The Completion of a Metric Space

Let (X, d) be a metric space. The goal of these notes is to construct a complete metric space which contains X as a subspace and which is the "smallest" space with respect to these two properties. The resulting space will be denoted by $\overline{X}$ and will be called the *completion* of X with respect to d. The hard part is that we have nothing to work with except X itself, and somehow it seems we have to pull a larger space out of thin air. Indeed, the point is that the completion will be constructed using only X and the metric d, and will not rely on knowing that we already have a space "larger" than X.

First we describe the completion as a set. For any Cauchy sequence (x_n) in X, let $[(x_n)]$ denote the set of all Cauchy sequences (y_n) in X such that $d(x_n, y_n) \to 0$ as $n \to \infty$. So, $[(x_n)]$ consists of all Cauchy sequences in X whose terms are getting "closer" and "closer" to the terms in (x_n) itself. For example, the Cauchy sequence (x_n) is in $[(x_n)]$ since $d(x_n, x_n) = 0$ is the constant zero sequence and so definitely converges to 0.

The intuition is the following. Recall that if X is not complete, the Cauchy sequence (x_n) does not necessarily converge and a Cauchy sequence (y_n) such that $d(x_n, y_n) \to 0$ does not necessarily converge either. However, if these Cauchy sequences did converge, the condition $d(x_n, y_n) \to 0$ would imply that they converged to the same thing. Similarly, if (x_n) and (y_n) were Cauchy sequences which did converge and had the same limit, then the sequence (y_n) would be in the set $[(x_n])$ we have defined above. The point is that $[(x_n)]$ consists of all Cauchy sequences which, if they did converge, would converge to the same thing as (x_n) if (x_n) converged as well. In other words, we are grouping the Cauchy sequences in X according to whether or not they "should" converge to the same thing.

Remark. In fact, if you have taken a course in abstract algebra before, this construction should sound somewhat similar to the construction of the *quotient group* of a group by a normal subgroup. In that case, we group the elements of a group according to whether or not they define the same *cosets*, and then look at the set consisting of those cosets themselves. This is analogous to the idea we are using here, and indeed both constructions are instances of considering the *equivalence classes* of an *equivalence relation*. We will not use this terminology here, but you can check other sources—say Wikipedia—for further details.

We call the set $[(x_n)]$ the *class* represented by (x_n). If (y_n) and (z_n) are both in $[(x_n)]$, it follows from the triangle inequality that $d(y_n, z_n) \to 0$, so that (z_n) is in the class $[(y_n)]$ as well. In particular, since (x_n) is in $[(x_n)]$, (x_n) itself is in $[(y_n)]$. The point is that different Cauchy sequences (x_n) and (y_n) can still give rise to the same class $[(x_n)] = [(y_n)]$, and this happens precisely when (x_n) is in $[(y_n)]$ and vice-versa:

Proposition. *Two Cauchy sequences (x_n) and (y_n) in (X, d) represent the same class if and only if (x_n) is in the class $[(y_n)]$.*

Definition. The set $\overline{X}$ is the set of classes of Cauchy sequences in X:

$$\overline{X} := \{[(x_n)] \mid (x_n) \text{ is a Cauchy sequence in } X\}.$$

This will be the underlying set in the completion of (X, d). Note that if we had simply defined $\overline{X}$ to be the set of Cauchy sequences of X, without grouping them into "classes", we would end up with a scenario where multiple elements of $\overline{X}$ represented the same "point" so we wouldn't be uniquely describing the elements of the completion. The point of grouping Cauchy sequences into classes as above is to eliminate this ambiguity, which is essential for making our claims about the completion actually work out.

Next we will define a metric on $\overline{X}$ and show that we can view X as a subspace of the resulting metric space. Finally, we will show that $\overline{X}$ with the metric we construct is in fact a *complete* metric space. But first, let us consider a simple and yet standard example:

Example. Consider the rational numbers $\mathbb{Q}$ with the Euclidean metric. In this case, two Cauchy sequences (x_n) and (y_n) of rational numbers satisfy the condition that $|x_n - y_n| \to 0$ if and only if they converge to the same real number x in $\mathbb{R}$. Thus this makes clear—at least in this example—the general statement above that two Cauchy sequences are in the same class if they "should" converge to the same thing.

Indeed, in this case we can *identify* the set of classes of Cauchy sequences of rational numbers with the set of real numbers itself: to a class $[(x_n)]$ of Cauchy sequences of rational numbers we associate the real number x which is the common limit of the Cauchy sequences in $[(x_n)]$ when viewed as sequences in $\mathbb{R}$, and to a real number x we associate the class of all Cauchy sequences which converge to x in $\mathbb{R}$. This gives a *bijection* between the set of real numbers and the set $\overline{\mathbb{Q}}$ of classes of Cauchy sequences of rational numbers as defined above. The construction we describe in the rest of these notes applied to this example is how we can *construct* the real numbers from the rational numbers via Cauchy sequences.

Returning to the construction of the completion, we now define a metric on $\overline{X}$. Given two classes $[(x_n)], [(y_n)] \in \overline{X}$, we can form the sequence $(d(x_n, y_n))$ of real numbers: that is, the n-th term in this sequence is the distance between the n-th terms in (x_n) and (y_n). We claim that $(d(x_n, y_n))$ is a Cauchy sequence in $\mathbb{R}$. To see this, let $\epsilon > 0$ and pick $N, N' \in \mathbb{N}$ such that

$$d(x_n, x_m) < \frac{\epsilon}{2} \text{ for } n, m \geq N \text{ and } d(y_n, y_m) < \frac{\epsilon}{2} \text{ for } n, m \geq N'.$$

Then for $n, m \geq M := \max\{N, N'\}$, we have (after playing around with the triangle inequality a bit):

$$|d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_n, y_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

Thus the sequence of real numbers $(d(x_n, y_n))$ is indeed Cauchy in $\mathbb{R}$. Since $\mathbb{R}$ is complete, this sequence converges and we denote the resulting limit by $\bar{d}([(x_n)], [(y_n)])$:

$$\bar{d}([(x_n)], [(y_n)]) := \lim_{n \to \infty} d(x_n, y_n).$$

This defines the metric $\bar{d}$ on $\overline{X}$ we are looking for. A few remarks are in order.

First, to construct $\bar{d}$ we use the fact that $\mathbb{R}$ is complete in a crucial way. So, you could object that this construction already depends on knowing what $\mathbb{R}$ is and on proving beforehand that $\mathbb{R}$ is complete. This is absolutely correct: the construction of $\mathbb{R}$ as the completion of $\mathbb{Q}$ must be carried out separately before the procedure we are using here works. To be precise, after we have "defined" $\mathbb{R}$ as the set of classes of Cauchy sequences in $\mathbb{Q}$, we must construct the metric on $\mathbb{R}$ and show it is complete in some other way and without using the technique we are outlining here. This is certainly possible, but we will skip the details in these notes. Or, another way to get around this is by first constructing $\mathbb{R}$ using *Dedekind cuts* in $\mathbb{Q}$ and then showing that $\mathbb{R}$ defined in this way is complete. This is also possible, but again we leave the details to other sources.

Second, note that our definition of $\bar{d}$ uses the specific Cauchy sequences (x_n) and (y_n), whereas we said earlier that different Cauchy sequences can represent the same class. In other words, say that (x'_n) was another Cauchy sequence representing the same class as (x_n): $[(x'_n)] = [(x_n)]$. Then the "distance" from this element of $\overline{X}$ to $[(y_n)]$ should not depend on whether we are using (x'_n)

or (x_n) to represent the class $[(x'_n)] = [(x_n)]$. To be precise, it should be true that even if (x'_n) and (x_n) are different Cauchy sequences representing the same class, the "distances"

$$\bar{d}([(x_n)], [(y_n)]) = \lim_{n\to\infty} d(x_n, y_n) \text{ and } \bar{d}([(x'_n)], [(y_n)]) = \lim_{n\to\infty} d(x'_n, y_n)$$

should be the same. This too can be proven using the inequality:

$$|d(x_n, y_n) - d(x'_n, y_n)| \leq d(x_n, x'_n)$$

which is derived from the triangle inequality, and is what mean by saying that $\bar{d}$ is *well-defined*: in other words, $\bar{d}$ only depends on classes of Cauchy sequences—i.e. on elements of $\overline{X}$—and not on the specific Cauchy sequence we use to represent a class.

After all this is cleared up, it is not hard to check that $\bar{d}$ indeed defines a metric on $\overline{X}$, which we do now. First, since $d(x_n, y_n)$ is always nonnegative, the limit of such numbers is nonnegative so

$$\bar{d}([(x_n)], [(y_n)]) \geq 0 \text{ for any } [(x_n)], [(y_n)] \in \overline{X}.$$

This expression is zero if and only if $\lim_{n\to\infty} d(x_n, y_n) = 0$, which is what it meant to say that the classes $[(x_n)]$ and $[(y_n)]$ are equal. (Note that here it is important that we are defining the elements of the completion as *classes* of Cauchy sequences and not simply as Cauchy sequences themselves.) This verifies the first property required of a metric. Since $d(x_n, y_n) = d(y_n, x_n)$, the limits of such expressions are the same, which verifies the second property required of a metric. Finally, suppose that $[(x_n)], [(y_n)], [(z_n)] \in \overline{X}$. Take the triangle inequality for X:

$$d(x_n, y_n) \leq d(x_n, z_n) + d(z_n, y_n),$$

and take limits throughout. The left side converges to $\bar{d}([(x_n)], [(y_n)])$ by definition, and the right converges to

$$\lim_{n\to\infty} d(x_n, z_n) + \lim_{n\to\infty} d(z_n, y_n),$$

which is $\bar{d}([(x_n)], [(z_n)]) + \bar{d}([(z_n)], [(y_n)])$. This gives the triangle inequality for $\bar{d}$.

Thus we have an honest metric on $\overline{X}$, and we have the definition:

Definition. Given a metric space (X, d), the *completion of X with respect to d* is the set $\overline{X}$ of classes of Cauchy sequences in X with the metric $\bar{d}$ defined above.

Example. Returning to the example of the completion of $\mathbb{Q}$ with respect to the standard metric, it is not hard to see using the continuity of the absolute value function that under the previously-mentioned identification of $\overline{\mathbb{Q}}$ with $\mathbb{R}$, the metric $\bar{d}$ on $\overline{\mathbb{Q}}$ indeed corresponds to the usual absolute value metric on $\mathbb{R}$.

Note that there is a natural way to view X as a subspace of $\overline{X}$: we simply associate to each $x \in X$ the constant sequence $x, x, x, \ldots$, and consider the class $[(x)]$ of this specific Cauchy sequence. This defines an injective mapping $X \to \overline{X}$, and one can check that the metric $\bar{d}$ evaluated on the constant Cauchy sequences $[(x)]$ and $[(y)]$ gives the same value as the distance $d(x, y)$ determined by the original metric d. So, we view (X, d) as the subspace of $(\overline{X}, \bar{d})$ consisting of classes of constant Cauchy sequences.

So, we have accomplished the first part of our goal: given a metric space (X, d), we have constructed a "larger" metric space $(\overline{X}, \bar{d})$ containing X as a subspace. Now, we claim that the metric space $(\overline{X}, \bar{d})$ constructed in this way is always complete:

Theorem. *Given any metric space (X, d), the completion $(\overline{X}, \bar{d})$ defined as above is complete. Moreover, the completion is the smallest complete metric space containing X in the sense that any other complete space Y containing X as a subspace also contains $\overline{X}$ as a subspace.*

The tricky part about this is in simply wrapping your head around what it means: a Cauchy sequence in $\overline{X}$ looks like (p_n) where each $p_n \in \overline{X}$, meaning that each p_n is itself a (class) of Cauchy sequences in X! So, (p_n) is something like a "Cauchy sequence of Cauchy sequences", and the claim is that such a thing always converges with respect to $\bar{d}$. Here are the details.

Proof. STILL TO COME □

Remark. If (X, d) was already complete to begin with, then the completion $\overline{X}$ constructed as above turns out to be X itself—or rather, there is a natural identification between $\overline{X}$ and X in this case. For example, the completion of the real numbers with respect to the standard metric is $\mathbb{R}$ itself with the standard metric.

The above remark should not be surprising given our original goal: to construct the "smallest" complete metric space containing a given metric space—if the original space is complete to begin with, it already is the smallest complete space containing itself!

We finish with some other interesting and important examples of completions; if you ever take an advanced number theory course, expect to see these examples again:

Example. Fix a prime number p. Define a metric d_p on $\mathbb{Q}$ as follows. First, we note that given any rational number $r \in \mathbb{Q}$, we can write it uniquely as a fraction:

$$r = \frac{m}{n} \text{ with } m, n \in \mathbb{N},\ n > 0, \text{ and with } m, n \text{ having no common factors.}$$

We can factor each m and n into a product of prime numbers, and then pull out the largest powers of p that occur in the so-obtained factorizations to write r as uniquely in the form

$$r = p^a \frac{m'}{n'} \text{with } m', n' \text{ having no common factors and neither divisible by } p.$$

Given two rational numbers s, s', write $s - s'$ in this form and define

$$d_p(s, s') = \frac{1}{p^a} \text{ where } a \text{ is the exponent of } p \text{ occuring in the expression for } s - s' \text{ outlined above.}$$

This defines a function $d_p : \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}$, which it turns out is a metric for any prime p called the *p-adic* metric on $\mathbb{Q}$.

For instance, let us compute the distance between $\frac{3}{5}$ and $\frac{3}{4}$ with respect to the 2-adic metric. We start by writing the difference of these two as

$$\frac{3}{5} - \frac{3}{4} = -\frac{3}{20} = -(2^{-2})\frac{3}{5}.$$

The "p^a" term in this case is 2^{-2}, and so the 2-adic distance between $\frac{3}{5}$ and $\frac{3}{4}$ is

$$d_2\left(\frac{3}{5}, \frac{3}{4}\right) = \frac{1}{2^{-2}} = 2^2 = 4.$$

Similarly, you can check that the 3-adic distance between $\frac{3}{5}$ and $\frac{3}{4}$ is $\frac{1}{3}$, and the 5-adic distance between them is 5. There are no other primes occurring in the prime factorization of $\frac{3}{5} - \frac{3}{4} = -\frac{3}{20}$, so the p-adic distance between $\frac{3}{5}$ and $\frac{3}{4}$ for another other prime p is $p^0 = 1$.

The completion of $\mathbb{Q}$ with respect to the p-adic metric is denoted by $\mathbb{Q}_p$ and is called the *field of p-adic numbers*. It turns out that this really is a field in the sense of Math 113: you can define addition and multiplication operations on $\mathbb{Q}_p$ extending those on $\mathbb{Q}$, and with these operations $\mathbb{Q}_p$ will be a field. Just as real number can be more simply characterized in terms of their decimal expansions instead of as Cauchy sequences of rational numbers, elements in $\mathbb{Q}_p$ can also be characterized in terms of certain expansions: an element $x \in \mathbb{Q}_p$ can be written as

$$x = \sum_{k=-n}^{\infty} a_k p^k \text{ where } n \in \mathbb{N} \text{ and } a_k \in \mathbb{Z},$$

and so has a "decimal expansion base p" of the form

$$x = \ldots a_3 a_2 a_1 . a_{-1} a_{-2} \ldots a_{-n}$$

where we have an infinite number of digits to the left of the decimal and only finitely many after—in contrast to decimal expansions of real numbers.

We will not go further into this construction here, but one can google "p-adic numbers" to find tons of other references. The key point for us is that this is nothing but an instance of the general construction of the completion of a metric space we have outlined in these notes.