Chapter 3. Compactness

Compactness is a sort of finiteness property that some spaces have and some do not. The rough idea is that spaces which are 'infinitely large' such as $\mathbb{R}$ or $[0,\infty)$ are not compact. Since compactness will be defined just in terms of open sets, any space homeomorphic to a noncompact space will also be noncompact, so finite intervals (a,b) and $[a,b)$ are also noncompact in spite of their 'finiteness'. On the other hand, closed intervals $[a,b]$ are compact — they cannot be stretched to be 'infinitely large'.

The definition of compactness is, at first glance, a little complicated:

Definition. A space X is *compact* if for each collection of open sets O_α in X whose union is X, there exist a finite number of these O_α's whose union is X.

More concisely, one says that every open cover of X has a finite subcover, where an *open cover* of X is a collection of open sets in X whose union is X, and a *finite subcover* is a finite subcollection whose union is still X.

For example, $\mathbb{R}$ is not compact because the cover by the open intervals $(-n,n)$ for $n = 1, 2, \cdots$ has no finite subcover, since infinitely many of these intervals are needed to cover all of $\mathbb{R}$. Another open cover which has no finite subcover is the collection of intervals $(n-1, n+1)$ for $n \in \mathbb{Z}$.

In a similar vein, the interval $(0,1)$ fails to be compact since the cover by the open intervals $(1/n, 1)$ for $n \geq 1$ has no finite subcover. Of course, there do exist open covers of $(0,1)$ which have finite subcovers, for example the cover by $(0,1)$ itself, or a little less trivially, the cover by all open subintervals of fixed length, say $1/4$, which has the finite subcover $(0,1/4)$, $(1/8,3/8)$, $(1/4,1/2)$, $(3/8,5/8)$, $(1/2,3/4)$, $(5/8,7/8)$, $(3/4,1)$. To be compact means that every possible open cover has a finite subcover. This could be difficult to check in individual cases, so we will develop general theorems to test for compactness.

Of course, spaces with only finitely many points are obviously compact, or more generally spaces whose topology has only finitely many open sets. However, such spaces are not very interesting, so we begin by showing that there are some less trivial compact spaces:

Theorem. *A closed interval $[a,b]$ is compact.*

Proof. This will be somewhat similar in flavor to the proof we gave that closed intervals are connected. The case $a = b$ is trivial, so we may assume $a < b$. Let a cover of $[a,b]$ by open sets O_α in $[a,b]$ be given. Since $a \in O_\alpha$ for some α, there exists $c > a$

such that the interval $[a,c]$ is contained in this O_α, and hence $[a,c]$ is contained in the union of finitely many O_α's. Let L be the least upper bound of the set of numbers $c \in [a,b]$ such that $[a,c]$ is contained in the union of finitely many O_α's. We know that $L > a$ by the preceding remarks, and by the definition of L we certainly have $L \le b$.

There is some O_α, call it O_β, that contains L. This O_β is open in $[a,b]$, so since $L > a$ there is an interval $[L-\varepsilon, L]$ contained in O_β for some $\varepsilon > 0$. By the definition of L there exist numbers $c < L$ arbitrarily close to L such that $[a,c]$ is contained in the union of finitely many O_α's. In particular, there are such numbers c in the interval $[L-\varepsilon, L]$. For such a c we can take a finite collection of O_α's whose union contains $[a,c]$ and add the set O_β containing $[L-\varepsilon, L]$ to this collection to obtain a finite collection of O_α's containing the interval $[a,L]$. If $L = b$ we would now be done, so it remains only to show that $L < b$ is not possible.

If $L < b$, the number ε could have been chosen so that not only is $[L-\varepsilon, L] \subset O_\beta$ but also $[L-\varepsilon, L+\varepsilon] \subset O_\beta$, since O_β is open in $[a,b]$. Then by adding O_β to the finite collection of O_α's whose union contains $[a,c]$, as in the preceding paragraph, we would have a finite collection of O_α's whose union contains $[a, L+\varepsilon]$. However, this means that L is not an upper bound for the set of c's such that $[a,c]$ is contained in a finite union of O_α's. This contradiction shows that $L < b$ is not possible, so we must have $L = b$. □

For a subspace A of a space X to be compact means of course that every open cover of A has a finite subcover. The open cover of A would consist of sets of the form $A \cap O_\alpha$ for O_α open in X. To say that $A = \bigcup_\alpha (A \cap O_\alpha)$ is equivalent to saying that $A \subset \bigcup_\alpha O_\alpha$. Thus for A to be compact means that for every collection of open sets in X whose union contains A, there is a finite subcollection whose union contains A. So it does no harm to interpret 'every open cover of A has a finite subcover' to mean precisely this.

Proposition. *A closed subset of a compact space is compact, in the subspace topology.*

Proof. Let $\{O_\alpha\}$ be a cover of A by open sets in X. We then obtain an open cover of X by adding the set $X - A$, which is open if A is closed. If X is compact this open cover of X has a finite subcover. The sets O_α in this finite subcover then give a finite cover of A since the set $X - A$ contributes nothing to covering A. □

As an example, the Cantor set is closed in $[0,1]$, so it is compact because $[0,1]$ is compact.

Here is another way to show that a space is compact:

Proposition. *If $f: X \to Y$ is continuous and onto, and if X is compact, then so is Y.*

Proof. Let a cover of Y by open sets O_α be given. Then the sets $f^{-1}(O_\alpha)$ form an open cover of X. If X is compact, this cover has a finite subcover. Call this finite subcover $f^{-1}(O_1), \cdots, f^{-1}(O_n)$. Assuming that f is onto, the corresponding sets $O_1, \cdots, O_n$ then cover Y since for each $y \in Y$ there exists $x \in X$ with $f(x) = y$, and this x will be in some set $f^{-1}(O_i)$ of the finite cover of X, so y will be in the corresponding set O_i. □

This implies for example that a circle is compact since it is the image of a continuous map $f: [0,1] \to \mathbb{R}^2$.

In order to expand our range of compact spaces we use the notion of product spaces, introduced in Chapter 1.

Theorem. *If X and Y are compact then so is their product $X \times Y$.*

By induction this implies that the product of any finite collection of compact spaces is compact.

Proof. Let a cover of $X \times Y$ by open sets O_α in $X \times Y$ be given. Each point $(x, y) \in X \times Y$ lies in some O_α, and this O_α is a union of basis sets $U \times V$, so there exists a basis set $U_{xy} \times V_{xy}$ containing (x, y) and contained in some O_α.

Suppose we choose a fixed x and let y vary. Then the sets $U_{xy} \times V_{xy}$ cover $\{x\} \times Y$, so the sets V_{xy} with fixed x and varying y form an open cover of Y. Since Y is compact, this cover has a finite subcover $V_{xy_1}, \cdots, V_{xy_n}$, where n may depend on x. The intersection $U_x = \bigcap_{j=1}^n U_{xy_j}$ is then an open set containing x with two key properties: The sets $U_x \times V_{xy_1}, \cdots, U_x \times V_{xy_n}$ cover $U_x \times Y$, and each $U_x \times V_{xy_j}$ is contained in some O_α.

Now we let x vary. The sets U_x form an open cover of X, so since X is compact there is a finite subcover $U_{x_1}, \cdots, U_{x_m}$. The products $U_{x_i} \times V_{x_i y_j}$ of the sets U_{x_i} with the corresponding sets $V_{x_i y_j}$ chosen earlier then form a finite cover of $X \times Y$. Each set in this finite cover is contained in some O_α, so by choosing an O_α containing each $U_{x_i} \times V_{x_i y_j}$ we obtain a finite cover of $X \times Y$. □

We can use this result to determine exactly which subspaces of $\mathbb{R}^n$ are compact. The result is usually called the Heine-Borel Theorem.

Theorem. *A subspace $X \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

For a subset $X \subset \mathbb{R}^n$ to be bounded means that it lies inside some ball of finite radius centered at the origin.

Proof. First let us assemble previously-proved results to show the 'if' half of the theorem. If we assume X is bounded, then it lies in a ball of finite radius and hence in some closed cube $[-r,r] \times \cdots \times [-r,r]$. This cube is compact, being a product of closed intervals which are compact. Since X is a closed subset of a compact space, it is also compact.

Now for the converse, suppose X is compact. The collection of all open balls in $\mathbb{R}^n$ centered at the origin and of arbitrary radius forms an open cover of X, so there is a finite subcover, which means X is contained in a single ball of finite radius, the largest radius of the finitely many balls covering X. Hence X is bounded.

To show X is closed if it is compact, suppose x is a limit point of X that is not in X. Then every neighborhood of x contains points of X. In particular each open ball $B_r(x)$ of radius r centered at x contains points of X, so the same is true also for the closed balls $\overline{B}_r(x)$. The complements $\mathbb{R}^n - \overline{B}_r(x)$ form an open cover of X as r varies over $(0,\infty)$ since their union is $\mathbb{R}^n - \{x\}$ and $x \notin X$. This open cover of X has no finite subcover since each $\overline{B}_r$ contains points of X. Thus we have shown that if X is not closed, it is not compact. ◻

Hausdorff Spaces

We showed earlier that a compact subspace of $\mathbb{R}^n$ is closed. This is not something that remains true when $\mathbb{R}^n$ is replaced by an arbitrary space X. For example if X is a finite set with any topology then every subspace of X is compact since a cover of X can have only finitely many sets, but not every subset of X will be closed unless X has the discrete topology. Fortunately there is a fairly simple and natural condition to impose on a topology which will guarantee that compact subspaces are closed:

Definition. A topological space X is a *Hausdorff space* if for each pair of distinct points $x, y \in X$ there exist open neighborhoods U of x and V of y that are disjoint.

For example, every metric space X is a Hausdorff space since if x and y are distinct points in X then the open balls of radius ε centered at x and y will be disjoint if ε is less than half the distance between x and y. This follows from the triangle inequality, since if there were a point z in the intersection of these two balls, then the distances between x, y, and z would satisfy $d(x,y) \leq d(x,z) + d(z,y) < \varepsilon + \varepsilon = 2\varepsilon < d(x,y)$, a contradiction.

The trivial topology on a space with more than one point is not Hausdorff. A more interesting example of a non-Hausdorff space is $\mathbb{R}$ with the topology having as its nonempty open sets the complements of finite sets, since in this topology any two nonempty open sets have a nonempty intersection. Many more examples of non-Hausdorff spaces can be constructed but they do not arise all that often 'in nature'. Most non-Hausdorff spaces are in one way or another artificial. It is probably not a great exaggeration to say that most interesting spaces are Hausdorff.

Here are three nice properties of Hausdorff spaces:

Proposition. *(a) In a Hausdorff space, points are closed subsets. (b) A subspace of a Hausdorff space is Hausdorff. (c) A product of two Hausdorff spaces is Hausdorff.*

Proof. (a) If X is Hausdorff and x is a point in X, then for any other point y there is an open neighborhood V_y of y that is disjoint from a neighborhood of x and hence disjoint from x itself. The union of all these open sets V_y as y ranges over $X - \{x\}$ is $X - \{x\}$, which is therefore open, so $\{x\}$ is closed.
(b) Let A be a subspace of X and let $x, y \in A$. If X is Hausdorff these two points have disjoint open neighborhoods U and V in X, so $A \cap U$ and $A \cap V$ are disjoint open neighborhoods of x and y in A.
(c) Suppose X and Y are Hausdorff and we have two distinct points (x_1, y_1) and (x_2, y_2) in $X \times Y$, so either $x_1 \neq x_2$ or $y_1 \neq y_2$. In the former case, if U_1 and U_2 are disjoint neighborhoods of x_1 and x_2 in X then $U_1 \times Y$ and $U_2 \times Y$ are disjoint neighborhoods of (x_1, y_1) and (x_2, y_2) in $X \times Y$. The opposite case $y_1 \neq y_2$ is handled similarly. □

Proposition. *A compact subspace of a Hausdorff space is closed.*

Proof. Let X be Hausdorff with $A \subset X$ compact. The main work will be to show that for each $x \in X - A$ there exist disjoint open sets U and V in X such that $x \in U$ and $A \subset V$. This will imply that A is closed since no point x in the complement of A can be a limit point of A since it has a neighborhood U disjoint from A.

Now we construct U and V with the desired properties. The point $x \in X - A$ and an arbitrary point $a \in A$ have disjoint open neighborhoods U_x and V_a in X since X is Hausdorff. As a ranges over A the open sets V_a form an open cover of A. Since A is compact, this open cover has a finite subcover. Call this subcover $V_1, \cdots, V_n$, and let the sets U_x that correspond to these V_i's be labelled $U_1, \cdots, U_n$. The intersection

$U = U_1 \cap \cdots \cap U_n$ is an open neighborhood of x that is disjoint from each V_i and hence also from the union $V = V_1 \cup \cdots \cup V_n$, which is an open set containing A. □

Corollary. *A map $f : X \to Y$ from a compact space X to a Hausdorff space Y is a homeomorphism if it is continuous, one-to-one, and onto.*

Proof. All that needs to be checked is that f^{-1} is continuous, and to do this it suffices to show that if $C \subset X$ is closed in X then $f(C)$ is closed in Y. But since C is a closed subset of a compact space, it is compact, hence $f(C)$ is compact and therefore also closed since Y is Hausdorff. □

For example, consider continuous maps $[0,1] \to [0,1] \times [0,1]$ defining paths in the square. Intuition would suggest that no such map can be onto, but in fact this intuition is incorrect, and there do exist continuous maps $[0,1] \to [0,1] \times [0,1]$ whose image is the whole square. However, no such map can be one-to-one, otherwise the Corollary would say it is a homeomorphism, but $[0,1]$ and $[0,1] \times [0,1]$ are not homeomorphic since $[0,1]$ has cutpoints but $[0,1 \times [0,1]$ does not.

The Corollary implies that the topology on a space X that is both compact and Hausdorff has a certain optimality property. Notice first that any topology that is finer than a Hausdorff topology is still Hausdorff, while a topology that is coarser than a compact topology is still compact.

- If X is compact and Hausdorff with respect to the topology $\mathcal{O}$, then any other topology that is strictly finer than $\mathcal{O}$ is Hausdorff but not compact, while any other topology that is strictly coarser than $\mathcal{O}$ is compact but not Hausdorff.

To see this, just apply the Corollary to the identity map $X \to X$, which is continuous if the topology in the domain is finer than the topology in the range. If the topology in the domain is compact Hausdorff and the topology in the range is strictly coarser, it would be compact but not Hausdorff, otherwise the Corollary would imply the identity map was a homeomorphism so the two topologies would be the same. Similarly, if the topology in the range is compact Hausdorff and the topology in the domain is strictly finer, it would be Hausdorff but not compact, otherwise the Corollary would say the identity map was a homeomorphism.

Normal Spaces

Often in a Hausdorff space X there is a stronger form of the Hausdorff condition that is satisfied: If A and B are disjoint closed sets in X, then there are disjoint open sets U and V in X with $A \subset U$ and $B \subset V$. In short, disjoint closed sets have disjoint open neighborhoods. A Hausdorff space with this property is called a *normal* space. We include the Hausdorff condition as part of the definition of a normal space to guarantee that points are closed, so that the condition of disjoint closed sets having disjoint open neighborhoods includes the condition that distinct points have disjoint open neighborhoods. The reason for restricting A and B to be closed is that one would not expect arbitrary disjoint subsets to have disjoint open neighborhoods, since for example in $\mathbb{R}$ the sets $\{0\}$ and $(0,1]$ do not have disjoint open neighborhoods.

In some cases the Hausdorff condition automatically implies normality:

Proposition. *If X is a compact Hausdorff space then X is normal.*

Proof. Note first that A and B are compact since they are closed subsets of a compact space. By the argument in the proof of the preceding Proposition we know that for each $x \in A$ there exist disjoint open sets U_x and V_x with $x \in U_x$ and $B \subset V_x$. Letting x vary over A, we have an open cover of A by the sets U_x, so there is a finite subcover. Let U be the union of the sets U_x in this finite subcover and let V be the intersection of the corresponding sets V_x. □

One naturally wonders whether all Hausdoff spaces are normal, so here is an artificially constructed example to show that this need not be true:

Example. Let X be the subset of $\mathbb{R}^2$ consisting of all points (x,y) with $y \geq 0$. We construct a topology on X that is finer than the usual topology in the following way. Let $X' \subset X$ be the complement of the x-axis, the points (x,y) with $y > 0$. Then a basis for the new topology on X consists first of all of the usual open disks $D_r(x)$ of radius r and center x that are contained in X', and then in addition for each point x in the x-axis we take the sets $D'_r(x) = \{x\} \cup (X' \cap D_r(x))$ obtained by adding x to an open half-disk centered at x. It is easy to check that the condition for having a basis is satisfied, and we leave this for the reader to do. The Hausdorff condition is also easily verified since any two distinct points of X are contained in disjoint basis sets. Notice that any subset of the x-axis is closed in this topology since its complement is open, being a union of basis sets. However, if we let A be the rational numbers in the x-axis and B the irrational numbers, then A and B are disjoint closed sets in X that do not have disjoint open neighborhoods, since any open set U containing A would contain basis sets $D'_r(a)$ for $a \in A$ that intersect all the basis sets $D'_s(b)$ for $b \in B$ with $|a-b| < r$, hence U would intersect any open set V containing B. Thus X is not normal.

Notice that the subspace topology on the x-axis is the discrete topology since the intersection of the open set $D'_r(x)$ with the x-axis is $\{x\}$, making $\{x\}$ an open set in this subspace. On the other hand, the subspace topology on X', the complement of the x-axis, is the usual topology. So the topology on X somehow mingles these two rather different sorts of topologies.

Proposition. *Metric spaces are normal.*

Before proving this we need a preliminary fact. Let X be a metric space with metric d. Given a subset $A \subset X$ define the distance $d(x,A)$ from a point $x \in X$ to A to the the greatest lower bound of the set of distances $d(x,a)$ from x to points $a \in A$. Note that $d(x,A) \geq 0$, and $d(x,A) = 0$ if and only if x is in the closure of A since $d(x,A) = 0$ is equivalent to saying that every ball $B_r(x)$ contains points of A.

Lemma. *$d(x,A)$ is a continuous function of x.*

Proof. Suppose x and y are points in X with $d(x,y) < \varepsilon$. For each $\delta > 0$ there exists $a \in A$ with $d(x,a) < d(x,A) + \delta$. Then $d(y,a) < d(y,x) + d(x,a) < \varepsilon + \delta + d(x,A)$. Since $d(y,a) \leq d(y,A)$ this implies $d(y,A) < \varepsilon + \delta + d(x,A)$. Since δ was any positive number, it follows that $d(y,A) \leq \varepsilon + d(x,A)$, or $d(y,A) - d(x,A) \leq \varepsilon$. Switching the roles of x and y in this inequality, we have also $d(x,A) - d(y,A) \leq \varepsilon$, so $|d(x,A) - d(y,A)| \leq \varepsilon$, under the original assumption that $d(x,y) < \varepsilon$. This implies that $d(x,A)$ is a continuous function of x. $\square$

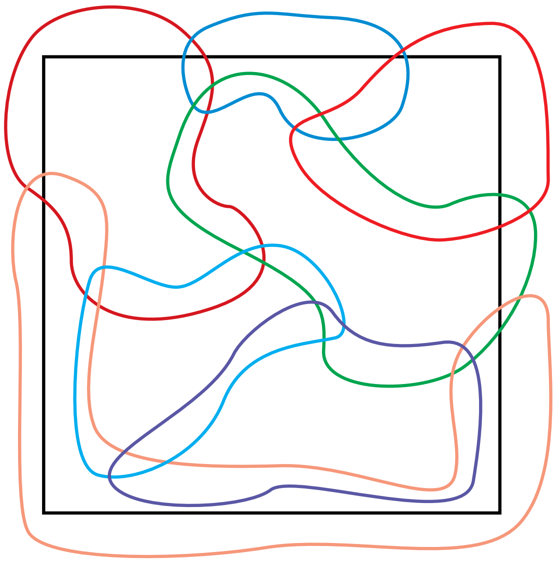
Proof of the Proposition. If A and B are disjoint closed sets in the metric space X, we claim that the sets $U = \{\, x \mid d(x,A) < d(x,B) \,\}$ and $V = \{\, x \mid d(x,B) < d(x,A) \,\}$ are disjoint open neighborhoods of A and B, respectively. Certainly they are disjoint, and they are open since for the difference function $d(x,A) - d(x,B)$, which is continuous by the Lemma, U and V are the inverse images of the open sets $(-\infty, 0)$ and $(0, \infty)$. To see that $A \subset U$ note that $d(x,A) = 0$ for points $x \in A$ and $d(x,A) > 0$ for point $x \notin A$ since A is closed. Similarly we have $B \subset V$ using the assumption that B is closed. $\square$

Lebesgue Numbers

Given a cover of a metric space X by subsets A_α, a *Lebesgue number* for the cover is a number $\varepsilon > 0$ such that every ball $B_\varepsilon(x)$ in X is contained in at least one set A_α of the cover.

Theorem. *Every open cover of a compact metric space X has a Lebesgue number.*

For example, what would a Lebesgue number for the following open cover of a square be?



Proof. Since X is compact, each open cover has a finite subcover, and there is no loss of generality to find a Lebesgue number for this subcover. Thus we may assume the cover is by open sets $U_1, \cdots, U_n$.

For a ball $B_\varepsilon(x)$ to be contained in U_i is equivalent to the requirement that $d(x, X - U_i) \geq \varepsilon$. Thus we are led to consider the functions $d_i(x) = d(x, X - U_i)$, and to ask whether there exists an $\varepsilon > 0$ such that for each $x \in X$ at least one of the numbers $d_1(x), \cdots, d_n(x)$ is at least ε. Phrased differently, we are asking whether there is an $\varepsilon > 0$ such that the function $f(x) = \max\{d_1(x), \cdots, d_n(x)\}$ satisfies $f(x) \geq \varepsilon$ for all x. Each function $d_i(x)$ is continuous by the preceding Lemma, and we leave it as an exercise for the reader to verify that the maximum of a finite set of continuous functions is again continuous. (By induction it suffices to do the case of two functions.) So f is continuous. Since U_i is open, its complement $X - U_i$ is closed and hence $d_i(x) > 0$ for all $x \in U_i$. Thus $f(x) > 0$ for all $x \in X$ since the U_i's cover X. Since X is compact, the image $f(X)$ is compact in $(0, \infty)$, so there is a lower bound $\varepsilon > 0$ for the values of f on X. By the remarks at the beginning of this paragraph, this ε is a Lebesgue number for the cover. □

Corollary. *Let $f: X \to Y$ be a continuous map from a compact metric space X to a metric space Y. Then f is uniformly continuous: For each $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(f(x), f(x')) < \varepsilon$ for all $x, x' \in X$ with $d(x, x') < \delta$.*

Proof. Given $\varepsilon > 0$, cover X by the open sets $f^{-1}(B_{\varepsilon/2}(y))$ as y ranges over Y. Let δ be a Lebesgue number for this cover of X. Then f takes each ball $B_\delta(x)$ to some ball $B_{\varepsilon/2}(y)$, so if $d(x,x') < \delta$ then $d(f(x),f(x')) \le d(f(x),y) + d(y,f(x')) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Infinite Products

There are times when one is interested in the product $X_1 \times X_2 \times \cdots$ of an infinite sequence of spaces $X_1, X_2, \cdots$. As a set, this product consists of all sequences $(x_1, x_2, \cdots)$ with $x_i \in X_i$ for each i. We will also use the notation $\prod_i X_i$ for this product. As an example, the Cantor set can be viewed as the product $\prod_i X_i$ where each X_i is the set $\{0,2\}$ by identifying an infinite base 3 decimal $.a_1a_2\cdots$ with the sequence $(a_1, a_2, \cdots)$.

A first guess for how to define a topology on an infinite product $\prod_i X_i$ would be to do the same thing as for finite products, taking as a basis the products $U_1 \times U_2 \times \cdots$ of open sets $U_i \subset X_i$. The same argument as for finite products shows that this does in fact define a topology, called the *box topology* on $\prod_i X_i$. However, if one tries to work with this topology, one finds it has some undesirable features. For example, many maps $Z \to \prod_i X_i$ that one would expect to be continuous are not continuous if the box topology is used on the infinite product. To take a concrete instance, if Z and each X_i is $\mathbb{R}$ then the harmless-looking map $x \mapsto (x, x, \cdots)$ fails to be continuous, since the inverse image of the product $(-1/2, 1/2) \times (-1/3, 1/3) \times (-1/4, 1/4) \times \cdots$ is just $\{0\}$, which is not open.

Fortunately there is another topology on infinite products that avoids defects like this. Instead of taking a basis to consist of all products $U_1 \times U_2 \times \cdots$ of open sets $U_i \subset X_i$, one only takes products where $U_i = X_i$ for all except finitely many values of i. This guarantees that a map $f : Z \to \prod_i X_i$, $f(z) = (f_1(z), f_2(z), \cdots)$, is continuous if and only if each f_i is continuous, since we have $f^{-1}(U_1 \times U_2 \times \cdots) = f_1^{-1}(U_1) \cap f_2^{-1}(U_2) \cap \cdots$, and this will be open if each f_i is continuous, as it is really only a finite intersection since $f_i^{-1}(U_i) = f_i^{-1}(X_i) = Z$ for all but finitely many values of i.

This topology on $\prod_i X_i$ is called the *product topology* since it turns out to be the one that is used for infinite products almost all the time. Notice that for finite products it gives the same thing as the product topology defined earlier.

Example. The Cantor set is homeomorphic to the product $\{0,2\} \times \{0,2\} \times \cdots$ where each $\{0,2\}$ is given the discrete topology. [Proof: Use the basis for the Cantor set given by the sets $O(a_1,\cdots,a_n)$ consisting of all decimals whose first n digits are $a_1,\cdots,a_n$ and whose remaining digits are arbitrary.]

Another advantage of the product topology over the box topology is that it preserves compactness:

Theorem. *An infinite product $\prod_{i=1}^{\infty} X_i$ with the product topology is compact if each X_i is compact.*

This would not be true for the box topology. For example, if each X_i is a finite set having at least two elements, with the discrete topology, then the box topology on the infinite product is the discrete topology, which is never compact for an infinite set.

Proof. We will argue by contradiction. Suppose there is an open cover $\mathcal{O}$ of $\prod_i X_i$ that has no finite subcover. Assuming this, we claim first that there exists a point $x_1 \in X_1$ such that no basis set $U_1 \times X_2 \times X_3 \times \cdots$ with $x_1 \in U_1$ is covered by finitely many sets in $\mathcal{O}$. For otherwise if for every x_1 there was a basis set $U_1 \times X_2 \times X_3 \times \cdots$ covered by finitely many sets in $\mathcal{O}$, with $x_1 \in U_1$, then the collection of all these U_1's as x_1 varied over X_1 would be an open cover of X_1, so since X_1 is compact this cover would have a finite subcover, and it would follow that the cover $\mathcal{O}$ had a finite subcover, contrary to assumption.

Having chosen the point x_1, we claim next that there is a point $x_2 \in X_2$ such that no basis set $U_1 \times U_2 \times X_3 \times \cdots$ with $(x_1,x_2) \in U_1 \times U_2$ is covered by finitely many sets in $\mathcal{O}$. For otherwise if for every x_2 there was such a basis set $U_1 \times U_2 \times X_3 \times \cdots$, then the collection of these U_2's as x_2 varied over X_2 would be an open cover of X_2, and compactness of X_2 would give a finite subcover. Letting U_1' be the intersection of the corresponding finite set of U_1's, we would then have a basis set $U_1' \times X_2 \times X_3 \times \cdots$ covered by finitely many sets in $\mathcal{O}$, with $x_1 \in U_1'$, contrary to how x_1 was chosen in the preceding paragraph.

Repeating this same argument, we can choose an infinite sequence of points $x_i \in X_i$ for $i = 1,2,\cdots$, such that for each n, no basis set $U_1 \times \cdots \times U_n \times X_{n+1} \times \cdots$ with $(x_1,\cdots,x_n) \in U_1 \times \cdots \times U_n$ is covered by finitely many sets in $\mathcal{O}$. But this is impossible because the point $(x_1,x_2,\cdots)$ has to lie in some set in $\mathcal{O}$ and since this set is open, it will contain a basis set $U_1 \times \cdots \times U_n \times X_n \times \cdots$ containing $(x_1,x_2,\cdots)$, so in particular this basis set will be covered by a single set in $\mathcal{O}$. □

We have considered the product of a countable collections of spaces X_n but it is also possible to define the product of an arbitrary infinite collection of spaces X_α

where α ranges over any index set I. If I is uncountable we cannot write elements of the product $\prod_\alpha X_\alpha$ as sequences $(x_1, x_2, \cdots)$. Instead, elements of $\prod_\alpha X_\alpha$ are regarded as functions $\alpha \mapsto x_\alpha \in X_\alpha$, assigning an element of X_α to each index α, just as sequences $(x_1, x_2, \cdots)$ can be viewed as functions $i \mapsto x_i \in X_i$. One can write an element of $\prod_\alpha X_\alpha$ briefly as (x_α), it being understood that this is not just a single x_α but a family of them parametrized by $\alpha \in I$. As in the case of countable products, the product topology on $\prod_\alpha X_\alpha$ has as basis the products $\prod_\alpha U_\alpha$ of open sets $U_\alpha \subset X_\alpha$ such that $U_a = X_\alpha$ for all but finitely many values of α. It is again a theorem that $\prod_\alpha X_\alpha$ is compact if each X_α is compact, but the proof requires a fancier form of induction that depends on a preliminary discussion of set theory topics called the 'well-ordering principle' and the 'axiom of choice', so we will not go into this here.