It is natural to ask under what circumstances is a subspace of a complete space complete. The following theorem provides the answer.

Theorem. *Let (X, d) be a complete metric space, and Y a closed subset of X. If d_Y is the metric on Y induced by d, then the metric space (Y, d_Y) is complete. Moreover, if Z is a subset of X that is not closed then the space (Z, d_Z) is not complete.*

Proof. Let (a_n) be a Cauchy sequence in Y. Then (a_n) is also a Cauchy sequence in X. Since X is complete there exists an $a \in X$ such that $a_n \to a$ as $n \to \infty$. But $a_n \in Y$ for all n, and Y is closed; so, by a proposition proved in Lecture 8, $a = \lim_{n\to\infty} a_n \in Y$. So (a_n) converges in Y, and since (a_n) was an arbitrary Cauchy sequence in Y, this shows that Y is complete.

Given that Z is not closed, there exists a point $a \in X$ such that $a \in \overline{Z}$ but $a \notin Z$. The proposition from Lecture 8 also tells us that there is a sequence (a_n) in Z such that $a_n \to a$ as $n \to \infty$. It is trivial that all convergent sequences are Cauchy; so (a_n) is Cauchy, and it does not have a limit in Z since its unique limit in X is a, which is not an element of Z. $\square$

Completions of incomplete spaces

We have seen that $\mathbb{Q}$, with its usual metric, is not complete. As another example of this consider the sequence s whose n-th term s_n consists of the first n figures of the decimal expansion of π. Thus, $s = (3, 3.1, 3.14, 3.141, 3.1415, 3.14159, \ldots)$. This is a sequence of rational numbers which in $\mathbb{R}$ converges to π. Since $\pi \notin \mathbb{Q}$, this shows that the subset $\mathbb{Q}$ of $\mathbb{R}$ is not closed. It also shows that $\mathbb{Q}$ is not complete. To see this, observe first that s is a Cauchy sequence, since for any $\varepsilon > 0$ there exists an N such that $|s_n - \pi| < \varepsilon/2$ for all $n > N$, and so if $n, m > N$ then $|s_n - s_m| \leq |s_n - \pi| + |s_m - \pi| < \varepsilon$. Now since there is no $q \in \mathbb{Q}$ such that $s_n \to q$ as $n \to \infty$ (since s cannot have two limits in $\mathbb{R}$) it follows that s is a Cauchy sequence in $\mathbb{Q}$ which does not have a limit in $\mathbb{Q}$.

The idea that real numbers are limits of sequences of rational numbers is quite familiar to us all: it is the very basis of the decimal notation for real numbers. To say that $\pi = 3.14159\ldots$ is surely to say exactly that π is the limit of the sequence s above! The same idea can be used generally to construct a complete metric space from an incomplete one. To do so, one should adjoin extra elements that, by definition, are limits of Cauchy sequences. There is one complicating feature: it may be the case that two different Cauchy sequences should have the same limit, and so we must not simply create one new element for each Cauchy sequence that does not have a limit in the original space. Rather, the elements we create must correspond to classes of Cauchy sequences that are equivalent, in some sense.† To be precise, for this purpose we should consider Cauchy sequences (a_n) and (b_n) to be equivalent if $\lim_{n\to\infty} d(a_n, b_n) = 0$. The new metric space will have one element for each equivalence class of Cauchy sequences in the original space.

† In the construction of the reals from the rationals this complication is almost averted by choosing a rather special class of Cauchy sequences—those sequences $s = (s_1, s_2, s_3, \ldots)$ such that $s_1 \in \mathbb{Z}$ and $s_{i+1} = s_i + a_i 10^{-i}$ with $a_i \in \mathbb{Z}$ and $0 \leq a_i \leq 9$ (for all $i \geq 1$)—and showing subsequently that every other Cauchy sequence in $\mathbb{Q}$ is equivalent to one of these. But the fact that different sequences can have the same limit is not totally bypassed by this device, since, for example, the decimal expansions 2.8000... and 2.7999... represent the same real number.

The theorem that is derived from these considerations is as follows.

Theorem. *For every metric space (X, d) there is a metric space $(\widehat{X}, \hat{d})$ such that*
(1) $(\widehat{X}, \hat{d})$ is complete,
(2) (X, d) is a subspace of $(\widehat{X}, \hat{d})$, and
(3) X is dense in $\widehat{X}$.

The space $(\widehat{X}, \hat{d})$ is called the *completion* of the space (X, d). We shall defer for a while the task of proving the theorem.

Note that item (2) of the statement says that $X \subseteq \widehat{X}$ and d coincides with the metric on X induced by the metric $\hat{d}$ on $\widehat{X}$. (That is, d is the restriction of $\hat{d}$.) In fact, as always with such constructions, the space $\widehat{X}$ that we construct does not, strictly speaking, have X as a subset. Rather, it has a subset X' which is in one to one correspondence with X in an obvious way, and if this one to one correspondence is written as $x \leftrightarrow x'$, then $d(x, y) = \hat{d}(x', y')$ for all x, $y \in X$. After this, most authors say something like "we can use the correspondence $x \leftrightarrow x'$ to identify X with X', and this completes the proof". If you regard this as cheating then you can construct a space which genuinely does have the original X as a subset, by choosing any set S which is disjoint from X and in one to one correspondence with $\widehat{X} \setminus X'$, so that there is a one to one correspondence between $X \cup S$ and $\widehat{X}$ that extends the correspondence $x \leftrightarrow x'$ between X and X', and using this to transfer the metric on $\widehat{X}$ to a metric on $X \cup S$. Authors seldom do this, since, frankly, it is not worth the effort.

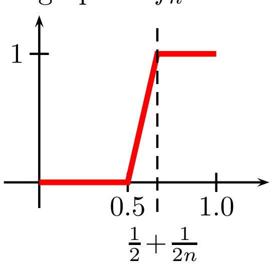
Note that item (3) of the theorem statement says that the closure of X in $\widehat{X}$ is the whole of $\widehat{X}$. In view of our characterization of closures (from Lecture 7), this says that every element of $\widehat{X}$ is the limit of a sequence in X (and thus a limit of a Cauchy sequence in X, since convergent sequences are necessarily Cauchy sequences).

Examples of complete and incomplete spaces

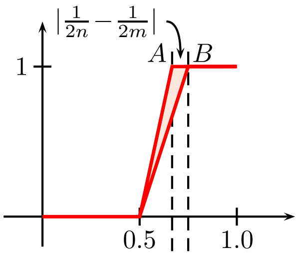
Let us look at some further examples of complete and incomplete spaces, starting with an incomplete one. Let $\mathcal{C}$ be the set of all continuous real-valued functions on $[0, 1]$, with metric d given by $d(f, g) = \int_0^1 |f(x) - g(x)|\, dx$. (Recall, from Lecture 3, that this is known as the L^1 metric on $\mathcal{C}$. It can be interpreted as saying that the distance between f and g is the area between their graphs.) For all $n \in \mathbb{Z}^+$ define

$$f_n(x) = \begin{cases} 0 & \text{if } 0 \le x \le \frac{1}{2}, \\ 2n(x - \frac{1}{2}) & \text{if } \frac{1}{2} < x \le \frac{1}{2} + \frac{1}{2n}, \\ 1 & \text{if } \frac{1}{2} + \frac{1}{2n} < x \le 1. \end{cases}$$

It is easy to check that $f_n \in \mathcal{C}$. The graph of f_n looks something like this:



If n, m are any positive integers then, as the diagram below illustrates, $d(f_n, f_m)$ can be described as the area of a certain triangle. The base AB of the triangle has length $|\frac{1}{2n} - \frac{1}{2m}|$, and the height is 1; so the area is $\frac{1}{2}|\frac{1}{2n} - \frac{1}{2m}|$.



Let $\varepsilon > 0$, and put $N = 1/\varepsilon$. For $n, m > N$ we have $0 < \frac{1}{2n} < \frac{1}{2N} = \varepsilon/2$, and similarly $0 < \frac{1}{2m} < \varepsilon/2$. Thus $\frac{1}{2n}, \frac{1}{2m} \in (0, \varepsilon/2)$, and so $|\frac{1}{2n} - \frac{1}{2m}| < \varepsilon/2$. It follows that

$$d(f_n, f_m) = \tfrac{1}{2}|\tfrac{1}{2n} - \tfrac{1}{2m}| < \varepsilon/4 < \varepsilon$$

whenever $n, m > N$. Thus $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence in $(\mathcal{C}, d)$.

We shall prove that $(\mathcal{C}, d)$ is not complete by showing that (f_n) does not have a limit in $\mathcal{C}$. We shall make use of the following fact:

if a function f is continuous on $[a,b]$ and $f(x) \geq 0$ for all $x \in [a,b]$, then $\int_a^b f(x)\,dx = 0$ only if $f(x) = 0$ for all $x \in [a,b]$.

(Note that this property is used in the proof that d is a metric: one needs it to show that $d(f,g) = 0$ implies $f = g$. See Question 4 of Tutorial 2.) A second fact that we shall employ is that if $a_n \to a$ as $n \to \infty$ then $\lim\limits_{n\to\infty} \int_{a_n}^b f(x)\,dx = \int_a^b f(x)\,dx$, provided that f is continuous on an interval which contains a, b and all the a_n.

Returning now to the investigation of the squence of functions f_n defined above, suppose, for a contradiction, that $f_n \to f$ as $n \to \infty$, where $f \in \mathcal{C}$. That is, f is continuous on $[0,1]$, and $\lim_{n\to\infty} d(f_n, f) = 0$. By the definition of the metric d, this says that $\int_0^1 |f_n(x) - f(x)|\,dx \to 0$ as $n \to \infty$. Now for all $n \in \mathbb{Z}^+$,

$$\begin{aligned} 0 \leq \int_0^{1/2} |f(x)|\,dx &= \int_0^{1/2} |f_n(x) - f(x)|\,dx \qquad (\text{as } f_n(x) = 0 \text{ for } x \in [0,1/2]) \\ &\leq \int_0^1 |f_n(x) - f(x)|\,dx \qquad (\text{since the integrand is nonnegative}) \\ &\longrightarrow 0 \quad \text{as } n \to \infty. \end{aligned}$$

This forces the constant $\int_0^{1/2} |f(x)|\,dx$ to be zero. Hence, by the first of the two facts whose use we foreshadowed, $f(x) = 0$ for all $x \in [0, 1/2]$.

Similarly, for all $n \in \mathbb{Z}^+$,

$$\begin{aligned} 0 \leq \int_{\frac{1}{2}+\frac{1}{2n}}^1 |1 - f(x)|\,dx &= \int_{\frac{1}{2}+\frac{1}{2n}}^1 |f_n(x) - f(x)|\,dx \qquad (\text{as } f_n(x) = 1 \text{ for } x \in [\tfrac{1}{2} + \tfrac{1}{2n}, 1]) \\ &\leq \int_0^1 |f_n(x) - f(x)|\,dx \qquad (\text{since the integrand is nonnegative}) \\ &\longrightarrow 0 \quad \text{as } n \to \infty, \end{aligned}$$

and thus it follows that $\int_{\frac{1}{2}}^1 |1 - f(x)|\,dx = \lim\limits_{n\to\infty} \int_{\frac{1}{2}+\frac{1}{2n}}^1 |1 - f(x)|\,dx = 0$. So $1 - f(x) = 0$ for all $x \in [0, \frac{1}{2}]$. But now we have shown that the function f satisfies $f(x) = 0$ for $0 \leq x \leq \frac{1}{2}$ and $f(x) = 1$ for $\frac{1}{2} \leq x \leq 1$. This is a contradiction, since $f(1/2)$ cannot simultaneously be both 0 and 1. So there is no such function $f \in \mathcal{C}$, and so $\mathcal{C}$ is not complete for the L^1 metric d.

It turns out that, on the other hand, if we define D to be the sup metric, then $(\mathcal{C}, D)$ is complete. Before proving this let us consider the set $\mathcal{B}$ of all bounded functions on $[0,1]$. Recall that $(\mathcal{B}, D)$ is a metric space and that D is defined by the formula

$$D(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|$$

for all f, $g \in \mathcal{B}$. (Note that we cannot use this formula to define a metric on the set of all functions on $[0,1]$, since the supremum will be undefined if $\{\, f(x) - g(x) \mid x \in [0,1] \,\}$ is unbounded.) We show that $(\mathcal{B}, D)$ is complete.

Let (f_n) be a Cauchy sequence in $\mathcal{B}$. As a first step we show that for each $t \in [0,1]$ the sequence $(f_n(t))$ is a Cauchy sequence in $\mathbb{R}$. To see this, let $t \in [0,1]$ and let ε be an arbitrary positive number. Since (f_n) is a Cauchy sequence in $\mathcal{B}$ there exists an integer N such that $D(f_n, f_m) < \varepsilon$ for all n, $m > N$. Now for all n, $m > N$ we have

$$|f_n(t) - f_m(t)| \le \sup_{x\in[0,1]} |f_n(x) - f_m(x)| = D(f_n, f_m) < \varepsilon,$$

and because N depends only on ε, which was arbitrary, this establishes the claim that the sequence $(f_n(t))$ is Cauchy.

Since $\mathbb{R}$ is complete, the Cauchy sequence $(f_n(t))$ has a limit in $\mathbb{R}$. This is applies for all $t \in [0,1]$. So we may define a function f on $[0,1]$ by setting $f(t)$ equal to $\lim_{n\to\infty} f_n(t)$. Thus f is the pointwise limit of the function sequence (f_n). We need to prove that (f_n) converges to f relative to the metric D; that is, we must show that (f_n) converges uniformly to f on $[0,1]$. We must also show that $f \in \mathcal{B}$.

Choose an integer M such that $D(f_n, f_m) \le 1$ for all n, $m \ge M$. Since $f_N \in \mathcal{B}$ there exists $B \in \mathbb{R}$ such that $f_N(t) \le B$ for all $t \in [0,1]$. So if $t \in [0,1]$ and $n \ge N$ then

$$|f_n(t)| \le |f_n(t) - f_N(t)| + |f_N(t)| \le \sup_{x\in[0,1]} |f_n(x) - f_N(x)| + B = D(f_n, f_N) + B \le 1 + B,$$

and therefore $\lim_{n\to\infty} |f_n(t)| \le 1 + B$. That is, $|f(t)| \le 1 + B$, and this shows that $f \in \mathcal{B}$.

Let $\varepsilon > 0$ and choose N such that $D(f_n, f_m) < \varepsilon/2$ for all all n, $m > N$. Fix $n > N$, and let m vary. For all $t \in [0,1]$ we have

$$|f_n(t) - f(t)| = \lim_{m\to\infty} |f_n(t) - f_m(t)| \le \varepsilon/2,$$

since if the terms of a convergent sequence lie in the set $A = [0, \varepsilon/2)$, its limit must lie in the closure $\overline{A} = [0, \varepsilon/2]$. We deduce that

$$D(f_n, f) = \sup_{t\in[0,1]} |f_n(t) - f(t)| \le \varepsilon/2 < \varepsilon$$

for all $n > N$, and therefore $f_n \to f$ as $n \to \infty$ in the space $(\mathcal{B}, D)$. Since (f_n) was an arbitrary Cauchy sequence in $(\mathcal{B}, D)$, we have shown that $(\mathcal{B}, D)$ is complete.

To show that the space $\mathcal{C}$ of continuous functions on $[0,1]$ is complete with respect to the uniform metric D it suffices to show that $\mathcal{C}$ is a closed subset of $\mathcal{B}$, since we have proved the general result that a closed subset of a complete space is complete. We have also proved that the closure of a subset of a metric space consists of all points which are limits of sequences in the subset. So if we can show that the limit of a convergent sequnce whose terms are in $\mathcal{C}$ is also in $\mathcal{C}$ then it will follow that $\mathcal{C}$ is its own closure, hence closed, and hence complete. So we need to show that the limit of a uniformly convergent sequence of continuous functions is continuous. This standard result is Question 4 of Tutorial 4.