FUNCTIONAL ANALYSIS

CHRISTIAN REMLING

1. Metric and topological spaces

A *metric space* is a set on which we can measure distances. More precisely, we proceed as follows: let $X \neq \emptyset$ be a set, and let $d : X \times X \to [0, \infty)$ be a map.

Definition 1.1. (X, d) is called a *metric space* if d has the following properties, for arbitrary $x, y, z \in X$:

(1) $d(x, y) = 0 \iff x = y$
(2) $d(x, y) = d(y, x)$
(3) $d(x, y) \leq d(x, z) + d(z, y)$

Property 3 is called the *triangle inequality*. It says that a detour via z will not give a shortcut when going from x to y.

The notion of a metric space is very flexible and general, and there are many different examples. We now compile a preliminary list of metric spaces.

Example 1.1. If $X \neq \emptyset$ is an arbitrary set, then

$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

defines a metric on X.

Exercise 1.1. Check this.

This example does not look particularly interesting, but it does satisfy the requirements from Definition 1.1.

Example 1.2. $X = \mathbb{C}$ with the metric $d(x, y) = |x - y|$ is a metric space. X can also be an arbitrary non-empty subset of $\mathbb{C}$, for example $X = \mathbb{R}$.

In fact, this works in complete generality: If (X, d) is a metric space and $Y \subseteq X$, then Y with the same metric is a metric space also.

Example 1.3. Let $X = \mathbb{C}^n$ or $X = \mathbb{R}^n$. For each $p \geq 1$,

$$d_p(x, y) = \left(\sum_{j=1}^{n} |x_j - y_j|^p \right)^{1/p}$$

defines a metric on X. Properties 1, 2 are clear from the definition, but if $p > 1$, then the verification of the triangle inequality is not completely straightforward here. We leave the matter at that for the time being, but will return to this example later.

An additional metric on X is given by

$$d_\infty(x, y) = \max_{j=1,\ldots,n} |x_j - y_j|$$

Exercise 1.2. (a) Show that (X, d_∞) is a metric space.
(b) Show that $\lim_{p\to\infty} d_p(x, y) = d_\infty(x, y)$ for fixed $x, y \in X$.

Example 1.4. Similar metrics can be introduced on function spaces. For example, we can take

$$X = C[a, b] = \{f : [a, b] \to \mathbb{C} : f \text{ continuous } \}$$

and define, for $1 \le p < \infty$,

$$d_p(f, g) = \left(\int_a^b |f(x) - g(x)|^p \, dx \right)^{1/p}$$

and

$$d_\infty(f, g) = \max_{a\le x\le b} |f(x) - g(x)|.$$

Again, the proof of the triangle inequality requires some care if $1 < p < \infty$. We will discuss this later.

Exercise 1.3. Prove that (X, d_∞) is a metric space.

Actually, we will see later that it is often advantageous to use the spaces

$$X_p = L^p(a, b) = \{f : [a, b] \to \mathbb{C} : f \text{ measurable}, \int_a^b |f(x)|^p \, dx < \infty\}$$

instead of X if we want to work with these metrics. We will discuss these issues in much greater detail in Section 2.

On a metric space, we can define convergence in a natural way. We just interpret "$d(x, y)$ small" as "x close to y", and this naturally leads to the following definition.

Definition 1.2. Let (X, d) be a metric space, and $x_n, x \in X$. We say that x_n *converges to* x (in symbols: $x_n \to x$ or $\lim x_n = x$, as usual) if $d(x_n, x) \to 0$.

Similarly, we call x_n a *Cauchy sequence* if for every $\epsilon > 0$, there exists an $N = N(\epsilon) \in \mathbb{N}$ such that $d(x_m, x_n) < \epsilon$ for all $m, n \ge N$.

We can make some quick remarks on this. First of all, if a sequence x_n is convergent, then the limit is unique because if $x_n \to x$ and $x_n \to y$, then, by the triangle inequality,

$$d(x, y) \leq d(x, x_n) + d(x_n, y) \to 0,$$

so $d(x, y) = 0$ and thus $x = y$. Furthermore, a convergent sequence is a Cauchy sequence: If $x_n \to x$ and $\epsilon > 0$ is given, then we can find an $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon/2$ for $n \geq N$. But then we also have

$$d(x_m, x_n) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \qquad (m, n \geq N),$$

so x_n is a Cauchy sequence, as claimed.

The converse is wrong in general metric spaces. Consider for example $X = \mathbb{Q}$ with the metric $d(x, y) = |x - y|$ from Example 1.2. Pick a sequence $x_n \in \mathbb{Q}$ that converges in $\mathbb{R}$ (that is, in the traditional sense) to an irrational limit ($\sqrt{2}$, say). Then x_n is a Cauchy sequence in (X, d) because it is convergent in the bigger space $(\mathbb{R}, d)$, so, as just observed, x_n must be a Cauchy sequence in $(\mathbb{R}, d)$. But then x_n is also a Cauchy sequence in $(\mathbb{Q}, d)$ because this is actually the exact same condition (only the distances $d(x_m, x_n)$ matter, we don't need to know how big the total space is). However, x_n can not converge in $(\mathbb{Q}, d)$ because then it would have to converge to the same limit in the bigger space $(\mathbb{R}, d)$, but by construction, in this space, it converges to a limit that was not in $\mathbb{Q}$.

Please make sure you understand exactly how this example works. There's nothing mysterious about this divergent Cauchy sequence. The sequence really wants to converge, but, unfortunately, the putative limit fails to lie in the space.

Spaces where Cauchy sequences do always converge are so important that they deserve a special name.

Definition 1.3. Let X be a metric space. X is called *complete* if every Cauchy sequence converges.

The mechanism from the previous example is in fact the only possible reason why spaces can fail to be complete. Moreover, it is always possible to *complete* a given metric space by including the would-be limits of Cauchy sequences. If this is done as economically as possible, then the resulting larger space is unique, up to the appropriate notion of isomorphism. It is called the *completion* of X.

We will have no need to apply this construction, so I don't want to discuss the (somewhat technical) details here. In most cases, the completion is what you think it should be; for example, the completion of $(\mathbb{Q}, d)$ is $(\mathbb{R}, d)$.

Exercise 1.4. Show that $(C[-1,1], d_1)$ is not complete.

Suggestion: Consider the sequence

$$f_n(x) = \begin{cases} -1 & -1 \le x < -1/n \\ nx & -1/n \le x \le 1/n \\ 1 & 1/n < x \le 1 \end{cases} .$$

A more general concept is that of a topological space. By definition, a *topological space* X is a non-empty set together with a collection $\mathcal{T}$ of distinguished subsets of X (called open sets) with the following properties:

(1) $\emptyset, X \in \mathcal{T}$
(2) If $U_\alpha \in \mathcal{T}$, then also $\bigcup U_\alpha \in \mathcal{T}$.
(3) If $U_1, \ldots, U_N \in \mathcal{T}$, then $U_1 \cap \ldots \cap U_N \in \mathcal{T}$.

This structure allows us to introduce some notion of closeness also, but things are fuzzier than on a metric space. We can zoom in on points, but there is no notion of one point being closer to a given point than another point.

We call $V \subseteq X$ a *neighborhood* of $x \in X$ if $x \in V$ and $V \in \mathcal{T}$. (*Warning:* This is sometimes called an open neighborhood, and it is also possible to define a more general notion of not necessarily open neighborhoods. We will always work with open neighborhoods here.) We can then say that x_n converges to x if for every neighborhood V of x, there exists an $N \in \mathbb{N}$ such that $x_n \in V$ for all $n \ge N$. However, on general topological spaces, sequences are not particularly useful; for example, if $\mathcal{T} = \{\emptyset, X\}$, then (obviously, by just unwrapping the definitions) every sequence converges to every limit.

Here are some additional basic notions for topological spaces. Please make sure you're thoroughly familiar with these (the good news is that we won't need much beyond these definitions).

Definition 1.4. Let X be a topological space.
(a) $A \subseteq X$ is called *closed* if A^c is open.
(b) For an arbitrary subset $B \subseteq X$, the *closure* of $B \subseteq X$ is defined as

$$\overline{B} = \bigcap_{A \supseteq B; A \text{ closed}} A;$$

this is the smallest closed set that contains B (in particular, there always is such a set).
(c) The *interior* of $B \subseteq X$ is the biggest open subset of B (such a set exists). Equivalently, the complement of the interior is the closure of the complement.
(d) $K \subseteq X$ is called *compact* if every open cover of K contains a finite

subcover.
(e) $\mathcal{B} \subseteq \mathcal{T}$ is called a *neighborhood base* of X if for every neighborhood V of some $x \in X$, there exists a $B \in \mathcal{B}$ with $x \in B \subseteq V$.
(f) Let $Y \subseteq X$ be an arbitrary, non-empty subset of X. Then Y becomes a topogical space with the *induced* (or *relative*) topology

$$\mathcal{T}_Y = \{U \cap Y : U \in \mathcal{T}\}.$$

(g) Let $f : X \to Y$ be a map between topological spaces. Then f is called *continuous* at $x \in X$ if for every neighborhood W of $f(x)$ there exists a neighborhood V of x such that $f(V) \subseteq W$. f is called continuous if it is continuous at every point.
(h) A topological space X is called a *Hausdorff space* if for every pair of points $x, y \in X$, $x \neq y$, there exist disjoint neighborhoods V_x, V_y of x and y, respectively.

Continuity on the whole space could have been (and usually is) defined differently:

Proposition 1.5. *f is continuous (at every point $x \in X$) if and only if $f^{-1}(V)$ is open (in X) whenever V is open (in Y).*

Exercise 1.5. Do some reading in your favorite (point set) topology book to brush up your topology. (Folland, Real Analysis, Ch. 4 is also a good place to do this.)

Exercise 1.6. Prove Proposition 1.5.

Metric spaces can be equipped with a natural topology. More precisely, this topology is natural because it gives the same notion of convergence of sequences. To do this, we introduce balls

$$B_r(x) = \{y \in X : d(y, x) < r\},$$

and use these as a neighborhood base for the topology we're after. So, by definition, $U \in \mathcal{T}$ if for every $x \in U$, there exists an $\epsilon > 0$ so that $B_\epsilon(x) \subseteq U$. Notice also that on $\mathbb{R}$ or $\mathbb{C}$ with the absolute value metric (see Example 1.2), this gives just the usual topology; in fact, the general definition mimics this procedure.

Theorem 1.6. *Let X be a metric space, and let $\mathcal{T}$ be as above. Then $\mathcal{T}$ is a topology on X, and $(X, \mathcal{T})$ is a Hausdorff space. Moreover, $B_r(x)$ is open and*

$$x_n \xrightarrow{d} x \quad \iff \quad x_n \xrightarrow{\mathcal{T}} x.$$

Proof. Let's first check that $\mathcal{T}$ is a topology on X. Clearly, $\emptyset, X \in \mathcal{T}$. If $U_\alpha \in \mathcal{T}$ and $x \in \bigcup U_\alpha$, then $x \in U_{\alpha_0}$ for some index α_0, and since

U_{α_0} is open, there exists a ball $B_r(x) \subseteq U_{\alpha_0}$, but then $B_r(x)$ is also contained in $\bigcup U_\alpha$.

Similarly, if $U_1, \ldots, U_N$ are open sets and $x \in \bigcap U_j$, then $x \in U_j$ for all j, so we can find N balls $B_{r_j}(x) \subseteq U_j$. It follows that $B_r(x) \subseteq \bigcap U_j$, with $r := \min r_j$.

Next, we prove that $B_r(x) \in \mathcal{T}$ for arbitrary $r > 0$, $x \in X$. Let $y \in B_r(x)$. We want to find a ball about y that is contained in the original ball. Since $y \in B_r(x)$, we have $\epsilon := r - d(x, y) > 0$, and I now claim that $B_\epsilon(y) \subseteq B_r(x)$. Indeed, if $z \in B_\epsilon(y)$, then, by the triangle inequality,

$$d(z, x) \leq d(z, y) + d(y, x) < \epsilon + d(y, x) = r,$$

so $z \in B_r(x)$, as desired.

The Hausdorff property also follows from this, because if $x \neq y$, then $r := d(x, y) > 0$, and $B_{r/2}(x)$, $B_{r/2}(y)$ are disjoint neighborhoods of x and y, respectively.

Exercise 1.7. It seems intuitively obvious that $B_{r/2}(x)$, $B_{r/2}(y)$ are disjoint. Please prove it formally.

Finally, we discuss convergent sequences. If $x_n \xrightarrow{d} x$ and V is a neighborhood of x, then, by the way $\mathcal{T}$ was defined, there exists $\epsilon > 0$ such that $B_\epsilon(x) \subseteq V$. We can then find an $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for $n \geq N$, or, equivalently, $x_n \in B_\epsilon(x)$ for $n \geq N$. So $x_n \in V$ for large enough n. This verifies that $x_n \xrightarrow{\mathcal{T}} x$.

Conversely, if this is assumed and $\epsilon > 0$ is given, then we can just take $V = B_\epsilon(x)$ as our neighborhood of x, and we then know that $x_n \in V$ or, equivalently, $d(x_n, x) < \epsilon$ for all large n. This says that $x_n \xrightarrow{d} x$. □

In metrizable topological spaces (that is, topological spaces where the topology comes from a metric, in this way) we can always work with sequences. This is a big advantage over general topological spaces.

Theorem 1.7. *Let (X, d) be a metric space, and introduce a topology $\mathcal{T}$ on X as above. Then:*
(a) $A \subseteq X$ is closed $\iff$ If $x_n \in A$, $x \in X$, $x_n \to x$, then $x \in A$.
(b) Let $B \subseteq X$. Then

$$\overline{B} = \{x \in X : \text{There exists a sequence } x_n \in B, x_n \to x\}.$$

(c) $K \subseteq X$ is compact precisely if every sequence $x_n \in K$ has a subsequence that is convergent in K.

These statements are false in general topological spaces (where the topology does not come from a metric).

Proof. We will only prove part (a) here. If A is closed and $x \notin A$, then, since A^c is open, there exists a ball $B_r(x)$ that does not intersect A. This shows that if $x_n \in A$, $x_n \to x$, then we also must have $x \in A$.

Conversely, if the condition on sequences from A holds and $x \notin A$, then there must be an $r > 0$ such that $B_r(x) \cap A = \emptyset$ (if not, pick an x_n from $B_{1/n}(x) \cap A$ for each n; this gives a sequence $x_n \in A$, $x_n \to x$, but $x \notin A$, contradicting our assumption). This verifies that A^c is open and thus A is closed.

Exercise 1.8. Prove Theorem 1.7 (b), (c).

□

Similarly, sequences can be used to characterize continuity of maps between metric spaces. Again, this doesn't work on general topological spaces.

Theorem 1.8. *Let (X, d), (Y, e) be metric spaces, let $f : X \to Y$ be a function, and let $x \in X$. Then the following are equivalent:*
(a) f is continuous at x (with respect to the topologies induced by d, e).
(b) For every $\epsilon > 0$, there exists a $\delta > 0$ such that $e(f(x), f(t)) < \epsilon$ for all $t \in X$ with $d(x, t) < \delta$.
(c) If $x_n \to x$ in X, then $f(x_n) \to f(x)$ in Y.

Proof. If (a) holds and $\epsilon > 0$ is given, then, since $B_\epsilon(f(x))$ is a neighborhood of $f(x)$, there exists a neighborhood U of x such that $f(U) \subseteq B_\epsilon(f(x))$. From the way the topology on a metric space is defined, we see that U must contain a ball $B_\delta(x)$, and (b) follows.

If (b) is assumed and $\epsilon > 0$ is given, pick $\delta > 0$ according to (b) and then $N \in \mathbb{N}$ such that $d(x_n, x) < \delta$ for $n \geq N$. But then we also have $e(f(x), f(x_n)) < \epsilon$ for all $n \geq N$, that is, we have verified that $f(x_n) \to f(x)$.

Finally, if (c) holds, we argue by contradiction to obtain (a). So assume that, contrary to our claim, we can find a neighborhood V of $f(x)$ such that for every neighborhood U of x, there exists $t \in U$ with $f(t) \notin V$. In particular, we can then pick an $x_n \in B_{1/n}(x)$ for each n, such that $f(x_n) \notin V$. Since V is a neighborhood of $f(x)$, there exists $\epsilon > 0$ such that $B_\epsilon(f(x)) \subseteq V$. Summarizing, we have found a sequence $x_n \to x$, but $e(f(x_n), f(x)) \geq \epsilon$; in particular, $f(x_n) \not\to f(x)$. This contradicts (c), and so we have to admit that (a) holds. □

The following fact is often useful:

Proposition 1.9. *Let (X, d) be a metric space and $Y \subseteq X$. As above, write $\mathcal{T}$ for the topology generated by d on X.*

Then (Y, d) is a metric space, too (this is obvious and was also already observed above). Moreover, the topology generated by d on Y is the relative topology of Y as a subspace of $(X, \mathcal{T})$.

Exercise 1.9. Prove this (this is done by essentially chasing definitions, but it is a little awkward to write down).

We conclude this section by proving our first fundamental functional analytic theorem. We need one more topological definition: We call a set $M \subseteq X$ *nowhere dense* if $\overline{M}$ has empty interior. If X is a metric space, we can also say that $M \subseteq X$ is nowhere dense if $\overline{M}$ contains no (non-empty) open ball.

Theorem 1.10 (Baire)**.** *Let X be a complete metric space. If the sets $A_n \subseteq X$ are nowhere dense, then $\bigcup_{n\in\mathbb{N}} A_n \neq X$.*

Completeness is crucial here:

Exercise 1.10. Show that there are (necessarily: non-complete) metric spaces that are countable unions of nowhere dense sets.

Suggestion: $X = \mathbb{Q}$

Proof. The following proof is similar in spirit to Cantor's famous diagonal trick, which proves that $[0, 1]$ is uncountable. We will construct an element that is not in $\bigcup A_n$ by avoiding these sets step by step.

First of all, we may assume that the A_n's are closed (if not, replace A_n with $\overline{A_n}$; note that these sets are still nowhere dense).

Then, since A_1 is nowhere dense, we can find an $x_1 \in A_1^c$. In fact, A_1^c is also open, so we can even find an open ball $B_{r_1}(x_1) \subseteq A_1^c$, and here we may assume that $r_1 \leq 2^{-1}$ (decrease r_1 if necessary).

In the next step, we pick an $x_2 \in B_{r_1/2}(x_1) \setminus A_2$. There must be such a point because A_2 is nowhere dense and thus cannot contain the ball $B_{r_1/2}(x_1)$. Moreover, we can again find $r_2 > 0$ such that

$$B_{r_2}(x_2) \cap A_2 = \emptyset, \quad B_{r_2}(x_2) \subseteq B_{r_1/2}(x_1), \quad r_2 \leq 2^{-2}.$$

We continue in this way and construct a sequence $x_n \in X$ and radii $r_n > 0$ with the following properties:

$$B_{r_n}(x_n) \cap A_n = \emptyset, \quad B_{r_n}(x_n) \subseteq B_{r_{n-1}/2}(x_{n-1}), \quad r_n \leq 2^{-n}$$

These properties guarantee that x_n is a Cauchy sequence: indeed, if $m \geq n$, then x_m lies in $B_{r_n/2}(x_n)$, so

$$d(x_m, x_n) \leq \frac{r_n}{2}. \tag{1.1}$$

Since X is complete, $x := \lim x_n$ exists. Moreover,

$$d(x_n, x) \leq d(x_n, x_m) + d(x_m, x)$$

for arbitrary $m \in \mathbb{N}$. For $m \geq n$, (1.1) shows that $d(x_n, x_m) \leq r_n/2$, so if we let $m \to \infty$, it follows that

$$d(x_n, x) \leq \frac{r_n}{2}. \tag{1.2}$$

By construction, $B_{r_n}(x_n) \cap A_n = \emptyset$, so (1.2) says that $x \notin A_n$ for all n. □

Baire's Theorem can be (and often is) formulated differently. We need one more topological definition: We call a set $M \subseteq X$ *dense* if $\overline{M} = X$. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$. Similarly, $\mathbb{Q}^c$ is also dense in $\mathbb{R}$. However, note that (of course) $\mathbb{Q} \cap \mathbb{Q}^c = \emptyset$.

Theorem 1.11 (Baire). *Let X be a complete metric space. If U_n ($n \in \mathbb{N}$) are dense open sets, then $\bigcap_{n\in\mathbb{N}} U_n$ is dense.*

Exercise 1.11. Derive this from Theorem 1.10.

Suggestion: If U is dense and open, then $A = U^c$ is nowhere dense (prove this!). Now apply Theorem 1.10. This will not quite give the full claim, but you can also apply Theorem 1.10 on suitable subspaces of the original space.

An immediate consequence of this, in turn, is the following slightly stronger looking version. By definition, a *G_δ set* is a countable intersection of open sets.

Exercise 1.12. Give an example that shows that a G_δ set need not be open (but, conversely, open sets are of course G_δ sets).

Theorem 1.12 (Baire). *Let X be a complete metric space. Then a countable intersection of dense G_δ sets is a dense G_δ set.*

Exercise 1.13. Derive this from the previous theorem.

Given this result, it makes sense to interpret dense G_δ sets as big sets, in a topological sense, and their complements as small sets. Theorem 1.12 then says that even a countable union of small sets will still be small. Call a property of elements of a complete metric space *generic* if it holds at least on a dense G_δ set.

Theorem 1.12 has a number of humoristic applications, which say that certain unexpected properties are in fact generic. Here are two such examples:

Example 1.5. Let $X = C[a,b]$ with metric $d(f,g) = \max|f(x) - g(x)|$ (compare Example 1.4). This is a complete metric space (we'll prove this later). It can now be shown, using Theorem 1.12, that *the generic continuous function is nowhere differentiable.*

Example 1.6. *The generic coin flip refutes the law of large numbers.*

More precisely, we proceed as follows. Let $X = \{(x_n)_{n\ge 1} : x_n = 0 \text{ or } 1\}$ and $d(x,y) = \sum_{n=1}^{\infty} 2^{-n}|x_n - y_n|$. This is a metric and X with this metric is complete, but we don't want to prove this here. In fact, this metric is a natural choice here; it generates the product topology on X.

From probability theory, we know that if the x_n are independent random variables and the coin is fair, then, with probability 1, we have that $S_n/n \to 1/2$, where $S_n = x_1 + \ldots + x_n$ is the number of heads (say) in the first n coin tosses.

The generic behavior is quite different: For a generic sequence $x \in X$,

$$\liminf_{n\to\infty} \frac{S_n}{n} = 0, \quad \limsup_{n\to\infty} \frac{S_n}{n} = 1.$$

Since these examples are for entertainment only, we will not prove these claims here.

Baire's Theorem is fundamental in functional analysis, and it will have important consequences. We will discuss these in Chapter 3.

Exercise 1.14. Consider the space $X = C[0,1]$ with the metric $d(f,g) = \max_{0\le x\le 1}|f(x) - g(x)|$ (compare Example 1.4). Define $f_n \in X$ by

$$f_n(x) = \begin{cases} 2^n x & 0 \le x \le 2^{-n} \\ 1 & 2^{-n} < x \le 1 \end{cases}.$$

Work out $d(f_n, 0)$ and $d(f_m, f_n)$, and deduce from the results of this calculation that $S = \{f \in X : d(f,0) = 1\}$ is not compact.

Exercise 1.15. Let X, Y be topological spaces, and let $f : X \to Y$ be a continuous map. True or false (please give a proof or a counterexample):
(a) $U \subseteq X$ open $\Longrightarrow f(U)$ open
(b) $A \subseteq Y$ closed $\Longrightarrow f^{-1}(A)$ closed
(c) $K \subseteq X$ compact $\Longrightarrow f(K)$ compact
(d) $L \subseteq Y$ compact $\Longrightarrow f^{-1}(L)$ compact

Exercise 1.16. Let X be a metric space, and define, for $x \in X$ and $r > 0$,

$$\overline{B}_r(x) = \{y \in X : d(y,x) \le r\}.$$

(a) Show that $\overline{B}_r(x)$ is always closed.
(b) Show that $\overline{B_r(x)} \subseteq \overline{B}_r(x)$. (By definition, the first set is the closure of $B_r(x)$.)
(c) Show that it can happen that $\overline{B_r(x)} \neq \overline{B}_r(x)$.