Math 55a: Honors Advanced Calculus and Linear Algebra

Metric topology III: Introduction to functions and continuity

NB we diverge here from the order of presentation in Rudin, where continuity is postponed until Chapter 4.

Continuity of functions between metric spaces. In a typical mathematical theory one is concerned not only with the objects and their properties, but also (and perhaps even more importantly) with the relationships among them, and with how they connect with already familiar mathematical objects — a mathematical theory concerned only with objects and properties, not their relationships, would likely be as sterile as a language with nouns and adjectives but no verbs. Typically the "verbs" are *functions* (a.k.a. *maps*) between the theory's objects. In the theory of metric spaces, we have already encountered two classes of functions: distance functions, and isometries between metric spaces. For most uses we need a much richer class of functions, the *continuous functions.*

Let X, Y be metric spaces, E a subset of X, and f some function from E to Y. For any $p \in E$, we shall say that f is *continuous at* p if, for a variable point $q \in E$, we can guarantee that $f(q)$ is as close as desired to $f(p)$ by making q close enough to p; formally, if[1]

> for each real number $\epsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(p), f(q)) < \epsilon$ for all $q \in E$ such that $d_X(p, q) < \delta$.

The function f is said to be *continuous* if it is continuous at p for all $p \in E$.

Some basic examples follow. A constant function is continuous, since δ may be chosen arbitrarily. Any isometry is continuous; for any metric space X and any fixed $p_0 \in X$, the function $X \to \mathbf{R}$, $p \mapsto d(p, p_0)$ is continuous. In both these examples one may take $\delta = \epsilon$. Likewise, the distance function $d : X \times X \to \mathbf{R}$ is continuous, this time with $\delta = \epsilon/2$ (why?). In each of these cases, δ depends on ϵ but not on the choice of p; when this happens, the function is said to be *uniformly continuous* — we shall say much more about uniform continuity in the coming lectures. An example of a function that is continuous but not uniformly continuous is the map $x \mapsto 1/x$ from $(0, 1)$ to $\mathbf{R}$ (see below). If X has the discrete metric then any f is continuous (why?).

A topological reformulation of continuity. Continuity can be profitably stated in terms of neighborhoods and open sets. This takes several steps to accomplish, which we do next.

To begin with, we immediately translate the definition into the following equivalent form:

[1] Because of the traditional choice of variables ϵ, δ in this fundamental definition, rigorous derivations of facts about continuous functions are often called "epsilon-delta proofs".

> $f : E \to Y$ is continuous at p if and only if for each neighborhood $N_\epsilon(f(p))$ of $f(p)$ there exists a neighborhood $N_\delta(p) \cap E$ of p such that $q \in N_\delta(p) \cap E \Rightarrow f(q) \in N_\epsilon(f(p))$.

Note that of necessity we use the "relative neighborhood" $N_\delta(p) \cap E$ of p, i.e. the δ-neighborhood of p considered as an element of the metric space E, not X. A "relatively open" subset of E is defined similarly; see appendix below.

We next rephrase this new definition in terms of "inverse images". For any function $f : X \to Y$ and any subset $S \subseteq Y$ we define the *inverse image* $f^{-1}(S)$ as the set of all points of X whose image under f is contained in S:

$$f^{-1}(S) := \{x \in X : f(X) \in S\}.$$

Notice that f need not be a bijection: inverse images are defined for any function. Thus our definition of continuity becomes

> $f : E \to Y$ is continuous at p if and only if for each neighborhood $N_\epsilon(f(p))$ of $f(p)$ the preimage $f^{-1}(N_\epsilon(f(p)))$ contains a relative neighborhood of p.

We deduce the following characterization of continuity. For simplicity we first obtain it when $E = X$:

Theorem. [Rudin, 4.8, p.86–87] *A function $f : X \to Y$ between metric spaces is continuous if and only if $f^{-1}(V)$ is open in X for every open $V \subseteq Y$.*

Proof: ($\Rightarrow$) We must show that each $p \in f^{-1}(V)$ is an interior point of $f^{-1}(V)$. Since V is open and $f(p) \in V$, there exists $\epsilon > 0$ such that $N_\epsilon(f(p)) \subseteq V$. Since f is continuous at p, the preimage $f^{-1}(N_\epsilon(f(p)))$ of this neighborhood contains a neighborhood of p. This neighborhood is thus contained in $f^{-1}(V)$, and we are done.
($\Leftarrow$) We show that f is continuous at each $p \in X$. For any $\epsilon > 0$, the ϵ-neighborhood $N_r(f(p))$ is open in Y; thus by hypothesis, its preimage is open in X. This preimage contains p, and thus contains some neighborhood of p. $\square$

In the general case we obtain:

Let X, Y be metric spaces and $E \subset X$. A function $f : E \to Y$ is continuous if and only if $f^{-1}(V)$ is relatively open in E for every open $V \subseteq Y$.

Proof: Consider E as a metric space in its own right, and apply the previous theorem. $\square$

Applications of the topological formulation of continuity. There are several advantages to this formulation of continuity. One is that it expresses continuity as a topological notion, i.e. in terms of open sets without mentioning the distance function. Thus it gives a natural extension of the notion of continuity to the category of topological spaces. A more immediate advantage is that several results now become almost obvious, notably the following:

Theorem. [Rudin, 4.7, p.86] *Let $f : X \to Y$, $g : Y \to Z$ be maps between metric spaces, and $h = (g \circ f) : X \to Z$ the composite map. Then if f and g are continuous, so is h.*

Proof: For any open subset $V \subseteq Z$, the preimage $g^{-1}(V)$ is open in Y because g is continuous; thus so is its preimage $f^{-1}(g^{-1}(V))$ in X. But $f^{-1}(g^{-1}(V)) = (g \circ f)^{-1}(V) = h^{-1}(V)$. Thus h is continuous. $\square$

[We have used here the following convenient fact about inverse images: if $f : X \to Y$ and $g : Y \to Z$ then for any $S \subseteq Z$ we have $(g \circ f)^{-1}(S) = f^{-1}(g^{-1}(S))$. A further useful fact of this nature is that the inverse image of an arbitrary union or intersection is the union or intersection, respectively, of the inverse images; symbolically, $f^{-1}(\cup_\alpha S_\alpha) = \cup_\alpha f^{-1}(S_\alpha)$ and $f^{-1}(\cap_\alpha S_\alpha) = \cap_\alpha f^{-1}(S_\alpha)$. You should verify all these claims. Curiously inverse images behave more nicely than (direct) images, defined similarly by $f(S) := \{f(x) \in X : x \in S\}$. Which of the above properties of inverse images fails for direct images, and why?]

We can now use the fact that the composition of continuous functions is continuous as a powerfool tool for generating new continuous functions from previously known ones. For instance, consider

Theorem. [Rudin, 4.9, p.87] *Let $f, g : X \to \mathbf{C}$ be continuous functions on a metric space X. Then $f + g$, $f - g$, fg are continuous. If $g(x) \neq 0$ for all $x \in X$ then the function f/g is continuous.*

[$f + g$ is the function taking any $x \in X$ to $f(x) + g(x)$; likewise $f - g$, fg, f/g.]

This would follow from the following facts: *i) If X, Y, Z are metric spaces and $f : X \to Y$, $g : X \to Z$ are continuous functions, then the map $(f, g) : X \to Y \times Z$ defined by $(f, g)(p) = (f(p), g(p))$ is continuous.*
ii) The functions from $\mathbf{C}^2$ to $\mathbf{C}$ taking (z_1, z_2) to $z_1 + z_2$, $z_1 - z_2$, $z_1 z_2$ are continuous. The function from $\mathbf{C} \times \mathbf{C}^*$ to $\mathbf{C}$ taking (z_1, z_2) to z_1/z_2 is continuous.

[In (i), we put the sup metric on $Y \times Z$. In (ii), $\mathbf{C}^*$ is $\mathbf{C} - \{0\}$, which inherits its metric structure from $\mathbf{C}$.]

Supposing we can both prove the general fact (i) and show that the specific functions in (ii) are continuous. Then the Theorem is an easy consequence. For instance we obtain fg by composing the function $(f, g) : X \to \mathbf{C}^2$ with the function from $\mathbf{C}^2$ to $\mathbf{C}$ taking (z_1, z_2) to $z_1 z_2$.

Now (i) is straightforward. Given p, ϵ we use the continuity of f, g to find positive δ_f, δ_g such that $d(p, q) < \delta_f \Rightarrow d(f(p), f(q)) < \epsilon$, $d(p, q) < \delta_g \Rightarrow d(g(p), g(q)) < \epsilon$, and let $\delta = \min(\delta_f, \delta_g)$. Conversely, one readily shows (do it!) that if (f, g) is continuous then each of f, g is. More generally, a function f from X to a Cartesian product $Y_1 \times \cdots \times Y_n$ is just an ordered n-tuple $(f_1, \ldots, f_n)$ with each $f_i : X \to Y_i$; if X, Y_i are metric spaces then f is continuous if and only if each f_i is.

As to (ii), we must first specify the metric on $\mathbf{C}$, and we choose $d(p, q) := |p - q|$. Recall that the absolute value $|z|$ of a complex number $z = x + iy$ is by definition $\sqrt{x^2 + y^2} = (z\bar{z})^{1/2}$, $\bar{z}$ being the complex conjugate $x - iy$ of z. We must verify that this $d(\cdot, \cdot)$ is actually a metric; see [Rudin, 1.33, p.14–15], or the Appendix below. This immediately yields the continuity of $z_1 \pm z_2$. As for $z_1 z_2$, let $z_1' = z_1 + \delta_1$, $z_2' = z_2 + \delta_2$. Then

$$|z_1' z_2' - z_1 z_2| = |z_1 \delta_2 + \delta_1 z_2 + \delta_1 \delta_2| \leq |z_1||\delta_2| + |\delta_1||z_2| + |\delta_1||\delta_2|.$$

Fix $\epsilon > 0$. We shall show that, for sufficiently small $\delta > 0$, each of the three terms $|z_1||\delta_2|, |\delta_1||z_2|, |\delta_1||\delta_2|$ is $< \epsilon/3$ provided $|\delta_1| < \delta$ and $|\delta_2| < \delta$. Indeed, it is enough that δ be the smallest of $\epsilon/3|z_1|$, $\epsilon/3|z_2|$, $(\epsilon/3)^{1/2}$. (If z_1 or z_2 is zero, the corresponding bound may be ignored, because then $z_1\delta_2$ or $\delta_1 z_2$ respectively does not contribute to $z_1' z_2' - z_1 z_2$.) This shows that the function $(z_1, z_2) \mapsto z_1 z_2$ is continuous.

This leaves only z_1/z_2. To show that this last function is continuous we need only prove the continuity of the function from $\mathbf{C}^*$ to $\mathbf{C}$ taking z to $1/z$ (why?). Suppose $\epsilon > 0$ and $z \neq 0$ are given. Let $z' = z + \delta_1$ as before. Then

$$\left| \frac{1}{z'} - \frac{1}{z} \right| = \left| -\frac{\delta_1}{zz'} \right| = \frac{|\delta_1|}{|z||z'|}.$$

Let δ be the smaller of $|z|/2$ and $(|z|^2/2)\epsilon$. Then if $|\delta_1| < \delta$ then $|z'| > |z|/2$ (triangle inequality), so

$$\frac{|\delta_1|}{|z||z'|} < \frac{\delta}{\frac{1}{2}|z|^2} < \epsilon,$$

and we are done!

Appendix: relatively open sets; the metric on C. Having encountered relatively open subsets, we shall find the following characterization useful:

Theorem. [Rudin, 2.30, p.36] *Let X be a metric space, and $E \subseteq Y \subseteq X$. Then E is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.*

Proof: ($\Rightarrow$) For each $p \in E$ there exists $r_p > 0$ such that E contains the r_p-neighborhood V_p of p. Let $G = \cup_p V_p$. Then $E = Y \cap G$ (why?), and G is open because it is a union of open sets. [NB It can be useful to know that an *arbitrary* union of open sets is open!]
($\Leftarrow$) Since G is open in X and contains E, each $p \in E$ has a neighborhood V_p contained in G. Thus the relative neighborhood $V_p \cap Y$ is contained in $G \cap Y = E$. We have shown that every $p \in E$ has a relative neighborhood contained in E, so E is open as claimed. $\square$

Here is the proof that $d(p,q) = |p-q|$ is a metric on $\mathbf{C}$. First, from $|z| = (z\bar{z})^{1/2}$ we deduce that $|z| \geq 0$, and $|z| = 0 \iff z = 0$. Thus $d(p,q) \geq 0$ with equality if and only if $p = q$. As with the standard metric on $\mathbf{R}$, symmetry follows from $|z| = |-z|$, and the triangle inequality from $|w+z| \leq |w| + |z|$. This last is nontrivial, but can be proved as follows. First note that for any $z = x + iy$ we have $|z| = |\bar{z}|$ and $|z|^2 = x^2 + y^2 \geq x^2$, so $|z| > |x|$. Next, for any $w, z \in \mathbf{C}$ we have

$$|wz| = (wz \cdot \overline{wz})^{1/2} = (w\bar{w}\, z\bar{z})^{1/2} = (w\bar{w})^{1/2}(z\bar{z})^{1/2} = |w||z|.$$

Finally, we calculate

$$\begin{aligned}
|w+z|^2 &= (w+z)(\bar{w}+\bar{z}) = w\bar{w} + w\bar{z} + z\bar{w} + z\bar{z} \\
&= |w|^2 + |z|^2 + (w\bar{z} + \overline{w\bar{z}}) \\
&= |w|^2 + |z|^2 + 2\,\mathrm{Re}(w\bar{z}) \\
&\leq |w|^2 + |z|^2 + 2|w\bar{z}| \\
&= |w|^2 + 2|w||z| + |z|^2 = (|w| + |z|)^2
\end{aligned}$$

so $|w+z| \leq |w| + |z|$ as claimed.