $n \geq N$. This is precisely the same notion of convergence you would have seen for sequences in $\mathbb{R}$ in an analysis course if you replace the arbitrary open set U with one of the form $(p - \epsilon, p + \epsilon)$. One key difference, as we'll see later, is that in general topological spaces limits of sequences are NOT necessarily unique, in that a sequence can converge to possibly more than one point.

With this we can now justify the claim we finished with last time, namely that the sequence

$$\left(\tfrac{1}{n}, \tfrac{1}{n}, \tfrac{1}{n}, \ldots\right)$$

in $\mathbb{R}^\omega$ does not converge to $(0, 0, 0, \ldots)$ with respect to the box topology. The set

$$(-1, 1) \times \left(-\tfrac{1}{2}, \tfrac{1}{2}\right) \times \left(-\tfrac{1}{3}, \tfrac{1}{3}\right) \times \cdots$$

is open in the box topology and contains $(0, 0, 0, \ldots)$. Thus if the given sequence did converge to $(0, 0, 0, \ldots)$, there would have to exist $N \in \mathbb{N}$ such that

$$\left(\tfrac{1}{n}, \tfrac{1}{n}, \tfrac{1}{n}, \ldots\right) \in (-1, 1) \times \left(-\tfrac{1}{2}, \tfrac{1}{2}\right) \times \left(-\tfrac{1}{3}, \tfrac{1}{3}\right) \times \cdots$$

for $n \geq N$. But since all terms in this sequence are the same, this would require that

$$\tfrac{1}{n} \in \left(-\tfrac{1}{i}, \tfrac{1}{i}\right) \text{ for } n \geq N$$

for all $i \in \mathbb{N}$. In particular, all of these intervals would have $\frac{1}{N}$ in their intersection, which is nonsense because the intersection only consists of 0:

$$\bigcap_{i=1}^{\infty} \left(-\tfrac{1}{i}, \tfrac{1}{i}\right) = \{0\}.$$

Thus $(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots)$ does not converge to $(0, 0, 0, \ldots)$ with respect to the box topology as claimed.

Product topology on $\mathbb{R}^\omega$. We motivated the definition of the product topology on $\mathbb{R}^\omega$ via the characterization of the product topology in the finite case as being the coarsest one satisfying some property. To be clear, the question is: what is the coarsest topology on $\mathbb{R}^\omega$ with the property that

$$pr_i^{-1}(U) \text{ is open in } \mathbb{R}^\omega \text{ whenever } U \text{ is open in } \mathbb{R}$$

for every projection $pr_i : \mathbb{R}^\omega \to \mathbb{R}$? First, note that such a preimage concretely looks like

$$pr_i^{-1}(U) = \mathbb{R} \times \cdots \times \mathbb{R} \times \underbrace{U}_{i\text{-th location}} \times \mathbb{R} \times \cdots .$$

Such a set would have to be open in the coarsest topology we are looking for. But then the intersection of finitely many such sets would also have to be open, and such intersections look like

$$pr_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap pr_{i_k}^{-1}(U_{i_k}) = \text{product with } U_{i_t} \text{ in the } i_t\text{-th location and } \mathbb{R}\text{'s elsewhere.}$$

Such products form a basis, and the topology they generated is what we call the *product* topology on $\mathbb{R}^\omega$. To emphasize again, this is the coarsest topology on $\mathbb{R}^\omega$ satisfying the condition given above in terms of preimages. (Later we will see that this condition in terms of preimages is precisely what it means to say that each projection map $pr_i : \mathbb{R}^\omega \to \mathbb{R}$ is continuous, so this is saying that the product topology is the coarsest one relative to which all projections are continuous.)

The key difference between this and the box topology is that, while in the box topology anything product of the form

$$U_1 \times U_2 \times U_3 \times \cdots ,$$

where each U_i is open in $\mathbb{R}$, is open, in the product topology such products are open only when all but *finitely many* factors are actually $\mathbb{R}$ itself (or, only finitely many factors are not all of $\mathbb{R}$). In the case of $\mathbb{R}^\omega$, this can also be phrased as saying that

$$U_1 \times U_2 \times U_3 \times \cdots ,$$

is open if there exists N such that $U_n = \mathbb{R}$ for $n \geq N$. Thus,

$$(-1, 1) \times \left(-\tfrac{1}{2}, \tfrac{1}{2}\right) \times \left(-\tfrac{1}{3}, \tfrac{1}{3}\right) \times \cdots$$

is not open in the product topology on $\mathbb{R}^\omega$, so the argument we gave for why $(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots)$ does not converge to $(0, 0, 0, \ldots)$ does not work here. In fact, this sequence *does* converge to $(0, 0, 0, \ldots)$ in the product topology, which will be shown on a discussion problem. The thing which makes this work is that having only finitely many U_i's in a product

$$U_1 \times U_2 \times U_3 \times \cdots ,$$

be not all of $\mathbb{R}$ makes it possible to take a *maximum* of indices. More generally, the fact (which will be on a homework) is that convergence in the product topology is the same as *component-wise* convergence: a sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ in $\mathbb{R}^\omega$, where each $\mathbf{x}_i$ denotes a sequence of real numbers, converges to $\mathbf{y} = (y_1, y_2, y_2, \ldots) \in \mathbb{R}^\omega$ if and only if for each i, the sequence x_{ni} (where x_{ni} denotes the i-th component of $\mathbf{x}_n$) converges to y_i in $\mathbb{R}$ as n varies. This is analogous to saying that, for instance in $\mathbb{R}^3$, the sequence

$$(a_n, b_n, c_n) \text{ converges to } (a, b, c)$$

if and only if $a_n \to a, b_n \to b$, and $c_n \to c$. The product topology is the *finest* one in which convergence is the same as component-wise convergence in this sense.

Note that the requirement that "all but finitely many U_i are $\mathbb{R}$ itself" automatically holds in the case of finite products, since there are only finitely many factors to begin with in that case. Thus, the box and product topologies on finite products are the same.

Lecture 7: Arbitrary Products, Closed Sets

Warm-Up. Denote by $\mathbb{R}^\infty$ the set of elements in $\mathbb{R}^\omega$ which are *eventually zero*, meaning that past a certain index all terms are 0:

$$\mathbb{R}^\infty = \{(x_1, x_2, x_3, \ldots) \in \mathbb{R}^\omega \mid \text{there exists } N \in \mathbb{N} \text{ such that } x_n = 0 \text{ for } n \geq N\}.$$

We show that $\mathbb{R}^\infty$ is closed in $\mathbb{R}^\omega$ under the box topology but not under the product topology. We do so by showing that its complement

$$\mathbb{R}^\omega - \mathbb{R}^\infty = \{\mathbf{x} \in \mathbb{R}^\omega \mid \mathbf{x} \text{ is not eventually zero}\}$$

is or is not open. To be clear, to say that $\mathbf{x} = (x_1, x_2, \ldots)$ is not eventually zero means that it contains infinitely many nonzero terms, since if there were only finitely many nonzero terms going beyond all of these would put you in a spot where all remaining terms were zero.

To show that $\mathbb{R}^\omega - \mathbb{R}^\infty$ is open in the box topology, we show that any point in this complement is contained in an open set which remains fully within this complement. (The complement will then be the union of these open sets, and so will be open itself.) Let $\mathbf{x} \in \mathbb{R}^\omega - \mathbb{R}^\infty$. Since $\mathbf{x}$ is not eventually zero, it contains infinitely many nonzero terms, say

$$x_{i_k} \neq 0 \text{ for } k = 1, 2, 3, \ldots.$$

For each of these nonzero terms, we can find an interval (a_{i_k}, b_{i_k}) in $\mathbb{R}$ containing it which excludes zero:

$$x_{i_k} \in (a_{i_k}, b_{i_k}) \text{ but } 0 \notin (a_{i_k}, b_{i_k}).$$

Take the open sets U_n which are these intervals for n equal to one of the i_k, and $\mathbb{R}$ otherwise:

$$U_{i_k} = (a_{i_k}, b_{i_k}) \text{ and } U_n = \mathbb{R} \text{ for } n \text{ not equal to any } i_k.$$

The product

$$U_1 \times U_2 \times \cdots$$

is then open in the box topology and contains $\mathbf{x}$. However, since any element in this product contains infinitely many nonzero terms, since in particular the terms coming from one of the $U_{i_k} = (a_{i_k}, b_{i_k})$ is nonzero. Thus any such element is not eventually zero, so $U \subseteq \mathbb{R}^\omega - \mathbb{R}^\infty$. Hence $\mathbb{R}^\omega - \mathbb{R}^\infty$ is open in $\mathbb{R}^\omega$ under the box topology, so $\mathbb{R}^\infty$ is closed.

Now, the argument given above does not apply when we have the product topology, since the product

$$U_1 \times U_2 \times \cdots$$

defined above is not open in the product topology since infinitely many factors are strictly smaller than $\mathbb{R}$ itself. Indeed, for $\mathbf{x} \in \mathbb{R}^\omega - \mathbb{R}^\infty$, let

$$V_1 \times V_2 \times \cdots$$

be a basic open set under the product topology containing it. Then only finitely many V_i are not $\mathbb{R}$, so

$$V_n = \mathbb{R} \text{ for } n \text{ past some index } N.$$

Define the element $\mathbf{y} \in \mathbb{R}^\omega$ by taking any terms from $V_1, \ldots, V_N$ as the first N components, but then taking 0 as the component in V_n for $n > N$. (Here we use the fact that $V_n = \mathbb{R}$ for $n > N$ to guarantee that V_n contains zero.) Then

$$\mathbf{y} \in V_1 \times V_2 \times \cdots$$

and $\mathbf{y}$ is eventually zero, so the basic open set $V_1 \times V_2 \times \cdots$ is not contained in the complement $\mathbb{R}^\omega - \mathbb{R}^\infty$. Since any open set must contain one of these basic ones, we conclude that no open set around U under the product topology is contained fully within $\mathbb{R}^\omega - \mathbb{R}^\infty$. Hence $\mathbb{R}^\omega - \mathbb{R}^\infty$ is not open, so $\mathbb{R}^\infty$ is not closed in the product topology on $\mathbb{R}^\omega$.

Arbitrary products. An arbitrary product $\prod_{\alpha \in I} X_\alpha$ (so the product of the sets X_α indexed by α in some index set I) should intuitively consist of tuples $(x_\alpha)_{\alpha \in I}$ of elements, one from each X_α. (Concretely, $x_\alpha \in X_\alpha$.) Thinking about an arbitrary product in this way is fine, and is what we'll do for most purposes, but note that there is subtlety we should be aware of: just how exactly do you make the notion of an arbitrary "tuple" indexed by elements of I precise? This might be clearer in the case of finite products (like $\mathbb{R}^n$), or maybe even products indexed by $\mathbb{N}$ (like $\mathbb{R}^\omega$), but is not so clear when I is some random (uncountable) index set.

Here is the way this is usually made precise, based on the function approach to defining $\mathbb{R}^\mathbb{R}$ we mentioned last time. An element $(x_\alpha)_{\alpha \in I}$ of $\prod X_\alpha$ should be a choice of an element $x_\alpha \in X_\alpha$ for each $\alpha \in I$, which we can think of as characterizing a *function* from I to the X_α's, namely the function sending $\alpha \in I$ to $x_\alpha \in I$. Concretely, this gives a function

$$f : I \to \bigcup_\alpha X_\alpha$$

where $\alpha \in I$ is specifically sent to an element of X_α, as opposed to a function which might send $\alpha \in I$ to something in a differently-indexed X_β. This says that the function f should have the property that

$$f(\alpha) \in X_\alpha \text{ for each } \alpha \in I,$$

so that the element of $\bigcup_\alpha X_\alpha$ which corresponds to α comes from X_α itself. Thus, we can *define* the given product to the be the set of all such functions:

$$\prod_{\alpha \in I} X_\alpha = \left\{ f : I \to \bigcup_{\alpha \in I} X_\alpha \;\middle|\; f(\alpha) \in X_\alpha \text{ for each } \alpha \in I \right\}.$$

This gives a precise way of thinking of a tuple $(x_\alpha)_{\alpha \in I}$, which is then just the tuple encoding the values of a function f as above: the "α-th" element in the tuple is $x_\alpha = f(\alpha)$, which is the element in X_α which f assigns to α.

Let's make sure that this definition makes sense in the settings with which we're already familiar. First, how do we view $\mathbb{R}^n$ from this perspective? $\mathbb{R}^n$ is a product n many copies of $\mathbb{R}$, which we can think of as a product indexed by the finite set $\{1, 2, \ldots, n\}$, where the set occurring at each index i is just $\mathbb{R}$ itself. Elements of this product should thus correspond to functions

$$f : \{1, 2, \ldots, n\} \to \bigcup_{i=1}^n \mathbb{R}$$

satisfying $f(i) \in \mathbb{R}$ for each i. To simplify this, we note that the union on the right is simply $\mathbb{R}$ in this case, so all we are looking at are functions

$$f : \{1, 2, \ldots, n\} \to \mathbb{R}.$$

Such a function is fully characterized by the values $f(1), f(2), \ldots, f(n)$, which thus describe an n-tuple of the form $(f(1), f(2), \ldots, f(n))$, which is how we normally view an element of $\mathbb{R}^n$. Thus our definition of an arbitrary product reduces to the one we're already used to in the case of $\mathbb{R}^n$. More generally, an element in a finite product

$$X_1 \times X_2 \times \cdots \times X_n,$$

viewed as product indexed by $\{1, \ldots, n\}$, corresponds to a function

$$f : \{1, \ldots, n\} \to \bigcup_{i=1}^{n} X_i$$

such that $f(i) \in X_i$. This condition just says that in the n-tuple $(f(1), f(2), \ldots, f(n))$ encoding the values of f, the i-th component $f(i)$ should come from X_i itself, as opposed to having, for instance, the first component $f(1)$ comes from X_2. Hence again, the definition of a product given above reduces to the one we expect in the finite case.

For any X and Y, an element of $Y^X = \prod_{x \in X} Y$, which is the product of "X-many" copies of Y, is formally defined as a function

$$f : X \to \bigcup_{x \in X} Y = Y,$$

which is how we get that Y^X is just the set of functions from X to Y. In particular, $\mathbb{R}^\omega$ is the same as $\mathbb{R}^\mathbb{N}$, which is the set of functions from $\mathbb{N}$ to $\mathbb{R}$; a function $\mathbb{N} \to \mathbb{R}$ is indeed a precise way of defining the notion of a *sequence* in $\mathbb{R}$.

But, in the end, thinking of an element of an arbitrary product $\prod_{\alpha \in I} X_\alpha$ as a tuple $(x_\alpha)_{\alpha \in}$ of elements of the various X_α's indexed by elements of I will do us no harm, and is what we'll usually do. The box topology on this product is then the one generated by the basis consisting of sets of the form

$$\prod_\alpha U_\alpha \text{ where } U_\alpha \text{ is open in } X_\alpha,$$

and the product topology is generated by similar things only with the additional stipulation that

$$U_\alpha = X_\alpha \text{ for all but finitely many } \alpha,$$

so that you can't have infinitely many of the U_α be unequal to the corresponding X_α. Concretely, in the $\mathbb{R}^\mathbb{R}$ case, thinking of this set as the set of functions $f : \mathbb{R} \to \mathbb{R}$, a basic open set in the product topology is one consisting of functions such that for some fixed finitely many open subsets $U_{x_1}, \ldots, U_{x_n}$ of $\mathbb{R}$, we require that

$$f(x_i) \in U_{x_i} \text{ for each } i = 1, \ldots, n$$

with no additional constraints on the values of f at points that aren't among $x_1, \ldots, x_n$. (So, for some finite numbers of points, f should send these points into some specified open sets.) Here, the values $f(x)$ of f are the "components" of the tuple $(f(x))_{x \in \mathbb{R}}$, which is what leads to the realization that convergence in the product topology is the same as pointwise convergence in this setting; i.e. here "pointwise" means "componentwise".

Function Spaces. We'll consider more general infinite products next time, but for now we consider the space $\mathbb{R}^\mathbb{R}$, which we think of as being the product of "$\mathbb{R}$-many" copies of $\mathbb{R}$. An element of this space consists of a collection of real numbers indexed by the real numbers themselves:

$$(x_\alpha)_{\alpha \in \mathbb{R}} \in \mathbb{R}^\mathbb{R},$$

so in particular each such element consists of uncountably many real numbers. (As opposed to an element of $\mathbb{R}^\omega$, which consists of countably many real numbers.) The question is how to make this notion of "a collection of real numbers indexed by real numbers" precise. The key is to rephrase this concept in terms of another we're more familiar with, namely that of a *function*. To specify an element $(x_\alpha)_{\alpha\in\mathbb{R}}$ of $\mathbb{R}^\mathbb{R}$ intuitively as above requires that we associate to each real number α (the index) a real number x_α (the term occurring at the given index), but such an association precisely describes a function from $\mathbb{R}$ to $\mathbb{R}$. Indeed, such a function gives for each $\alpha \in \mathbb{R}$ a number $f(\alpha)$, which we interpret as the term $x_\alpha = f(\alpha)$ occurring at index α. Thus, we can make the uncountably infinite product $\mathbb{R}^\mathbb{R}$ precise by *defining* it to be the set of all functions from $\mathbb{R} \to \mathbb{R}$:

$$\mathbb{R}^\mathbb{R} = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is a function}\}.$$

More generally, Y^X denotes the set of functions from X to Y, which we can thus think of as the product of "X-many" copies of Y; an element of Y^X can be thought of as a collection $(y_\alpha)_{\alpha\in X}$ of elements y_α of Y indexed by elements α of X, which can be more precisely viewed as defining the function from X to Y which associates to $\alpha \in X$ the element $x_\alpha \in Y$. Thus, once we define the product topology on infinite products in general next time, we'll immediately have a topology we can put on a set of functions. For instance, the product topology on $\mathbb{R}^\mathbb{R}$ will be one in which the notion of convergence corresponds to what is normally called *pointwise convergence* of a sequence of functions: a sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ defines a sequence $(f_n(\alpha))_{\alpha\in\mathbb{R}}$ in $\mathbb{R}^\mathbb{R}$, and convergence of this latter sequence in the product topology on $\mathbb{R}^\mathbb{R}$ corresponds precisely to pointwise convergence of f_n. If you haven't seen pointwise convergence before, here is the definition: to say that a sequence of functions f_n converges pointwise to the function f means that for each $\alpha \in \mathbb{R}$, the sequence of real numbers $f_n(\alpha)$ (with n varying) converges to the real number $f(\alpha)$.

Why do we care about topologies? Up until this point in the course we've given many examples of topologies, but so far they might have seemed esoteric or constructed only to illustrate a certain property and not really things which would show up in "practice". I hope that at least it might be clearer why we should care about metric spaces in general (metric spaces are essentially the types of spaces where analysis takes place), but we have not given a reason why we should care about topological spaces which aren't metrizable yet.

Here is the example which first convinced me as an undergrad why we should care about such things: there is no metric on the set of functions from $\mathbb{R}$ to $\mathbb{R}$ with respect to which convergence means the same thing as pointwise convergence, but now we're saying that there is a *topology* on this set of functions relative to which this is true, namely the product topology. Indeed, much of the practical uses of topology in other areas of mathematics come from wanting "good" topologies on sets of functions, where what counts as "good" depends on what application you have in mind. We'll look at various other examples of such topologies later on.

Definition. closure, interior

Characterization of elements in closure.

Closure example. The Warm-Up showed that $\mathbb{R}^\infty$ was closed in $\mathbb{R}^\omega$ under the box topology, so the closure of $\mathbb{R}^\infty$ in $\mathbb{R}^\omega$ under the box topology is $\mathbb{R}^\infty$ itself. We left the question as to what this closure should be under the product topology unanswered, but we'll come back to this next time.

Lecture 8: Hausdorff Spaces

Warm-Up 1. We claim that the closure of $\mathbb{R}^\infty$ in $\mathbb{R}^\omega$ under the product topology is all of $\mathbb{R}^\omega$. Indeed, this is essentially what we proved in the Warm-Up last time when showing that $\mathbb{R}^\infty$ was not closed in $\mathbb{R}^\omega$ under the product topology, only we didn't phrase it at the time in terms of closures. Let $\mathbf{y} \in \mathbb{R}^\omega$ and let

$$U_1 \times U_2 \times \cdots$$

be a basic neighborhood of $\mathbf{y}$ with respect to the product topology. Since this is open in the product topology, there exists N such that $U_i = \mathbb{R}$ for $i \geq N$. But now define $\mathbf{x} \in \mathbb{R}^\omega$ by taking any possible elements from $U_1, \ldots, U_{N-1}$ as the first $N-1$ components of $\mathbf{x}$ and setting

$$x_i = 0 \text{ for } i \geq N.$$

Then $\mathbf{x} \in U_1 \times U_2 \times \cdots$ and $\mathbf{x}$ is eventually 0, so $\mathbf{x} \in \mathbb{R}^\infty$. Hence any neighborhood of $\mathbf{y}$ contains an element of $\mathbb{R}^\infty$, so $\mathbf{y}$ is in the closure of $\mathbb{R}^\infty$ under the product topology.

Denseness. We record here the definition of "dense" since we introduced it earlier than the book does: a subset A of a space X is *dense* in X if $\overline{A} = X$. This is saying that any open subset whatsoever of X contains an element of A. The Warm-Up above shows that $\mathbb{R}^\infty$ is dense in $\mathbb{R}^\omega$ under the product topology; it is not dense under the box topology as the Warm-Up from last time now shows. The most common example of a dense subset is no doubt $\mathbb{Q}$ in $\mathbb{R}$, which plays an important role in analysis. We'll see later why denseness is important.

Warm-Up 2. Suppose A is a subset of X and B a subset of Y. We show that

$$\overline{A \times B} = \overline{A} \times \overline{B},$$

so that the closure of a product is the product of closures. (In fact, this is true for more general products as well, and the argument in general is very similar to the argument we'll give here.) First, note that the set on the right contains $A \times B$ and is closed in $X \times Y$ since it is the product of closed sets. Hence

$$\overline{A \times B} \subseteq \overline{A} \times \overline{B}$$

simply because $\overline{A} \times \overline{B}$ is the one of the things being intersected when constructing $\overline{A \times B}$.

Now, let $(p, q) \in \overline{A} \times \overline{B}$. Then $p \in \overline{A}$ and $q \in \overline{B}$. Let $U \times V$ be a basic neighborhood of (p, q). Since U is a neighborhood of p and p belongs to the closure of A, U contains an element of A, say $a \in A$. Similarly, V is a neighborhood of q and q belongs to the closure of B, so V contains an element b of B. Thus (a, b) is an element of $A \times B$ contained in $U \times V$, so every neighborhood of (p, q) intersects $A \times B$, meaning that $(p, q) \in \overline{A \times B}$. Hence $\overline{A} \times \overline{B} \subseteq \overline{A \times B}$, so we conclude that $\overline{A \times B} = \overline{A} \times \overline{B}$ as claimed.

Definition. defn of Hausdorff

Uniqueness of limits.

Definition.

Cofinite is coarsest T_1.

Example. T_1 but not Hausdorff

T_1 **but non-unique limits.** We started with the following example. The cofinite topology on $\mathbb{R}$ gives an example of a T_1-space which is not Hausdorff: it is T_1 since given $x \neq y$, $\mathbb{R} - \{x\}$ is a neighborhood of y which contains x (implying that $\{x\}$ is closed), and it is not Hausdorff since any two nonempty open sets intersect in infinitely many points because open sets can only exclude finitely many points. The observation is that the sequence

$$1,\ 2,\ 3,\ 4,\ 5,\ \ldots$$

in this space converges to *every* $x \in \mathbb{R}$, so limits of sequences in a T_1-space need not be unique. To see that any possible $x \in \mathbb{R}$ can serve as a limit of this sequence, fix $x \in \mathbb{R}$ and consider any neighborhood U of x in the cofinite topology. Then U is not empty and only excludes finitely many points of $\mathbb{R}$, so it can only exclude finitely many terms from the given sequence. Thus for n large enough (i.e. past some index), $n \in U$, showing that the given sequence converges to x.

Separation axioms. To put the notion of a T_1-space (i.e. a space satisfying the T_1-axiom) and that of a Hausdorff space into the right context, we note that we will eventually consider other so-called *separation axioms*, which describe the extent to which objects in a space can be "separated" from one another. The Hausdorff axiom is also known as the T_2-axiom, and later we will discuss the T_3-axiom (what it means for a space to be *regular*) and the T_4-axiom (what it means for a space to be *normal*). We might also talk about the "$T_{3\frac{1}{2}}$"-axiom! The T_1-condition says that "any point can be separated from any other point" and the Hausdorff (T_2) condition says that "points can be separated from one another".

Lecture 9: Continuous Functions

Warm-Up. We claim that $\mathbb{R}^2_{Zar}$ (i.e. $\mathbb{R}^2$ with the Zariski topology) is not Hausdorff. In fact, we show that any two nonempty open sets must always intersect, so the types of disjoint open sets required in the Hausdorff condition cannot exist. Since any open set contains a basic open set of the form

$$D(f) = \mathbb{R}^2 - V(f),$$

where f is a single polynomial in two variables, it is enough to show that such basic open sets always intersect. Suppose f and g are nonzero (otherwise $D(f), D(g)$ are empty) polynomials and recall from Homework 1 that

$$D(f) \cap D(g) = D(fg).$$

Since f and g are not the zero polynomials, neither is fg. Hence there exists $(x, y) \in \mathbb{R}^2$ such that

$$f(x, y)g(x, y) \neq 0.$$

This point is then in $D(f) \cap D(g) = D(fg)$, so this intersection is not empty as required.

Motivating continuity. Recall that ϵ-δ definition of continuity for a function $f : \mathbb{R} \to \mathbb{R}$: f is continuous if for every $a \in \mathbb{R}$ and every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|x - a| < \delta \text{ implies } |f(x) - f(a)| < \epsilon.$$

Intuitively, this says "given a measure of how close we want to end up near $f(a)$, there exists a measure of how close we should get to a in order to guarantee we end up within the prescribed measure of closeness to $f(a)$". The point is that this definition can be phrased solely in terms of open sets. First, in terms of intervals we get the condition

$$x \in (a-\delta, a+\delta) \text{ implies } f(x) \in (f(a)-\epsilon, f(a)+\epsilon).$$

Second, in terms of preimages we get the condition

$$x \in (a-\delta, a+\delta) \text{ implies } x \in f^{-1}((f(a)-\epsilon, f(a)+\epsilon)).$$

Finally, in terms of subsets we get the condition

$$x \in (a-\delta, a+\delta) \subseteq f^{-1}((f(a)-\epsilon, f(a)+\epsilon).$$

Thus, the ϵ-δ definition says that given x in the preimage of $(f(a)-\epsilon, f(a)+\epsilon)$, there exists an open set around it which remains within the preimage. Since $(f(a)-\epsilon, f(a)+\epsilon)$ is open and any open subset of $\mathbb{R}$ is a union of such intervals, we get the condition that the preimage of any open subset of $\mathbb{R}$ is itself open in $\mathbb{R}$, which is the sought-after characterization of "continuous" in terms of open sets.

Thus it makes sense to define continuity in the setting of arbitrary topological spaces as the book does: $f : X \to Y$ is *continuous* if for every U open in Y, the preimage $f^{-1}(U)$ is open in X. (You can also find a topological definition of "continuous at a point" in the book, but the global notion of continuous without reference to a specific point is the one we'll find most useful.) To give some intuition in general behind this definition, compare again to the ϵ-δ definition: if we interpret an open set in an arbitrary topological space as providing its own measure of "closeness" (i.e. saying that $p, q \in U$ means that p and q are "near" each other "relative" to U), the definition of continuous indeed says that "given a measure U of how close we want to end up in Y, there exists a measure $f^{-1}(U)$ of how close we have to be in X in order to guarantee we end up within the prescribed measure of closeness U in Y".

Definition.

Product topology in terms of continuity.

Maps from discrete, or into trivial.

Maps into a discrete space. As some examples we looked at (and which are in the book) show, whether or not a function is continuous depends heavily on the topologies in question. To emphasize this, we asked the question as to which functions $\mathbb{R} \to \mathbb{R}_d$ were continuous, where the domain has the standard topology and where $\mathbb{R}_d$ denotes $\mathbb{R}$ with the discrete topology. First, any constant function is continuous. (The book proves a general version of fact.)

Now, suppose $f : \mathbb{R} \to \mathbb{R}_d$ is continuous and that $f(p), f(q)$ are distinct points in the image, so that f is not constant. Then $\{f(p)\}$ and $\mathbb{R} - \{f(p)\}$ are both open (since everything is open in the discrete topology), so

$$\mathbb{R}_d = \{f(p)\} \cup (\mathbb{R} - \{f(p)\})$$

expresses $\mathbb{R}_d$ as the union of disjoint nonempty (note that $f(q)$ is in the latter) sets. Since f is continuous, the preimages of each of these are open in $\mathbb{R}$, so

$$\mathbb{R} = f^{-1}(f(p)) \cup f^{-1}(\mathbb{R} - \{f(p)\})$$

expresses $\mathbb{R}$ as the union of two disjoint nonempty open sets. You may recall from an analysis course that this is not possible since $\mathbb{R}$ is *connected*. (We'll talk about connected spaces soon enough, so no worries if this is not a notion you recall all the details about.) Thus it is not possible to have two distinct points in the image of f, so f is constant and we conclude that the only continuous functions $\mathbb{R} \to \mathbb{R}_d$ are the constant ones. (In general, the only continuous functions *into* a discrete space are the "locally constant" ones, which is a notion we'll look at later. By contrast, *any* function *from* a discrete space into something else is always continuous.)

Jump discontinuities. Consider a map $f : \mathbb{R} \to \mathbb{R}$ with a "jump discontinuity". We pointed out that such a function indeed became continuous if we changed the topology on the domain to that of the lower limit topology. This, and the consideration of one-sided limits, is how the lower limit topology shows up in certain applications.

Equivalent characterizations of continuity. We gave the following equivalent formulations of continuity, which can also be found in the book: $f : X \to Y$ is continuous if and only if the preimage of any *closed* set in Y is closed in X, and also $f : X \to Y$ is continuous if and only if

$$f(\overline{A}) \subseteq \overline{f(A)} \text{ for any } A \subseteq X.$$

We proved the forward direction of this latter claim in class, and will prove the other direction next time. (Both of these are good exercises in getting accustomed to unwinding definitions.) For now, we point out that the characterization in terms of closures is the topological analog of the characterization of continuous given in terms of sequences you would have seen in analysis.

Indeed, in the case of $\mathbb{R}$, saying that $f(p) \in f(\overline{A})$ means $f(p)$ is obtained by applying f to the limit p of a sequence p_n in A; if it is true that $f(p) \in \overline{f(A)}$ as the closure-characterization would imply, then $f(p)$ should also be the limit of the sequence $f(p_n)$ in A, so f "sends convergent sequences to convergent sequences", agreeing with the sequential definition of continuity given in analysis. To say it another way, continuous means that points which are "arbitrarily" close to A are sent to points which are "arbitrarily" close to $f(A)$.

The issue is that in the general topological setting we must phrase this in terms of closures instead of sequences since there may not be "enough" sequences available to accurately capture continuity; to be precise, is it NOT true in general that a function $f : X \to Y$ with the property that $f(p_n) \to f(p)$ in Y whenever $p_n \to p$ in X must be continuous. This is one of the instances in which thinking about topological concepts solely in terms of sequences is not enough—the notion of compactness will give us another such instance. (There is a generalization of the notion of a sequence known as a *net*, and a corresponding notion of convergence for nets. In that setting it is true that a function is continuous if and only if it sends convergent nets to convergent nets, but this is not something we'll explore in this course.)

Lecture 10: More on Continuity

Warm-Up. A map $Y \to \prod X_\alpha$ is given by a collection of maps $Y \to X_\alpha$, one for each α. We proved as a Warm-Up that $Y \to \prod X_\alpha$ is continuous with respect to the product topology if and only if each component map $Y \to X_\alpha$ is continuous. This is proved in the book, and provides another characterization of the product topology: the product topology is the finest one relative to which this is true.

Restrictions and extensions. Given a continuous functions $f : X \to Y$, restricting the domain to a subset A of X still gives a continuous functions (this restriction is usually denoted by $f|_A : A \to Y$), and restricting the codomain to a smaller subset of Y which still contains the image of X also gives a continuous functions. So, restrictions never alter continuity. Similarly, *extending* the codomain does not alter continuity, meaning that if $f : X \to Y$ is continuous where $Y \subseteq Z$ has the subspace topology, then $f : X \to Z$ is still continuous.

The question as to when a given continuous function can be *extended* to one on a larger domain is subtle, and is one we'll come back to later on. Extensions are not always possible, but it turns out that it will be possible under some mild topological assumptions; this is the content of *Urysohn's lemma* and the *Tietze extension theorem*.

Homeomorphisms. We finished with defining the notion of a homeomorphism, and gave an example of a continuous bijection whose inverse was not continuous, which explains why we need to assume both a function and its inverse are continuous in the definition of homeomorphism. This example can be found in the book as well. This is different than other types of "isomorphisms" you night have seen in an abstract algebra of linear algebra course, where the inverse of a group homomorphism in the former case is automatically a homomorphism, and the inverse of a linear transformation in the latter case is automatically linear.

Examples. spheres, ellipsoids, square, circle

Cantor space.

Lecture 11: Quotient Spaces

Warm-Up 1. Suppose $f, g : X \to Y$ are continuous and agree on a dense subset A of X, meaning that the restrictions $f|_A$ and $g|_A$ are equal. (Recall that A being dense in X means that $\overline{A} = X$.) If Y is Hausdorff, we show that $f = g$ on all of X, so that continuous functions into Hausdorff spaces are completely determined by their behavior on a dense subset of the domain. This is something you likely saw in an analysis course, where continuous functions $\mathbb{R} \to \mathbb{R}$ are determined by their action on $\mathbb{Q}$ for instance, but in the general topological setting we need the codomain to be Hausdorff. For an example of where this doesn't work if the codomain isn't Hausdorff, let L denote the line with two origins and consider the functions $f, g : \mathbb{R} \to L$ defined by $f(x) = g(x) = x$ for $x \neq 0$ but with $f(0)$ being one origin in L and $g(0)$ the other; these two functions are continuous and agree on the dense subset $\mathbb{R} - \{0\}$ of $\mathbb{R}$, but are not the same on all of $\mathbb{R}$.

Suppose $p \in X$. We want to show that $f(p) = g(p)$. If instead $f(p) \neq g(p)$, we can find disjoint open sets U and V of Y which separate them since Y is Hausdorff. Then $f^{-1}(U)$ and $g^{-1}(V)$ are both open in X, so $f^{-1}(U) \cap g^{-1}(V)$ is open as well. This intersection thus contains an element $a \in A$ since A is dense in X. But this gives

$$f(a) \in U \text{ and } g(a) \in V,$$

which, since $f(a) = g(a)$ because $f|_A = g|_A$, contradicts the fact that U and V were supposed to be disjoint. Hence $f(p) = g(p)$ as claimed, so f and g agree on all of X.

Warm-Up 2. Cantor space

Spaces obtained by gluing. For our purposes, thinking of an *equivalence relation* on a space X as a way of specifying which elements should be thought of as being the "same" will be good enough, meaning we won't need to recall the formal definition of an equivalence relation as a relation which is reflexive, symmetric, and transitive. Given an equivalence relation $\sim$ on X, the *quotient space* $X/{\sim}$ is the set of equivalence classes, where the equivalence class containing $p \in X$ is by definition of the set of all elements of X which are equivalent to p; this quotient space is, intuitively, the space obtained after gluing elements in an equivalence class to one another. We are interested in putting a natural topology on this quotient space which reflects this intuitive "gluing" idea.

Example. Consider the equivalence relation on $\mathbb{R}$ defined by saying $x \sim y$ if $x - y \in \mathbb{Z}$. Thus, x and y are equivalent if and only if they have the same "decimal part". For instance, all integers are equivalent to one another, 2.32345 if equivalent to 0.32345 and -7.32345, and so on. The upshot is that any element of $\mathbb{R}$ is equivalent to a unique element of $[0, 1)$, so that we can think of the quotient as being this interval, only that we should consider the endpoints 0 and 1 to be the "same" since they belong to the same equivalence class. Thus, after gluing, we again should get a circle.

Quotient topology. Finally we define the quotient topology on $X/\sim$. If there is any justice in the world this should be a topology which makes the obvious map

$$\pi : X \to X/\sim,$$

sending a point to the equivalence class containing it, continuous. We *define* the quotient topology on $X/\sim$ to be the finest topology we can put on $X/\sim$ to make this true. Concretely, a subset U of $X/\sim$ is open in the quotient topology if and only if its preimage $\pi^{-1}(U)$ under the *quotient map* π is open in X. This preimage concretely is the union of all equivalence classes contained in U.

Example 2. Take X to be the union of the lines $y = 0$ and $y = 1$ in $\mathbb{R}^2$ equipped with the subspace topology and define an equivalence relation on X by saying $(x, 0) \sim (x, 1)$ for $x \neq 0$. The quotient space $X/\sim$ is the line with two origins. Indeed, as a set this quotient is just a line only with $(0, 0)$ and $(0, 1)$ representing different points since these were not declared to be equivalent. If (a, b) in this quotient does not contain 0, its preimage under the quotient map is the union of the corresponding intervals on the lines $y = 0$ and $y = 1$, which is open in X. If (a, b) contains one origin, its preimage is an open on one of the lines $y = 0$ or $y = 1$, which is still open. Thus such subsets of $X/\sim$ are open, which gives the topology one the line with two origins we've described previously.

Lecture 12: More on Quotients

Warm-Up. Consider the quotient space obtained from $\mathbb{R}$ by declaring all integers to be equivalent to one another. Give $\mathbb{R}/\sim$ the quotient topology and let $\pi : \mathbb{R} \to \mathbb{R}/\sim$ denote the natural map sending a point to its equivalence class. We show that this map is not open, meaning that it does not send open sets to open sets. Concretely, we can visualize $\mathbb{R}/\sim$ as a "bouquet" of countably many circles; indeed, this quotient is obtained by gluing the endpoints of each interval $[n, n + 1]$ to get a circle, and then gluing all of these circles together at a common point corresponding to all integers.

Recall that the quotient topology on $\mathbb{R}/\sim$ is defined by declaring $U \subseteq \mathbb{R}/\sim$ to be open if and only if $\pi^{-1}(U)$ is open in $\mathbb{R}$. We claim the the image of $(-1/2, 1/2)$ under π is not open in $\mathbb{R}/\sim$. To see that this image $\pi((-1/2, 1/2))$ is not open, we determine its preimage under π. This preimage consists of the same interval $(-1/2, 1/2)$, but also all things which get mapped to the same thing as 0, meaning all integers since all integers map to the same thing under π. Thus

$$\pi^{-1}(\pi((-1/2, 1/2))) = (-1/2, 1/2) \cup \mathbb{Z},$$

which is not open in $\mathbb{R}$. By definition of the quotient topology, this means that $\pi((-1/2, 1/2))$ is not open in $\mathbb{R}/\sim$, so π is not an open map.

Hawaiian Earring. compare with Hawaiian earring

Real Projective Line. The *real projective line* is the set $\mathbb{R}P^1$ of lines in $\mathbb{R}^2$ which pass through the origin. (The idea is that in "projective geometry" we replace "points" by "directions".) Let $p : S^1 \to \mathbb{R}P^1$ (where S^1 is the unit circle in $\mathbb{R}^2$) be the map which sends a point on S^1 to the line passing through it and the origin. Then p is surjective, and we give $\mathbb{R}P^1$ the resulting quotient topology.

We claim that under this quotient topology the map $p : S^1 \to \mathbb{R}P^1$ is open. Indeed, suppose U is open in S^1. (So, U is the intersection of S^1 with an open subset of $\mathbb{R}^2$.) Then $p(U) \subseteq \mathbb{R}P^1$ consists of all lines in $\mathbb{R}^2$ passing through the origin and an element of U. To see that this is open we must consider $p^{-1}(p(U))$. The map $p : S^1 \to \mathbb{R}P^1$ is 2-to-1, where a point $p \in S^1$ and its corresponding *antipodal* point $a(p)$ ($a : S^1 \to S^1$ is the map $(x, y) \mapsto (-x, -y)$) get sent to the same thing, so we get that

$$p^{-1}(p(U)) = U \cup a(U),$$

which is a union of open sets in S^1. Hence this preimage is open, so $p(U)$ is open in $\mathbb{R}P^1$ by definition of the quotient topology. Thus p is an open map.

Also Hausdorff and homeomorphic to S^1.

Other Projective Lines.

Lecture 13: Connected Spaces

Warm-Up. Suppose $f : X \to Y$ is continuous. Restricting to the image gives a continuous surjective map $f : X \to f(X)$, and we can thus consider the quotient topology on $f(X)$. On the other hand, $f(X) \subseteq Y$ can be given the subspace topology. We are interested in how these topologies on $f(X)$ relate to one another.

We claim that the subspace topology is coarser than the quotient topology. Indeed, if $U \subseteq f(X)$ is open in the subspace topology, then $f^{-1}(U)$ is open in X since $f : X \to f(X)$ is continuous. But saying that $f^{-1}(U)$ is open in X is precisely what it means for U to be open in $f(X)$ under the quotient topology, which shows that the subspace topology is coarser than the quotient topology.

In general, the quotient topology is not coarser than the subspace topology. For instance, take $f : [0, 1) \cup (1, 2] \to \mathbb{R}$ to be the map defined by

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ -x + 3 & 1 < x \leq 2. \end{cases}$$

This is continuous and has image $[0, 2)$. The set $[1, 2)$ is not open in the image under the subspace topology, but its preimage under f is $(1, 2]$, which is open in $[0, 1) \cup (1, 2]$, meaning that $[1, 2)$ is open in $[0, 2)$ under the subspace topology.

Fun example. As described on the homework, the quotient of the unit square $[0, 1] \times [0, 1]$ under the equivalence relation where we identify $(x, 0)$ with $(1 - x, 1)$ and $(0, y)$ with $(1, 1 - y)$ is homeomorphic to the *real projective plane*, which is the space of lines through the origin in $\mathbb{R}^3$ equipped with the quotient topology arising from the map $S^2 \to \mathbb{R}P^2$ sending a point on the unit sphere to the line passing through it and the origin. We can take another quotient of the unit square by identifying $(x, 0)$ with $(1 - x, 1)$ and $(0, y)$ with $(1, y)$. (Visually the difference is that in this new quotient we only twist one edge when gluing as opposed to two edges as in the case of $\mathbb{R}P^2$.) The resulting quotient is known as the *Klein bottle*.

We claim that the Klein bottle really is a new space, in the sense that it is not homeomorphic to $\mathbb{R}P^2$. Up to this point we don't have good ways of showing that spaces aren't homeomorphic, apart from being to identify some specific property one space has (T_1, Hausdorff, etc.) that the other one doesn't. The point of introducing this example now is to give a brief glimpse into the subject of *algebraic topology*, which gives us new ways of studying spaces using algebra. Consider a *triangulation* of $\mathbb{R}P^2$, which is, as the name suggests, a way of breaking $\mathbb{R}P^2$ up into a collection of triangular regions. Under a possible homeomorphism between $\mathbb{R}P^2$ and the Klein bottle, this triangulation would get sent to a triangulation of the Klein bottle, and the number of triangles, edges, and vertices in such a triangulation would be preserved. The problem is that in $\mathbb{R}P^2$, it turns out that taking

$$\#(\text{vertices}) - \#(\text{edges}) + \#(\text{triangles})$$

always gives the value 1, whereas in the Klein bottle it gives the value 0, and a homeomorphism would in fact have to preserve this value. Thus $\mathbb{R}P^2$ and the Klein bottle cannot be homeomorphic. The value described above is known as the *Euler characteristic* of a space, and is a concept which would be defined more precisely in a course in algebraic topology. The spring quarter of this course would touch on this a bit.

Definition. disconnected, connected

Examples. first examples

Union of connected sets. union of connected with point in common

Finite products of connected sets.

Example. $\mathbb{R}^\omega$ in product topology (using closure properties)

Example. $\mathbb{R}^\omega$ in box topology

Lecture 14: More on Connectedness

Warm-Up 1. intervals connected, $\mathbb{R}$

Warm-Up 2. closure of connected

$\mathbb{R}^n$ **for different** n. Using the fact that continuous maps send connected sets to connected sets, we can show that $\mathbb{R}$ is not homeomorphic to $\mathbb{R}^n$ for $n > 1$. Suppose $f : \mathbb{R} \to \mathbb{R}^n$ was a homeomorphism. This would then give a homeomorphism $\mathbb{R} - \{0\} \to \mathbb{R}^n - \{f(0)\}$, which is not possible since the inverse of this would have to send the connected space $\mathbb{R}^n - \{f(0)\}$ to the disconnected space $\mathbb{R} - \{0\}$.

The same argument does not work for showing that higher dimensional Euclidean spaces are not homeomorphism to others of different dimensions. In the case of $\mathbb{R}^2$ vs $\mathbb{R}^3$ you could try to look at a similar argument where you remove a line from $\mathbb{R}^2$ instead of a single point, but the problem is that it is in fact possible for a line to be sent under a homeomorphism to a 2-dimensional region, so the same trick does not work here. Indeed, showing that $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}^m$ for $m \neq n$ in general is a much harder problem and requires deeper techniques; you'll see one using the notion of *homology* in the spring quarter.

More examples. The line with two origins Y is connected. Indeed, suppose $Y = U \cup V$ were a valid separation. Since any open set containing one origin intersects any open set containing the other origin, it must be that both origins belong to U or both belong to V. This implies that U and V (or rather their analogs in $\mathbb{R}$) would then give a valid separation of $\mathbb{R}$ with the standard topology, which is not possible since $\mathbb{R}$ is connected. (The fact that the line with two origins is connected also follows from the fact that it is path connected, which is a notion we'll look at next time.)

The space $GL_n(\mathbb{R})$ is invertible $n \times n$ matrices is disconnected. Indeed, the subsets GL_n^+ and GL_n^- of matrices with positive and negative determinant respectively form a separation. Note that these sets are open since they are preimages of $(0, \infty)$ and $(-\infty, 0)$ respectively under the map $GL_n(\mathbb{R}) \to \mathbb{R}$ sending a matrix to its determinant.

Image of connected under continuous. include examples of $\mathbb{R}P^n$

Topologist's Sine Curve.

Definition. path connected, path connected implies connected

Examples. The line with two origins Y is path connected. Indeed, for points p and q, at least one of which is not an origin, the same type of line segment which connects them in $\mathbb{R}$ will still connect them in Y. To connect one origin to the other, we can take a segment which starts at one origin and moves to the right, and then moves back left only ending at the other origin.

Lecture 15: Local Connectedness

Warm-Up. We showed that the topologist's sine curve $\overline{S}$ is not path connected. This is in the book, although we gave a slightly different argument. Suppose $\gamma : [a, b] \to \overline{S}$ is a continuous path connecting $(0, 0)$ to $(1, sin 1)$. As in the book, we may assume that $\gamma(t)$ has positive x-coordinate for $a < t$, so that $\gamma(a) = (0, 0)$ is the only point on this curve which is on the y-axis. For any basic neighborhood $[a, \epsilon)$ around a in $[a, b]$, its image under the composition $\pi_1 \circ \gamma$, where π_1 is projection onto the x-coordinate, is a connected subset of the x-axis since continuous functions send connected sets to connected sets. Thus this image must be an interval $[0, d)$, meaning that all points in $[0, d)$ arise as x-coordinates of points along γ. This implies that γ cannot be continuous: for any small open ball (say of radius $1/2$) around the origin, there is no open neighborhood $[a, \epsilon)$ around a which remains in the preimage since there is always a value in such a neighborhood which maps to a point with y-coordinate equal to 1, which thus falls outside the given open ball. Hence there is no continuous path connecting $(0, 0)$ to $(1, \sin 1)$.

Components. The book defines the notion of a (connected) component in terms of an equivalence relation. Here is an alternate definition: a *connected component* of a space X is a maximally connected subset, meaning a connected subset C such that if S is any connected subset of X containing C, then $S = C$. In other words, a connected component is a connected subset which is not contained in any larger connected subset. The equivalence between this definition and the book's definition comes from the fact that, in the book's definition, connected components are always disjoint. This property also follows from our definition: if C_1 and C_2 are two components which are not disjoint, then $C_1 \cup C_2$ is connected as well, so that C_1 and C_2 would not have been maximally connected.

Examples. The components are $\mathbb{Q}$ are the singleton sets. Indeed, if S is a subset of $\mathbb{Q}$ with at least two elements $p < q$, pick an irrational x such that $p < x < q$. Then

$$S = [S \cap (-\infty, x)] \cup [S \cap (x, \infty)]$$

is a separation of S, so that S is not connected. Hence no subset of $\mathbb{Q}$ with more than one element is connected, but one element sets are certainly connected. This means that $\mathbb{Q}$ is what's called *totally disconnected*, meaning precisely that the only connected subsets are singletons.

The space $GL_n(\mathbb{R})$ of invertible $n \times n$ matrices has two components: the subset of matrices with positive determinant and the subset of matrices with positive determinant. Showing that these two subsets are indeed connected takes a bit of work and requires some linear algebra, so we'll skip the proof here. But here is another important observation, which applies to other "groups" (in the sense of abstract algebra) of matrices as well: the connected component of the identity matrix is precisely the set of matrices which can be written as products of exponentials of other matrices. This fact and its generalization to other groups is a crucial fact in various applications of matrix group to geometry and physics.

Locally connected spaces. The definition of what it means for a space to be *locally connected* (or locally path connected) can be found in the book. Here we just give a succinct way of stating this definition using the notion of a "local basis", which is a concept will see coming up a few times going forward. A *local basis* at $x \in X$ is a collection $\{U_\alpha\}$ of neighborhoods of x such that for any other neighborhood V of x, there exists U_α contained in V. In a sense, the sets in a local basis at x "generate" all other neighborhoods of x. Then, we can say that X is locally connected if each point has a local basis of connected neighborhoods. Intuitively, a space is locally connected if it appears connected when zooming in closely enough on any given point.

Example. As a final example, $\mathbb{R}_\ell$ is totally disconnected. For any subset S with two elements $x < y$,

$$S = [S \cap (-\infty, y)] \cup [S \cap [y, \infty)]$$

is a separation of S, so S is not connected. Hence only singleton sets are connected in $\mathbb{R}_\ell$. This then implies that $\mathbb{R}_\ell$ is nowhere locally connected, since no neighborhood of any point can be connected.

Lecture 16: Compact spaces

Warm-Up 1. A map $f : X \to Y$ is said to be *locally constant* if any point of X has a neighborhood on which f is constant. We claim that if $f : X \to Y$ is locally constant, then f is actually constant on each component of X. Let C be a component of X and fix $p \in C$. Let S be the subset of C consisting of all $q \in C$ for which $f(q) = f(p)$. First, if $s \in S$, pick a neighborhood U of s on which f is constant. For any $x \in U$, we then have $f(x) = f(s) = f(p)$, so $x \in S$. Hence $s \in U \subseteq S$, showing that S is open in C. Similarly, if $c \in C - S$, pick a neighborhood V of c on which f is constant. Then $f(x) = f(c) \neq f(p)$ for any $x \in V$, so $V \subset C - S$ and hence $C - S$ is open, so S is closed in C. Thus S is clopen in C, so $S = C$ since C is connected, showing that f is constant on C as claimed.

The converse of the result above holds when X is locally connected, which follows from the fact that components in a locally connected space are actually open. For an example showing the converse fails when X is not locally connected, consider the identity map $\mathbb{Q} \to \mathbb{Q}$ where $\mathbb{Q}$ has the standard topology. This map is constant on each component since each component only contains a single point, but it is not locally constant since it is not constant on any $(a, b) \cap \mathbb{Q}$.

Warm-Up 2. We give an example of a surjective continuous map from a locally connected space to one which is not locally connected, which shows that "local connectedness" is not preserved by continuity. The identity map $\mathbb{Q} \to \mathbb{Q}$, where the domain has the discrete topology and the codomain the standard topology, works. Indeed, $\mathbb{Q}$ is locally connected in the discrete topology since for any $r \in \mathbb{Q}$, $\{r\}$ is itself a connected neighborhood of r, but $\mathbb{Q}$ is not locally connected in the standard topology as explained at the end of the previous Warm-Up.

Further topics. Just to illustrate how some of these definitions show up in practice, we briefly introduce the idea of a *universal cover*. A *covering space* of X is a space C with a continuous surjection $p : C \to X$ such that every $p \in X$ has a neighborhood U for which $p^{-1}(U)$ is a disjoint union of open sets in C which are each mapped homeomorphically onto U by p. The idea is that U is "covered" by multiple copies of itself up in the covering space C. A universal cover of X is a covering space from which, in a sense we won't define, all other covers can be derived. It turns out that in order to guarantee a universal cover exists we must assume X has various levels of connectedness, for instance that it is connected, locally path connected, and what's called "semi-locally simply connected". These are concepts you will learn about in the spring quarter of topology, where you'll see that covering spaces are fundamental tools in algebraic topology.

Compactness. The definition of compact can be found in the book, as can all properties we saw: closed subspaces of compact spaces are compact, continuous images of compact spaces are compact, and that compact subsets of Hausdorff spaces are closed. Note that this final property is not necessarily true without the Hausdorff condition: the set $[-1, 1]$ containing one of the origins is compact in the line with two origins but not closed, since its complement contains the singleton non-open set containing the other origin. As the book states in a lemma, the real takeaway in the proof that compact subsets of Hausdorff spaces are closed is the result that compact sets and points in Hausdorff spaces can be *separated by open sets*, meaning that for any compact K and $x \notin K$, there exists disjoint open sets U and V containing K and x respectively. We'll see other types of a "separation properties" soon.

The intuition is that compactness allows one to replace an infinite amount of data with a finite amount of data; in a vague sense, compactness is an infinite analog of finiteness. The proof that closed intervals $[a, b]$ in $\mathbb{R}$ are compact is in the solutions to the Discussion 5 Problems. Note that this proof using only open covers, and not sequences; in general topological spaces, sequences are not enough to characterize compactness.

Lecture 17: More on Compactness

Warm-Up 1. We showed that a continuous bijection from a compact space to a Hausdorff space is always a homeomorphism, a result which can be found in the book. Note that this doesn't require assuming the domain if Hausdorff nor that the codomain is compact ahead of time, but both of these facts are consequences. This gives at least one instance in which we don't have to think about whether an inverse is continuous separately.

Warm-Up 2. The *Cantor set* (or to be precise, the standard middle-thirds Cantor set), is the subset C of $\mathbb{R}$ defined as follows. Set $C_0 = [0, 1]$, then

$$C_1 = [0, 1/3] \cup [2/3, 1], \;\; C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1],$$

and in general C_n is obtained by removing from each interval making up C_{n-1} its middle third portion. Then $C = \bigcap_n C_n$ is the Cantor set.

We claim that with the subspace topology inherited from $\mathbb{R}$, C is homeomorphic to the product space $\{0,2\}^\omega$, where $\{0,2\}$ has the discrete topology. The key point is that elements of C can be also be characterized as those elements of $[0,1]$ whose base-3 decimal expansions only consist of 0's and 2's, a fact we will take for granted. (This is why I'm using $\{0,2\}^\omega$ instead of $\{0,1\}^\omega$; of course, C is also then homeomorphic to $\{0,1\}^\omega$.) The 0's and 2's in the base-3 expansion of an element of C then tell you whether to move to the left or to the right at each step in the Cantor set construction: a 0 in the n-th decimal location means to take the interval in C_{n-1} containing the given element and then go into the left interval obtained after removing the middle third, and a 2 means to go into the right interval, which in the end describes which interval among those making up C_n the given element is in.

The map $C \to \{0,2\}^\omega$ defined by

$$0.x_1x_2x_3\ldots \mapsto (x_1, x_2, x_3, \ldots)$$

is then the required homeomorphism. This is clearly surjective, and since C is compact (it is a closed subset of the compact set $[0,1]$) and $\{0,2\}^\omega$ is Hausdorff, showing that it is continuous is enough to show that it is a homeomorphism by the first Warm-Up. To see that it is continuous take a basic nonempty open subset

$$U_1 \times U_2 \times \cdots \times U_n \times \{0,2\} \times \{0,2\} \times \cdots$$

of $\{0,2\}^\omega$ in the product topology. Then each U_i is either $\{0\}, \{2\}$, or $\{0,2\}$. The preimage of this consists of the elements of C contained in C_n (there is no restriction on the decimal digits after the n-th one since the sets in the product above are $\{0,2\}$ after the n-th term) belonging to those intervals determined by moving left and right in the manner described above; if $U_i = \{0\}$ you move left, if it is $\{2\}$ you move right, and if it is $\{0,2\}$ you consider both possibilities. Hence this preimage is just C intersect a union of some (or all) of the closed intervals making up C_n, and each of such intersections are open in C since these closed intervals can all be surrounded by an open interval which intersects none of the other closed intervals. Thus this preimage is open, so the given map is indeed continuous.

Heine-Borel. Using the fact finite products of compact spaces are compact, we gave a proof of the Heine-Borel Theorem, which says that a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded. For the forward direction, if $K \subseteq \mathbb{R}^n$ is compact, it is certainly closed and can be covered by finitely many balls centered at $\mathbf{0}$ (since the open balls of radius n, with n varying, centered at $\mathbf{0}$ form an open cover), which implies that K is bounded.

Conversely, suppose $K \subseteq \mathbb{R}^n$ is closed and bounded. Since it is bounded, it is contained in some box

$$[a_1, b_1] \times \cdots \times [a_n, b_n].$$

Such a box is the product of compact sets, so it is compact itself, and thus K, being a closed subset of this compact set, is compact as claimed.

Variations on compactness. We finished by commenting on some variations of the definition of compactness. A space is *Lindelöf* if every open cover of it has a countable subcover. (Thus, Lindelöf spaces are ones where one can replace an uncountable amount of data with a countable amount.) Any compact space is Lindelöf, as is $\mathbb{R}^n$. We'll briefly touch on Lindelöf spaces later when discussing the countability axioms.

The other variation on compactness worth mentioning here is the notion of a space being *paracompact*. The precise definition is a little technical, but *essentially* it says that given any open cover, any point has a neighborhood which intersects only finitely many of the sets in that open cover, or in other words, any point has a neighborhood which is covered by finitely many elements of the given open cover. Thus, paracompactness is a type of local variation of compactness. One of the most important consequences of a space being paracompact is the existence of "partitions of unity", which provide a key tool in various aspects of geometry and analysis. We won't look at such things in this course, but you can check later sections in the book for more details.

Lecture 18: Local Compactness

Warm-Up. We say that a function $f : X \to \mathbb{R}$ is *locally bounded* if every point of x has a neighborhood on which f is bounded. We claim that if X is compact, any locally bounded function is actually bounded. Indeed, for $p \in X$ let U_p be a neighborhood on which f is bounded, so there exists $M_p > 0$ such that $|f(x)| \leq M_p$ for all $x \in U_p$. The sets $\{U_p\}_{p \in X}$ form an open cover of X, so since X is compact finitely many of them, say $U_1, \ldots, U_n$, still cover X. Then $M = \max\{M_1, \ldots, M_n\}$ is global bound on X, for if $x \in X$, x belongs to some U_i so that $|f(x)| \leq M_i \leq M$, showing that M bounds on f on all of X.

Local compactness. We are now interested in the question as to when a space X sits inside of a compact Hausdorff space. If this is the case, then X must itself be Hausdorff. It turns out that the only additional condition we need in order to guarantee that X sits inside of such space is that X be *locally compact*, which means for any $p \in X$ is contained in a compact set which contains a neighborhood of p. ONLY EQUIVALENT IN HAUSDORFF CASE Equivalently, X is locally compact if every point has a neighborhood with compact closure; the equivalence comes from the fact that if U is open inside a compact K, then $\overline{U} \subset K$ is closed in a compact set, so it is itself compact. Intuitively, X is locally compact if appears compact when you zoom in closely enough on a given point.

One point compactifications. To motivate the construction of a compact space containing a given space, we first considered the case of $\mathbb{R}$. Of course, $\mathbb{R}$ is not compact, but by taking the "ends" of $\mathbb{R}$, bringing them together, and gluing these ends at a single point, we can imagine $\mathbb{R}$ as being a subspace of S^1. To be concrete, $\mathbb{R}$ is homeomorphic to the space obtained by deleting the "north pole" of S^1, where the required homeomorphism is given by *stereographic projection*: for $p \in S^1$ which is not the north pole, the stereographic projection of p onto $\mathbb{R}$ is the point on the x-axis where the line through p and the north pole intersects the x-axis. The north pole is then regarded as a "point at infinity", and is the additional point we need to include in $\mathbb{R}$ in order to construct the "one point compactification" S^1.

The topology on $S^1 = \mathbb{R} \cup \{\infty\}$ can be described as follows. First, any set which is open in $\mathbb{R}$ to begin with is still open in S^1. Now, an open set around the point at infinity (i.e. the north pole of S^1) is one which under stereographic projection corresponds to a subset of $\mathbb{R}$ of the form

$$(-\infty, -m) \cup (m, \infty).$$

(Intuitively, as you go to ∞ in either direction of $\mathbb{R}$ you approach the point at infinity.) The key observation is that such a set is simply the complement of a compact subset of $\mathbb{R}$, namely $[-m, m]$. Thus, the neighborhood of the point at infinity are complements of compact sets in $\mathbb{R}$.

Similarly, we can imagine $\mathbb{R}^2$ as sitting inside the compact space S^2. In this case, the higher-dimensional analog of stereographic projection gives a homeomorphism between S^2 with the north pole excluded and $\mathbb{R}^2$, viewed as the xy-plane in $\mathbb{R}^3$. We again think of the north pole as thus being a "point at infinity" in relation to $\mathbb{R}^2$ and call $S^2 = \mathbb{R}^2 \cup \{\infty\}$ the one point compactification of $\mathbb{R}^2$. As in the case of $\mathbb{R}$, neighborhoods of ∞ correspond to complements of compact sets in $\mathbb{R}^2$ under stereographic projection.

In general, given a locally compact space X, the *one point compactification* of X is $Y = X \cup \{\infty\}$ equipped with the topology where open sets not containing ∞ are simply open subsets of X and neighborhoods of ∞ are complements of compact subsets of X. As the book shows, Y is then compact and Hausdorff. To be clear, X being locally compact is required in order to show that Y is Hausdorff; for non-locally compact spaces, Y will still be compact, but it won't be Hausdorff.

Lecture 19: More on Local Compactness

Warm-Up 1. We showed that one point compactifications are unique, in the sense that if Y and Y' are two compact Hausdorff spaces containing X such that $Y - X$ and $Y' - X$ are both single points, then Y and Y' are homeomorphic. This can be found in the book.

Warm-Up 2. We claim that if X is locally compact and Hausdorff, then for any closed subset A and $x \in X - A$, there exist disjoint open sets U and V such that $x \in U$ and $A \subseteq V$. (This property is what it means to say that X is *regular*, so the point of this Warm-Up is to show that locally compact Hausdorff spaces are always regular. We will look at the notion of regularity in more detail later on.)

Let Y denote the one point compactification of X. Let $\overline{A}$ denote the closure of A in Y, which is compact since Y is compact. Since $X - A$ is a neighborhood of x which does not contain an element of A, $x \notin \overline{A}$. Thus $\{x\}$ and $\overline{A}$ are disjoint compact sets in Y, so since Y is Hausdorff by a problem on the homework there disjoint open sets U' and V' in Y containing x and A respectively, and then $U = X \cap U'$ and $V = X \cap V'$ are the required sets in the stated problem.

Compactifications. In general, a *compactification* of a locally compact Hausdorff space X is a compact Hausdorff Y having X as a dense subspace. The one point compactification of X is the simplest compactification, but there are others. Perhaps the most important compactification is the *Stone-Cech* compactification. We won't study this compactification in this course, but you can more information about it later on in the book. Essentially, this is the compactification from which all other compactifications can be derived.

Examples of non-locally compact spaces. Both $\mathbb{Q}$ with its standard topology and $\mathbb{R}_\ell$ are not locally compact. To see that $\mathbb{Q}$ is not locally compact, we note that no compact subset of $\mathbb{Q}$ can contain an open set of the form $(a, b) \cap \mathbb{Q}$ with a, b irrational, which it would have to if it where to contain an open subset of $\mathbb{Q}$. If so, this would imply that any $[a, b] \cap \mathbb{Q}$ was compact, which is not true: intervals (c, d) with rational endpoints approaching a and b give (after intersecting with $\mathbb{Q}$) an open cover of $[a, b] \cap \mathbb{Q}$ (recall that a, b are irrational), but no finite number of these can still cover $[a, b] \cap \mathbb{Q}$.

To see that $\mathbb{R}_\ell$ is not locally compact, we show that any compact subset of $\mathbb{R}_\ell$ must be countable. This will then imply that no neighborhood of the form (a, b) can be contained in a compact set, so $\mathbb{R}_\ell$ is not locally compact. Suppose C is a compact subset of $\mathbb{R}_\ell$. For each $x \in C$, the sets

$$(-\infty, x - \frac{1}{n}) \cup [x, \infty)$$

cover C, so we get a finite subcover; the right endpoints $x - \frac{1}{n}$ of the first portions making up the sets in this finite cover have a maximum, so we can find a rational a_x such that $(a_x, x]$ contains no point of C apart from X. For different $x \in C$ these $(a_x, x]$ are thus disjoint, so the map $C \to \mathbb{Q}$ defined by $x \mapsto a_x$ is an injection, showing that C is countable.

Local compactness revisited. As the book shows, for a Hausdorff space local compactness can be rephrased as the property that for any point p and for every neighborhood U of that point, there exists a neighborhood V of p whose closure is compact and contained in U. This phrasing of local compactness for Hausdorff spaces is closer to the form the definition of "locally connected" takes. Indeed, using the notion of a local basis we mentioned previously. in a locally compact Hausdorff space, every point has a local basis of neighborhoods with compact closure.

Lecture 20: Countability Axioms

Warm-Up. We give an example of a continuous map from a locally compact space whose image is not locally compact, thus showing that local compactness (as opposed to compactness) is not a property preserved by continuous functions. Consider the identity function $\mathbb{Q}_d \to \mathbb{Q}$, where $\mathbb{Q}_d$ denotes $\mathbb{Q}$ with the discrete topology and where the codomain has the standard topology. This map is continuous (as all maps with discrete domain are), and $\mathbb{Q}_d$ is locally compact since for any $r \in \mathbb{Q}$, $\{r\}$ is a compact neighborhood of r. However, the image $\mathbb{Q}$ is not locally compact as shown last time.

To guarantee that the image of a locally compact space is locally compact we have to assume that the map, in addition to being continuous, is also open. Indeed, suppose $f : X \to Y$ is continuous and open where X is locally compact. Pick $y \in f(X)$ and $x \in X$ such that $f(x) = y$. Since X is locally compact, there exists a compact set $K \subseteq X$ containing a neighborhood U of x. Since f is open $f(U)$ is then a neighborhood of $y = f(x)$ contained in the compact set $f(K)$, so Y is locally compact.

First countability and sequences. We proved the following properties of first countable spaces mentioned in the book without proof. These properties show that sequences are enough to characterize limit points and continuity when a space is first countable, generalizing properties of $\mathbb{R}$ seen in an analysis course.

Suppose X is first countable. Then:

(i) For any $A \subseteq X$, $x \in \overline{A}$ if and only if there is a sequence of points in A converging to x.

(ii) A function $f : X \to Y$ (where Y is any space) is continuous if and only if whenever $x_n \to x$ in X, we have $f(x_n) \to f(x)$ in Y.

Proof of (a). The backwards direction of (a) is true in any topological space, since if $a_n \to x$ where each $a_n \in A$, then any neighborhood of x will contain all a_n past some index, so any neighborhood of x contains a point of A and hence $x \in \overline{A}$. For the forward direction suppose $x \in \overline{A}$ and let $\{U_n\}$ be a local basis at x. For each n, $U_1 \cap \cdots \cap U_n$ is a neighborhood of x so there exists $a_n \in A$ such that

$$a_n \in U_1 \cap \cdots \cap U_n.$$

We claim that $a_n \to x$. To see this, let V be any neighborhood of x. Since the U_n form a local basis at x, there exists N such that $x \in U_N \subseteq V$. Then for $n \geq N$ we have

$$a_n \in U_1 \cap \cdots \cap U_n = U_1 \cap \cdots \cap U_N \cap \cdots \cap U_n \subseteq U_n \subseteq V,$$

so $a_n \to x$ as claimed. □

Proof of (b). The forward direction is true in general without the assumption that X is first countable. Indeed, suppose $x_n \to x$ and let V be a neighborhood of $f(x)$. Then $f^{-1}(V)$ is a neighborhood of x, so since $x_n \to x$ there exists N such that $x_n \in f^{-1}(V)$ for $n \geq N$, which implies that $f(x_n) \in V$ for $n \geq N$ as well. Hence $f(x_n) \to f(x)$.

Conversely suppose $f(x_n) \to f(x)$ in Y whenever $x_n \to x$ in X. To show that f is continuous it is equivalent to show that for any $A \subseteq X$, we have

$$f(\overline{A}) \subseteq \overline{f(A)}.$$

Let $A \subseteq X$ and let $y \in f(\overline{A})$. Pick $x \in \overline{A}$ such that $f(x) = y$. By (a) there exists a sequence a_n in A such that $a_n \to x$. By our assumption we then have $f(a_n) \to f(x) = y$. Since $f(a_n) \in f(A)$, this implies that $y \in \overline{f(A)}$ as claimed. $\square$

Compact metric spaces are second countable. We showed that any compact metric space X is second countable. Fix $n \in \mathbb{N}$ and consider the collection $\{B_{1/n}(p)\}_{p\in X}$ of all open balls in X of radius $\frac{1}{n}$. These cover X since, in particular, $p \in B_{1/n}(p)$, so by compactness of X there exist some

$$B_{1/n}(p_{n,1}), \ldots, B_{1/n}(p_{n,k_n})$$

covering X.

Let $\mathcal{B} = \{B_{1/n}(p_{n,k_j})\}$ be the collection of all such finite covers with varying n. This is countable since it is a countable union of finite sets, and we claim that it is a countable basis of X. To see this, let $q \in X$ and pick any neighborhood V of q. Then there exists some $B_r(q)$ contained in V. Pick $n \in \mathbb{N}$ such that $\frac{1}{n} < \frac{r}{2}$. Since the sets

$$B_{1/n}(p_{n,1}), \ldots, B_{1/n}(p_{n,k_n})$$

cover X, one, say $B_{1/n}(p_{n,k_j})$ contains q. If $x \in B_{1/n}(p_{n,k_j})$, we have:

$$d(x,q) \leq d(x, p_{n,k_j}) + d(p_{n,k_j}, q) < \frac{1}{n} + \frac{1}{n} < \frac{r}{2} + \frac{r}{2} = r.$$

Thus $x \in B_r(q)$, so $q \in B_{1/n}(p_{n,k_h}) \subseteq B_r(q) \subseteq V$, showing that the sets $B_{1/n}(p_{n,k_j})$ form a basis for X as claimed.

Lecture 21: Regular Spaces

Warm-Up. We showed that any second countable space is separable and Lindelöf, which is a result proved in the book.

$\mathbb{R}_\ell$ is not metrizable. For metric spaces, being second countable is equivalent to being separable. The forward implication was in the Warm-Up, and the backwards implication is on the homework. We thus have another way of showing that certain spaces are not metrizable, meaning having topologies which are not induced by a metric. For instance, $\mathbb{R}_\ell$ is separable (since $\mathbb{Q}$ is dense) but not second countable, so it is not metrizable.

To see that $\mathbb{R}_\ell$ is not second countable, suppose $\mathcal{B}$ is any basis for $\mathbb{R}_\ell$. For each $x \in \mathbb{R}_\ell$, pick a basis element $B_x \in \mathcal{B}$ such that

$$x \in B_x \subseteq [x, x+1).$$

Note that this implies $\inf B_x = x$, since B_x must contain its minimum. Thus if $x \neq y$, $B_x \neq B_y$, showing that there are uncountably many such B_x, so $\mathcal{B}$ cannot be a countable basis.

Examples of regularity. We point out there that we showed locally compact Hausdorff spaces are regular in the second Warm-Up of Lecture 19. The book shows, as we did, that $\mathbb{R}_\ell$ is regular as well. (Actually, the book shows more, in that it shows $\mathbb{R}_\ell$ is normal, whereas we only did the regular case in class.) Also, the book contains the example that $\mathbb{R}_K$ is not regular. Note that this argument is essentially the same as the argument given in a homework problem that $\mathbb{R}_K$ is connected.

Lecture 22: Normal spaces

Warm-Up. A space X is *completely regular* if for any closed set A and any $x \in X - A$, there exists a continuous function $f : X \to [0, 1]$ such that $f(x) = 0$ and $f(A) = \{1\}$. We show that any completely regular space is regular. (Complete regularity is known as the $T_{3\frac{1}{2}}$-axiom since it sits between T_3, regular, and T_4, normal. The fact that normal spaces are completely regular will follow from *Urysohn's lemma*.)

Let A, x, and f be as in the definition of completely regular. Then $f^{-1}([0, 1/2))$ is a neighborhood of x and $f^{-1}((1/2, 1])$ is an open set in X containing A. These two open sets are disjoint since $[0, 1/2)$ and $(1/2, 1]$ are disjoint, so they give the required sets in the definition of regular.

Foreshadowing Urysohn's Lemma. It is not true that a regular space must be completely regular, but examples showing this are difficult to describe. (A starred homework problem in the book goes through a description of one example.) If in the definition of completely regular above we place x by a closed set B disjoint from A, the analogous property will show that X must then be normal. The amazing fact is that in this case the converse is true, a result which is known as Urysohn's lemma, and which we'll look at next time.

Warning. As the book states, subspaces and products are not well-behaved with respect to normality, in that a subspace of a normal space need not be normal and the product of normal spaces need not be normal either. The book shows that $\mathbb{R}_\ell \times \mathbb{R}_\ell$ is not normal, even though $\mathbb{R}_\ell$ is normal. Examples of non-normal subspaces of normal spaces are harder to find, but here is one: the space $[0, 1]^{\mathbb{R}}$ is normal since it is compact and Hausdorff (compactness will follow from Tychonoff's theorem that products of compact spaces are always compact), and such spaces, as the book shows, are always normal, but the subspace $(0, 1)^{\mathbb{R}}$ is normal. The book has a difficult homework problem showing that $\mathbb{R}^{\mathbb{R}}$ is not normal, and the fact that $(0, 1)^{\mathbb{R}}$ is not normal then follows from the fact that $(0, 1)$ and $\mathbb{R}$ are homeomorphic.

Spaces which are normal. As the book shows, there are various types of spaces which are always normal: metric spaces, compact Hausdorff spaces, and regular second countable spaces. The proofs of these facts can be found in the book, but show that many spaces which show up in practice are indeed normal.

Lecture 23: Urysohn's Lemma

Warm-Up. Suppose X has the property that any closed sets A, B such that $\overline{A} \cap B = \emptyset$ and $A \cap \overline{B} = \emptyset$ can be separated by disjoint open sets. We showed that then any subspace of X is normal. In fact, if any subspace of X is normal, X will have this given property. This was Problem 1 on Homework 8, so we omit the proof here.

A space with this property is said to be *completely normal*, which is known as the T_5-axiom.

Urysohn's lemma. The statement of Urysohn's lemma is the following: if X is normal and A, B are disjoint closed subsets of X, then there exists a continuous function $f : X \to [0, 1]$ which is 0 on A and 1 on B. We say that A and B can be *separated by a function*. Think of this as an extension problem: the constant zero function on A and the constant function 1 on B describe a real-valued continuous function on $A \cup B$, and Urysohn's lemma says that this can be extended to a continuous function on all of X. Indeed, Urysohn's lemma is at the core of the Tietze extension theorem we'll soon look at, which considers the question of extending an arbitrary (i.e. non-constant) continuous function on a closed subset of a normal space.

The proof of Urysohn's lemma can be found in the book.

Lecture 24: More on Urysohn

Warm-Up. As a Warm-Up we proved the forward direction of Exercise 33.4 in the book, which is part of Problem 3 on Homework 8, so we omit the proof here. The property based on this given in Exercise 33.5 of the book (which was also on Homework 8) is what it means for X to be *perfectly normal*, which is known as the T_6-axiom.

Complete regularity. We showed that products of completely regular spaces are completely regular, a proof which can be found in the book.

Urysohn metrization theorem. Urysohn's metrization theorem states that any second countable regular space is metrizable. The proof can be found in the book; in class we did not give the full proof, but only described the use of Urysohn's lemma in the proof and gave the idea behind the rest of the proof.

Manifold imbeddings. An n-dimensional *manifold* is a second countable Hausdorff space where every point has a neighborhood homeomorphic to $\mathbb{R}^n$. Manifolds are fundamental objects of study in geometry and topology, and a key fact is that they can always be realized as subsets of some Euclidean space of large enough dimension. The proof of this in the case of compact manifolds is in the book; again, in class we did not look at the actual proof but only briefly spoke about where Urysohn's lemma comes up.

Lecture 25: Tietze Extension Theorem

Warm-Up. Our Warm-Up this day dealt with the existence of so-called *partitions of unity*, at least in the case of finite covers. Showing the existence of such things is where Urysohn's lemma shows up in the construction of the manifold imbeddings mentioned at the end of last time. None of this will be on our final exam, so I'll omit all the details for now. The existence of partitions of unity in the case of infinite covers depends on the notion of *paracompactness*, which is something we mentioned a while back and which manifolds always possess. Again, we'll omit this all for now.

Tietze extension theorem. The Tieteze extension theorem says that if X is normal and A a closed subset, then any continuous function $f : A \to \mathbb{R}$ can be extended to a continuous function $\tilde{f} : X \to A$. The proof, of course, uses Urysohn's lemma in a nice way and can be found in the book.

Lecture 26: Tychonoff's Theorem

Our final goal is to prove Tychonoff's Theorem, which states that the product of an arbitrary number of compact spaces is compact in the product topology. We'll prove this using what's known as *Alexander's Subbase Theorem*; the proof of Tychonoff's Theorem itself is then a fairly short consequence, although it is easy to get lost in the notation. Proving Alexander's Subbase Theorem is where the real difficulty lies, and this is where we'll need to use some hardcore set theory; we'll come back to this next time.

Subbases. To setup Alexander's Subbase Theorem, we need to briefly review the notion of a *subbasis* of a topology, which is step below the notion of a basis. Subbases were introduced in the book back when bases where, but we didn't need them until now. A *subbasis* for a topology on X is a collection of sets $\mathcal{B}$ whose union is X. From this we get a basis (in the sense we've been using all along) by taking intersections of finitely many things in the subbasis; that is, a basic open set is defined to be one of the form

$$V_1 \cap \cdots \cap V_n$$

where each $V_i \in \mathcal{B}$ is a subbasis element. An arbitrary open set in the topology generated by this subbasis is then a union of these basic open sets, so is of the form

$$\bigcup_\alpha (V_{\alpha,1} \cap \cdots \cap V_{\alpha,n_\alpha}) \text{ where each } V_{\alpha,i} \in \mathcal{B}.$$

Key for us is that a subbasis for the product topology on $\prod_\alpha X_\alpha$ is given by preimages of the form

$$pr_\beta^{-1}(U_\beta) = \prod_\alpha U_\alpha \text{ where } U_\alpha = X_\alpha \text{ for } \alpha \neq \beta$$

and where U_β is open in X_β. Back when deriving the characterization of the product topology as the coarsest one relative to which all projections were continuous, we indeed showed that a basic open set is one which can be written as the intersection of finitely many such preimages, so these preimages do form a subbasis.

Alexander's Subbase Theorem. The statement is:

> Suppose X is a topological space with subbasis $\mathcal{B}$. If every open cover of X by subbase elements has a finite subcover, then X is compact.

The point is that when checking compactness, we need only consider open covers consisting of subbasis elements: if such open covers always have finite subcovers, it turns out that *all* open covers will as well. This is good, since usually subbasic open sets are simpler to work with than arbitrary open sets, as we'll now see in Tychonoff's Theorem.

Tychonoff's Theorem. Suppose $\{X_\alpha\}$ is a collection of compact spaces. Then $\prod_\alpha X_\alpha$ is compact with respect to the product topology.

Proof. By Alexander's Subbase Theorem, it is enough to show that any open cover of $\prod X_\alpha$ consisting of sets of the form $pr_\beta^{-1}(U_\beta)$ for some β, where $U_\beta \subseteq X_\beta$ is open, has a finite subcover. Thus, suppose $\mathcal{U}$ is an open cover consisting of such sets. For each α, set

$$\mathcal{U}_\alpha = \left\{U_\alpha \text{ open in } X_\alpha \mid pr_\alpha^{-1}(U_\alpha) \in \mathcal{U}\right\}.$$

Note that each element of $\mathcal{U}$ is the preimage under some projection of a set in some $\mathcal{U}_\alpha$. The point is that we are grouping all the U_β's whose preimages show up in $\mathcal{U}$ according the space X_β from which they come.

Now, we claim that for at least one $\beta, \mathcal{U}_\beta$ is an open cover of X_β. If not, then for any α the union $\bigcup \mathcal{U}_\alpha$ of all the sets in $\mathcal{U}_\alpha$ is a proper subset of X_α, so there exists $x_\alpha \in X_\alpha$ such that $x_\alpha \notin \bigcup \mathcal{U}_\alpha$. But then the element (x_α) of $\prod X_\alpha$ cannot be in any element of $\mathcal{U}$ since, if so, we would have $(x_\alpha) \in pr_\gamma^{-1}(U_\gamma)$ for some γ and $U_\gamma \subseteq X_\gamma$, meaning that $x_\gamma \in U_\gamma \subseteq \bigcup \mathcal{U}_\gamma$, contradicting the choice of x_γ. Hence for some $\beta, \mathcal{U}_\beta$ covers X_β.

Since X_β is compact, we then get a finite subcover $\{U_{\beta,1}, \ldots, U_{\beta,n}\}$ of $\mathcal{U}_\beta$. The preimages

$$pr_\beta^{-1}(U_{\beta,1}), \ldots, pr_\beta^{-1}(U_{\beta,n})$$

then give a finite subcover of $\mathcal{U}$; indeed, each such preimage looks like

$$pr_\beta^{-1}(U_{\beta,i}) = \prod V_\alpha \text{ where } V_\beta = U_{\beta,i} \text{ and } V_\alpha = X_\alpha \text{ for } \alpha \neq \beta,$$

so their union is $\prod W_\alpha$ where $W_\alpha = X_\alpha$ for $\alpha \neq \beta$ and $W_\beta = U_{\beta,1} \cup \cdots \cup U_{\beta,n} = X_\beta$, and is thus all of $\prod X_\alpha$. Hence we conclude that $\prod X_\alpha$ is compact as claimed. $\square$

Axiom of Choice. So, Tychonoff's Theorem is actually quick to prove, at least assuming Alexander's Theorem and once we wrap our head around the notation. Again, the point is that by focusing solely on subbasic open sets, we can direct our focus to only one index (the final β used in the proof above), use compactness in this index to get a finite cover for the corresponding space, and then take preimages to get a finite cover of the entire product. But now what remains is to prove Alexander's Theorem, and this is more involved. We'll do this next time, where we'll have to make use of the Axiom of Choice in a novel way.

Recall that the Axiom of Choice says that the product of nonempty sets is always nonempty: if $\{X_\alpha\}$ is a collection of nonempty sets, then $\prod_\alpha X_\alpha$ is nonempty. This might seem like such an obvious fact that it's not worth mentioning, but let's think about what it actually means for this product to be nonempty. In the simpler case of a product of two nonempty sets X_1 and X_2, to show that $X_1 \times X_2$ is nonempty all we do is say:

> Pick $x_1 \in X_1$, which can be done since $X_1 \neq \emptyset$, and pick $x_2 \in X_2$, which can be done since $X_2 \neq \emptyset$. Then (x_1, x_2) is in $X_1 \times X_2$, so $X_1 \times X_2 \neq \emptyset$.

In a similar way, it is easy to show that the product of n nonempty sets is nonempty: we just sit down and pick one element from each of our finitely many nonempty sets, and use them as components for an element of the product. Of course, as the number of sets increases it "takes longer" to pick an element from each set, but it can be done.

The issue arises when we try to do the same thing for the product of infinitely many sets. The point is that we cannot literally sit down and pick an element from each of component set as would be needed to describe an element of the product since it would take an infinite amount of time! This is whole crux of the matter: choosing finitely many things at a time is always doable with the Axiom of Choice, but making an infinite number of choices simultaneously is where the Axiom of Choice is required. Now, this is not to say that choice is always needed when showing that certain products are nonempty; for instance, $\mathbb{R}^\omega$ is nonempty since we can write down $(0, 0, 0, 0, \ldots)$ as an explicit element. The point here is that we have explicitly written down the element we want from each component, namely 0 in this case. The Axiom of Choice would be needed if we didn't have a set way of choosing these components, which is the issue we have when dealing with an infinite number of random sets we have no information about.

Tychonoff implies choice. To give a sense of the relation between the Axiom of Choice and Tychnoff's Theorem, we now show that Tychonoff's Theorem implies the Axiom Choice. The fact that the Axiom of Choice implies Tychonoff's Theorem, so that the two are actually equivalent, follows from the proof of Tychonoff's Theorem we gave above and from what we will do next time; the logic is "Axiom of Choice implies Zorn's Lemma, which implies the Alexander Subbasis Theorem, which implies Tychonoff's Theorem".

Suppose Tychonoff's Theorem holds. For the sake of clean notation, we'll only prove the Axiom of Choice in the case of a countably infinite collection (usually called the "Axiom of Countable Choice"), but the general case follows the same reasoning with a slight modification. So, suppose $\{X_1, X_2, X_3, \ldots\}$ is a collection of countably many nonempty sets. We aim to show there exists something in the product $X_1 \times X_2 \times \cdots$. For each n set $Y_n = X_n \cup \{\infty_n\}$, where ∞_n denotes some new point, and give Y_n the topology whose open sets are

$$\emptyset, Y_n, X_n, \{\infty_n\}.$$

Since there are only finitely many open sets, any open cover of Y_n is automatically finite so each Y_n is compact. By Tychonoff's Theorem, $Y_1 \times Y_2 \times Y_3 \times \cdots$ is compact as well.

Now, define the open subsets U_n of $Y_1 \times Y_2 \times Y_3 \times \cdots$ by:

$$\begin{aligned} U_1 &= \{\infty_1\} \times Y_2 \times Y_3 \times \cdots \\ U_2 &= Y_1 \times \{\infty_2\} \times Y_3 \times \cdots \\ U_3 &= Y_1 \times Y_2 \times \{\infty_3\} \times \cdots \end{aligned}$$

and so on. We claim that these sets do not cover all of $Y_1 \times Y_2 \times Y_3 \times \cdots$. Before showing this, note what this means: we get that $\bigcup_n U_n$ is a proper subset of $\prod_n Y_n$, meaning that there must exist some $\mathbf{y} = (y_1, y_2, y_3, \ldots) \in Y_1 \times Y_2 \times Y_3 \times \cdots$ which is not in this union. But to say that this element is not in this union means that $y_1 \neq \infty_1$ (since $\mathbf{y} \notin U_1$), $y_2 \neq \infty_2$ (since $\mathbf{y} \notin U_2$), and so on. Thus it must be the case that each y_n comes from the X_n part of $Y_n = X_n \cup \{\infty_n\}$, so $\mathbf{y} = (y_1, y_2, \ldots) \in X_1 \times X_2 \times \cdots$ is the element we are trying to show exists in order to say that the product $X_1 \times X_2 \times \cdots$ is nonempty.

To show that the U_n's all together do not cover $\prod_n Y_n$, we show that no finite number among them can cover $\prod_n Y_n$; since we know $\prod_n Y_n$ is compact, this suffices since if the U_n's did cover the product, they would necessarily need to have a finite subcover. For any N we take, pick elements $x_i \in X_i$ for each $1 \leq i \leq N$. Then

$$\mathbf{x} = (x_1, x_2, \ldots, x_N, \infty_{N+1}, \infty_{N+2}, \infty_{N+3}, \ldots)$$

is in $Y_1 \times Y_2 \times \cdots$ but is not in $U_1 \cup \cdots U_N$ since for $1 \leq i \leq N$, $\mathbf{x}$ has an i-th component which is not ∞_i. Thus $U_1 \cup \cdots \cup U_N \neq Y$ for all N, so $\{U_1, U_2, \ldots\}$ has no finite subcover of $\prod_n Y_n$ among it, so this collection itself is not a cover of $\prod_n Y_n$ as required.

One important point: in order to make the above proof work, we needed to know that we can pick elements from $X_1, \ldots, X_N$ all at once for any N, but since this just requires making a finite number of choices at a time, the Axiom of Choice is not required. Specifying all the remaining elements of $\mathbf{x}$ to be $\infty_{N+1}, \infty_{N+2}, \ldots$ also does not require choice since we are explicitly saying here which elements from $Y_{N+1}, Y_{N+2}, \ldots$ should be chosen.

Towards Zorn. The Axiom of Choice takes on many equivalent forms, and the one we'll actually need is called *Zorn's Lemma*. We'll save the statement for next time, but be prepared to see a statement which will take a bit of effort to digest. The amazing fact is that, while Zorn's Lemma will seem to be fairly complicated at first, it is actually equivalent to the more obvious Axiom of Choice and has some quite powerful applications.

Lecture 27: Alexander Subbase Theorem

Our final goal is to prove the Alexander Subbase Theorem, on which our proof of Tychnoff's Theorem relied. As mentioned last time, the proof requires an equivalent form of the Axiom of Choice known as *Zorn's Lemma*, so we begin by explaining what goes into this result. Even though this is called a "lemma", Zorn's Lemma is an incredibly important and useful result in mathematics, mainly because it gives a way to show that various objects exist in situations where constructing them explicitly would be impossible.

Zorn's Lemma. Suppose P is a nonempty, partially-ordered set in which every chain has an upper bound. Then P has a maximal element.

Partial orders. There are various possibly unfamiliar terms in the statement of Zorn's Lemma, so we first clarify the statement itself. A *partial order* on a set P is a relation $\leq$ satisfying:

- $a \leq a$ for all $a \in P$,
- if $a \leq b$ and $b \leq c$, then $a \leq c$, and
- if $a \leq b$ and $b \leq a$, then $a = b$.

Here, $\leq$ is purely a symbol we use to denote the given relation, but the point is that these properties suggest $\leq$ behaves as it if was an actual "ordering" on elements of P: anything should be "less than or equal to" itself, the "less than or equal to" relation should be transitive, and the only way in which two things can be "less than or equal to" each other is if they are the actually the same. We also use the strict notation $a < b$ to mean that $a \leq b$ and $a \neq b$.

Two key examples are the usual "less than or equal to" relation on $\mathbb{R}$, where $x \leq y$ literally means that x is less than or equal to y, and the partial order on a collection of subsets of a set given by $\subseteq$, where we interpret $A \subseteq B$ as saying that A is "less than or equal to" B. However, these examples have one important difference: in the case of $\mathbb{R}$, all elements are comparable to one another in the sense that given any $x, y \in \mathbb{R}$, it is true that $x \leq y$ or $y \leq x$, but this is not necessarily true when considering collections of subsets. A *chain* in P is a subset whose elements are all comparable to one another in this way. (A partial order in which all elements are comparable is called a *total order*, so a chain in P is then a totally-ordered subset of P.) The term "chain" comes from the idea that you can order all elements from "smaller" to "larger", which in the countable case looks like:

$$\ldots \leq a \leq b \leq c \leq \ldots.$$

An *upper bound* of a subset S of P is an element $u \in P$ such that $s \leq u$ for all $s \in S$, which is the same way the term "upper bound" is used, say, in analysis. Finally, a *maximal* element of P is one for which there is nothing strictly larger: $a \in P$ is maximal if whenever $a \leq b$ for some $b \in P$, we have $a = b$. The usual (total) ordering on all of $\mathbb{R}$ has no maximal elements, but subsets of $\mathbb{R}$ might have maximal elements; if we take all subsets of a set S, then under $\subseteq$ the only maximal element is S itself, but a collection of only certain subsets might have none, one, or more maximal elements.

Zorn's Lemma thus says that as long we know that any totally-ordered subset can be bounded above by something, then we can conclude that at least one maximal element exists. In the type of situation we care about, Zorn's Lemma will be applied in the following way. Take P to be a collection of subsets of some set. Suppose further P has the property that for any subcollection

$C \subseteq P$ of sets such that any two are comparable via $\subseteq$, meaning that given A and B in C it is always true that either $A \subseteq B$ or $B \subseteq A$, we have that the union $\bigcup C$ of all things in C also belongs to P. Then we can conclude that there is a set S in P which is not strictly contained within any larger element of P. Here, the partial ordering on P is given by $\subseteq$, C describes a chain in P with $\bigcup C$ being its upper bound in P, and the resulting S is a maximal element of P. Such maximal elements, as we'll see, often have important properties we care about.

Choice implies Zorn. We now give a sense as to where Zorn's Lemma comes from, and how it relates to the Axiom of Choice. Specifically, we give a very rough sketch of the proof that the Axiom of Choice implies Zorn's Lemma. Zorn's Lemma is actually equivalent to the Axiom of Choice, but the direction we look at there (choice implies Zorn) is the one we need to take us from the Axiom of Choice to Tychonoff's Theorem. Our proof sketch is quite rough since we will get to a point where we would need to know much more advanced set theory—in particular properties of *cardinal* and *ordinal numbers*—to make it precise, but the basic idea will come across.

Suppose P is a nonempty, partially-ordered set in which every chain has an upper bound, and aiming for a contradiction suppose P did not contain any maximal elements. Then for any $a \in P$, we can always find some $b \in B$ such that $a < b$. Using the Axiom of Choice we can thus pick such an element $f(a)$ for any $a \in P$. (Using the "nonempty product" interpretation of the Axiom of Choice, this comes form considering, for any $a \in P$, the nonempty set U_a of all elements of P which are strictly larger than a and picking an element $(f(a))_a$ from the nonempty product $\prod_a U_a$.) Fix $a \in P$, so that $a < f(a)$. But by this construction we also have $f(a) < f(f(a))$, and so on we get:

$$a < f(a) < f(f(a)) < f(f(f(a))) < \cdots .$$

This list gives a chain in P, so by the assumption of Zorn's Lemma this chain has an upper bound, call it a_1:

$$a < f(a) < f(f(a)) < f(f(f(a))) < \cdots \leq a_1.$$

But now we can consider the chain

$$a_1 < f(a_1) < f(f(a_1)) < f(f(f(a_1))) < \cdots ,$$

which itself has an upper bound a_2:

$$a_1 < f(a_1) < f(f(a_1)) < f(f(f(a_1))) < \cdots \leq a_2.$$

Continuing in this way over and over (and over and over!) again gives a bunch of elements of P:

$$a < f(a) < \cdots \leq a_1 < \cdots \leq a_2 < \cdots \leq a_3 < \cdots \leq a_4 < \cdots .$$

In fact, there would be so many elements of P listed here that this would imply (and this is the part which requires some pretty deep stuff which we will in no way attempt to make precise here) that the cardinality of P would be larger than that of any other set, and in particular P would have cardinality (strictly) larger that of P itself (or also of its power set), which is nonsense. Thus we conclude that P must have had a maximal element after all.

The big three. As stated above, the Axiom of Choice not only implies but is actually implied by Zorn's Lemma, so that they are equivalent. Just for the sake of interest, we give the statement of one more equivalent form of either of these: the *Well-Ordering Theorem.* A *well-ordering* on a set P is a total order in which every nonempty subset of P has a least (i.e. smallest) element. For instance, the usual ordering on $\mathbb{N}$ is a well-ordering, whereas the usual ordering on $\mathbb{R}$ is not.

The Well-Ordering Theorem says that *every* set can in fact be well-ordered. In the case of $\mathbb{R}$, the point is that the usual order is *not* the one which works, but that there is *some* way to "order" the elements of $\mathbb{R}$ so that every nonempty subset does have a least element.

This is pretty surprising indeed, and the well-ordering on $\mathbb{R}$ which works would actually have no relation to the usual ordering. An explicit such well-ordering on $\mathbb{R}$ is not possible to write down, but nonetheless we know it must exist (if we accept the Axiom of Choice) since the Axiom of Choice, the Well-Ordering Theorem, and Zorn's Lemma are all equivalent to one another. These types of surprising results are the main reason why the Axiom of Choice—as obvious as it may seem—is viewed as quite controversial by many mathematics: it has some seemingly paradoxical consequences which often say that a certain objects exists without giving any sense as to how to actually *construct* said object. There's an old joke that says: the Axiom of Choice is clearly true, the Well-Ordering Theorem is clearly false, and who knows about Zorn's Lemma? The joke, of course, is that the first of these seems obvious, the second seems like it could not possibly be true (since we cannot even imagine what a well-ordering of $\mathbb{R}$ would actually look like), and the third (Zorn) is such a complicated looking statement that no one really has any idea what it even means, and yet all three are actually saying the same thing in the end.

$\mathbb{R}^\omega$ **has a basis.** Before proving the Alexander Subbase Theorem, we give one application of Zorn's Lemma in linear algebra. Consider $\mathbb{R}^\omega$ equipped with vector addition and scalar multiplication defined as one would expect:

$$(x_1, x_2, \ldots) + (y_1, y_2, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots) \text{ and } r(x_1, x_2, \ldots) = (rx_1, rx_2, \ldots).$$

We aim to show that $\mathbb{R}^\omega$ has a *basis* in the sense of linear algebra: a linearly independent subset of $\mathbb{R}^\omega$ which spans all of $\mathbb{R}^\omega$. Now, the trouble is that it is not actually possible to write down an explicit basis (!), so our proof is non-constructive. This is in stark contrast to the case of $\mathbb{R}^n$, where bases are easy to write down. Note that the obvious candidate of taking the vectors $\mathbf{e}_i$ which have a 1 in the i-th location and 0 everywhere else (which work in the $\mathbb{R}^n$ case) do not work in $\mathbb{R}^\omega$, since it is not true that anything in $\mathbb{R}^\omega$ can be written as a linear combination of *finitely many* of these $\mathbf{e}_i$, which is a technical requirement in the definition of "span" in the setting of infinite dimensions; the issue is that any linear combination of finitely many of the $\mathbf{e}_i$'s must eventually end in all zeroes! So in fact, the $\mathbf{e}_i$ vectors only span the subspace $\mathbb{R}^\infty$ of $\mathbb{R}^\omega$.

Let I denote the collection of all linearly independent subsets of vectors in $\mathbb{R}^\omega$. Take any chain $C \subseteq I$. Then $\bigcup C$ is still a collection of linearly independent vectors in $\mathbb{R}^\omega$, and so is an upper bound for this chain in I. To see that $\bigcup C$ is still linearly independent, take any finite number of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \bigcup C$. (To say that a set of vectors is "linearly independent" technically means that any *finite* number of vectors taken from that set are linearly independent.) Each $\mathbf{v}_i$ comes from some $C_i \in C$. The fact that C is a chain implies that there exists $C_0 \in C$ which contains each of $C_1, \ldots, C_n$, so $\mathbf{v}_1, \ldots, \mathbf{v}_n \in C_0 \subseteq I$ must be linearly independent. Hence $\bigcup C$ is a linearly independent collection of vectors as claimed.

By Zorn's Lemma there thus exists a maximally linearly independent set B of vectors in $\mathbb{R}^\omega$. If these vectors did not span $\mathbb{R}^\omega$, picking $\mathbf{x} \in \mathbb{R}^\omega$ not in their span gives a linearly independent collection $B \cup \{\mathbf{x}\}$ which is strictly larger than B, contradicting maximality of B. Thus B must span $\mathbb{R}^\omega$, so that B is a basis of $\mathbb{R}^\omega$ as desired.

In general, the same reasoning shows that any *vector space*, even an infinite dimensional one, has a basis—a fact which is actually equivalent to the Axiom of Choice. Many other facts you might have seen elsewhere turn out to also be applications of the same idea: the fact that any ideal in a nontrivial ring with unity is contained in a maximal ideal, the fact that any field has an algebraic closure, etc.

Alexander's Subbase Theorem. Finally we prove the Alexander Subbase Theorem, thereby completing the proof of Tychonoff's Theorem. Recall the statement: suppose X is a topological space with subbasis $\mathcal{B}$; if every open cover of X by subbase elements has a finite subcover, then X is compact. The proof works by contradiction: use Zorn's Lemma to get a maximal open cover with some given property, and then use the maximality itself to show that this could not actually exist after all.

Proof. Aiming for a contradiction, suppose X is not compact, so that there exists an open cover of X with no finite subcover. Let

$$\mathcal{F} = \{\text{open covers of } X \text{ with no finite subcover}\}$$

be the nonempty collection of all such things. Equip $\mathcal{F}$ with the partial order $\subseteq$ given by set containment. We claim that $\mathcal{F}$ satisfies the assumptions of Zorn's Lemma. Indeed, suppose $\{E_\alpha\}$ is a chain in $\mathcal{F}$ and let $E = \bigcup_\alpha E_\alpha$ denote the union of everything in this chain. Clearly E will be an upper bound for this chain once we know that E is actually in $\mathcal{F}$. Since any E_α is already an open cover of X, E is as well. Take any finite number of things $U_1, \dots, U_n$ in E. Then each U_i is an element of some E_{α_i}. Since $\{E_\alpha\}$ is totally ordered, there is some E_β which contains all of $E_{\alpha_1}, \dots, E_{\alpha_n}$. Then $U_1, \dots, U_n$ are all in E_β, so $U_1, \dots, U_n$ cannot cover of all X because if they did they would make up a finite subcover of E_β, contradicting the fact that $E_\beta \in \mathcal{F}$. Thus no finite number of things in $E = \bigcup_\alpha E_\alpha$ can cover X, so $E \in \mathcal{F}$ as required.

Thus $\mathcal{F}$ satisfies the assumptions of Zorn's Lemma, so there exists a maximal element $\mathcal{M}$ in $\mathcal{F}$; that is, $\mathcal{M}$ is a open cover of X with no finite subcover which is maximal among such open covers. The contradiction we are after will arise from showing that $\mathcal{M}$ must actually have a finite subcover after all. Consider $\mathcal{M} \cap \mathcal{B}$, which is made up of the open sets in the cover $\mathcal{M}$ which are actually subbasis elements. This collection cannot cover all of X since, if so, it would necessarily have a finite subcover by the assumption of the Alexander Subbase Theorem, which would then also be a finite subcover of $\mathcal{M}$, contradicting $\mathcal{M} \in \mathcal{F}$. Thus there exists $x \in M$ such that $x \notin \bigcup(\mathcal{M} \cap \mathcal{B})$, which denotes the union of all things in $\mathcal{M} \cap \mathcal{B}$. But $\mathcal{M}$ does cover all of X, so there exists $U \in \mathcal{M}$ such that $x \in U$, and hence by the definition of a subbasis there exists a basic open set $V_1 \cap \ldots \cap V_n$, where each $V_i \in \mathcal{B}$, such that

$$x \in V_1 \cap \cdots \cap V_n \subseteq U.$$

Now, none of the V_i can be in $\mathcal{M}$, since $x \in V_1 \cap \cdots \cap V_n \subseteq V_i$ would then imply that x was already covered by the elements of $\mathcal{M} \cap \mathcal{B}$, but we chose x to not be in $\bigcup(\mathcal{M} \cap \mathcal{B})$.

Thus for each i, $\mathcal{M} \cup \{V_i\}$ is a cover of X which is strictly larger than $\mathcal{M}$. Since $\mathcal{M}$ is meant to be maximal with respect to those open covers with no finite subcover, this larger cover must then have a finite subcover, say

$$U_{i,1}, \dots, U_{i,n_i}, V_i \in \mathcal{M} \cup \{V_i\},$$

where concretely each $U_{i,j}$ comes from $\mathcal{M}$. (Note that V_i must be included in this subcover since $\mathcal{M}$ alone does not have a finite subcover.) Thus for each i we have

$$X \subseteq \bigcup_j U_{i,j} \cup V_i$$

and hence:

$$X \subseteq \bigcap_i \left(\bigcup_j U_{i,j} \cup V_i \right) \subseteq \bigcup_{i,j} U_{i,j} \cup (V_1 \cap \cdots \cap V_n) \subseteq \bigcup_{i,j} U_{i,j} \cup U,$$

which says that the sets $U_{i,j}$ and U all together cover X. Since each of these sets comes from $\mathcal{M}$ and there are finitely many (there are finitely many $i = 1, \ldots, n$ and then for each of these finitely many $j = 1, \ldots, n_i$), these would give a finite subcover of $\mathcal{M}$, again contradicting the fact that $\mathcal{M} \in \mathcal{F}$ was meant to have no finite subcover.

Thus no maximal open cover of X with no finite subcover can exist after all, so this final contradiction shows that our original assumption that $\mathcal{F}$ is nonempty must have been false, so $\mathcal{F}$ is indeed empty, meaning that there does not exist an open cover of X without a finite subcover, or equivalently that every open cover of X has a finite subcover. Hence X is compact as claimed. $\square$

Hallelujah! Thanks for reading!