Mathematics 4530
Tychonoff's Theorem
Ken Brown, Cornell University, October 2013

1. Introduction

Tychonoff's theorem asserts that the product of an arbitrary family of compact spaces is compact. This is proved in Chapter 5 of Munkres, but his proof is not very straightforward. The proof I'll give below follows a paper by David Wright (*Proceedings of the American Mathematical Society* **120** (1994), 985–987). As a warmup, let's start with two factors.

2. The Baby Tychonoff Theorem

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

Note that this immediately extends to arbitrary finite products by induction on the number of factors. This yields:

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

Proof of the corollary. If $X \subseteq \mathbb{R}^n$ is compact, then it is closed and bounded by the same proof we used in class when $n = 1$. Conversely, if X is closed and bounded, then X is a closed subset of a rectangle $R = I_1 \times \cdots \times I_n$, where each I_i is a closed interval in $\mathbb{R}$. Since each I_i is compact, so is the product R; the closed subset X of R is therefore also compact. $\square$

To prove a space X is compact, one usually proves that if $\mathcal{U}$ is a family of open sets that covers X, then a finite subcollection covers X. Sometimes, however, it is more convenient to prove the contrapositive: If $\mathcal{U}$ is a family of open sets such that no finite subcollection covers X, then $\mathcal{U}$ does not cover X. This is what we will do below.

Proof of the theorem. Let $\mathcal{W}$ be a collection of open subsets of $X \times Y$ such that no finite subcollection covers $X \times Y$; we will show that $\mathcal{W}$ does not cover $X \times Y$.

Claim 1. There exists $x_0 \in X$ such that no open tube $U \times Y$ with $x_0 \in U$ is finitely covered by $\mathcal{W}$.

Proof. If this is false, then every $x \in X$ has a neighborhood U_x such that $U_x \times Y$ is finitely covered. By compactness of X, finitely many of these sets U_x cover X, so finitely many of the tubes $U_x \times Y$ cover $X \times Y$. This contradicts the assumption that $X \times Y$ is not finitely covered.

Claim 2. There exists $y_0 \in Y$ such that no open rectangle $U \times V$ containing (x_0, y_0) is finitely covered by $\mathcal{W}$.

Proof. If this is false, then for every $y \in Y$ there is a finitely covered open rectangle $U_y \times V_y$ containing (x_0, y). By compactness of Y, there is a finite subset $F \subseteq Y$ such that $Y = \bigcup_{y \in F} V_y$. Set $U := \bigcap_{y \in F} U_y$. Then U is a neighborhood of x_0, and the tube

$$U \times Y = \bigcup_{y \in F} U \times V_y \subseteq \bigcup_{y \in F} U_y \times V_y$$

is finitely covered, contradicting Claim 1.

Claim 2 immediately implies the theorem. Indeed, we have a point $z := (x_0, y_0)$ such that no basic open set containing z is finitely covered by $\mathcal{W}$. In particular, no basic open set containing z can be contained in a set $W \in \mathcal{W}$, so $z \notin \bigcup_{W\in\mathcal{W}} W$. This shows that $\mathcal{W}$ does not cover $X \times Y$. □

Remark. We did not have to use the contrapositive of the condition for compactness. In fact, it is more straightforward to just use the original definition, and to base the proof on the tube lemma; see Munkres, pp. 167–168. (The tube lemma asserts that in a product $X \times Y$ with Y compact, any neighborhood of a slice $x \times Y$ contains a tube $U \times Y$, where U is a neighborhood of x; the proof of this is very similar to the proof of Claim 2.) The advantage of our convoluted proof, however, is that it extends easily to infinitely many factors.

3. Countably Many Factors

In preparation for treating infinitely many factors, we record a slight generalization of the argument used in the proof of Claim 2 above.

Lemma. *Let $\mathcal{W}$ be a family of open sets in a product $X \times Y \times Z$. Assume there is a point $x_0 \in X$ such that no open set $U \times Y \times Z$ with $x_0 \in U$ is finitely covered by $\mathcal{W}$. If Y is compact, then there is a point $y_0 \in Y$ such that no open set $U \times V \times Z$ with $(x_0, y_0) \in U \times V$ is finitely covered by $\mathcal{W}$.*

Proof. Suppose no such y_0 exists. Then for every $y \in Y$ there is a finitely covered open set $U_y \times V_y \times Z$ with $x_0 \in U_y$ and $y \in V_y$. By compactness of Y, there is a finite subset $F \subseteq Y$ such that $Y = \bigcup_{y\in F} V_y$. Set $U := \bigcap_{y\in F} U_y$. Then U is a neighborhood of x_0, and the set

$$U \times Y \times Z = \bigcup_{y\in F} U \times V_y \times Z \subseteq \bigcup_{y\in F} U_y \times V_y \times Z$$

is finitely covered, contradicting the hypothesis. □

We can now easily generalize the "baby" argument to treat countably many factors:

Theorem. *Let $X = \prod_{i=1}^{\infty} X_i$, where each X_i is compact. Then X is compact.*

Proof. Let $\mathcal{W}$ be a family of open sets that does not finitely cover X; we will construct a point $x = (x_1, x_2, \dots)$ such that no neighborhood of x is finitely covered. Note first that there is a point $x_1 \in X_1$ such that no open tube $U \times X_2 \times X_3 \times \cdots$ with $x_1 \in U$ is finitely covered. The proof of this assertion is the same as the proof of Claim 1 above, with $X_2 \times X_3 \times \cdots$ playing the role of Y. Next, we can find $x_2 \in X_2$ such that no open rectangle $U \times V \times X_3 \times X_4 \times \cdots$ with $(x_1, x_2) \in U \times V$ is finitely covered. This follows from the lemma if we view X as $X_1 \times X_2 \times (X_3 \times \cdots)$. Continuing in this way, we inductively define $x_1, x_2, x_3, \dots$ such that for each n, no basic open set of the form $U_1 \times \cdots \times U_n \times X_{n+1} \times \cdots$, with $x_i \in U_i$ for $i \leq n$, is finitely covered. For the inductive step, view X as $(X_1 \times \cdots \times X_{n-1}) \times X_n \times (X_{n+1} \times \cdots)$ and apply the lemma.

We now have a point $x = (x_1, x_2, \dots) \in X$ such that no basic neighborhood of x is finitely covered by $\mathcal{W}$. Thus $x \notin \bigcup_{W\in\mathcal{W}} W$. □

4. Grown-up Tychonoff

Finally, here is the full-fledged Tychonoff theorem:

Theorem. *Given an arbitrary family $(X_\alpha)_{\alpha\in J}$ of compact spaces, their product $X := \prod_{\alpha\in J} X_\alpha$ is compact.*

Proof. Once again, we let $\mathcal{W}$ be a family of open sets that does not finitely cover X, and we construct $x = (x_\alpha)_{\alpha\in J}$ such that no neighborhood of x is finitely covered. We may assume by the well-ordering theorem that the index set J is well-ordered, and we construct x_α by transfinite induction so that no basic open set of the form

$$\prod_{\beta\leq\alpha} U_\beta \times \prod_{\beta>\alpha} X_\beta,$$

with $x_\beta \in U_\beta$ for all $\beta \leq \alpha$, is finitely covered. For the inductive step, assume that x_β has been defined with the desired property for all $\beta < \alpha$, and apply the lemma, viewing X as $\left(\prod_{\beta<\alpha} X_\beta\right) \times X_\alpha \times \left(\prod_{\beta>\alpha} X_\beta\right)$. The details are essentially the same as in the countable case. □

Remark. The proof used the axiom of choice, in the form of the well-ordering theorem. This is unavoidable. Indeed, logicians have shown that Tychonoff's theorem cannot be proved without the axiom of choice.

5. Application: Invariant Means

Having gone to all the trouble of proving the well-ordering theorem and the general form of Tychonoff's theorem, we now give a (somewhat weird) application of it. As is typical of results that require the axiom of choice, it proves the existence of something that one could never hope to actually construct concretely. We begin with a naïve question:

Is there a sensible way of associating an average value to every bounded, doubly infinite sequence of real numbers, such as

$$\ldots, 0, 1, 0, 1, 0, 1, \ldots? \tag{1}$$

To phrase the question precisely, consider sequences $a = (a_n)_{n\in\mathbb{Z}}$, i.e., elements of $\mathbb{R}^{\mathbb{Z}}$. We denote by $\mathcal{B}$ the set of all such sequences that are bounded. It is a real vector space. We will not make use of any topology on $\mathcal{B}$, but we recall in passing that there is a standard one, different from the product topology, that is useful in analysis: One puts a norm on $\mathcal{B}$ by setting $\|a\| = \sup_{n\in\mathbb{Z}}|a_n|$, and this yields a metric $d(a,b) := \|a-b\|$ and hence a topology.

A *mean* on $\mathcal{B}$ is a linear map $\mu\colon \mathcal{B} \to \mathbb{R}$ such that

$$\inf_n a_n \leq \mu(a) \leq \sup_n a_n$$

for all $a \in \mathcal{B}$. Means exist in great abundance. For example, we can take a finite subset $F \subseteq \mathbb{Z}$ and define a mean by averaging over F:

$$\mu(a) := \frac{1}{|F|}\sum_{n\in F} a_n.$$

More generally, we can assign a weight $w_n \geq 0$ to each $n \in \mathbb{Z}$, with $\sum_n w_n = 1$, and use the weighted average

$$\mu(a) := \sum_n a_n w_n.$$

The *shift operator* $a \mapsto a^s$ on $\mathcal{B}$ is defined by $(a^s)_n = a_{n+1}$. A mean is said to be *invariant*, or *shift-invariant*, if $\mu(a^s) = \mu(a)$ for all $a \in \mathcal{B}$. It is reasonable to expect a sensible averaging procedure to be shift-invariant; this just says that we can talk about the average of a sequence like the one in (1) above, without having to choose arbitrarily which element is considered to be in postion 0. But no one knows how to construct an invariant mean. (In particular, the weighted averages above are *not* invariant). The best we can do is to construct approximately invariant means and then deduce, in a nonconstructive way, that invariant means exist. Here is the constructive part:

Lemma. *There is a sequence $(\mu_n)_{n\geq 1}$ of means such that for all $a \in \mathcal{B}$,*

$$\lim_{n\to\infty} |\mu_n(a^s) - \mu_n(a)| = 0.$$

Proof. Let $\mu_n(a) := (1/n)\sum_{i=1}^n a_i$. Then

$$|\mu_n(a^s) - \mu_n(a)| = |a_{n+1} - a_1|/n \leq 2\|a\|/n,$$

which tends to 0 as $n \to \infty$. □

There are several known methods for doing the nonconstructive part of the proof that invariant means exist. The one we will use is to topologize the set $\mathcal{M}$ of all means in such a way that it becomes a compact space (via Tychonoff). It will then follow that the approximately invariant sequence μ_n has a cluster point μ, which is invariant. Here are the details.

View $\mathcal{M}$ as a subset of $\mathbb{R}^{\mathcal{B}}$ by identifying a mean μ with the indexed family $(\mu(a))_{a\in\mathcal{B}}$. We then give $\mathbb{R}^{\mathcal{B}}$ the product topology and $\mathcal{M}$ the subspace topology. Thus our topology is cooked up to make the evaluation map $\mu \mapsto \mu(a)$ a continuous map $\mathcal{M} \to \mathbb{R}$ for each $a \in \mathcal{B}$. Note that the definition of "mean" actually makes $\mathcal{M}$ a subset of a product of closed intervals:

$$\mathcal{M} \subseteq \prod_{a\in\mathcal{B}} [m(a), M(a)],$$

where $m(a) = \inf_n a_n$ and $M(a) = \sup_n a_n$.

Proposition. $\mathcal{M}$ *is compact.*

Proof. We know that $X := \prod_{a\in\mathcal{B}}[m(a), M(a)]$ is compact, so it suffices to show that $\mathcal{M}$ is a closed subset of X. Now an element $\mu \in X$ is a mean if and only if it satisfies

$$\mu(a + b) - \mu(a) - \mu(b) = 0$$

for all $a, b \in \mathcal{B}$ and

$$\mu(\lambda a) - \lambda\mu(a) = 0$$

for all $a \in \mathcal{B}$ and $\lambda \in \mathbb{R}$. This exhibits $\mathcal{M}$ as an intersection of sets of the form $F^{-1}(0)$ for various continuous maps $F\colon X \to \mathbb{R}$, so $\mathcal{M}$ is closed. □

We now appeal to the "Bolzano–Weierstrass" property of compact spaces:

Proposition. *If X is a compact space, then every sequence $(x_n)_{n\geq 1}$ in X has a cluster point, i.e., there is a point $x \in X$ such that every neighborhood of x contains x_n for infinitely many n.*

Proof. For each x that is not a cluster point, there is a neighborhood U_x of x that contains x_n for only finitely many n. Clearly no finite subcollection of the U_x can cover X, so the family of all U_x does not cover X. Since $\bigcup_x U_x$ contains all points that are not cluster points, it follows that there must exist at least one cluster point. $\square$

We can now prove our promised existence theorem:

Theorem. *There exists an invariant mean* $\mu\colon \mathcal{B} \to \mathbb{R}$.

Proof. Take a sequence (μ_n) as in the lemma, and let μ be a cluster point. To show that μ is invariant, let $a \in \mathcal{B}$ be arbitrary. We will argue that for a suitable large n,

$$\mu(a^s) \approx \mu_n(a^s) \approx \mu_n(a) \approx \mu(a). \tag{2}$$

More precisely, fix $\epsilon > 0$ and let

$$U := \left\{\nu \in \mathcal{M} \;\middle|\; |\nu(a) - \mu(a)| < \epsilon \text{ and } |\nu(a^s) - \mu(a^s)| < \epsilon\right\}.$$

Then U is a neighborhood of μ in $\mathcal{M}$. [In fact, it is a basic neighborhood, defined by imposing conditions on two "coordinates" of a general ν.] We therefore have $\mu_n \in U$ for infinitely many n. Take such an n large enough that $|\mu_n(a^s) - \mu_n(a)| < \epsilon$. Then the three approximations in (2) are all valid, with error less than ϵ, whence $|\mu(a^s) - \mu(a)| < 3\epsilon$. Since ϵ is arbitrary, it follows that $\mu(a^s) = \mu(a)$. $\square$