The Heine-Borel and Arzela-Ascoli Theorems

David Jekel

February 8, 2015

This paper explains two important results about compactness, the Heine-Borel theorem and the Arzela-Ascoli theorem. We prove them first in $\mathbb{R}^d$. Then (for the more curious) we explain how they generalize to the more abstract setting of metric spaces.

The Heine-Borel Theorem

The goal of this section is to prove the following theorem which shows that different definitions of compactness are equivalent. This is also proved in Folland's appendix, using a slightly different method.

Theorem 1. *Let $S \subset \mathbb{R}^d$. The following are equivalent:*

a. Every open cover of S has a finite subcover.

b. S is closed and bounded.

c. Every sequence in S has a subsequence converging to a point in S.

Part 1: Prove (a) implies (b).

Assume (a) holds. First, we show S is bounded. Let $B(0,k)$ be the open ball of radius k centered at the origin. The collection $\{B(0,k)\}_{k=1}^\infty$ is an open cover of S. By (a) it has a finite subcover, so $S \subset \bigcup_{j=1}^J B(0,k_j)$. The union of these balls is equal to the largest ball since they are nested. Thus, S is contained in some ball, so it is bounded.

To prove S is closed, it suffices to show that S contains all its limit points. Suppose that $\{x_n\}$ is a sequence in S converging to a point x_0. Suppose for the sake of contradiction that $x_0 \notin S$. Let $O_k = \{x : |x - x_0| > 1/k\}$. Then the union of the O_k's is $\mathbb{R}^d \setminus \{x_0\}$, which contains S. Thus, $\{O_k\}$ is an open cover of S, so there is a finite subcover. Since the O_k's are nested, we have $S \subset O_k$ for some k. By definition of O_k, this means that $|x - x_0| > 1/k$ for all $x \in S$. In particular, $|x_n - x_0| > 1/k$. But this contradicts $x_n \to x_0$. Therefore, what we assumed was false and $x_0 \in S$ as desired.

Part 2: Prove (b) implies (c).

Assume (b) holds. Let $\{x_n\}$ be a sequence in S. We must construct a convergent subsequence. Since S is bounded, it is contained in some cube $Q_1 = [-N, N]^d$. Divide Q_1 into 2^d cubes with half the length. Since there are infinitely many indices n in the sequence, one of the smaller cubes must contain x_n for infinitely many n. Call this smaller cube Q_2. Now subdivide Q_2 and find a cube $Q_3 \subset Q_2$ of half the length which also contains x_n for infinitely many n. Continue inductively to construct a sequence of nested cubes $Q_1 \supset Q_2 \supset Q_3 \supset \dots$ such that each cube has half the length of the previous one, and each cube contains infinitely many points of the sequence.

Now we choose our subsequence as follows: Let $x_{n_0} = x_0$. Let n_2 be the first index such that $x_{n_2} \in Q_2$. Let n_3 be the first index greater than n_2 such that $x_{n_3} \in Q_3$; such an n_3 must exist because Q_3 contains infinitely points of the sequence. Let n_j be the first index greater than n_{j-1} such that $x_{n_j} \in Q_j$. This defines a subsequence $\{x_{n_j}\}_{j=1}^{\infty}$.

We claim $\{x_{n_j}\}$ is Cauchy. Choose $\epsilon > 0$. Choose J large enough that the diagonal of the cube Q_J has length less than ϵ (specifically, choose J such that $\sqrt{d}N2^{-J} < \epsilon$). If $j, k \geq J$, then x_{n_j} and x_{n_k} are in Q_{n_J}, and therefore, $|x_{n_j} - x_{n_k}| < \epsilon$. This shows $\{x_{n_j}\}$ is Cauchy and hence it converges.

Part 3: Prove (c) implies (a).

This is the hardest part, and has several steps. First, note that if (c) holds, then S must be bounded. Indeed, if it was not bounded, there would exist a sequence $\{x_n\}$ in S such that $|x_n| \to \infty$. This sequence could not possibly have a convergent subsequence.

The next step is the following lemma, which is interesting in itself:

Lemma 1 ("Lebesgue number lemma")**.** *Suppose $S \subset \mathbb{R}^d$ and (b) holds. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover S. There exists a $\delta > 0$ such that any set $E \subset S$ with* $\operatorname{diam} E < \delta$ *is contained in a single U_α.*

Remarks: Here the *diameter* of a set $E \subset \mathbb{R}^d$ is

$$\operatorname{diam} E = \sup_{x,y \in E} |x - y|.$$

Also, A is just some set which we use to index the open sets in our cover. The number δ is called a *Lebesgue number* for $\{U_\alpha\}$.

Proof of Lemma 1. Suppose for the sake of contradiction that no such δ exists. That means that in particular that $\delta = 1/k$ does not work for any integer k. So for any integer k, there exists $E_k \subset S$ with $\operatorname{diam} E_k < 1/k$ such that E_k is not contained in any single U_α. Choose a point $x_k \in E_k$. Since we assumed (b), the sequence $\{x_k\}$ has a subsequence $\{x_{k_j}\}$ converging to a point $x_0 \in S$.

Then x_0 is in some U_α, and since U_α is open, there is a ball $B(x_0, r) \subset U_\alpha$. Choose j large enough that $1/k_j < r/2$ and $|x_{k_j} - x_0| < r/2$. Then $\operatorname{diam} E_k < r/2$. So if $y \in E_k$, we have $|y - x_{k_j}| < 1/k_j < r/2$. But $|x_{k_j} - x_0| < r/2$, so by the triangle inequality, $|y - x_0| < r$. Therefore, $E_k \subset B(x, r) \subset U_\alpha$, and this contradicts our choice of E_k. $\square$

Completing the proof of Part 3. Let U_α be an open cover of S. We must find a finite subcover. Let $\delta > 0$ be a Lebesgue number for the open cover, as constructed in the Lemma. Since S is bounded, it is contained in some cube. By subdividing the cube into small pieces, we can cover S by finitely many cubes $Q_1, \ldots, Q_K$ such that $\operatorname{diam} Q_k < \delta$. Then $\operatorname{diam}(S \cap Q_k) \leq \operatorname{diam} Q_k < \delta$, and therefore each $S \cap Q_k$ is contained in some U_{α_k} by the definition of δ. Since the $S \cap Q_k$'s cover S, we know that the U_{α_k}'s cover S, and hence $U_{\alpha_1}, \ldots, U_{\alpha_K}$ is the desired finite subcover. $\square$

The Arzela-Ascoli Theorem in $\mathbb{R}^d$

Any bounded sequence of real numbers has a convergent subsequence. However, the same is not true for sequences of functions, especially if "convergence" means uniform convergence. Consider the example $f_n(x) = \sin nx$. I claim that no subsequence converges uniformly on $[0, 2\pi]$. Suppose for the sake of contradiction that a subsequence $\{f_{n_j}\}_{j=1}^\infty$ converges uniformly on $[0, 2\pi]$. Then it is uniformly Cauchy, so there is an J such that

$$|f_{n_j}(x) - f_{n_k}(x)| < 1 \text{ for all } x \in [0, 2\pi], \text{ for all } j, k \geq J.$$

For a given $j \geq J$, there is an interval $(a, b) \subset [0, 2\pi]$ on which $\sin n_j x > 0$. Now choose k large enough that $2\pi/n_k < b - a$. Then (a, b) must contain a full period of $\sin n_k x$, and hence contains a point x_0 with $\sin n_k x = -1$. Then

$$|\sin n_j x_0 - \sin n_k x_0| > 1,$$

contradicting the above assertion that the subsequence is uniformly Cauchy.

Thus, for a sequence of functions to have a convergent subsequence, we will need stronger hypotheses than merely boundedness. We will assume that the sequence is *equicontinuous*, which means that for any $x_0 \in \mathbb{R}^d$ and $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|x - x_0| < \delta \text{ implies } |f_n(x) - f_n(x_0)| < \epsilon \text{ for all } n.$$

Equicontinuity of a sequence of functions is stronger than just continuity, because it requires that *the same* δ work for *all* of the functions. We can see from the above example that this does not happen automatically: The sequence $\sin nx$ is not equicontinuous; if we take $x_0 = 0$ and $\epsilon < 1$, then for any δ, there is an n such that $\sin nx = 1$ for some x with $|x - x_0| < \delta$, so we cannot get the same δ to work for all the f_n's.

We will also assume that the sequence $\{f_n\}$ is *pointwise bounded*, which means that for each x_0, the sequence of real numbers $\{f_n(x_0)\}_{n=1}^\infty$ is bounded.

Theorem 2 (Arzela-Ascoli)**.** *Suppose $\{f_n\}$ is a sequence of functions $\mathbb{R}^n \to \mathbb{R}$ which is equicontinuous and pointwise bounded. Then there exists a subsequence f_{n_j} which converges uniformly on compact sets to a continuous function f.*

"Uniform convergence on compact sets" means the following: For each compact set S and each $\epsilon > 0$, there exists a K such that

$$|f_{n_j}(x) - f_{n_k}(x)| < 1 \text{ for all } x \in S, \text{ for all } j, k \geq K.$$

Here, K depends on S and ϵ.

Our strategy will be construct a subsequence f_{n_j} such that $\{f_{n_j}(x)\}$ converges whenever x is a rational number. Then we will have to prove that in fact $f_{n_j}(x)$ converges for all real numbers, and that the convergence is uniform on compact sets. To do this, we will need the following lemma.

Lemma 2. *Let $\{f_n\}$ be an equicontinuous and pointwise bounded sequence of functions $\mathbb{R}^d \to \mathbb{R}$. If $S \subset \mathbb{R}^d$ is compact, then S is uniformly equicontinuous on S.*

"Uniformly equicontinuous on S" mean that for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|x - y| \text{ implies } |f_n(x) - f_n(y)| < \epsilon \text{ for all } x, y \in S, \text{ for all } n.$$

This is a stronger condition than equicontinuity: In equicontinuity, the δ is allowed to depend on ϵ and x, but not n. In uniform equicontinuity, the δ is allowed to depend on ϵ, but not n or x.

Proof of Lemma. The proof of the lemma is exactly the same as the proof that a continuous function is uniformly continuous on any compact set: Choose $\epsilon > 0$. By equicontinuity, for each z, there exists a δ_z such that

$$|z - y| < \delta_z \text{ implies that } |f_n(z) - f_n(y)| < \epsilon/2 \text{ for all } n.$$

The balls $\{B(z, \delta_z)\}_{z \in S}$ form an open cover of S. By the Lebesgue number lemma, there exists a δ such that any $E \subset S$ with $\operatorname{diam} E < \delta$ is contained in some open set from our cover. Then if $|x - y| < \delta$, the set $\{x, y\}$ has diameter less than δ, so x and y are both contained in $B(x, \delta_z)$ for some $z \in S$. This implies that

$$|f_n(x) - f_n(z)| < \epsilon/2 \text{ and } |f_n(y) - f_n(z)| < \epsilon/2 \text{ for all } n.$$

Hence, by the triangle inequality, $|f_n(x) - f_n(y)| < \epsilon$ for all n. Since x and y were arbitrary, we know that

$$|x - y| < \delta \text{ implies } |f_n(x) - f_n(y)| < \epsilon \text{ for all } x, y \in S, \text{ for all } n. \qquad \square$$

Proof of Arzela-Ascoli Theorem. We first find a subsequence f_{n_j} such that for each x with rational coordinates, the sequence of numbers $\{f_{n_j}(x)\}_{j=1}^{\infty}$ converges. The argument is a version of Cantor's famous "diagonalization" proof. The rational numbers can be listed in a sequence $\{a_k\}_{k=1}^{\infty}$. The "pointwise boundedness" assumption guarantees that $\{f_n(a_1)\}_{n=1}^{\infty}$ is a bounded sequence of real numbers, hence there is a subsequence $\{f_{n_j}\}$ such that $\{f_{n_j}(x)\}$ which converges. To avoid horrible notation later, denote this subsequence by $\{f_{1,j}\}_{j=1}^{\infty}$. Now note that $\{f_{1,j}(a_2)\}$ is a bounded sequence, and therefore, we can choose a subsequence $\{f_{2,j}\}_{j=1}^{\infty}$ such that $\{f_{2,j}(a_2)\}$ converges. We continue inductively: Once the subsequence $\{f_{k,j}\}_{j=1}^{\infty}$ has been chosen, we choose a subsequence of *that* called $\{f_{k+1,f}\}_{j=1}^{\infty}$ such that $\{f_{k+1}(a_k)\}$ converges.

The end result is a sequence of sequences,

$$\begin{aligned}
&f_{1,1}, f_{1,2}, f_{1,3}, \dots \\
&f_{2,1}, f_{2,2}, f_{2,3}, \dots \\
&f_{3,1}, f_{3,2}, f_{3,3}, \dots \\
&\dots
\end{aligned}$$

where each row is a subsequence of the previous row and they are all subsequences of $\{f_n\}$. Consider the diagonal sequence $\{f_{j,j}\}_{j=1}^{\infty}$. For each k, the sequence $\{f_{j,j}\}_{j=k}^{\infty}$ is a subsequence of $\{f_{k,j}\}_{j=k}^{\infty}$. Since $\{f_{k,j}(a_k)\}_{j=1}^{\infty}$ converges, we know $\{f_{j,j}(a_k)\}_{j=k}^{\infty}$ converges. This implies $\{f_{j,j}(a_k)\}_{j=1}^{\infty}$ converges (since this only depends on the behavior for $j \geq k$).

Therefore, $\{f_{j,j}\}$ is a subsequence of f_n which converges at all the rational points. Let's rename $\{f_{j,j}\}$ as $\{f_{n_j}\}$. We now have to prove that $\{f_{n_j}\}$ converges uniformly on each compact set S. Since S is contained in some ball B. Since we do not yet have a candidate for a limit function, we will do this by proving it is uniformly Cauchy on S. Choose $\epsilon > 0$. By the Lemma, there is a δ such that

$$|x - y| < \delta \text{ implies } |f_n(x) - f_n(y)| < \epsilon/3 \text{ for all } x, y \in S, \text{ for all } n.$$

The balls $\{B(y, \delta/2)\}_{y \in S}$ are an open cover of S, and thus, we can cover S by finitely many balls $\{B(y_\ell, \delta/2)\}_{\ell=1}^{L}$. Let z_ℓ be a rational number in $B(y_\ell, \delta/2)$. Since $\{f_{n_j}(z_\ell)\}$ converges, it is Cauchy, and there is a K_ℓ such that

$$|f_{n_j}(z_\ell) - f_{n_k}(z_\ell)| < \epsilon/3 \text{ for all } j, k \geq K_\ell.$$

Let $K = \max(K_1, \dots, K_L)$ (here it is essential that we had finitely many z_ℓ's). I claim that

$$|f_{n_j}(x) - f_{n_k}(x)| < \epsilon \text{ for all } x \in S \text{ for all } j, k \geq K.$$

Any $x \in S$ is contained in some $B(y_\ell, \delta/2)$, which implies $|x - z_\ell| \leq |x - y_\ell| + |z_\ell - y_\ell| < \delta$. Thus, by our choice of δ,

$$|f_n(x) - f_n(z_\ell)| < \epsilon/3 \text{ for all } n.$$

For $j, k \leq K$, we have

$$\begin{aligned}
|f_{n_j}(x) - f_{n_k}(x)| &\leq |f_{n_j}(x) - f_{n_j}(z_\ell)| + |f_{n_j}(z_\ell) - f_{n_k}(z_\ell)| + |f_{n_k}(z_\ell) - f_{n_k}(z)| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon
\end{aligned}$$

as claimed.

Thus, $\{f_{n_j}\}$ is uniformly Cauchy on S and therefore it converges uniformly on S to a continuous function. This implies that for any x, $\{f_{n_j}(x)\}$ converges (by taking $S = \{x\}$). Thus, there is a well-defined limit function $f(x) = \lim_{j\to\infty} f_{n_j}(x)$. On any compact set S, f_{n_j} converges uniformly to f. This f is continuous on any ball B because on $\overline{B}$ it is a uniform limit of continuous functions. $\square$

Remark: The same proof will work if f_n is defined on any open set $\Omega \subset \mathbb{R}^d$ rather than $\mathbb{R}^d$ (with slight modification).

Exercise 1. *Let $\{f_n\}$ be a sequence of continuous functions $\mathbb{R}^d$ and f a continuous function. Prove that $f_n \to f$ uniformly on compact sets, if and only if every subsequence of $\{f_n\}$ has a subsubsequence which converges to f uniformly on compact sets.*

Exercise 2 (Peano's existence theorem)**.** *Consider the differential equation $y' = F(t, y)$, where F is continuous on $[-a, a] \times [-b, b]$ for some $a, b > 0$. Prove that it has a solution $y(t)$ with $y(0) = 0$ defined in some neighborhood of the origin by the following steps:*

a. *Note F is bounded. If $|F(t, y)| \leq M$, then choose δ such that $M\delta \leq b$.*

b. *(Integral equation with a time delay): For each n, prove that there is a $y_n : [-\delta/n, \delta] \to \mathbb{R}$ with $y_n(0) = 0$ and*

$$y_n(t) = \int_0^t F(s - \delta/n, y(s - \delta/n))\, ds \text{ for } t \in [0, \delta].$$

Hint: Set $y_n(t) = 0$ for $t \in [-\delta/n, 0]$, then define y_n on $[0, \delta/n]$ by the integral equation above. Show that $|y_n(t)| \leq b/n$ on this interval. Next, extend y_n to $[\delta/n, 2\delta/n]$ and show it is bounded by $2b/n$ on this interval. Continue inductively.

c. *Show that the sequence $\{y_n\}$ is equicontinuous and pointwise bounded on $[0, \delta]$. Conclude there is a subsequence converging uniformly to some $y : [0, \delta] \to \mathbb{R}$.*

d. *Prove that y satisfies the integral equation:*

$$y(t) = \int_0^t F(s, y(s))\, ds \text{ for } t \in [0, \delta].$$

e. *By a symmetrical argument, construct a solution on $[-\delta, 0]$. Paste these solutions together to obtain a solution of the integral equation on $[-\delta, \delta]$ and conclude that it solves the differential equation.*

f. *Using the example $y' = y^{1/3}$, show that the solution is not necessarily unique.*

Metric Spaces

In $\mathbb{R}^d$, we had a distance function $d(x,y) = |x-y|$ that satisfies the following important properties:

- Nonnegativity: $d(x,y)$ is a nonnegative real number.
- Non-degeneracy: $d(x,y) = 0$ if and only if $x = y$.
- Symmetry: $d(x,y) = d(y,x)$.
- Triangle inequality: $d(x,z) \leq d(x,y) + d(y,z)$.

For many basic results about continuity, compactness, connectedness, etc., all we needed was the ability to measure distance. So we could have done the proofs on some abstract set in which all we assumed was that a distance function exists. So we make the following general definition: A *metric space* (X,d) is a set X with a distance function d that satisfies the four properties listed above.

Metric spaces are everywhere. Here are a few important examples:

- $X = \mathbb{R}^d$ is a metric space with $d(x,y) = |x-y|$.
- Take $X = C([0,1])$, the set of continuous functions $f : [0,1] \to \mathbb{R}$. We define $d(f,g) = \sup_{x\in[0,1]} |f(x)-f(y)|$. You should verify that this satisfies the triangle inequality and other properties for a distance function.
- Take $X = C([0,1])$, but with a different distance function: $d(f,g) = \int_0^1 |f-g|$. Verify that this is actually is a distance function.
- If X is a metric space, and $Y \subset X$, then Y is also a metric space with the same distance function (if $d : X \times X \to [0,\infty)$, then the distance function for Y is the restriction to $Y \times Y$).

Open and Closed Sets, Continuity

The ideas of balls, open sets, and continuity generalize to metric spaces. Suppose (X,d) is a metric space. If $x \in X$ and $r > 0$, we define $B(x,r) = \{y \in X : d(x,y) < r$. We say $U \subset X$ is *open* if for any $x \in U$, there exists $r > 0$ such that $B(x,r) \subset U$. Observe/verify the following properties:

- X and $\varnothing$ are open.
- Any union of open sets is open.
- A finite intersection of open sets is open.
- Balls are open. (The proof uses the triangle inequality.)

We say $C \subset X$ is *closed* if C^c is open.

If (X, d) and (Y, d') are metric spaces, we say $f : X \to Y$ is *continuous* if the following holds: If $V \subset Y$ is open, then $f^{-1}(V) \subset X$ is open. Equivalently, $f : X \to Y$ is continuous if for any $x_0 \in X$ and any $\epsilon > 0$, there exists a $\delta > 0$ such that $d(x, x_0) < \delta$ implies $d'(f(x), f(x_0)) < \epsilon$. (You should verify that these two definitions of continuity are equivalent; the proof is exactly the same as the proof for $X = \mathbb{R}^d$.)

Sequences, Completeness, and Compactness

If $\{x_n\}$ is a sequence in X, we say $x_n \to x$ if $d(x_n, x) \to 0$. If f is continuous and $x_n \to x$, then $f(x_n) \to f(x)$. As an exercise, prove

Proposition 1. *The following are equivalent:*

- $Y \subset X$ *is closed.*
- *If* $\{y_n\}$ *is a sequence in* Y *and* $y_n \to y_0$*, then* $y_0 \in Y$.

We say $\{x_n\}$ is Cauchy if for any $\epsilon > 0$, there exists N such that $d(x_n, x_m) < \epsilon$ for all $m, n \geq N$. In $\mathbb{R}^n$, we saw that all Cauchy sequences converge. However, this is not the case in general. For example, $\mathbb{Q}$ is a metric space with $d(x, y) = |x - y|$, but not every Cauchy sequence of rational numbers converges to a rational number. If it happens that every Cauchy sequence in X converges to a point in X, then we say X is *complete*. For example,

- The real numbers are complete.
- If X is complete, then $Y \subset X$ is complete if and only if it is closed. Here completeness of Y means that every Cauchy sequence in Y converges *to a point in* Y, not just a point in X.
- $C([0, 1])$ is complete in the metric $d(f, g) = \sup_{x \in [0,1]} |f(x) - f(y)|$. (See below.)
- However, it is not complete in the metric $d(f, g) = \int_0^1 |f - g|$. (You might have to think about this a while before you can find an example of a Cauchy sequence which does not converge to a continuous function.)

In $\mathbb{R}^d$, we showed that compact sets satisfy a covering property. In the general case, we will take this covering property as the definition: We $S \subset X$ is *compact* if every open cover of X has a finite subcover. Just as in $\mathbb{R}^d$, continuous functions map compact sets to compact sets.

Proposition 2. *Let* X *and* Y *be metric spaces and* $f : X \to Y$ *a continuous function. If* $S \subset X$ *is compact, then* $f(S)$ *is compact.*

Proof. Let $\{V_\alpha\}_{\alpha \in A}$ be an open cover of $f(S)$. Let $U_\alpha = f^{-1}(V_\alpha)$. Then $\{U_\alpha\}_{\alpha \in A}$ is an open cover of S, because for any $x \in S$, $f(x)$ must be in some V_α, and then $x \in U_\alpha$. Since S is compact, there is a finite subcover $U_{\alpha_1}, \ldots, U_{\alpha_K}$. Then $V_{\alpha_1}, \ldots, V_{\alpha_K}$ are a finite subcover of $f(S)$; for if $y \in f(S)$ and $y = f(x)$, then x is in some U_{α_k}, which means $f(x) \in V_{\alpha_k}$. $\square$

For any compact metric space X, let $C(X)$ be the space of continuous functions $X \to \mathbb{R}$ (or $X \to \mathbb{C}$ if you prefer), with the distance given by

$$d(f,g) = \|f-g\|_u = \sup_{x\in X} |f(x)-g(x)|.$$

The supremum must be achieved because $f-g$ is a continuous function on a compact set X, so the image must be a compact subset of $\mathbb{R}$. Thus, the distance function is well-defined. Now let's prove

Proposition 3. *$C(X)$ is complete.*

Proof. Suppose $\{f_n\}$ is a Cauchy sequence in $C(X)$. For any $x \in X$, $\{f_n(x)\}$ is a Cauchy sequence of real numbers because

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_u .$$

Thus, $f(x) = \lim_{n\to\infty} f_n(x)$ exists. Now we must show that $\|f_n - f\|_u \to 0$. Choose $\epsilon > 0$. There exists N such that

$$\|f_n - f_m\|_u \le \epsilon \text{ for all } n,m \ge N.$$

So for any x, $|f_n(x) - f_m(x)| \le \epsilon$ for $n,m \ge N$. Taking $m \to \infty$ shows that $|f_n(x) - f(x)| \le \epsilon$ for all $n \ge N$. Since x was arbitrary, we have $\|f_n - f\|_u \le \epsilon$. Thus, $f_n \to f$ uniformly. Finally, we must show f is continuous. Fix $x_0 \in X$ and $\epsilon > 0$. Choose n large enough that $\|f_n - f\|_u < \epsilon/3$. Since f_n is continuous, there exists a δ such that

$$d(x,x_0) < \delta \text{ implies } |f_n(x) - f_n(x_0)| < \epsilon/3.$$

By the triangle inequality, $d(x,x_0)$ implies

$$\begin{aligned}|f(x) - f(x_0)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &< \epsilon.\end{aligned}$$

Thus, f is continuous and $f_n \to f$ in the $C(X)$ metric. $\square$

Generalization of the Heine-Borel Theorem

S is *sequentially compact* if every sequence in S has a subsequence converging to a point in S. As in $\mathbb{R}^d$, compactness and sequential compactness are equivalent. However, a closed and bounded set might not be compact, for two reasons.

First, if our metric space X is not complete, we could have a Cauchy sequence $\{x_n\}$ contained in S which does not converge to anything; in that case, it would not be possible to find a convergent subsequence, so S would not be sequentially compact. Thus, we need to assume that S is complete even if the larger metric space isn't.

Second, the proof of the Heine-Borel theorem used the fact that any bounded set in $\mathbb{R}^d$ can be covered by finitely many cubes of diameter $< \delta$. In other metric spaces, a bounded set might not be coverable by finitely many sets of diameter $< \delta$. So we use a stronger condition called *total boundedness*: S is *totally bounded* if for any $r > 0$, S can be covered by finitely many balls of radius r. The center of each ball is assumed to be in S. We can now generalize the Heine-Borel theorem:

Theorem 3. *Let X be a metric space and $S \subset X$. The following are equivalent:*

a. S is compact.

b. S is complete and totally bounded.

c. S is sequentially compact.

Proof. (a) implies (b). First, we show S is complete. Suppose $\{x_n\}$ is a Cauchy sequence in S. Let

$$\delta_n = \sup_{m \geq n} d(x_m, x_n).$$

Since the sequence is Cauchy, δ_n must be finite and approach zero as $n \to \infty$ (verify). Let

$$O_n = \{x \in X : d(x, x_n) > \delta_n\}.$$

This set is open (verify). Any finite subcollection $O_{n_1}, \ldots, O_{n_k}$ cannot cover S; for if $N = \max(n_1, \ldots, n_k)$, then we have $N \geq n_k$ for each k, and hence

$$d(x_{n_k}, x_N) \leq \sup_{m \geq n_k} d(x_m, x_{n_k}) = \delta_{n_k}.$$

But $d(x_N, x_{n_k}) \leq \delta_{n_k}$ means that $x_N \notin O_{n_k}$, but $x_N \in S$. Thus, $O_{n_1}, \ldots, O_{n_K}$ do not cover S. Therefore, $\{O_n\}_{n=1}^{\infty}$ must not be a cover of S, since otherwise we could find a finite subcover $O_{n_1}, \ldots, O_{n_K}$, which we just proved was impossible. So there is an $x_0 \in S$ which is not in any of the O_n's. This implies that $d(x_0, x_n) \leq \delta_n \to 0$, and therefore $x_n \to x_0$.

Now we show S is totally bounded. Choose $r > 0$. Then $\{B(x, r)\}_{x \in S}$ is an open cover of S, so there is a finite subcover $B(x_1, r), \ldots, B(x_K, r)$. So S can be covered by finitely many balls of radius r.

(b) implies (c). Let $\{x_n\}$ be a sequence in S. We construct a Cauchy subsequence using a similar diagonalization argument to the proof of Arzela-Ascoli earlier. We define subsequences $\{x_{k,n}\}_{n=0}^{\infty}$ by induction on k. Let $x_{0,n} = x_n$. Now suppose $\{x_{k,n}\}$ has been chosen. By (b), we can cover S by finitely many balls of radius $1/k$. Since there are only finitely many balls, at least one of them must contain $x_{k,n}$ for infinitely many values of n. Call it $B(y_k, 1/k)$, and let $\{x_{k+1,n}\}$ be the subsequence of $\{x_{k,n}\}$ consisting of all the points in $B(y_k, 1/k)$. Then let $x_{n_j} = x_{j,j}$. To show $\{x_{n_j}\}$ is Cauchy, choose $\epsilon > 0$, and let $2/K < \epsilon$. For $j, k \geq K$, we know $x_{n_j} = x_{j,j}$ is an element of the sequence $\{x_{j,n}\}_{n=1}^{\infty}$ and the same holds for x_{n_k}. Thus, x_{n_j} and x_{n_k} are both in $B(y_K, 1/K)$, and so by the triangle inequality

$$d(x_{n_j}, x_{n_k}) < 2/K < \epsilon \text{ for all } j, k \geq K.$$

So $\{x_{n_j}\}$ is Cauchy and converges to a point in S by (b).

(c) implies (a). Let $\{U_\alpha\}$ be an open cover. As before, our strategy is to prove first that there is a Lebesgue number δ such that any set of diameter less than δ is contained in a single U_α, and second that S can be covered by finitely many sets of diameter less than δ.

First, let's show $\{U_\alpha\}$ has a Legesgue number δ. Suppose not. Then there exists E_n with $\operatorname{diam} E_n < 1/n$ such that E_n is not contained in any U_α. Let $x_n \in E_n$. By (c), there is a convergent subsequence $x_{n_j} \to x_0 \in S$. Then x_0 is in some U_{α_0}, and there is a ball $B(x_0, r) \subset U_{\alpha_0}$. Choose j large enough that $1/n_j < r/2$ and $d(x_{n_j}, x_0) < r/2$. Then if $y \in E_{n_j}$, we have

$$d(y, x_0) \leq d(y, x_{n_j}) + d(x_{n_j}, x_0) < 1/n_j + r/2 < r,$$

and therefore, $y \in B(x_0, r)$ and $y \in U_{\alpha_0}$. This shows $E_{n_j} \subset U_{\alpha_0}$ contrary to our assumption. So there is a Lebesgue number for $\{U_\alpha\}$.

Next, we show that S can be covered by finitely many balls of radius $\epsilon < \delta/2$. Suppose not. Choose some $x_1 \in S$. By assumption, $B(x_1, \epsilon)$ does not cover S, so we can choose an $x_1 \in S \backslash B(x_1, \epsilon)$. We construct a sequence $\{x_n\}$ inductively: Once $x_1, \ldots, x_n$ are chosen, S is not covered by $B(x_1, \epsilon), \ldots, B(x_n, \epsilon)$, so there is an $x_{n+1} \in S$ that is not in any of these balls. Then if $j < k$, we have $d(x_j, x_k) \geq \epsilon$ since $x_k \notin B(x_j, \epsilon)$. Therefore, $\{x_n\}$ cannot have a convergent subsequence, a contradiction.

Therefore, S can be covered by finitely many balls $B(x_1, \epsilon), \ldots, B(x_K, \epsilon)$. Since $\operatorname{diam} B(x_k, \epsilon) \leq 2\epsilon < \delta$, we know each $B(x_k, \epsilon)$ is contained in some U_{α_k}. Therefore, $U_{\alpha_1}, \ldots, U_{\alpha_K}$ is the desired finite subcover of S. $\square$

Generalized Arzela-Ascoli Theorem

The Arzela-Ascoli theorem generalizes to continuous functions on a compact metric space X. The definitions of equicontinuous and pointwise bounded are exactly the same as for functions on $\mathbb{R}^d$, and we could state a version of the Arzela-Ascoli theorem that sounds exactly the same as Theorem 2. However, here we will look at the theorem more topologically–as a characterization of compact subsets of $C(X)$. The proof deals with the three metric spaces X, $\mathbb{R}$, and $C(X)$, and roughly speaking relates balls in $C(X)$ to balls in X and $\mathbb{R}$.

Theorem 4. *Let X be a compact metric space, and let $\mathcal{F} \subset C(X)$.*

1. *$\mathcal{F}$ is totally bounded if and only if it is equicontinuous and pointwise bounded.*

2. *$\mathcal{F}$ is compact if and only if it is closed, equicontinuous, and pointwise bounded.*

3. *Any equicontinuous and pointwise bounded sequence in $C(X)$ has a convergent subsequence.*

4. *Any convergent sequence in $C(X)$ is equicontinuous and pointwise bounded.*

Proof. (1). Suppose $\mathcal{F}$ is totally bounded. We can choose finitely many $C(X)$-balls $B(f_1, 1), \ldots, B(f_K, 1)$ which cover $\mathcal{F}$. Let $M = \max(\|f_1\|_u, \ldots, \|f_k\|_u)+1$. Any $f \in \mathcal{F}$ is in contained in some $B(f_k, 1)$, which implies

$$\|f\|_u \leq \|f_k\|_u + \|f - f_k\|_u \leq M.$$

So $|f(x)| \leq M$ for all x, and $\{f(x)\}_{f \in \mathcal{F}}$ is bounded. (This proof in fact shows $\mathcal{F}$ is *uniformly bounded* since the M did not depend on x.)

To prove $\mathcal{F}$ is equicontinuous, choose $x_0 \in X$ and $\epsilon > 0$. By assumption, $\mathcal{F}$ can be covered by finitely many balls $B(f_1, \epsilon/3), \ldots, B(f_K, \epsilon/3)$. For each f_k, there is a δ_k such that

$$d(x, x_0) < \delta_k \text{ implies } |f_k(x) - f_k(x_0)| < \epsilon/3.$$

Let $\delta = \min(\delta_1, \ldots, \delta_K)$. Any $f \in \mathcal{F}$ is contained in some $B(f_k, \epsilon/3)$, which implies $\|f_k - f\| < \epsilon/3$. If $d(x, x_0) < \delta \leq \delta_k$, then we have

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(x_0)| + |f_k(x_0) - f(x_0)| \\ &\leq \|f - f_k\|_u + |f_k(x) - f_k(x_0)| + \|f - f_k\|_u \\ &< \epsilon. \end{aligned}$$

Since the same δ works for all f, we have equicontinuity.

Conversely, suppose $\mathcal{F}$ is equicontinuous and pointwise bounded, and we prove it is totally bounded. Choose $r > 0$. For each $x \in X$, there is an δ_x such that $d(y, x) < \delta_x$ implies that $|f(y) - f(x)| < r/4$. Then $\{B(x, \delta_x)\}_{x \in X}$ is an open cover of X, so there is a finite subcover $B(x_1, \delta_1), \ldots, B(x_J, \delta_J)$. By pointwise boundedness, we know $\{f(x_j)\}_{f \in \mathcal{F}}$ is bounded for each n. Say $|f(x_j)| \leq M_j$, and let $M = \max(M_1, \ldots, M_J)$. Now $[-M, M]$ is a bounded set in $\mathbb{R}$ and hence can be covered by finitely many balls $B(y_1, r/4), \ldots, B(y_K, r/4)$.

Let $A = \{x_1, \ldots, x_J\}$ and $B = \{y_1, \ldots, y_K\}$. Then B^A (the set of functions from A to B) is a finite set because A and B are finite sets. We will cover $\mathcal{F}$ by finitely many $C(X)$-balls indexed by B^A. If $\phi : A \to B$, then let

$$\mathcal{F}_\phi = \{f \in \mathcal{F} : f(x_j) \in B(\phi(x_j), r/4) \text{ for each } j\}.$$

(Here $\phi(x_j)$ one of the y_k's.) Every $f \in \mathcal{F}$ is in one of the $\mathcal{F}_\phi$'s because $f(x_j)$ must be in some $B(y_k, r/4)$ for each j. Choose one f_ϕ from each $\mathcal{F}_\phi$. Suppose $f \in \mathcal{F}_\phi$. Any x is contained in some $B(x_j, r_j)$, which implies

$$|f(x) - f(x_j)| < r/4, \qquad |f_\phi(x) - f_\phi(x_j)| < r/8.$$

But $f(x) \in B(\phi(x_j), r/4)$ and $f_\phi(x) \in B(\phi(x_j), r/4)$, which implies

$$|f(x) - \phi(x_j)| < r/4, \qquad |f_\phi(x) - \phi(x_j)| < r/4.$$

By the triangle inequality, $|f(x) - f_\phi(x)| < r$, and therefore $\|f - f_\phi\|_u < r$ (the inequality remains strict because the maximum of $|f - f_\phi|$ is achieved). Therefore, $f \in B(f_\phi, r)$. And since f was arbitrary, $\mathcal{F}_\phi \subset B(f_\phi, r)$. Therefore, the balls $\{B(f_\phi, r)\}_{\phi \in B^A}$ cover $\mathcal{F}$, so $\mathcal{F}$ can be covered by finitely many balls of radius r, which completes (1).

(2) now follows easily. By Theorem 3, $\mathcal{F}$ is compact if and only if it is complete and totally bounded. Because $C(X)$ is complete, $\mathcal{F}$ is closed if and only if it is complete. And by part (1), it is totally bounded if and only if it equicontinuous and pointwise bounded.

(3). Let $\{f_n\}$ be an equicontinuous and pointwise bounded sequence. Let $S \subset C(X)$ be the set of f_n's. By (1), S is totally bounded, and this implies $\overline{S}$ is totally bounded (exercise). So $\overline{S}$ is compact by (2). So any sequence in $\overline{S}$ has a convergent subsequence and hence $\{f_n\}$ has a convergent subsequence.

(4). Let $\{f_n\}$ be a sequence in $C(X)$ converging to f, and let S be the set of f_n's together with f. I claim S is totally bounded. Choose $r > 0$. There is an N such that $\|f_n - f\|_u < r$ for $n \geq N$, so $f_n \in B(f, r)$. Then the finitely many balls $B(f_1, r), B(f_2, r), \ldots, B(f_N, r)$ and $B(f, r)$ cover S. So S is totally bounded, hence equicontinuous and pointwise bounded by (1). $\square$

Exercise 3. *Let X be a compact metric space, Y a metric space, and $f : X \to Y$ continuous. Prove that f is uniformly continuous.*

Exercise 4. *By adapting the proof of (1), show that any totally bounded $\mathcal{F} \subset C(X)$ is uniformly equicontinuous. (The δ does not depend on x.)*

Exercise 5. *Prove (4) directly from the definition without using (1) or (2).*