## 42. ARZELÀ-ASCOLI THEOREM

We will now move to the most well-known conclusion of the topic of equicontinuity known as the Arzelà-Ascoli Theorem.

## 42.1 Selection of a pointwise-convergent subsequence.

Notice that all of our previous conclusions assumed a sequence of functions that converges in a pointwise sense; the work focused on upgrading the mode of convergence to uniform. However, what if we do not know that the sequence converges pointwise to begin with? Is there a way to resort to a convergent subsequence? As it turns out, the answer is positive under relatively mild assumptions. We start with:

**Lemma 42.1** (A selection principle) Let A be a (non-empty) finite or countable set and let  $\{f_n\}_{n \in \mathbb{N}}$  be a sequence of functions  $A \to X$  (with domain equal to A) where X is a metric space with metric denoted by  $\rho$ . Then

$$(X,\rho) \text{ compact } \Rightarrow \exists \{n_k\}_{k\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}} \colon \left(n_k \to \infty \land \forall x \in A \colon \lim_{k \to \infty} f_{n_k}(x) \text{ exists}\right) \quad (42.1)$$

*Proof.* This is proved by a repeated use of Cantor's diagonal argument. Indeed, repeating points if necessary, let  $\{x_j\}_{j\in\mathbb{N}} \in A^{\mathbb{N}}$  contains all points of A. The compactness of X ensures the existence of a subsequence  $\{n_k^{(0)}\}_{k\in\mathbb{N}}$  such that  $\{f_{n_k^{(0)}}(x_0)\}_{k\in\mathbb{N}}$  converges, which then contains another subsequence  $\{n_k^{(1)}\}_{k\in\mathbb{N}}$  such that also  $\{f_{n_k^{(1)}}(x_1)\}_{k\in\mathbb{N}}$  converges. Taking the diagonal sequence  $\{\hat{n}_k\}_{k\in\mathbb{N}}$  defined by  $\hat{n}_k := n_k^{(k)}$  which, except for a finite number of terms, is a subsequence of all of the above subsequences we get that  $\{f_{\hat{n}_k}(x_j)\}_{k\in\mathbb{N}}$  converges for all  $j \in \mathbb{N}$ . Since  $\hat{n}_k \ge k$ , we get the claim.

In order to write this formally and at the same time efficiently, we will use the following observation: For  $f, g: A \rightarrow X$  let

$$d(f,g) := \sum_{j=0}^{\infty} 2^{-j} \rho_X(f(x_j), g(x_j))$$
(42.2)

where the sum converges because *X*, being compact, has finite diameter — in fact,  $d(f,g) \in [0, M]$  where  $M := 2 \operatorname{diam}(X)$ . The idea behind this definition is that, for any sequence of functions  $\{h_n\}_{n \in \mathbb{N}} \in (X^A)^{\mathbb{N}}$  and any function  $h \in X^A$ ,

$$h_n \to h$$
 pointwise on  $A \Leftrightarrow d(h_n, h) \to 0$  (42.3)

whose checking we will leave to the reader. As *d* satisfies the properties of the metric, it metrizes pointwise convergence — i.e., casts it in the form of metric space convergence — on the space of functions  $A \rightarrow X$ .

We now claim

$$(X, \rho_X)$$
 totally bounded  $\Rightarrow (X^A, d)$  totally bounded (42.4)

Indeed, our goal is to find, for each r > 0, a finite collection of open balls of *d*-radius r in  $X^A$  that cover the whole space. Let  $k \in \mathbb{N}$  be such that  $2^{-k}M < r/2$ . We start by noting that, since  $\{B(y,r): y \in X\}$  cover X, there exists  $n \in \mathbb{N}$  and  $y_0, \ldots, y_n \in X$  such

Preliminary version (subject to change anytime!)

that  $X = \bigcup_{i=0}^{n} B(y_i, r/4)$ . Now consider the following set of functions

$$\mathcal{F} := \left\{ f \in X^A \colon \left( \forall j \le k \colon f(x_j) \in \{y_0, \dots, y_n\} \right) \land \left( \forall j > k \colon f(x_j) = y_0 \right) \right\}$$
(42.5)

Given a function  $g \in X^A$  and  $j \in \{0, ..., k\}$ , set  $i_j := \min\{i \leq n : g(x_j) \in B(y_i, r/2) \text{ and}$ then take  $f \in \mathcal{F}$  such that  $\forall j = 0, ..., k : f(x_j) = y_{i_j}$ . This *f* then obeys

$$\forall j = 0, \dots, k: \ \rho_X\big(g(x_j), f(x_j)\big) < r/4 \tag{42.6}$$

and so we have

$$d(f,g) \leq \sum_{j=0}^{k} 2^{-j} r/4 + \sum_{j=k+1}^{\infty} 2^{-j} M \leq \frac{1}{2} r + 2^{-k} M < r$$
(42.7)

It follows that  $X^A = \bigcup_{f \in \mathcal{F}} \{g \in X^A : d(g, f) < r\}$  proving that  $X^A$  is totally bounded.

Since  $(X, \varrho)$  is also closed, we readily check that also  $(X^A, d)$  is closed. The claim now follows from the sequential characterization of compactness.

We remark that the previous proof touches on a deep result in topology, called the *Tychonoff Theorem*, which states the an arbitrary product of compact topological spaces is compact in the product topology. For countable products of metric spaces, the product topology is metrized by a metric of the kind (42.2) and so our conclusion can indeed be regarded as a corollary to the Tychonoff Theorem.

It is also worth noting that the assumption that *A* is at most countable is essential for the "selection" of the subsequence in Lemma 42.1 to be possible. Indeed, consider the following example. Let  $f_n: \{0,1\}^{\mathbb{N}} \to \{0,1\}$  be defined by

$$f_n(\{\sigma_i\}_{i=0}^\infty) := \sigma_n. \tag{42.8}$$

Endow {0,1} with discrete metric. Given a strictly increasing sequence  $\{n_k\}_{k \in \mathbb{N}} \in \mathbb{N}$ , let  $\tilde{\sigma} \in \{0,1\}^{\mathbb{N}}$  defined by

$$\tilde{\sigma}_n := \begin{cases} 1, & \text{if } n \in \{n_{2k} \colon k \in \mathbb{N}\}, \\ 0, & \text{else.} \end{cases}$$
(42.9)

Then  $\{f_{n_k}(\tilde{\sigma})\}_{k\in\mathbb{N}}$  oscillates between zero and one and hence does NOT converge. As such a  $\tilde{\sigma}$  exists for every subsequence, it follows that the sequence  $\{f_n\}_{n\in\mathbb{N}}$  does NOT admit *any* pointwise-convergent subsequence whatsoever.

Clearly, the same example works even if *A* just *contains* a copy of  $\{0, 1\}^{\mathbb{N}}$ , which is true whenever *A* is of the cardinality of the continuum or larger, and *X* contains at least two points. (Indeed, extend  $f_n$  by the "zero" value outside the copy of  $\{0, 1\}^{\mathbb{N}}$  while noting that the restriction of any metric on *X* to  $\{0, 1\}$  is equivalent to the discrete metric.) This is related to the fact that on uncountable spaces, pointwise convergence cannot be phrased via a metric. We state and prove this in:

**Lemma 42.2** Let A be an uncountable set and X a metric space with at least two points. Then the pointwise convergence on the space of functions  $A \to X$  is not metrizable. More precisely, there exists no metric d on  $X^A$  such that  $d(f_n, f) \to 0$  is equivalent to  $f_n \to f$  pointwise. *Proof.* Let *A* be an uncountable set. Since *X* is contains at least two points, it suffices to treat the case that  $X = \{0, 1\}$  endowed with the discrete metric. The proof relies on the Axiom of Choice that can be avoided if *A* and  $\{0, 1\}^A$  can be totally ordered.

Writing  $\underline{0}$  for the function equal to zero everywhere, suppose, for the sake of contradiction, that there does exist a metric d on  $\{0,1\}^A$  such that  $d(f_n,\underline{0}) \to 0$  is equivalent to  $f_n \to \underline{0}$  pointwise. Let  $B_k := \{f \in \{0,1\}^A : d(f,\underline{0}) < 1/k\}$  be the open ball of radius 1/kcentered at  $\underline{0}$ . Then also

$$f_n \to \underline{0}$$
 pointwise  $\Leftrightarrow \forall k \in \mathbb{N} \colon \{n \in \mathbb{N} \colon f_n \notin B_k\}$  is finite (42.10)

Next we claim that

$$\forall x \in A: \left\{ n \in \mathbb{N}: B_n \subseteq \left\{ h \in \{0,1\}^A: h(x) = 0 \right\} \right\} \neq \emptyset$$
(42.11)

Indeed, if this failed for some x, then (using the Axiom of Choice) we could pick  $\{f_n\}_{n \in \mathbb{N}}$  such that  $\forall n \in \mathbb{N} : f_n \in B_n \land f_n(x) = 1$ . But that would contradict (42.10) because, by the fact that the  $B_i$ 's are nested,  $\Leftarrow$  there would force  $f_n \rightarrow \underline{0}$  pointwise which cannot hold simultaneously with  $f_n(x) = 1$  for all  $n \in \mathbb{N}$ .

For each  $x \in A$ , the above shows that

$$n_0(x) := \inf \left\{ n \in \mathbb{N} \colon B_n \subseteq \left\{ h \in \{0, 1\}^A \colon h(x) = 0 \right\} \right\}$$
(42.12)

obeys  $n_0(x) \in \mathbb{N}$ . Since *A* is uncountable and  $n_0$  takes values in  $\mathbb{N}$  which is countable, there must exist  $k \in \mathbb{N}$  such that

$$\{x \in A \colon n_0(x) = k\} \text{ is infinite}$$

$$(42.13)$$

Taking the smallest *k* with this property, we can then find (again, using the Axiom of Choice) a sequence  $\{x_i\}_{i \in \mathbb{N}}$  such that  $\forall i \in \mathbb{N}$ :  $n_0(x_i) = k$  which means

$$\forall i \in \mathbb{N} : \ B_k \subseteq \{h \in \{0, 1\}^A : h(x_i) = 0\}$$
(42.14)

and such that the  $x_i$ 's are distinct, meaning  $\forall i, j \in \mathbb{N} : i \neq j \Rightarrow x_i \neq x_j$ . Now define functions  $\{f_n\}_{n \in \mathbb{N}}$  by

$$f_n(x) := \begin{cases} 1, & \text{if } x \in \{x_i : i \ge n\}, \\ 0, & \text{else.} \end{cases}$$
(42.15)

Then (by the fact that all the  $x_i$ 's are distinct)  $f_n \to \underline{0}$  pointwise and so, by (42.10) and the fact that the  $B_j$ 's are nested, there is  $i_0 \in \mathbb{N}$  such that  $\forall i \ge i_0$ :  $f_i \in B_k$ . But (42.14) then forces  $f_i(x_i) = 0$  for all  $i \ge i_0$ , in contradiction with (42.15). No such metric d may thus exist after all.

In a homework assignment, I claimed that the above can be proved from the fact that, in metric spaces, a sequence  $\{x_n\}_{n\in\mathbb{N}}$  converges if and only if, for some x, every subsequence of  $\{x_n\}_{n\in\mathbb{N}}$  contains a subsubsequence converging to x. Unfortunately, this equivalence does hold for pointwise convergence of sequences of functions and so, to get the desired conclusion, one has to use a different argument. (The stated characterization does rule out pointwise convergence of functions  $\mathbb{R} \to \mathbb{R}$  almost everywhere — which means: except on a set of zero length.)

## 42.2 Arzelà-Ascoli Theorem.

Having discussed the pointwise convergence at length, we now return to the main line of presentation and state and prove:

**Theorem 42.3** (Arzelà-Ascoli Theorem) Let X and Y be metric spaces and  $\{f_n\}_{n \in \mathbb{N}}$  functions  $X \to Y$ . Assume

X compact 
$$\land$$
 Y compact  $\land$   $\{f_n\}_{n \in \mathbb{N}}$  equicontinuous (42.16)

Then

$$\exists \{n_k\}_{k \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}} \exists f \in Y^X \colon n_k \to \infty \land f_{n_k} \to f \text{ uniformly}$$
(42.17)

*Proof.* The proof is short as it builds on the various facts proved earlier. First, *X* being compact implies that *X* is separable, which means that there exists  $A \subseteq X$  countable and dense. Thanks to the compactness of *Y*, Lemma 42.1 shows that there is  $\{n_k\}_{k\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$  with  $n_k \to \infty$  such that  $\{f_{n_k}(x)\}_{k\in\mathbb{N}}$  converges for all  $x \in A$ . But  $\{f_n\}_{n\in\mathbb{N}}$  being equicontinuous along with *X* being compact show that  $\{f_n\}_{n\in\mathbb{N}}$  is uniformly equicontinuous (see Lemma 41.10) and Lemma 41.7 then shows that  $\{f_n\}_{n\in\mathbb{N}}$  admits a pointwise limit *f* everywhere and  $f_n \to f$  uniformly.

The Arzelà-Ascoli Theorem is a very useful tool in various applications of analysis. For instance, it can be used to prove Peano's Theorem on the existence of solutions to first-order ordinary differential equations. We will demonstrate its use on the following problem from variational calculus.

Given a continuous  $f: [a, b] \to \mathbb{R}$  which is differentiable on (a, b), set

$$\Phi(f) := \int_{a}^{b} \left[ f'(x)^{2} + V(f(x)) \right] \mathrm{d}x$$
(42.18)

where  $V : \mathbb{R} \to \mathbb{R}$  is a continuous function. The map  $f \mapsto \Phi(f)$  assigns a real number to a function and is often referred to as a *functional*. (One interpretation of the above  $\Phi$  is the energy of a vibrating string or a power-line suspended between two poles.)

Denoting

$$\mathcal{F} := \left\{ f \in \mathbb{R}^{[a,b]} : \text{ continuous on } [a,b] \land \text{ differentiable on } (a,b) \right\}$$
(42.19)

under the assumption that V is bounded from below,

$$\inf\{\Phi(f)\colon f\in\mathcal{F}\}\tag{42.20}$$

exists in  $\mathbb{R}$ . The question that one often needs to resolve is: Does there exist a minimizing  $f \in \mathcal{F}$ ? A natural first attempt in this vain is to consider a "minimizing sequence" — i.e., a sequence  $\{f_n\}_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$  such that  $\Phi(f_n)$  tends to the infimum — and extract, by whatever means, a limit thereof. The following lemma is then useful:

**Lemma 42.4** Assume  $V: \mathbb{R} \to \mathbb{R}$  continuous and bounded from below and let  $\{f_n\}_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$ be such that  $\sup_{n \in \mathbb{N}} |f_n(a)| < \infty$  and  $\sup_{n \in \mathbb{N}} \Phi(f_n) < \infty$ . Then  $\{f_n\}_{n \in \mathbb{N}}$  is equicontinuous and so there exists  $\{n_k\}_{k \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$  and  $f: [a, b] \to \mathbb{R}$  such that

$$n_k \to \infty \land f_{n_k} \to f \text{ uniformly}$$

$$(42.21)$$

Preliminary version (subject to change anytime!)

Typeset: June 1, 2023

*Proof.* Note that

$$\forall n \in \mathbb{N} \colon \int_{a}^{b} f'_{n}(x)^{2} \mathrm{d}x = \Phi(f_{n}) - \int_{a}^{b} V(f(x)) \mathrm{d}x \leq \Phi(f_{n}) - \inf_{z \in \mathbb{R}} V(z)$$
(42.22)

From the assumption that  $\sup_{n \in \mathbb{N}} \Phi(f_n) < \infty$  we thus get

$$C := \sup_{n \in \mathbb{N}} \int_{a}^{b} f'_{n}(x)^{2} \mathrm{d}x < \infty$$
(42.23)

But then the Fundamental Theorem of Calculus along with the Cauchy-Schwarz inequality show

$$\left|f_n(x) - f_n(y)\right| = \left|\int_x^y f'(t)dt\right| \le \left(\int_x^y 1dt\right)^{1/2} \left(\int_x^y f'(t)^2 dt\right)^{1/2} \le \sqrt{C}|x-y|^{1/2}$$
(42.24)

thus proving that  $\{f_n\}_{n \in \mathbb{N}}$  is (uniformly) equicontinuous. In light of the assumption  $\sup_{n \in \mathbb{N}} |f_n(a)| < \infty$  we we also get

$$M := \sup_{n \in \mathbb{N}} \sup_{x \in [a,b]} |f_n(x)| \leq \sup_{n \in \mathbb{N}} |f_n(a)| + \sqrt{C} |b-a|^{1/2} < \infty$$

$$(42.25)$$

The functions thus effectively map the compact set [a, b] into the compact set [-M, M]. By the Arzelà-Ascoli Theorem, there exists a uniformly convergent subsequence.

Finding the uniformly convergent minimizing convergence  $f_{n_k} \to f$  is not the end of the story; indeed, one then needs to show that, in fact,  $f \in \mathcal{F}$  and that  $\Phi(f_{n_k}) \to \Phi(f)$  which requires more tricks from metric space theory. Indeed, we need the concept of weak convergence which hinges on the following variant of a selection principle:

**Lemma 42.5** Let a < b be reals and  $\{h_n\}_{n \in \mathbb{N}}$  a uniformly bounded sequence of continuous functions  $h_n: [a, b] \to \mathbb{R}$ . Then there exists a strictly increasing  $\{n_k\}_{k \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$  such that

$$\forall g \in C([a,b]): \lim_{k \to \infty} \int_{a}^{b} g(t)h_{n_{k}}(t)dt \text{ exists}$$
 (42.26)

*Proof.* For each  $g \in C([a, b])$  define  $\phi_n \colon C([a, b]) \to \mathbb{R}$  by

$$\phi_n(g) := \int_a^b g(t)h_n(t)\mathrm{d}t \tag{42.27}$$

The estimates for the integral ensure

$$\left|\phi_{n}(g)\right| \leq \left(\sup_{x \in [a,b]} \left|h_{n}(x)\right|\right) \left(\sup_{x \in [a,b]} \left|g(x)\right|\right)$$
(42.28)

and

$$\left|\phi_{n}(g) - \phi_{n}(\tilde{g})\right| \leq \left(\sup_{x \in [a,b]} \left|h_{n}(x)\right|\right) \left(\sup_{x \in [a,b]} \left|g(x) - \tilde{g}(x)\right|\right)$$
(42.29)

Since  $\{h_n\}_{n \in \mathbb{N}}$  are uniformly bounded, we conclude that the family of functions  $\{\phi_n\}_{n \in \mathbb{N}}$  is uniformly bounded and also Liptschitz and thus uniformly equicontinuous (as functions on the metric space  $C([0, 1], \mathbb{R})$ ).

The space  $C([a, b], \mathbb{R})$  is separable (which can be shown by approximating continuous functions using piecewise linear functions taking rational values at rationals) and so Lemma 42.1 gives existence of a sequence  $n_k \rightarrow \infty$  such that  $\phi_{n_k}$  converges pointwise

on a dense subset of  $C([a, b], \mathbb{R})$ . Lemma 41.7 (along with the fact that  $\mathbb{R}$  is complete) extends the pointwise convergence to all of  $C([a, b], \mathbb{R})$ .

We now say:

**Definition 42.6** (Weak convergence) Let a < b be reals and let  $\{h_n\}_{n \in \mathbb{N}}$  be a uniformly bounded sequence of continuous functions  $h_n: [a, b] \to \mathbb{R}$ . Let  $h: [a, b] \to \mathbb{R}$ . We say that  $h_n \to h$  weakly if h is Riemann integrable and

$$\forall g \in C([a,b]): \lim_{n \to \infty} \int_{a}^{b} g(t)h_{n}(t)dt = \int_{a}^{b} g(t)h(t)dt$$
(42.30)

Of particular interest is the situation under which such a "weak-limit" function h exists. This can be guaranteed by varying g to approximate a step function  $1_{[a,x]}$  which is the key argument underlying another beautiful result from this area of mathematics: the *Riesz representation theorem*. We will not have time to discuss this in any level of detail but let us just say that it is here where the Stieltjes integration becomes extremely useful.

As to the above optimization problem, the concept of weak convergence is actually applied to the sequence of derivatives  $\{f'_n\}_{n\in\mathbb{N}}$  of the minimizing sequence  $\{f_n\}_{n\in\mathbb{N}}$ . The point of having the weak limit of the derivatives rests in the observation that, using additional arguments, one can show that the derivative of the limit function f actually exists and coincides with the weak limit of  $\{f'_n\}_{n\in\mathbb{N}}$ . This then produces an actual minimizer of  $f \mapsto \Phi(f)$ .

## 42.3 Helly's selection theorem.

We will continue talking about uniform convergence in the next sections. The take-home message of the above developments is that, while the Cantor diagonal argument often permits us to extract a subsequential limit on a countable dense set, further regularity (such as uniform equicontinuity) is needed to extend the convergence to the whole space and, under compactness, to uniform convergence.

Alternatives to this do exists which we demonstrate by stating a theorem that does not conform to this scheme. Indeed, the regularity is not supplied by continuity but rather by monotonicity:

**Theorem 42.7** (Helly's selection theorem) Let  $\{f_n\}_{n \in \mathbb{N}}$  be a sequence of uniformly bounded, non-decreasing functions  $\mathbb{R} \to \mathbb{R}$ . Then

$$\exists \{n_k\}_{k \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}} \colon \left( n_k \to \infty \land \forall x \in \mathbb{R} \colon \lim_{k \to \infty} f_{n_k}(x) \text{ exists} \right)$$
(42.31)

If, moreover, the limit function  $f(x) := \lim_{k\to\infty} f_k(x)$  is continuous then the convergence is uniform on compact subsets of  $\mathbb{R}$ .

We leave the proof (which is an easy variation on the argument from Lemma 42.1) to homework. The monotonicity of the functions is of course crucial to get subsequential convergence everywhere. The above result is extremely useful in probability where it provides an important tool in the concept of weak convergence of random variables — which is akin to, but not the same as, that defined in (42.30).