How to turn Palka's Chapter II into an introduction to

METRIC SPACES

Notwithstanding his having had to go Texas to get a job, Palka is clearly an inhabitant of the civilized world; his rather terse exposition of the topology of $\mathbb{R}^2$ is written from a nicely general viewpoint, and so it can be turned into a fine introduction to the (point-set) topology of metric spaces with very little work. Here is how to do it.[1]

1. *Basic Notation and Terminology.*

One has to take the real numbers for granted so that one can make the following

Definition: A **metric space** is a pair (X, d), where X is a set (just called the **space**) and $d : X \times X \to \mathbb{R}^+$ is a function from pairs of elements of X to nonnegative real numbers (called the **metric** of the space), that satisfies the following axioms:

(1) For every x, $y \in X$ the value $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$;

(2) For every x, $y \in X$ one has $d(x, y) = d(y, x)$;

(3) For every x, y, $z \in X$ one has

$$d(x, y) \leq d(x, z) + d(z, y) \ .$$

(The extremely useful (3) is called the **triangle inequality**; to see why, consider the case where $X = \mathbb{R}$ or $\mathbb{C}$ and $d(z, w) = |z - w|$, or more generally consider the "usual metric" on $\mathbb{R}^n$ below.) The function $d(\cdot, \cdot)$ can be thought of as "distance;" when it is understood what the distance function is in a particular context, one may simply speak of "the metric space X." The purpose of the axiomatic treatment is to bring to the fore those properties of the "distance-between-points-of-$\mathbb{R}^n$" function[2]

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2} = \|\mathbf{x} - \mathbf{y}\|$$

that relate to convergence and continuity (while discarding the ones that have more to do with the vector operations and dot product) and run a number of commonly-necessary arguments once and for all, rather than having to run the same arguments again and again, with trivial modifications, in many analogous situations.

---

[1] Other standard textbooks devote a chapter or a long chapter section to this material in the metric-space context; cf. Ahlfors's Ch. 3 and Conway's Ch. II. I suspect that Palka restrained himself in an effort to make the book seem less intimidating as an undergraduate textbook. (Moreover, had he not camouflaged its graduate-text possibilities in this way he would have been in quite direct competition with Conway and might thus never have had a chance to put a yellow cover around the book.)

[2] It is easy to see that this function satisfies the axioms (1)–(3); it is frequently called the *usual metric* or the *Euclidean norm metric* on $\mathbb{R}^n$.

It is fairly obvious that if (X, d) is a metric space and $A \subseteq X$ then $(A, d|A \times A)$ also satisfies the axioms: this space is said to be obtained by **relativizing the metric of** X **to** A, and the restriction of the metric to $A \times A$ is called the **relativized metric**. Thus one can (and does) regard any subset of a Euclidean space as being itself a metric space under the relativized usual metric.

It is loosely correct to say that "any property of Euclidean spaces that can be stated in terms of distances and inequalities can probably be defined for metric spaces." For example, the definitions of "discs" that Palka makes on the pp. 33–34 page turn make sense, although these sets are usually called "spheres" or "balls:" one says

The **open sphere** of radius $r > 0$ and center x in X is the set
$\Delta(x, r) = \{z \in X : d(x, z) < r\}$;

The **closed sphere** of radius $r > 0$ and center x in X is the set
$\overline{\Delta}(x, r) = \{z \in X : d(x, z) \leq r\}$;

The **punctured sphere** of radius $r > 0$ and center x in X is the set
$\Delta^*(x, r) = \{z \in X : 0 < d(x, z) < r\}$.

One does not usually define the "circle" or "surface of the ball" of radius $r > 0$ in the general situation, although of course we shall continue to use circles in the complex plane (and, later, genuine Euclidean spheres in $\mathbb{R}^n$).

We may use the definitions of Palka's §1.2 to define **interior points** and **open sets** exactly as they are defined in $\mathbb{C}$. His **Theorem 1.1** says that the metrically open sets of a metric space (X, d) form a **topology** on the set X in the sense of axiomatic general topology. By defining the **closed** subsets of a metric space to be those whose complements are open, and by defining **boundary**, **closure** and **interior** in terms of open spheres exactly as Palka does on p. 35, we extend these concepts to general metric spaces in a straightforward way. Note, however, that in the general setting the closure of $\Delta(x, r)$ may *not* be $\overline{\Delta}(x, r)$! For example, look at

$$X = \{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}^+\} \subset \mathbb{R}$$

in the relativized usual metric (*i.e.*, using the usual distances in $\mathbb{R}$). We have $\{1\} = \Delta(1, 1/2) = \overline{\Delta}(1, 1/3)$ so $\{1\}$ is simultaneously an open sphere and a closed sphere, of different radii; it has no boundary points and is thus simultaneously an open set and a closed set, and $\overline{\Delta}(1, 1/2) = \{1, 1/2\} \neq \overline{\{1\}} = \overline{\Delta(1, 1/2)}$. However, Palka's **Theorem 1.3** is still valid. Note, as he seems not to do, that the closure $\overline{A}$ of a set A in a metric space can be characterized as the **smallest closed set that contains** A; this fact is sometimes expressed by writing

$$\overline{A} = \bigcap \{F : F \supseteq A,\ F \text{ closed}\}$$

or its non-symbolic equivalents.

Palka has already been kind enough to define **sequences in** sets in general (p. 35, §1.5), so we don't have to do that. Where he uses angle brackets for sequences $\langle a_n \rangle_{n=1}^{\infty}$

we shall use the old-fashioned braces $\{a_n\}_{n=1}^{\infty}$, but he and we mean the same thing. **Convergence** of sequences, and other forms of limiting behavior, can be defined in a general metric space just as easily as they are in $\mathbb{C}$: the definition reads

A sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space X **converges to the limit** $x_0 \in X$, written

$$\lim_{n\to\infty} x_n = x_0 ,$$

if for every real $\epsilon > 0$ there exists some $N \in \mathbb{N}$ (depending, in general, on ϵ) such that

$$\text{for all } n \geq N, \quad x_n \in \Delta(x_0, \epsilon) .$$

That was stated in terms of discs/spheres to conform to Palka's discussion in §1.6; the usual "ϵ-N-definition" would have read

For every real $\epsilon > 0$ there exists some $N \in \mathbb{N}$ (depending, in general, on ϵ) such that

$$\text{for all } n \geq N, \quad d(x_0, x_n) < \epsilon$$

which of course says the same thing. A sequence can have at most one limit, because if $\{x_n\}_{n=1}^{\infty}$ converges both to x_0 and to $\tilde{x}_0$, then for every $n \in \mathbb{N}$

$$d(x_0, \tilde{x}_0) \leq d(x_0, x_n) + d(x_n, \tilde{x}_0)$$

and since each of the terms on the r. h. s. can be made $< \epsilon$ by taking n sufficiently large, we have $d(x_0, \tilde{x}_0) < 2\epsilon$ for every real $\epsilon > 0$, possible if and only if $d(x_0, \tilde{x}_0) = 0$ whence $x_0 = \tilde{x}_0$.

Palka's §1.7 definitions of **accumulation points** of complex sequences make sense in general metric spaces and carry over to them without any trouble. Since the complex plane $(\mathbb{C}, d(z, w) = |z - w|)$ is a metric space, his p. 38 examples work in the general context, as does his proof of his

Theorem (1.6): A sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space (X, d) has the point $x_0 \in X$ as an accumulation point if and only if there is a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ such that $\lim\limits_{k\to\infty} x_{n_k} = x_0$.

Similarly, his proofs carry over to the general setting to show that

Theorem (1.7): If a sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space (X, d) is convergent and has limit $x_0 \in X$, then $\lim_{k\to\infty} x_{n_k} = x_0$ for every subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$.

Theorem (1.8): The point $x_0 \in X$ belongs to the (metric) closure of a set $A \subseteq X$ in a metric space (X, d) if and only if there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in A with $\lim_{n\to\infty} x_n = x_0$.

Theorem (1.9): A subset $A \subseteq X$ is closed in a metric space (X, d) if and only if A contains every accumulation point of every sequence in A.

It is important for future reference to know that if $\{x_n\}_{n=1}^{\infty}$ is a **bounded** sequence in $\mathbb{R}$ in the usual sense, then there exists at least one $x_0 \in \mathbb{R}$ that is an accumulation point of $\{x_n\}_{n=1}^{\infty}$ in the usual metric (and therefore there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ that converges to x_0). This is the **Bolzano-Weierstraß Theorem**. One may choose for x_0 either the *limes inferior* or *limes superior* of the given sequence, defined respectively by

$$\liminf x_n = \sup\{\inf\{x_j : j \geq k\} : k \in \mathbb{N}\}; \quad \limsup x_n = \inf\{\sup\{x_j : j \geq k\} : k \in \mathbb{N}\} .$$

(Other accumulation points might of course exist.) For only a little more detail, see Palka's Appendix A, §2, p. 545 ff., or for the details that you like, see your favorite advanced calculus text.

2. *Continuity and Limits of Functions.*

The definitions that Palka gives for complex-valued functions of a complex variable carry over, word for word, to give definitions of continuity, limits, etc., for functions $f : A \to Y$ from a subset $A \subseteq X$ of one metric space (X, d) to another metric space (Y, ρ). While he is careful (for his pedagogical purposes) *not* to use ϵ-δ formulations of continuity, for our pedagogical purposes we want to observe that a function $f : A \to Y$ from a subset $A \subseteq X$ of one metric space (X, d) to another metric space (Y, ρ) is (and should be) **continuous at a point** $x_0 \in A$ if and only if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every $x \in A$ such that $d(x, x_0) < \delta$, it is true that $\rho(f(x), f(x_0)) < \epsilon$. **Continuity on a (sub-)set** in a metric space, or on the whole metric space, are then defined as you might expect (cf. the bottom of Palka's p. 40). The proofs valid in the plane then carry over to give us

Theorem (2.1): A function $f : A \to Y$ defined on a subset of a metric space (X, d) with values in a metric space (Y, ρ) is continuous at a point $x_0 \in A$ if and only if: for every sequence $\{x_n\}_{n=1}^{\infty}$ in A converging to x_0, the image sequence $\{f(x_n)\}_{n=1}^{\infty}$ in Y converges to $f(x_0)$.

The analogues of Palka's **Theorems 2.3** and **2.4** then go through without any trouble for real- or complex-valued functions (or for that matter, $\mathbb{R}^n$- or $\mathbb{C}^n$-valued functions for which the combinations are well defined) defined on subsets A of a ("general") metric space (X, d). One should pay particular attention to the "sequential" proof of the fact that continuity is preserved under composition of functions, **Theorem 2.3**, and at least mentally adapt the proof to the case of functions $f : A \to Y$ and $g : B \to W$, where (X, d), (Y, ρ) and (W, r) are metric spaces, $A \subseteq X$, $B \subseteq Y$, and f and g are continuous at $x_0 \in A$ and $y_0 = f(x_0) \in B$ respectively. Of course one should also be able to give an ϵ-δ proof of this important fact.

The notions of **limits of functions** that Palka has so carefully written up in his §2.2, p. 43 ff., carry over so straightforwardly to the metric-space setting that one may safely leave the job of transcription to the interested reader.

We should make a definition, and check a fact, that Palka seems to omit. Suppose (X, d) is a metric space and $\emptyset \neq A \subseteq X$. Then we may **define** a nonnegative-real-valued function on X, called the **distance to** A function, by setting

$$d(x, A) = \inf\{d(x, a) : a \in A\} \quad \text{for all } x \in X .$$

If x_1 and x_2 are points in X then for every $a \in A$ we have as a result of the triangle inequality for metric spaces that $d(x_1, a) \leq d(x_1, x_2) + d(x_2, a)$, therefore $d(x_1, A) \leq d(x_1, x_2) + d(x_2, a)$, or equivalently $d(x_1, A) - d(x_1, x_2) \leq d(x_2, a)$. Now since $a \in A$ is arbitrary and $d(x_2, A)$ is defined by a greatest lower bound, we have $d(x_1, A) - d(x_1, x_2) \leq d(x_2, A)$, or $d(x_1, A) - d(x_2, A) \leq d(x_1, x_2)$, whence by symmetry

$$|d(x_1, A) - d(x_2, A)| \leq d(x_1, x_2) \ .$$

It is immediate that $x \to d(x, A)$ is continuous on X—even *uniformly continuous on* X, if we only knew what that meant. Moreover,

Proposition: A point $x \in X$ belongs to $\overline{A}$ if and only if $d(x, A) = 0$.

Proof. On one hand, if $x \in \overline{A}$ then there exists a sequence $\{a_n\}_{n=0}^{\infty}$ of points in A converging to x, and then by continuity $d(x, A) = \lim_{n\to\infty} d(a_n, A) = 0$ (all the values $d(a_n, A) = 0$). On the other, if $d(x, A) = \inf\{d(x, a) : a \in A\} = 0$ then for each $n \in \mathbb{N}$ there must exist an $a_n \in A$ with $d(x, a_n) < \dfrac{1}{n}$, and then $\{a_n\}_{n=0}^{\infty}$ is a sequence of points in A converging to x, whence $x \in \overline{A}$.[3]

Let us make an observation now for use below: if (X, d) is a metric space, Y a subset of X that is made into a metric space under the relativized metric $d|Y \times Y$, and $A \subseteq Y$, then for points $y \in Y$ the definition of $d(y, A)$ gives the same value for y considered as an element of (X, d) that it does for y considered as an element of $(Y, d|Y \times Y)$.

3. *Connected Sets.*

At this point things take a rather counter-intuitive turn: the general notion of connectedness is quite subtle and takes some getting used to. In the first place it is defined by the *via negativa*: a set is **connected** if it is not **disconnected**. The definition of **disconnected** is fortunately the same in a metric space (or a general topological space) as it is in $\mathbb{C}$: a nonempty set A in a metric space (X, d) is **disconnected** if there are two open sets U, V in the big space X for which

(i) $U \cap V = \emptyset$;

(ii) $A \cap U \neq \emptyset$ and $A \cap V \neq \emptyset$;

(iii) $A = (A \cap U) \cup (A \cap V)$.

This definition appears to be dependent on the "ambient space" X, but in fact it is not: it is easy to show (the proof is essentially the one that Palka should have used to prove **Lemma 3.1**) that it is logically equivalent to say: a nonempty set A in a metric space

---

[3] The reader might find it amusing to give a non-sequential proof of this proposition. Note that it implies that any closed set in a metric space is exactly equal to the zero-set of an appropriately-chosen continuous real-valued function; this fact is not immediately obvious. Indeed, it implies that for any disjoint pair of closed sets in a metric space, there is a [0,1]-valued continuous function defined on the whole space whose value is 0 exactly on one of the sets and 1 exactly on the other.

(X,d) is **disconnected** if there are two sets $\tilde{U}$, $\tilde{V} \subseteq A$ that are *open sets in the relativized metric* of A and such that

(i) $\tilde{U} \cap \tilde{V} = \emptyset$;

(ii) $\tilde{U} \neq \emptyset$ and $\tilde{V} \neq \emptyset$;

(iii) $A = \tilde{U} \cup \tilde{V}$.

The proof of equivalence is trivial in one direction: if the open sets U and V exist in X, then one may simply put $\tilde{U} = A \cap U$ and $\tilde{V} = A \cap V$. Conversely, if the subsets $\tilde{U}$ and $\tilde{V}$ of A exist, then consider the nonnegative-real-valued functions

$$f_1(x) = d(x, \tilde{V}) \qquad \text{and} \qquad f_2(x) = d(x, \tilde{U}) \ .$$

Let $U = \{x \in X : f_1(x) > f_2(x)\}$ and $V = \{x \in X : f_2(x) > f_1(x)\}$. Since (for example) $f_1(x) > f_2(x)$ is equivalent to $f_2(x) - f_1(x) > 0$ and the l. h. s. of that inequality is a continuous function of $x \in X$, both these sets are open and they are obviously disjoint from one another. The fact that $\tilde{V}$ is the complement in A of the relatively open set $\tilde{U}$ means that it is a closed set in the relative metric; thus it equals its own closure *in* A, and so for every point $x \in \tilde{U}$ it is true that $d(x, \tilde{V}) > 0$. From this, it is obvious that $\tilde{U} \subseteq U$, and similarly $\tilde{V} \subseteq V$. Thus we have exhibited U and V satisfying the original definition of disconnectedness for A.

{This formulation may not look the same as that of Palka's Lemma 3.1, but it is: given relatively-open subsets $\tilde{U}$, $\tilde{V}$ of A that satisfy $\tilde{U} \cap \tilde{V} = \emptyset$, $\tilde{U} \neq \emptyset$ and $\tilde{V} \neq \emptyset$, and $A = \tilde{U} \cup \tilde{V}$, let

$$U^* = \bigcup \{\Delta(x,\epsilon) : \Delta(x,\epsilon) \cap \tilde{V} = \emptyset\}$$
$$V^* = \bigcup \{\Delta(x,\epsilon) : \Delta(x,\epsilon) \cap \tilde{U} = \emptyset\}$$

where the spheres are taken *in the "big space"* X. Evidently these sets are open. The set $\tilde{U} \subseteq U^*$ because $\tilde{U}$ is the complement of $\tilde{V}$ in A, so it is open in A and thus at each $x \in \tilde{U}$ there is a sphere centered at x that does not meet $\tilde{V}$; similarly $\tilde{V} \subseteq V^*$. It follows that Palka's condition *(ii)* that both U^* and V^* meet A is satisfied, and similarly his condition *(iii)* that $A \subseteq U^* \cup V^*$. His condition *(i)* is satisfied because if $x \in A$ is a point such that (say) $x \in \tilde{U}$ then it cannot meet the defining condition for V^*, so $x \notin U^* \cap V^*$, and similarly for the case where $x \in \tilde{V}$. Conversely, given U^* and V^* that are open sets in the "big space" X as in Palka's Lemma 3.1, the sets $\tilde{U} = A \cap U^*$ and $\tilde{V} = A \cap V^*$ have the properties we set forth in our second definition of disconnectedness above.}

So now we know what a disconnected set is, and therefore what a connected set is. Unfortunately, to continue with Palka's §3.2 we need a theorem that he doesn't prove until the very end of §3.4, as

Theorem 3.8: If (X,d) and (Y,ρ) are metric spaces, X is connected, and $f : X \to Y$ is a continuous function, then the image set $f[X]$ is a connected subset of Y.

The proof Palka gives goes over to the metric-space setting with no significant changes.

We need the metric-space version of **Theorem 3.8** early, because we need some new objects to take the place of the line segments of $\mathbb{C}$ before we go on. However, let us skip over Palka's **Theorem 3.2** for a moment and go on to **Theorem 3.3**: evidently this theorem is just as true in any metric space as it is in $\mathbb{C}$. Having done that, we can return to **Theorem 3.2** and its metric-space analogues.

Definition: A **path** in a metric space (X, d) is a continuous function $\gamma : [a, b] \to X$ defined on a non-trivial interval $(a < b)$ of $\mathbb{R}$. Its **initial point** is $\gamma(a)$ and its **terminal point** is $\gamma(b)$.

Definition: Let (X, d) be a metric space. A subset $A \subseteq X$ is said to be **path-connected** or **arcwise connected** if: given any pair of points x_0, $x_1 \in A$, there exists a path $\gamma : [a, b] \to A$ with $\gamma(a) = x_0$ and $\gamma(b) = x_1$. It is said to be **locally path-** or **arc-wise connected** if for every $x \in A$ and every $\epsilon > 0$ there exists an open (relative to A) arcwise connected set V with $x \in V \subseteq \Delta(x, \epsilon)$.

Evidently every open interval in $\mathbb{R}$ and open disc in $\mathbb{C}$ is arcwise connected, and therefore every open subset of $\mathbb{R}$ or $\mathbb{C}$ is locally arcwise connected. We needed **Theorem 3.8** in order to see that

Proposition: An arcwise connected set A in a metric space (X, d) is a connected set.

Proof. Without loss of generality we can assume that $A = X$ (otherwise, relativize the metric). If we could write $X = U \cup V$ with U and V open, nonempty, and disjoint, and took a path $\gamma : [a, b]$ for which $\gamma(a) \in U$ and $\gamma(b) \in V$, we would see that the sets $\gamma^{-1}[U]$ and $\gamma^{-1}[V]$ were open, nonempty, and disjoint and that $[a, b] = \gamma^{-1}[U] \cup \gamma^{-1}[V]$, contradicting the connectedness of $[a, b]$ that is a special case of Palka's **Theorem 3.2**, pp. 48–49. Thus it is impossible to disconnect A.

Palka's **Theorem 3.3**, p. 49, goes over to metric spaces with no changes in statement or proof. One should notice, by analogy with his sets "starlike with respect to a point z_0," that if x_0, x_1, x_2 are points in a metric space and $\gamma_0 : [a, b]$ is a path with initial point x_0 and terminal point x_1, while γ_1 is a path with initial point x_1 and terminal point x_2, then it is easy to cobble up a path with initial point x_0 and terminal point x_2: for example, consider the path defined on $[0, 1]$ by

$$\gamma(t) = \begin{cases} \gamma_0\left((1-2t)a + 2tb\right) & \text{for } 0 \le t \le \dfrac{1}{2} \ ; \\ \gamma_1\left((2-2t)c + (2t-1)d\right) & \text{for } \dfrac{1}{2} \le t \le 1 \ . \end{cases}$$

(Checking continuity [not to mention well-definedness] at $t = 1/2$ is easy.)

The arguments that Palka runs in his §3.4 can be run in a locally arcwise connected metric space, so let's look at them. Suppose (X, d) is such a space and $U \subseteq X$ is a nonempty open subset. Given an $x_0 \in U$, let

$$D(x_0) = \{x_1 \in U : \exists \text{ a path } \gamma \text{ in } U \text{ with initial point } x_0 \text{ and terminal point } x_1\}.$$

Just as in the case $X = \mathbb{C}$, this is a connected set called the **(arc-)component of** x_0 **in** U. It is open because if $x_1 \in D(x_0)$ then, because U is open, there is a sphere $\Delta(x_1, \epsilon) \subseteq U$ for some $\epsilon > 0$, and because X is locally arcwise connected there is an open V with $x_1 \in V \subseteq U$ such that every $x_2 \in V$ can be joined to x_1 by a path lying in V—but then, as we just saw, there is a path $\gamma : [0, 1] \to U$ with initial point x_0 and terminal point x_2, so $V \subseteq D(x_0)$. This shows that every $x_1 \in D(x_0)$ belongs to an open subset of $D(x_0)$, so every point of $D(x_0)$ is interior and it is thus an open set. Evidently each $D(x_0)$ is arcwise connected and thus connected; two arc-components that have any points in common are equal (why?)[4], and thus U is partitioned into open sets $D(x)$ by the arc-components of its points. The complement of each arc-component of U is the union of the other arc-components of U and hence open: therefore, each arc-component has an open complement, hence is closed, hence is open-and-closed. Thus if U is a nonempty *connected* open subset of a locally arcwise connected space it must be *arcwise connected*, *i.e.*, any two points of U can be joined by a path. This is the general version of Palka's **Theorem 3.6**. Similarly his **Theorem 3.7** generalizes: a connected subset A of an open set U of a locally arcwise connected space must be contained in one of the arc-components of U.

In order to give our first proof of the "fundamental theorem of algebra" we shall need the following fact:

Theorem: If $\emptyset \neq U \subseteq \mathbb{C}$ is a connected open set and $F \subset U$ is a finite set, then $U \setminus F$ is (arc-)connected.

Proof. We need the following lemma; the reader should check the details (it suffices to consider the unit circle). One should not take the trigonometric functions for granted, but use the function $x \to \sqrt{1 - x^2}$, since the properties of the square-root function on $\mathbb{R}^+$ may legitimately be assumed.

Lemma: Every circle $\partial\Delta(c, \rho) = \{z \in \mathbb{C} : |z - c| = \rho\}$ is arc-connected.

It suffices to prove the theorem for a one-point set $F = \{c\}$, since the general statement will then follow by a trivial induction. Let z_0 and z_1 be (distinct) points of $U \setminus \{c\}$, and let $\gamma : [a, b] \to U$ be a path with z_0 as initial and z_1 as terminal point. If $\gamma(t) \neq c$ for all $a \leq t \leq b$ there is nothing to prove. Otherwise, let $\rho > 0$ be small enough that $z_0, z_1 \notin \overline{\Delta}(c, \rho)$. Since $U \setminus \overline{\Delta}(c, \rho)$ is open, there exist numbers $e \in (a, b]$ for which $\gamma[[a, e]] \subseteq \big[U \setminus \overline{\Delta}(c, \rho)\big]$; let $\check{e}$ be their supremum. Again because $U \setminus \overline{\Delta}(c, \rho)$ is open, we cannot have $\gamma(\check{e}) \in \big[U \setminus \overline{\Delta}(c, \rho)\big]$, nor (because $\Delta(c, \rho)$ is open) can we have $\gamma(\check{e}) \in \Delta(c, \rho)$. It follows that $\gamma : [a, \check{e}] \to [U \setminus \Delta(c, \rho)]$ is a path joining z_0 to $\gamma(\check{e}) \in \partial\Delta(c, \rho)$. Similarly, we can find a number $\hat{e} \in [a, b)$ for which $\gamma(\hat{e}) \in \partial\Delta(c, \rho)$ and $\gamma[[\hat{e}, b]] \subseteq [U \setminus \Delta(c, \rho)]$. We know

[4] If we define a relation on A by saying that $x{\sim}y$ if and only if they are initial and terminal points respectively of a path in A, then we have for all practical purposes demonstrated that "$\sim$" is an equivalence relation. It therefore partitions A into equivalence classes, which are just its arc-components. The argument here (or in Palka) shows that for an open subset U of a locally arc-connected space, these equivalence classes—the arc-components of U—are open. But then the complement of each of these arc-components, being itself the union of the remaining arc-components, is open: the arc-component we started with, because its complement is open, is closed, hence it is open-and-closed. Conclusion: U is partitioned into its open-and-closed arc-components.

that $\partial\Delta(c,\rho)$ is arc-connected, so there is no problem finding a path $\gamma_1 : [\check{e}, \hat{e}] \to \partial\Delta(c,\rho)$ joining $\gamma(\check{e})$ to $\gamma(\hat{e})$. The path defined by

$$\Gamma = \begin{cases} \gamma(t) & \text{for } t \in [a, \check{e}] \cup [\hat{e}, b] \\ \gamma_1(t) & \text{for } t \in [\check{e}, \hat{e}] \end{cases}$$

then joins z_0 to z_1 in $U \setminus \{c\}$.

Remark: If one says that a set $F \subseteq U$ (where $U \subseteq \mathbb{C}$ is a connected open set) is **locally finite** (or that the points of F are **isolated**) if there is some disc centered on each point $c \in F$ that contains no point of F other than c, then it is easy to extend the result just proved to show that locally finite subsets of U do not disconnect it. This extension is quite handy in the context of holomorphic functions, since if $f : U \to \mathbb{C}$ is holomorphic and not identically zero, then each of its zeros is isolated.

4. *Compact Sets.*

Re Palka's §4.1: The first thing to realize about this material in general metric spaces is that *there is no adequate definition of "bounded" in general, and the Bolzano-Weierstraß theorem does not hold in general.* There is a notion of **totally bounded** set that will turn out to have the properties possessed by bounded sets in $\mathbb{R}$ or $\mathbb{C}$, the Bolzano-Weierstraß theorem is "almost true," and we shall investigate these matters shortly. Indeed, what happens is that the objects bearing famous names—Bolzano-Weierstraß theorem, Cauchy sequence, Heine-Borel theorem—become names of properties that some metric (or topological) spaces have (and some don't have).

For example (re §4.2), the notion of Cauchy sequence presents no problems: a sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space (X, d) is a **Cauchy sequence** (or has the **Cauchy property**) if for each $\epsilon > 0$ there is an $N \in \mathbb{N}$, in general depending on ϵ, such that $d(x_n, x_m) < \epsilon$ holds for all pairs of indices n, $m \geq N$. In $\mathbb{R}$ and $\mathbb{C}$, Cauchy sequences always converge, but $\mathbb{Q}$ in the metric relativized from $\mathbb{R}$ gives an example of a metric space in which that is not true: any of the familiar sequences of rational numbers that converges to $\sqrt{2}$ will be convergent-and-therefore-Cauchy in $\mathbb{R}$, therefore also Cauchy in $\mathbb{Q}$, but of course the limit does not exist in $\mathbb{Q}$. A metric space (X, d), or more generally a subset $A \subseteq X$, is said to be **complete** or **Cauchy-complete** if every Cauchy sequence in A converges to a limit belonging to A. It is easy to see that every convergent sequence is Cauchy, and therefore—since a set is closed if and only if it contains the limit of every convergent sequence in it—that a complete set must be closed, and that a closed subset of a complete space is itself complete. Moreover, in any setting, a Cauchy sequence that has an accumulation point must be convergent to that accumulation point: if $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in a metric space (X, d) and x_0 is an accumulation point of $\{x_n\}_{n=1}^{\infty}$, then given any $\epsilon > 0$ we may find $N_1 \in \mathbb{N}$ for which $n, m \geq N_1 \Rightarrow d(x_n, x_m) < \dfrac{\epsilon}{2}$, and then—because x_0 is an accumulation point—find an index $N \geq N_1$ for which $d(x_N, x_0) < \dfrac{\epsilon}{2}$. It is now clear that $n \geq N \Rightarrow d(x_n, x_0) \leq d(x_n, x_N) + d(x_N, x_0) < \dfrac{\epsilon}{2} + \dfrac{\epsilon}{2} = \epsilon$, as one would wish.

Palka's §4.3 presents us with onomasiological problems. As he points out, what he calls "compact" is called "sequentially compact" by just about everybody else, and the word "compact" is used to describe something having to do with open covers. These concepts are going to turn out to be equivalent in metric spaces anyhow, but we may as well put them straight immediately.

Definition: A set A in a metric space (X, d) is **sequentially compact** if every sequence in A has an accumulation point that belongs to A.

Definition: Let A be a set in a metric space X. An **open cover** of A is a family $\mathfrak{U}$ of open sets U such that $A \subseteq \bigcup\{U : U \in \mathfrak{U}\}$. A **subcover** of $\mathfrak{U}$ is a subfamily $\mathfrak{V} \subseteq \mathfrak{U}$ that is also a cover. A **refinement** of $\mathfrak{U}$ is a(nother) cover $\mathfrak{V}$ with the property that for every $U \in \mathfrak{U}$ there exists $V \in \mathfrak{V}$ such that $V \subseteq U$ (intuitively, the sets of $\mathfrak{V}$ are "smaller" than those of $\mathfrak{U}$). Obviously, by a **finite subcover** (or **finite refinement**) of $\mathfrak{U}$ one means a subcover or refinement, respectively, having only finitely many elements.

Definition: A set A in a metric space X is **compact** if every open cover of A contains a finite subcover of A.

This definition makes the conclusion of the Heine-Borel theorem—that every covering of a (finite) closed interval of $\mathbb{R}$ by open intervals contains a finite subcover—into a concept.

Definition: The **diameter** of a set A in a metric space (X, d) is the nonnegative extended real number $\sup\{d(a, b) : a, b \in A\}$, which may be $+\infty$. In accordance with standard convention, the diameter of the empty set is zero.

It is easy to see that a set of diameter $\delta < \infty$ is a subset of some sphere of radius, say, 2δ; on the other hand, if A fits into a sphere of radius δ then its diameter can be at most 2δ.

Definition: A set A in a metric space is **totally bounded** or **precompact** if for every $\epsilon > 0$ there exist subsets $A_1, \ldots, A_p$ (the choice and number of the subsets will in general depend on ϵ) of A, each of diameter $< \epsilon$, such that $A = A_1 \cup \cdots \cup A_p$. Equivalently, one may require that for every $\epsilon > 0$ there exist a finite cover of A by open (or closed) spheres of radius $< \epsilon$.

It is easy to see that *precompactness* is the property that generalizes boundedness of subsets of $\mathbb{R}$: if $A \subseteq \mathbb{R}$ is bounded, then since $\mathbb{R} = \bigcup_{n\in\mathbb{Z}} [n\epsilon, (n+1)\epsilon)$ can be chopped up into pieces of diameter ϵ it is clear that bounded A can be chopped up into finitely many such pieces. And indeed, one has the analogue of the Bolzano-Weierstraß theorem:

Proposition: A set A in a metric space (X, d) is totally bounded if and only if every sequence in A has a Cauchy subsequence. (Thus) A totally bounded set is complete if and only if it is sequentially compact.

Proof. Suppose A is totally bounded, and let $\{x_n\}_{n=1}^{\infty}$ be a sequence in A. We will produce a sequence of subsequences of the given sequence.[5] To start the process, let

[5] I am about to run the Cantor diagonal process. People who have already seen it can read the next several

$A = A_{1,1} \cup \cdots \cup A_{1,p_1}$ be a decomposition of A into subsets of diameter at most 1. Clearly there must be at least one of these sets A_{1,j_1}—we shall now call it A^1—with the property that $x_n \in A_{1,j_1}$ holds for infinitely many indices n; so we can pick out a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of the given sequence whose values all lie in A^1. We shall set $x_{1,k} = x_{n_k}$, so the sequence $\{x_{1,k}\}_{k=1}^{\infty}$ is a relabeled version of a subsequence of the originally-given sequence. Suppose we have produced m sequences

$$\begin{aligned} \{x_{1,k}\}_{k=1}^{\infty} &= x_{1,1},\, x_{1,2},\, \ldots \\ &\ldots \\ &\ldots \\ &\ldots \\ \{x_{m,k}\}_{k=1}^{\infty} &= x_{m,1},\, x_{m,2},\, \ldots \end{aligned}$$

such that each $\{x_{\ell+1,k}\}_{k=1}^{\infty}$ is a subsequence of $\{x_{\ell,k}\}_{k=1}^{\infty}$ (and therefore of the originally-given sequence) and the values of the ℓ-th are contained in a subset $A^{\ell} \subseteq A$ of diameter $1/\ell$ for $\ell = 1, \ldots, m$. Then we may chop A^m into pieces of diameter at most $1/(m+1)$ and select a subsequence $\{x_{m+1,k}\}_{k=1}^{\infty}$ of $\{x_{m,k}\}_{k=1}^{\infty}$ all of whose values lie in some $A^{m+1} \subseteq A^m$ with diameter at most $1/(m+1)$: the process is the same as the one that selected the first subsequence. By induction (and countable choice, but we're analysts and don't care) we thus extend our array of subseqences to have countably many rows:

$$\begin{aligned} \{x_{1,k}\}_{k=1}^{\infty} &= x_{1,1},\, x_{1,2},\, \ldots \\ &\ldots \\ &\ldots \\ &\ldots \\ \{x_{m,k}\}_{k=1}^{\infty} &= x_{m,1},\, x_{m,2},\, \ldots \\ \{x_{m+1,k}\}_{k=1}^{\infty} &= x_{m+1,1},\, x_{m+1,2},\, \ldots \\ &\ldots \\ &\ldots \\ &\ldots \end{aligned}$$

each of which is a subsequence of the sequence originally given *and* a subsequence of its predecessor. If we now consider the **diagonal subsequence** $\{x_{k,k}\}_{k=1}^{\infty}$, we see a subsequence of the originally-given sequence with the property that if $j, k \geq N$ then $d(x_j, x_k) \leq \dfrac{1}{N}$ (both values belong to a subset of A whose diameter is at most $1/N$). This subsequence is evidently Cauchy, so we have exhibited a Cauchy subsequence of $\{x_n\}_{n=1}^{\infty}$.

Conversely, suppose A is *not* totally bounded. Then there exists some $\epsilon > 0$ such that there exists no finite cover of A by spheres of radius ϵ. Take any $x_1 \in A$; then

lines very quickly. People who haven't already seen it should study this method a bit, since it tends to turn up in sequential compactness arguments, such as proofs of the Arzelá-Áscoli theorem.

$A \setminus \Delta(x_1, \epsilon) \neq \emptyset$ (otherwise A would have been covered by one sphere of radius ϵ), so one may choose $x_2 \in A \setminus \Delta(x_1, \epsilon)$. If $x_1, \ldots, x_n$ have been chosen, then one may choose $x_{n+1} \in \left[A \setminus \bigcup_{j=1}^{n} \Delta(x_j, \epsilon)\right] \neq \emptyset$ (otherwise A would have been covered by n spheres of radius ϵ). Proceeding inductively, one thus chooses a sequence $\{x_n\}_{n=1}^{\infty}$ in A each of whose elements is at distance $\geq \epsilon$ from all the others, and there is no hope of such a sequence having a Cauchy subsequence.

To finish things up, observe that if A is totally bounded and complete, then every sequence in A has a Cauchy subsequence—but Cauchy sequences converge to limits in A, so A is sequentially compact. Conversely, if A is sequentially compact then every Cauchy sequence in A has an accumulation point in A—but as we saw earlier, the Cauchy sequence then converges to that accumulation point, so A is complete. And since every sequence in A has a convergent subsequence it has *a fortiori* a Cauchy subsequence, so A must be totally bounded.

The passage from "sequentially compact" to "compact" in the open-covering sense is facilitated by the following theorem, which is among the handiest tools ever created for making compactness useful.

Theorem [Lebesgue Covering Lemma]: Let A be a sequentially compact subset of a metric space and $\mathfrak{U}$ an open cover of A. There exists a real number $\lambda > 0$, called the **Lebesgue number** of $\mathfrak{U}$, with the property that if $B \subseteq A$ has diameter $\leq \lambda$ then there exists a single set $U \in \mathfrak{U}$ such that $B \subseteq U$.[6] Equivalently (although the value of λ will in general be different) there exists a $\lambda > 0$ such that every sphere $\Delta(x, \epsilon)$ with $x \in A$ and $\epsilon \leq \lambda$ is contained in some single element $U \in \mathfrak{U}$.

Proof. It will clearly suffice to prove the latter equivalent proposition. If it were false, then for every $n \in \mathbb{N}$ one could find an $x_n \in A$ and a sphere $\Delta(x_n, 1/n)$ such that for no $U \in \mathfrak{U}$ was it true that $[\Delta(x_n, 1/n) \cap A] \subseteq U$. Since A is sequentially compact, there would exist a convergent subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ with limit $x_0 \in A$. Since $\mathfrak{U}$ is an open cover of A, there is some $U \in \mathfrak{U}$ with $x_0 \in U$; since each $U \in \mathfrak{U}$ is open, there exists a sphere $\Delta(x_0, \epsilon)$ for which $[\Delta(x_0, \epsilon) \cap A] \subseteq U$. But now, if $k \in \mathbb{N}$ is taken so large that $d(x_0, x_{n_k}) < \dfrac{\epsilon}{2}$ and $\dfrac{1}{n_k} < \dfrac{\epsilon}{2}$, then every point in $\Delta(x_{n_k}, 1/n_k) \cap A$ is at distance at most ϵ from x_0 and therefore $[\Delta(x_{n_k}, 1/n_k) \cap A] \subseteq [\Delta(x_0, \epsilon) \cap A] \subseteq U$, contradicting the choice of x_{n_k}. The contradiction proves the theorem.

It is now quite straightforward to show that in metric spaces, the notions of sequential compactness and (covering) compactness agree.

Proposition: A subset A of a metric space (X, d) is compact if and only if it is sequentially compact.

Proof. Suppose A is sequentially compact and let an open cover $\mathfrak{U}$ of A be given. Then A is totally bounded, and so one can write A as a union $A = A_1 \cup \cdots \cup A_n$ of

[6] Colloquially, "every sufficiently small subset of A fits inside a single element of $\mathfrak{U}$."

subsets each of which has diameter smaller than the Lebesgue number λ of the cover $\mathfrak{U}$. But then for each $j = 1, \ldots, n$ one may select some $U_j \in \mathfrak{U}$ for which $A_j \subseteq U_j$, so $A = A_1 \cup \cdots \cup A_n \subseteq U_1 \cup \cdots \cup U_n$ and thus $\{U_j\}_{j=1}^n$ is a finite subcover of $\mathfrak{U}$, establishing the compactness of A. The converse is easier: if A is not sequentially compact then there is a sequence $\{x_n\}_{n=1}^\infty$ in A that does not have an accumulation point, which means that for every point $a \in A$ there is a sphere $\Delta(a, \epsilon_a)$ centered at a such that $x_n \in \Delta(a, \epsilon_a)$ holds for only finitely many $n \in \mathbb{N}$. The family $\{\Delta(a, \epsilon_a) : a \in A\}$ is an open cover of A, but it cannot contain a finite subcover, because for any finite subfamily $\{\Delta(a_j, \epsilon_{a_j})\}_{j=1}^k$ one can have $x_n \in \bigcup_{j=1}^k \Delta(a_j, \epsilon_{a_j})$ for only finitely many indices $n \in \mathbb{N}$, so that the remaining x_n's must belong to $A \setminus \bigcup_{j=1}^k \Delta(a_j, \epsilon_{a_j})$. Thus A cannot be compact.

The covering definition of compactness makes possible a generalization of Palka's **Theorem 4.5** that is frequently handy. A family $\mathfrak{F}$ of subsets of a metric space X is said to have the **finite intersection property** if: for every finite subfamily $\{F_1, \ldots, F_n\} \subseteq \mathfrak{F}$, the intersection $\bigcap_{j=1}^n F_j \neq \emptyset$. Then

Proposition: A metric space (X, d) is compact if and only if: for each family $\mathfrak{F}$ of closed subsets of X that has the finite intersection property, the intersection $\bigcap \{F : F \in \mathfrak{F}\}$ of *all* the elements of the family is nonempty.

Proof. This is simply a "complement version" of the covering definition of compactness. Given a family of closed subsets $\mathfrak{F}$ with the finite intersection property, consider the family of their complements $\mathfrak{U} = \{X \setminus F : F \in \mathfrak{F}\}$, each of which is an open set. $\mathfrak{U}$ covers X if and only if $\bigcap \{F : F \in \mathfrak{F}\} = \emptyset$. Thus if X is compact and $\bigcap \{F : F \in \mathfrak{F}\} = \emptyset$, then some finite family $\{X \setminus F_j\}_{j=1}^n \subseteq \mathfrak{U}$ covers X, and so $\bigcap_{j=1}^n F_j = \emptyset$ and $\mathfrak{F}$ cannot have had the finite intersection property. Conversely, if X is not compact then there is some open cover $\mathfrak{U}$ of X that contains no finite subcover. That means that no finite family $\{X \setminus U_j\}_{j=1}^n$ of their complements can have empty intersection, *i.e.*, that the family of closed sets $\mathfrak{F} = \{X \setminus U : U \in \mathfrak{U}\}$ has the finite intersection property; however, $\bigcap \{F : F \in \mathfrak{F}\} = X \setminus \bigcup \{U : U \in \mathfrak{U}\} = \emptyset$ because $\mathfrak{U}$ is a cover of X.

The considerations just given "relativize" easily: instead of assuming that the whole space X is compact, to get a nonempty intersection it suffices to assume that *one* of the sets in a family $\mathfrak{F}$ with nonempty intersection is compact.[7] As a special case we have Palka's **4.5**: if $\{K_n\}_{n=1}^\infty$ is a sequence of nonempty compact subsets of a metric space X with $K_1 \supseteq \cdots K_n \supseteq K_{n+1} \supseteq \cdots$, then it patently has the finite intersection property since $\bigcap_{j=1}^n K_j = K_n$, and consequently $\bigcap_{n=1}^\infty K_n \neq \emptyset$.[8]

Palka's proof of **Theorem 4.6** carries over without substantial modification to give

[7] The reader should check this.

[8] Palka's 4.5 is usually called the *nested set theorem*. The usual statement of the Cantor Intersection Theorem requires the metric space to be complete but not necessarily compact, and the sets K_n to be closed (but not necessarily compact) and to have their diameters decrease to zero as $k \to \infty$. The conclusion—that the intersection is nonempty—then holds even in infinite-dimensional settings.

Proposition: Let (X, d) and (Y, ρ) be metric spaces, X be compact, and $f : X \to Y$ a continuous function; then the image $f[A]$ under f of any compact subset $A \subseteq X$ is compact in Y.

(The reader can check the details.) Similarly, his **4.7** becomes

Proposition: Suppose that A is a subset of a metric space (X, d), that $f : A \to \mathbb{R}$ is a continuous function, and that $K \subseteq A$ is compact. There exist points z_0 and w_0 in A such that $f(z_0) \leq f(z) \leq f(w_0)$ holds for every $z \in K$; *i.e.*, when restricted to K the function f attains both a minimum and a maximum value.

The proof is the same.

The proposition just asserted is handy when one wants to prove the metric-space version of Palka's **Lemma 4.4**:

Proposition: Suppose (X, d) is an open set in a metric space and that $K \subseteq U$ is compact. There exists a radius $r > 0$ with the property that for each point $z \in K$ the sphere $\Delta(z, r)$ is contained in U.

Civilized Proof. Consider the distance function $d(x, X \setminus U)$. As we saw on pp. 4–5 above, this is a continuous function of x, and since $X \setminus U$ is a closed set, $d(x, X \setminus U) = 0$ holds if and only if $x \in X \setminus U$. It follows that $d(z, X \setminus U) > 0$ holds for each $z \in K$, and since $d(z, X \setminus U)$ must attain a minimum value on the compact set K, there exist some $z_0 \in K$ and $r > 0$ for which $0 < r = d(z_0, X \setminus U) \leq d(z, X \setminus U)$ holds for every $z \in K$. But now if $z \in K$ then the sphere $\Delta(z, r)$ cannot touch $X \setminus U$: if there were to exist $x \in \Delta(z, r) \cap (X \setminus U)$ then one would have $d(z, X \setminus U) \leq d(z, x) < r$. It follows that $\Delta(z, r) \subseteq X \setminus (X \setminus U) = U$ for every $z \in K$.

Palka's §4.5 has as its sole result the famous theorem "a function continuous on a compact set is uniformly continuous there" of advanced calculus. This theorem also holds for metric spaces, with about the same proof: let us just review the definitions and why they are different. If (X, d) and (Y, ρ) are metric spaces, $A \subseteq X$ is nonempty, and $f : X \to Y$ is a function, then to say that "f is continuous on A" is to say

$$\forall \epsilon > 0 \ \forall x \in A \ \exists \delta > 0 \text{ such that } \forall z \in A\text{: if } d(x, z) < \delta \text{ then } \rho(f(x), f(z)) < \epsilon\ .$$

In the usual "two-person game" conceptual model, your opponent has to give you both $x \in A$ and $\epsilon > 0$ as information with which to determine your $\delta > 0$. However, to say that "f is uniformly continuous on A" is to say

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ such that } \forall x,\, z \in A\text{: if } d(x, z) < \delta \text{ then } \rho(f(x), f(z)) < \epsilon\ .$$

Thus, given $\epsilon > 0$ you must somehow come up with a $\delta > 0$ that works for every pair of points x, $z \in A$ simultaneously. Fortunately, the two problems are equivalent for compact A: if f is continuous on A, and if for a given $\epsilon > 0$ it were impossible to find a "$\delta > 0$" that worked for every pair of points x, z in A, then for every $n \in \mathbb{N}$ there would exist a pair z_n, w_n in A with $d(z_n, w_n) < 1/n$ but $\rho(f(z_n), f(w_n)) \geq \epsilon$. The compactness of A gives us a convergent subsequence $\{z_{n_k}\}_{k=1}^{\infty}$ with limit, say, $z_0 \in A$. By hypothesis f is continuous at z_0 so there is a $\delta > 0$ for which $d(x, z_0) < \delta \Rightarrow \rho(f(x), f(z_0)) < \epsilon/2$. But now if $k \in \mathbb{N}$ is so large that $1/n_k < \delta/2$ then both $d(z_0, z_{n_k}) < \delta$ and $d(z_0, w_{n_k}) < \delta$, thus both $\rho(f(z_0), f(z_{n_k})) < \epsilon/2$ and $\rho(f(z_0), f(w_{n_k})) < \epsilon/2$. The triangle inequality in Y then gives $\rho(f(z_{n_k}), f(w_{n_k})) \leq \rho(f(z_0), f(z_{n_k})) + \rho(f(z_0), f(w_{n_k})) < \epsilon$, contradicting the choice of z_{n_k} and w_{n_k} and showing that there must have existed a $\delta > 0$ for which $\forall x, z \in A$: if $d(x, z) < \delta$ then $\rho(f(x), f(z)) < \epsilon$.

One can use the Lebesgue Covering Lemma to give a proof of this theorem. To simplify notation, let us "relativize" and assume that $A = X$. The continuity of f implies that for each sphere $\Delta(y, \epsilon/2) \subseteq Y$, the f-preimage $f^{-1}[\Delta(y, \epsilon/2)] \subseteq X$ is open.[9] These sets form an open cover of X. If $\lambda > 0$ is such that every sphere of radius λ is contained in some single element of that cover (*i.e.*, if λ is a Lebesgue number of the cover), then whenever $d(z, w) < \lambda$ so $w \in \Delta(z, \lambda)$, we shall have $\rho(f(z), y) < \epsilon/2$ and $\rho(f(w), y) < \epsilon/2$ for some $y \in Y$ and therefore $\rho(f(z), f(w)) \leq \rho(f(z), y) + \rho(y, f(w)) < \epsilon$.

5. *Uniform Convergence of Sequences of Functions.*

Palka does not treat this matter until Ch. VII, p. 243 ff., but since it fits in well here, let's look at it now. Suppose (X, d) and (Y, ρ) are metric spaces and $\{f_n\}_{n=1}^{\infty}$ is a sequence of functions $f_n : X \to Y$. It may happen that for each $x \in X$ the sequence $\{f_n(x)\}_{n=1}^{\infty} \subseteq Y$ is convergent; if so, the limit defines a function $f : x \to \lim_{n\to\infty} f_n(x)$ from X to Y called the **pointwise limit** of the f_n's, and the sequence is said to be **pointwise convergent to** f. If Y is complete it is necessary and sufficient for the pointwise convergence of $\{f_n\}_{n=1}^{\infty}$ that each $\{f_n(x)\}_{n=1}^{\infty}$ be a Cauchy sequence; one says that the sequence is **pointwise Cauchy**. Written out with quantifiers, the definition of a pointwise convergent sequence looks like

$$\forall x \in X \ \ \forall \epsilon > 0 \ \ \exists N \in \mathbb{N} \text{ such that } n \geq N \Rightarrow \rho(f(x), f_n(x)) < \epsilon$$

with a similar definition for pointwise Cauchy. The choice of N is allowed to depend on both $x \in X$ and $\epsilon > 0$. However, it may happen that for every $\epsilon > 0$ one can choose the same N for all the x's in a certain subset $A \subseteq X$, *i.e.*, that the following stronger condition holds:

$$\forall \epsilon > 0 \ \ \exists N \in \mathbb{N} \text{ such that } n \geq N \Rightarrow \forall x \in A, \ \ \rho(f(x), f_n(x)) < \epsilon \ .$$

When this condition holds, one says that the sequence $\{f_n\}_{n=1}^{\infty}$ is **uniformly convergent to** f on A. There is a similar definition for sequences **uniformly Cauchy on** A:

$$\forall \epsilon > 0 \ \ \exists N \in \mathbb{N} \text{ such that } m, n \geq N \Rightarrow \forall x \in A, \ \ \rho(f_m(x), f_n(x)) < \epsilon \ .$$

The cases of most interest are those in which (i) $A = X$—**uniform convergence on** X; (ii) the condition holds for every compact $A \subseteq X$—**uniform convergence on compacta**;[10]

---

[9] Cf. Palka's Theorem 2.5, p. 43.

[10] Palka (pp. 246–248) calls this *normal convergence*, which is rather old-fashioned; the name we give is currently standard.

and (iii) A can be any finite subset of X (but that just gives pointwise convergence). An important property of uniform convergence is

Proposition: If $\{f_n\}_{n=1}^{\infty}$ is a sequence of *continuous* functions from (A,d) to (Y,ρ) that converges *uniformly to* $f(x)$ *for* $x \in A$, then $f : A \to Y$ is continuous.

Proof. Let $\epsilon > 0$ be given. For any points x, $x_0 \in A$ one can write for any $n \in \mathbb{N}$

$$(*) \qquad \rho(f(x_0), f(x)) \le \rho(f(x_0), f_n(x_0)) + \rho(f_n(x_0), f_n(x)) + \rho(f_n(x), f(x)) \ .$$

The uniform convergence of $\{f_n\}_{n=1}^{\infty}$ to f lets us find and fix $n \in \mathbb{N}$ for which the first and third terms of the r. h. s. of $(*)$ are $< \epsilon/3$, *independent of the choice of* $x \in A$. The continuity of f_n then lets us find $\delta > 0$ for which $d(x_0, x) < \delta$ and $x \in A$ imply $\rho(f_n(x_0), f_n(x)) < \epsilon/3$, and now $(*)$ tells us that that condition on x implies that $\rho(f(x_0), f(x)) < \epsilon$. That proves continuity of f at (the arbitrary point) $x_0 \in A$.

The most interesting cases of the Proposition are (i): if $\{f_n\}_{n=1}^{\infty}$ is a sequence of continuous functions that converges uniformly on X, then the (pointwise) limit function is continuous on X. (ii) If $\{f_n\}_{n=1}^{\infty}$ is a sequence of continuous functions that converges uniformly on A for every compact $A \subseteq X$ (though the "N" that goes with the "ϵ" may get larger as A does) then the pointwise limit function is continuous on X, because for every convergent sequence $\{x_k\}_{k=1}^{\infty}$ with limit x_0 in X we may take $A = \{x_k\}_{k=1}^{\infty} \cup \{x_0\}$, which is a compact set.[11] The proposition tells us that the limit function f is continuous on A; but then $\lim_{k\to\infty} f(x_n) = f(x_0)$. Since $\{x_k\}_{k=1}^{\infty}$ with limit x_0 was arbitrary, this shows that f satisfies the sequential condition for continuity at every $x_0 \in X$, so f is continuous on X.

As a final matter to consider, suppose $\{f_n : X \to Y\}_{n=1}^{\infty}$ is a sequence of functions that is uniformly Cauchy on a set $A \subseteq X$. Assuming (Y,ρ) is complete, the limit function $f(x) = \lim_{n\to\infty} f_n(x)$ is then well defined on the points $x \in A$ by taking the pointwise limit. But now we see that *this pointwise limit is attained uniformly on A:* given $\epsilon > 0$ we can find $N \in \mathbb{N}$ for which $m, n \ge N \Rightarrow \forall x \in A,\ d(f_m(x), f_n(x)) < \epsilon/2$. Letting $m \to \infty$ but holding $n \ge N$ fixed, we see that then $\forall x \in A,\ d(f(x), f_n(x)) \le \epsilon/2 < \epsilon$; so the defining condition for "$f_n \to f$ uniformly on A" is satisfied.

[11] Why? The reader should prove this.