

# The Arzelà-Ascoli Theorem

Ben Galin\*

July 1, 2006

In Euclidean space, the Heine-Borel Theorem provides us with an elegant method of determining whether a given set is compact: one need only check if the set is closed and bounded. This allows not only for a practical way of proving compactness, but also builds an intuitive understanding of what compactness means in Euclidean space. However, the Heine-Borel Theorem may fail in more general metric spaces. In particular, the theorem does not apply to the space of continuous real-valued functions on an interval,  $\mathcal{C}(I)$ , where the metric in consideration is the supremum metric.

Fortunately, the Arzelà-Ascoli Theorem allows for an analogous formulation of necessary and sufficient conditions for compactness in  $\mathcal{C}(I)$ . It states that the *closure* of any set  $\Phi$  in that space is compact provided that two conditions are satisfied: (i) the set  $\Phi$  is bounded in the supremum metric on the interval, and (ii) the set  $\Phi$  is equicontinuous, which is a stronger condition than a uniform continuity of each function.

The Arzelà-Ascoli Theorem has been put to much use in different fields of mathematics. Notably, the theorem can be utilized in the proof of Peano's Theorem, which asserts the existence of solutions for ordinary differential equations of continuous functions. In addition, there exist numerous generalizations of the theorem, as well as analogous theorems that were proven in a similar way. See for instance Tang and Li's [2] generalization to Dyson Perturbation Theory and Román-Flores and Rojas-Medar's [1] analogous proof in the context of level-continuous fuzzy sets on Banach spaces.

The version of the Arzelà-Ascoli Theorem we prove below is slightly weaker

---

\*Graded version. This work is licensed under the Creative Commons Attribution 2.5 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/2.5>. Source code with limited rights can be found at <http://www.bens.ws/professional.php>.

than the general statement, which allows for general metric spaces. However, our version provides both a constructive and an intuitive approach for the proof: we will construct a finite set of piecewise linear continuous functions that will approximate the functions in  $\Phi$ . Such a method may not be possible in proofs for a general metric space.

The formal phrasing of the theorem is as follows.

**Theorem** (Arzelà-Ascoli). *Let  $\Phi$  be a subset of  $\mathcal{C}(I)$ , the space of continuous real-valued functions on  $I = [0, 1]$ , equipped with the sup metric. Suppose that*

(a) *there is some  $L > 0$  such that  $|\varphi(x)| \leq L$  for all  $x \in I$  and all  $\varphi \in \Phi$ ,*

and

(b) *for every  $\varepsilon > 0$  there is a  $\delta > 0$  such that  $|\varphi(x) - \varphi(y)| < \varepsilon$  for all  $\varphi \in \Phi$  whenever  $|x - y| < \delta$ , and  $x, y \in I$ .*

*Then the closure of  $\Phi$  is compact. Equivalently, every sequence of functions in  $\Phi$  has a subsequence that converges in  $\mathcal{C}(I)$ .*

To prove the theorem, we will start by simplifying it: we show that the closure of  $\Phi$  is compact if  $\Phi$  itself is totally bounded. The proof of this lemma makes use of the fact that each function in the closure set is a limit point of  $\Phi$  and thus close to a function in  $\Phi$ . Hence, if  $\Phi$  is covered by a finite number of  $\varepsilon$ -tubes, the closure set is covered by ever so slightly wider tubes, centered about the same functions. In addition, the closure set is complete, being a closed subset of a complete space. The compactness of the closure set will follow immediately, since a complete totally bounded set is compact in any metric space.

It will remain to show, of course, that  $\Phi$  is totally bounded. To that end, we will show that each function in  $\Phi$  can be approximated by a piecewise linear continuous function, selected out of a finite set. Here, both hypotheses in the Arzelà-Ascoli theorem will be used: we will create a grid—spanning horizontally on the interval  $I$  and vertically on the interval contained by the uniform bounds of  $\Phi$ —such that any function in  $\Phi$  is close to a unique piecewise linear continuous function, running through vertices of the grid. It will follow that  $\Phi$  can be covered by a finite number of  $\varepsilon$ -tubes, each centered at one of the (piecewise linear) continuous functions. That is, the set  $\Phi$  is totally bounded. This will conclude the proof of the compactness of the closure set.

Lastly, the Bolzano-Weierstrass Theorem will establish that the compactness of the closure set implies that every sequence of functions in  $\Phi$  has a subsequence converging to a function in the space, which is the statement of the second conclusion. The converse will follow from the fact that elements in the closure set are limit points of  $\Phi$ . Thus, a sequence of accumulation points of  $\Phi$  is close to a sequence in  $\Phi$ , and the latter has a converging subsequence. It is then plausible that this subsequence in  $\Phi$  will drag a subsequence in the closure set to a limit. Applying the Bolzano-Weierstrass Theorem again would result in proving the compactness of the closure set, as needed.

As explained above, we begin by proving the following lemma:

**Lemma.** *Let  $\Phi$  be a subset of  $\mathcal{C}(I)$ , the space of continuous real-valued functions on  $I = [0, 1]$ , equipped with the supremum metric. If  $\Phi$  is totally bounded, then  $\text{cl}(\Phi)$ , its closure with respect to  $\mathcal{C}(I)$ , is compact.*

*Proof.* Suppose that  $\Phi$  is totally bounded. Then for any  $\varepsilon > 0$ , we can find a finite subset  $S = \{\xi_1, \dots, \xi_N\} \subseteq \mathcal{C}(I)$  such that  $\Phi$  is covered by the union of  $\varepsilon/2$ -balls, centered at  $\xi_i \in S$ . Equivalently, for any  $\varphi \in \Phi$  we can find  $\xi_i \in S$ , such that  $d(\varphi, \xi_i) < \varepsilon/2$ , where  $d$  is the supremum metric. Consider  $\psi \in \text{cl}(\Phi)$ . Then  $\psi \in \Phi$ , or  $\psi$  is an accumulation point of  $\Phi$ . If  $\psi \in \Phi$ , then it is clearly in one of the  $\varepsilon$ -balls centered at some  $\xi_i \in S$ . Otherwise, by definition there is  $\varphi \in \Phi$  with  $d(\varphi, \psi) < \varepsilon/2$ . By the triangle inequality we have

$$\begin{aligned} d(\psi, \xi_i) &\leq d(\psi, \varphi) + d(\varphi, \xi_i) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{aligned}$$

for some  $\xi_i \in S$ . Thus,  $\psi$  is in the  $\varepsilon$ -ball centered at  $\xi_i$ . It follows that  $\text{cl}(\Phi)$  is covered by a finite number of  $\varepsilon$ -balls. That is, the closure of  $\Phi$  is totally bounded. Since  $\mathcal{C}(I)$  is complete, any closed subset of it—in particular,  $\text{cl}(\Phi)$ —is complete. A totally bounded complete set is compact in any metric space, hence  $\text{cl}(\Phi)$  is compact, as needed.  $\square$

We now turn to proving the Arzelà-Ascoli Theorem:

*Proof.* Denote the closure of  $\Phi$  with respect to  $\mathcal{C}(I)$  by  $\text{cl}(\Phi)$ . By the Lemma, we only need to prove that  $\Phi$  is totally bounded. Let  $\varepsilon > 0$ , and consider the corresponding  $L$  and  $\delta$  as given by the two hypotheses. Let  $x_0 = 0$  and divide the

interval  $I$  into finitely many subintervals  $[x_i, x_{i+1}]$  of length less than  $\delta$  and where  $x_n = 1$ . Similarly, let  $y_0 = -L$  and divide the interval  $[-L, L]$  into finitely many subintervals  $[y_i, y_{i+1}]$  of length less than  $\varepsilon$  and where  $y_m = L$ .

Consider  $\varphi \in \Phi$ . Note that for any  $x_i$  in the division above, the value  $\varphi(x_i)$  is in  $[-L, L]$ . Since the subintervals  $[y_i, y_{i+1}]$  are of length less than  $\varepsilon$ , there exists at least one  $y_{f(i)}$  such that  $d(\varphi(x_i), y_{f(i)}) < \varepsilon$ . Choose the minimal such index  $f(i)$  and consider a graph which passes through all vertices  $(x_i, y_{f(i)})$  and is linear between any two vertices. (Note in particular that we are choosing the *minimal index* satisfying the inequality above, not the *index minimizing* the inequality.) We show that the function  $\psi$  corresponding to this graph is an element of  $\mathcal{C}(I)$ . Note that  $\psi$  is clearly real-valued and defined on  $I$ . To show that  $\psi$  is continuous, note that whenever  $x^* \neq x_i$  for some  $i \in \{0, \dots, n\}$ , the continuity of  $\psi$  at  $x^*$  follows from the linearity of  $\psi$ , and when  $x^* = x_i$  for such an  $i$ , the continuity of  $\psi$  at  $x^*$  follows from the fact that the two ends of piecewise linear segments meet at  $x^*$ .

Thus, we can construct a unique piecewise linear function  $\psi$  such that  $|\psi(x_k) - \varphi(x_k)| < \varepsilon$  for all  $k \in \{0, \dots, n\}$ . Note that by the second hypothesis we have  $|\varphi(x_k) - \varphi(x_{k+1})| < \varepsilon$  for  $k \in \{0, \dots, n-1\}$ . Thus, by the triangle inequality we have

$$\begin{aligned} |\psi(x_k) - \psi(x_{k+1})| &\leq |\psi(x_k) - \varphi(x_k)| + |\varphi(x_k) - \varphi(x_{k+1})| + |\varphi(x_{k+1}) - \psi(x_{k+1})| \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon, \end{aligned}$$

for all  $k \in \{0, \dots, n-1\}$ . Consider any  $x$  such that  $x_k \leq x \leq x_{k+1}$ . Then from the piecewise linearity of  $\psi$ , we must have

$$\psi(x_k) \leq \psi(x) \leq \psi(x_{k+1}) \quad \text{or} \quad \psi(x_k) \geq \psi(x) \geq \psi(x_{k+1}).$$

Subtract  $\psi(x_k)$  from all sides to obtain  $0 \leq |\psi(x_k) - \psi(x)| \leq |\psi(x_k) - \psi(x_{k+1})|$ . Thus,  $|\psi(x_k) - \psi(x)| < 3\varepsilon$ . In addition, we also have by hypothesis,  $|\varphi(x_k) - \varphi(x)| < \varepsilon$ . By the triangle inequality we have

$$\begin{aligned} |\psi(x) - \varphi(x)| &\leq |\psi(x) - \psi(x_k)| + |\psi(x_k) - \varphi(x_k)| + |\varphi(x_k) - \varphi(x)| \\ &< 3\varepsilon + \varepsilon + \varepsilon = 5\varepsilon. \end{aligned}$$

That is, for any  $\varphi \in \Phi$ , there exists  $\psi \in \mathcal{C}(I)$  such that  $d(\varphi, \psi) < 5\varepsilon$ . It follows that  $\Phi$  is covered by  $5\varepsilon$ -balls, centered at such  $\psi$  functions.

To conclude that  $\Phi$  is totally bounded, we need to show that there are only finitely many such  $\psi$  functions. There are  $(n + 1)(m + 1)$  vertices through which a function  $\psi$  can pass. Each function must pass through precisely one vertex  $(x_i, y_{f(i)})$  for each  $i \in \{0, \dots, n\}$ . Thus the total number of possible functions is at most  $(m + 1)^{n+1}$ . Thus,  $\Phi$  is totally bounded, and hence  $\text{cl}(\Phi)$  is compact, as needed for the first conclusion.

Recall that the Bolzano-Weierstrass Theorem states that a subset of a metric space is compact if and only if it is sequentially compact. Since  $\mathcal{C}(I)$  is a metric space, by establishing that  $\text{cl}(\Phi)$  is compact, we also have that it is sequentially compact. Thus every sequence in  $\text{cl}(\Phi)$  has a subsequence that converges to a point in the  $\text{cl}(\Phi)$ . Since we have the inclusions  $\Phi \subseteq \text{cl}(\Phi) \subseteq \mathcal{C}(I)$ , we can conclude in particular that every sequence in  $\Phi$  converges to a point in  $\mathcal{C}(I)$ , which is the statement of the second conclusion.

Conversely, suppose that every sequence of functions in  $\Phi$  has a subsequence that converges in  $\mathcal{C}(I)$ . In particular, that subsequence must converge in  $\text{cl}(\Phi)$ , which contains all limit points of  $\Phi$ . By the uniqueness of limits, this is the only converging point of the subsequence. Consider a sequence in  $\text{cl}(\Phi)$ . If the sequence has infinitely many terms in  $\Phi$ , we have a converging subsequence to a function in  $\text{cl}(\Phi)$ . Otherwise the sequence has infinitely many accumulation points of  $\Phi$ . Denote this subsequence by  $(\varphi_i)_{i \geq 1}$ . Let  $\varepsilon_i = 1/i$  and for each  $\varphi_i$  choose  $\xi_i \in \Phi$  with  $d(\varphi_i, \xi_i) < \varepsilon_i$ . By hypothesis, the sequence  $(\xi_i)_{i \geq 1}$  has a subsequence  $(\xi_{i_k})_{k \geq 1}$  converging to a function in  $\text{cl}(\Phi)$ . By the Sandwich Lemma we have that  $(\varphi_i)_{i \geq 1}$  has a subsequence converging to a function in  $\text{cl}(\Phi)$ , so that  $\text{cl}(\Phi)$  is sequentially compact. The Bolzano-Weierstrass Theorem implies that  $\text{cl}(\Phi)$  is compact. Thus, we have that the two conclusions are equivalent.  $\square$

We conclude this discussion by showing that the two hypothesis are indeed necessary. Suppose that the first hypothesis fails. That is, suppose that the set is not bounded. For instance, consider a set  $F = \{f_1, f_2, \dots\} \subseteq \mathcal{C}(I)$ , where  $f_k = k$ . Note that in this case  $\delta$  can be any positive number, so the second hypothesis is satisfied. However, for a positive  $\varepsilon < 1/2$ , if  $\psi \in \mathcal{C}(I)$  with  $d(\psi, f_k) < \varepsilon$ , then  $d(\psi, f_j) \geq \varepsilon$  for all  $j \neq k$ . That is, infinitely many functions are required to cover  $F$ . Thus, the closure of  $F$  is not compact.

Similarly, for the second hypothesis, consider a set  $G = \{g_1, g_2, \dots\} \subseteq \mathcal{C}(I)$ , where  $g_k(x) = x^k$ . Note that each function in  $G$  is bounded by  $L = 2$  for all  $x \in I$ ,

so the first hypothesis is satisfied. However, the second hypothesis is not satisfied: suppose toward a contradiction that for all  $\varepsilon > 0$  there exists a  $\delta > 0$  such that  $|g_k(x) - g_k(y)| < \varepsilon$  for all  $g_k \in G$  whenever  $|x - y| < \delta$  and  $x, y \in I$ . Let  $\varepsilon = 1/2$ . Since  $g_k(x) \rightarrow 0$  for  $x < 1$  and for all  $k$ , we can find some  $g_k$  such that  $g_k(x) < 1/2$  for some  $x \in (1 - \delta, 1)$ . That is,  $|g_k(x) - g_k(1)| > \varepsilon$ , while  $|x - 1| < \delta$ , a contradiction.

We need to show that there is a sequence of functions in  $G$  for which no subsequence converges in  $\mathcal{C}(I)$ . Consider the sequence  $(g_k)_{k \geq 1} \subseteq G$ . Recall that  $g_k(x) \rightarrow 0$  for  $x < 1$  and for all  $k$ . However,  $g_k(1) = 1$  for all  $k$ . Hence, the sequence converges pointwise to

$$g(x) = \begin{cases} 0 & x \neq 1 \\ 1 & x = 1 \end{cases}.$$

Since  $g(x)$  is not continuous, it cannot be a uniform limit of  $(g_k)_{k \geq 1}$ . By the uniqueness of limits, it follows that there is no uniform limit of the sequence. On the other hand, any converging subsequence of  $(g_k)_{k \geq 1}$  must converge to  $g(x)$ , which is not a function in  $\mathcal{C}(I)$ . The proof of the equivalence of the two conclusions in the Arzelà-Ascoli Theorem implies that the closure of  $G$  is not compact.

## References

- [1] Heriberto Eduardo Román-Flores and Marko Antonio Rojas-Medar, *Embedding of level-continuous fuzzy sets on Banach spaces*, Information Sciences **144** (2002), no. 1, 227–247(21).
- [2] Tai-Ming Tang and Zi-Ping Li, *A generalization of the Arzelá-Ascoli Theorem and its application to Dyson Perturbation Theory*, Letters in Mathematical Physics **43** (1998), no. 1, 55–63.