EXTENDING CALCULUS: LIMITS

ARICK SHAO

1. Limits in General Spaces

In first-year calculus, one encounters to some extent the basic definition of limits of functions. This is generally referred to as *"delta-epsilon"*, due to the nearly unanimous choice of Greek letters used in this definition. Here, we will discuss how one generalizes such definitions of limits to other spaces besides the real numbers. Examples of "other spaces" include $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^n$ for general n – higher dimensional spaces which are the settings of *multivariable calculus* – as well as many *infinite-dimensional spaces* (such as certain spaces of functions).

Consider first the familiar case of the real number line $\mathbb{R}$ (i.e., from single-variable calculus). This formal notion of limits and all the concepts that follow from this provide a vast analytical toolbox that one can use to study the real numbers and various functions involving real numbers. By generalizing this notion of limits, then one can essentially apply this "toolbox" in multi-dimensional or even infinite-dimensional settings. In many areas of physics and mathematics (e.g., quantum mechanics, partial differential equations), these notions of "limits" and "convergence" for infinite-dimensional spaces are fundamental.

1.1. **δ's and ϵ's.** First, we consider a function

$$f : (a, b) \to \mathbb{R},$$

and we fix a point $x_0 \in [a, b]$, that is, on (a, b) or its boundary. We say that

$$\lim_{x \to x_0} f(x) = L$$

iff for any $\epsilon > 0$, there exists $\delta > 0$ such that if

$$x \in (a, b), \qquad 0 < |x - x_0| < \delta,$$

then

$$|f(x) - L| < \epsilon.$$

In less formal terms, this means that $f(x)$ can become as close as one wants to L, as long as one takes $x \in (a, b)$ close enough to x_0.

Similarly, for a sequence $x_1, x_2, \dots$, we say that

$$\lim_{n \to \infty} x_n = L$$

iff for any $\epsilon > 0$, there exists some $N_0 \in \mathbb{N}$ such that if $n \geq N_0$, then

$$|x_n - L| < \epsilon.$$

The intuition is more or less the same. One can make x_n as close as one wants to L, as long as one takes n to be sufficiently large.

These are the familiar definitions found, if not adequately discussed, in nearly every calculus textbook. In fact, they form much of the foundations for all of calculus. As these are rather powerful concepts, one would like to extend and apply them to other potentially more interesting situations.

To see how one might extend these definitions, we look first at the intuitions involved. The statements $0 < |x - x_0| < \delta$ and $|f(x) - L| < \epsilon$ are generally interpreted as x being "δ-close" to x_0 and $f(x)$ being "ϵ-close" to L. Indeed, the absolute value of the difference of two numbers represents precisely the *distance* between these numbers on the real line. In fact, in our definitions of limits, the only intuition about $|x - x_0|$ and $|f(x) - L|$, for instance, that comes into play is the observation that they represent the distances between x and x_0 and between $f(x)$ and L, respectively. This suggests that our definitions of limits ought to be sensible if we replace the real number line and these absolute values of differences by some other "space" and some other "measure of distance" on that "space".

Consider, for example, the usual three-dimensional space,

$$\mathbb{R}^3 = \{(x, y, z) | x, y, z \in \mathbb{R}\}.$$

We can think of each element of $\mathbb{R}^3$, i.e., an ordered triple of real numbers, as a positional vector. Given two such vectors

$$\mathbf{r}_1 = (x_1, y_1, z_1), \qquad \mathbf{r}_2 = (x_2, y_2, z_2),$$

then the distance between them is the length of the line segment connecting the two points. This is precisely the magnitude of the difference of the vectors:

$$|\mathbf{r}_2 - \mathbf{r}_1| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Thus, on the "space" $\mathbb{R}^3$, we have a "measure of distance" given by the above.

This is, of course, directly extendible to the space $\mathbb{R}^n$ of n-tuples (of real numbers) for any n. In other words, on the space $\mathbb{R}^n$, we have the "measure of distance"

$$d(\mathbf{r}_1, \mathbf{r}_2) = |\mathbf{r}_2 - \mathbf{r}_1|, \qquad \mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^n.$$

Let us now take these "spaces" and "distances" back to our definitions of limits. Consider now a sequence $\mathbf{r}_1, \mathbf{r}_2, \dots$ of points in, say, $\mathbb{R}^3$. We then define

$$\lim_{n \to \infty} \mathbf{r}_n = \mathbf{L}, \qquad \mathbf{L} \in \mathbb{R}^3,$$

to mean that for any $\epsilon > 0$, there exists some $N_0 \in \mathbb{N}$ such that if $n \geq N_0$, then

$$|\mathbf{r}_n - \mathbf{L}| < \epsilon.$$

The intuition is exactly the same as before – one can make $\mathbf{r}_n$ as close as one wants to $\mathbf{L}$, as long as one takes n to be sufficiently large. Observe also that there is nothing special about the dimension 3, as this definition works just the same when "3" is replaced by any natural number n.

Similarly, for limits of functions, let $D \subseteq \mathbb{R}^m$ be some region in $\mathbb{R}^m$, let

$$\mathbf{F} : D \to \mathbb{R}^n, \qquad \mathbf{r}_0 \in D.$$

We say that [1]

$$\lim_{\mathbf{r} \to \mathbf{r}_0} \mathbf{F}(\mathbf{r}) = \mathbf{L}.$$

iff for any $\epsilon > 0$, there exists $\delta > 0$ such that if

$$\mathbf{r} \in D, \qquad 0 < |\mathbf{r} - \mathbf{r}_0| < \delta,$$

[1] Note this definition still works if $\mathbf{r}_0$ is on the boundary of D.

then

$$|\mathbf{F}(\mathbf{r}) - \mathbf{L}| < \epsilon.$$

The intuition is once again unchanged – $\mathbf{F}(\mathbf{r})$ can become as close as one wants to $\mathbf{L}$, as long as one takes $\mathbf{r} \in D$ close enough to $\mathbf{r}_0$.

Even if one avoids the technical intricacies of these "δ-ϵ" or "N_0-ϵ" definitions, which can seem rather intimidating at first, the above discussion still remains applicable. When one speaks of "limits", of something "approaching" or "converging to" something, one refers to *some measures of distance shrinking closer and closer to zero.* Other properties of the real line provide only extraneous detail. Thus, the above discussion of replacing the distance on $\mathbb{R}$ by other distances on other spaces remains as relevant even in the context of purely informal discussions.

Exercise 1. *Do the following limits exist? If so, what are the limits? Why?*

$$\lim_{(x,y)\to(0,0)} \frac{2x^2 - y^2}{x^2 + y^2}, \qquad \lim_{(x,y)\to(0,0)} \frac{\sin(x^2 + y^2)}{x^2 + y^2},$$

$$\lim_{n\to\infty} \left(2^{-n}, 1 - \frac{1}{n}\sin n, \frac{3n^2 + 4n}{-2n^2 + 2}\right), \qquad \lim_{(x,y,z)\to(1,3,4)} \left(y, \frac{xz - 4x + 2z - 8}{z - 4}\right),$$

$$\lim_{(x,y)\to(0,0)} \frac{x^2 \sin y}{x^2 + y^2}, \qquad \lim_{(x,y)\to(0,0)} \frac{x^2 \cos y}{x^2 + y^2},$$

$$\lim_{(x,y)\to(-1,1)} e^{-xy}\cos(x + y), \qquad \lim_{n\to\infty} \left(\sqrt{n^2 + n} - n\right)$$

Since the notion of "distance" extends far beyond just $\mathbb{R}^n$, there is no need for us to consider only these $\mathbb{R}^n$ when we speak of limits. For example, we can consider settings which are less "flat", like the sphere

$$\mathbb{S}^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}.$$

One can similarly define the distance between two points on the sphere. [2] Thus, as before, one can discuss limits of functions defined on the sphere, and of functions which take values on the sphere. The same can be said of other curved surfaces and also of higher (or lower) dimensional figures.

1.2. Metric Spaces.

Next, we take a step back and describe what, exactly, we mean by "distance". Thus far, we have left this notion undefined and appealed instead to our intuitions. However, as we deal with other settings that are more difficult or impossible to visualize, such as "infinite-dimensional" spaces, we will want to have a more formal characterization of "distance." In abstract mathematics, the concept that achieves this purpose is that of a *metric space.*

Consider some set X. Let $X \times X$ denote the set of ordered pairs, with each component being an element of X, i.e.,

$$X \times X = \{(x, y) \mid x, y \in X\}.$$

A nonnegative real-valued function

$$d : X \times X \to [0, \infty)$$

is called a *metric* on X iff the following conditions hold:

(1) $d(x, y) = 0$ if and only if $x = y$, for any $x, y \in X$.

[2] This distance is the length of the shorter arc on the great circle that contains the two points.

(2) $d(x, y) = d(y, x)$ for any $x, y \in X$.
(3) $d(x, z) \leq d(x, y) + d(y, z)$ for any $x, y, z \in X$.

If d is such a metric, then the pair (X, d) is generally referred to as a *metric space*. Intuitively, $d(x, y)$ represents the "distance" from x to y, while X is the "space" of all things for which this distance d can be measured.

These three conditions comprising the definition of a metric represent "reasonable" criteria that any measure of distance should satisfy. The condition (1) simply states that two points contain no distance between them if and only if they are one and the same. The symmetry condition (2) can be interpreted as that the distance traveled from x to y must be the same as the distance traveled from y to x. Finally, the *triangle inequality* (3) states that one cannot shorten the distance via additional detours – if one goes from x to y and then from y to z, this combined distance has to be at least as long the direct distance between x and z.

For example, for any $n \in \mathbb{N}$, then the function

$$d(\mathbf{r}_1, \mathbf{r}_2) = |\mathbf{r}_1 - \mathbf{r}_2|, \qquad \mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^n$$

is a metric on $\mathbb{R}^n$. The aforementioned distance on the sphere $\mathbb{S}^2$ is also a metric. Analogous distances on curves, curved surfaces, and higher-dimensional analogues (generally called *manifolds*) are also metrics, though a fair amount of differential geometry is required in order to prove this.

Thus far, we have still restricted ourselves to finite-dimensional objects. With a firmly defined notion of "distance", though, we can now move far beyond this. The main case of interest in mathematical analysis (and in mathematical physics) will be spaces of functions. In other words, we will want to analyze families of functions using many of the same tools that we have used to analyze real numbers.

For example, consider a closed interval $[a, b]$ in the real line, and let $C[a, b]$ denote the set of all continuous functions from $[a, b]$ into $\mathbb{R}$, i.e., of the form

$$f : [a, b] \rightarrow \mathbb{R}.$$

The *extreme value theorem* from basic calculus implies that any $f \in C[a, b]$ must be bounded, and that f must attain its maximum and minimum values. [3] The actual proof of this requires quite a bit of work, but we will take this fact for granted here.

With the above in mind, we can define the following metric on $C[a, b]$: given continuous functions $f, g : [a, b] \rightarrow \mathbb{R}$, we define

$$d_\infty(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)| = \max_{x \in [a,b]} |f(x) - g(x)|.$$

Some other commonly used "distances" for spaces of functions are listed below:

(1) The "square integral" distance:

$$d_2(f, g) = \left(\int_a^b |f(x) - g(x)|^2 dx \right)^{\frac{1}{2}}, \qquad f, g : [a, b] \rightarrow \mathbb{R}.$$

The corresponding "space" $L^2[a, b]$ is (roughly) that of all square integrable real-valued functions on $[a, b]$, i.e., functions $f : [a, b] \rightarrow \mathbb{R}$ such that

$$\int_a^b |f(x)|^2 dx < \infty.$$

[3] In other words, there is some $x_0 \in [a, b]$ with $f(x_0) = \sup_{x \in [a,b]} f(x)$, and similarly for inf.

(2) The integral distance:

$$d_1(f,g) = \int_a^b |f(x) - g(x)| dx, \qquad f, g : [a,b] \to \mathbb{R}.$$

The corresponding "space" $L^1[a,b]$ is (again roughly) that of all integrable real-valued functions on $[a,b]$, i.e., functions $f : [a,b] \to \mathbb{R}$ such that

$$\int_a^b |f(x)| dx < \infty.$$

(3) The "squared summation distance" for sequences:

$$d_2^*((x_n),(y_n)) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^2 \right)^{\frac{1}{2}}, \qquad (x_n),(y_n) \text{ are real-valued sequences.}$$

The corresponding "space" ℓ^2 is defined similarly to the preceding examples.

These are only some of the most basic examples. In practice, many other spaces of functions are used in many different areas of study.

Exercise 2. *Show that this function d_∞ is in fact a metric on $C[a,b]$.*

1.3. **Abstract Limits.** We have already established that all we need in order to have a viable notion of limits is some notion of "distance". Moreover, in our preliminary discussions involving $\mathbb{R}^n$, we already described how this abstract definition of limits would be constructed. Thus, all we need to do here is to plug in our abstract metric spaces into our existing intuitions.

Let (X, d_X) be a metric space (i.e., $d_X : X \times X \to [0,\infty)$ is a metric on X). Suppose (x_n) is a sequence, with each element in X. Then, we write

$$\lim_{n\to\infty} x_n = z, \qquad z \in X,$$

iff for any $\epsilon > 0$, there exists some $N_0 \in \mathbb{N}$ such that for any $n \geq N_0$,

$$d_X(x_n, z) < \epsilon.$$

For example, if we plug in $\mathbb{R}^n$ for X and the standard distance in $\mathbb{R}^n$ for d_X, then we obtain the previous definition of limits for $\mathbb{R}^n$-valued sequences.

Exercise 3. *Do the following sequences in $C[0,1]$ converge (with respect to the metric d_∞)? If so, then what is the limit?*

(1) $f_n : [0,1] \to \mathbb{R}$, *where* $f_n(x) = n^{-1} \sin x$.
(2) $f_n : [0,1] \to \mathbb{R}$, *where* $f_n(x) = x^n$.

Exercise 4. *Consider the n-tuples*

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n, \qquad \mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n,$$

and consider the following functions:

$$d_1(\mathbf{x},\mathbf{y}) = \sum_{k=1}^n |x_k - y_k|,$$

$$d_2(\mathbf{x},\mathbf{y}) = \sqrt{\sum_{k=1}^n |x_k - y_k|^2},$$

$$d_\infty(\mathbf{x},\mathbf{y}) = \max_{k=1}^n |x_k - y_k|.$$

Show that d_1, d_2, and d_∞ are metrics on $\mathbb{R}^n$ (note d_2 is the usual distance on $\mathbb{R}^n$). Furthermore, show that if a sequence $(\mathbf{x}_n)$ in $\mathbb{R}^n$ converges to $\mathbf{y}$ with respect to one of these metrics, then it also converges to $\mathbf{y}$ with respect to the other two metrics.

Similarly, we say that (x_n) is a *Cauchy sequence* iff for any $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that for any $n, m \geq N_0$, we have

$$d_X(x_n, x_m) < \epsilon.$$

Then, the metric space (X, d_X) is called *complete* iff every Cauchy sequence in X actually has a (d_X-)limit in X. This is the full formalization of the definition of completeness that was previously discussed.

One can show (with some effort) that the spaces $C[a,b]$, $L^2[a,b]$, $L^1[a,b]$, ℓ^2 described in the preceding section, with their associated metrics, form complete metric spaces. The first case, for example, is incredibly important in the basic theory of ordinary differential equations. One can generate a sequence of "approximate solutions" to some differential equation, with each element of this sequence in $C[a,b]$. By showing that this sequence is a Cauchy sequence, then one knows that it in fact has a limit, which we hope would be the actual solution to the equation. This is an extremely robust technique for proving the existence of solutions to differential equations, even when such solutions cannot be solved explicitly!

The definition for limits of functions is once again analogous. Let (Y, d_Y) be another metric space, and let $f : X \to Y$. Given $x_0 \in X$ and $y_0 \in Y$, we say that

$$\lim_{x \to x_0} f(x) = y_0$$

iff for any $\epsilon > 0$, there exists $\delta > 0$ such that if

$$0 < d_X(x, x_0) < \delta, \qquad x \in X,$$

then

$$d_Y(f(x), y_0) < \epsilon.$$

Once again, the intuitions behind limits in $\mathbb{R}$ and also in $\mathbb{R}^n$ can be ported directly to this abstract metric space setting.

Exercise 5. *From your knowledge of limits in $\mathbb{R}$, what would be a reasonable definition for a function $f : X \to Y$ on metric spaces to be continuous?*

Exercise 6. *Prove whether the following functions are continuous or not:*

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

$$g : \mathbb{R}^2 \to \mathbb{R}, \qquad g(x,y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Exercise 7. *Are the following functions continuous?*

$$\mathcal{E} : C[-1,1] \to \mathbb{R}, \qquad \mathcal{E}(f) = f(0),$$

$$\mathcal{I} : C[-1,1] \to \mathbb{R}, \qquad \mathcal{I}(f) = \int_{-1}^{1} f(x)dx.$$

References

1. W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, Inc., 1976.
2. J. Stewart, *Multivariable calculus*, 7 ed., Brooks/Cole, 2012.