MATHEMATICS 3103 (Functional Analysis)
YEAR 2012–2013, TERM 2

HANDOUT #2: COMPACTNESS OF METRIC SPACES

Compactness in metric spaces

The closed intervals $[a, b]$ of the real line, and more generally the closed bounded subsets of $\mathbb{R}^n$, have some remarkable properties, which I believe you have studied in your course in real analysis. For instance:

> **Bolzano–Weierstrass theorem.** Every bounded sequence of real numbers has a convergent subsequence.

This can be rephrased as:

> **Bolzano–Weierstrass theorem (rephrased).** Let X be any closed bounded subset of the real line. Then any sequence (x_n) of points in X has a subsequence converging to a point of X.

(*Why* is this rephrasing valid? Note that this property does *not* hold if X fails to be closed or fails to be bounded — why?) And here is another example:

> **Heine–Borel theorem.** Every covering of a closed interval $[a, b]$ — or more generally of a closed bounded set $X \subset \mathbb{R}$ — by a collection of open sets has a finite subcovering.

These theorems are not only interesting — they are also extremely useful in applications, as we shall see. So our goal now is to investigate the generalizations of these concepts to metric spaces.

We begin with some definitions: Let (X, d) be a metric space. A **covering** of X is a collection of sets whose union is X. An **open covering** of X is a collection of *open* sets whose union is X. The metric space X is said to be **compact** if every open covering has a finite subcovering.[1] This abstracts the Heine–Borel property; indeed, the Heine–Borel theorem states that closed bounded subsets of the real line are compact.

We can rephrase compactness in terms of closed sets by making the following observation: If $\mathcal{U}$ is an open covering of X, then the collection $\mathcal{F}$ of *complements* of sets in $\mathcal{U}$ is a collection of closed sets whose intersection is empty (why?); and conversely, if $\mathcal{F}$ is a collection of closed sets whose intersection is empty, then the collection $\mathcal{U}$ of complements of sets in $\mathcal{F}$ is an open covering. Thus, a space X is compact if and only if every collection of closed sets with an empty intersection has a finite subcollection whose intersection is also empty. Or, passing to the contrapositive, we can put it another way by making the following definition: a collection $\mathcal{F}$ of sets is said to have the **finite intersection property** if every finite subcollection of $\mathcal{F}$ has a nonempty intersection. We have then shown:

[1] Or more formally: If $(U_\alpha)_{\alpha \in I}$ (where I is some index set) is a collection of open sets of X satisfying $\bigcup\limits_{\alpha \in I} U_\alpha = X$, then there exists a finite subset $J \subseteq I$ such that $\bigcup\limits_{\alpha \in J} U_\alpha = X$.

Proposition 2.1 *A metric space X is compact if and only if every collection $\mathcal{F}$ of closed sets in X with the finite intersection property has a nonempty intersection.*

So far so good; but thus far we have merely made a trivial reformulation of the definition of compactness. Let us go farther by making another definition:

A metric space X is said to be **sequentially compact** if every sequence $(x_n)_{n=1}^{\infty}$ of points in X has a convergent subsequence. This abstracts the Bolzano–Weierstrass property; indeed, the Bolzano–Weierstrass theorem states that closed bounded subsets of the real line are sequentially compact.

And finally, let us make another definition: A metric space (X, d) is said to be **totally bounded** (or **precompact**) if, for every $\epsilon > 0$, the space X can be covered by a finite family of open balls of radius ϵ. (You could alternatively use closed balls and get the same concept — why?) Another way of saying this is: A metric space (X, d) is totally bounded if, for every $\epsilon > 0$, there exists a finite subset $A \subseteq X$ such that $d(x, A) < \epsilon$ for all $x \in X$. (Why is this equivalent?) Any such finite subset is called an **ϵ-net**.[2]

We then have the following fundamental theorem characterizing compact metric spaces:

Theorem 2.2 (Compactness of metric spaces) *For a metric space X, the following are equivalent:*

(a) X is compact, i.e. every open covering of X has a finite subcovering.

(b) Every collection of closed sets in X with the finite intersection property has a nonempty intersection.

(c) If $F_1 \supseteq F_2 \supseteq F_3 \supseteq \ldots$ is a decreasing sequence of nonempty closed sets in X, then $\bigcap\limits_{n=1}^{\infty} F_n$ is nonempty.

(d) X is sequentially compact, i.e. every sequence in X has a convergent subsequence.

(e) X is totally bounded and complete.

We have already proved the equivalence of (a) and (b). Let us now prove (b) $\Longrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (e) $\Longrightarrow$ (a).

Proof of (b) $\Longrightarrow$ (c). This is trivial, since a *decreasing* sequence of *nonempty* closed sets obviously has the finite intersection property. (Why? If $n_1, \ldots, n_k$ are given indices, what is $F_{n_1} \cap F_{n_2} \cap \ldots \cap F_{n_k}$?) $\square$

Proof of (c) $\Longrightarrow$ (d). Let (x_n) be a sequence of points in X, and let F_n be the closure of the set $\{x_n, x_{n+1}, x_{n+2}, \ldots\}$. The family of sets $\{F_n\}$ is decreasing (i.e. $F_1 \supseteq F_2 \supseteq F_3 \supseteq \ldots$), and all the sets F_n are nonempty and closed. Therefore, by (c), the set $\bigcap\limits_{n=1}^{\infty} F_n$ contains

[2] Some authors (e.g. Kolmogorov–Fomin) define an ϵ-net to be any finite subset $A \subseteq X$ such that $d(x, A) \leq \epsilon$ for all $x \in X$, i.e. with *non-strict* inequality. This changes slightly the statements of proofs but makes no essential difference — can you see why?

at least one point a. Then it is easy to see that (x_n) contains a subsequence converging to a: for instance, set $n_1 = 1$ and then let n_k be the smallest integer $> n_{k-1}$ such that $d(x_{n_k}, a) < 1/k$; such an integer exists because a belongs to all the sets F_n. (You should make sure you understand this last step.) $\square$

PROOF OF (D) $\implies$ (E). To prove that X is complete, let (x_n) be any Cauchy sequence in X. By (d), (x_n) contains a subsequence converging to some point $a \in X$. But then, by Proposition 1.27 (see last week's notes), the whole sequence (x_n) converges to a. This shows that X is complete.

Now suppose that X is not totally bounded, i.e. there exists a number $\alpha > 0$ such that X has no finite covering by open balls of radius α. Then we can define a sequence $(x_n)_{n=1}^{\infty}$ of points in X having $d(x_i, x_j) \geq \alpha$ for all $i \neq j$, by the following inductive construction: First let x_1 be any point in X. Then, supposing that $x_1, \ldots, x_{n-1}$ have been chosen, we know that the union of the open balls of center x_i $(1 \leq i \leq n-1)$ and radius α is not the whole space, hence we can choose a point x_n satisfying $d(x_i, x_n) \geq \alpha$ for all $i < n$. When we are done, we have $d(x_i, x_j) \geq \alpha$ for all $i \neq j$ (why?). On the other hand, the sequence (x_n) cannot have any convergent subsequence; for if it had a subsequence (x_{n_k}) converging to a, then there would exist an integer k_0 such that $d(x_{n_k}, a) < \alpha/2$ for all $k \geq k_0$, and hence by the triangle inequality $d(x_{n_k}, x_{n_{k'}}) < \alpha$ for all $k, k' \geq k_0$, contrary to the definition of the sequence (x_n).[3]

PROOF OF (E) $\implies$ (A). Suppose that X is not compact, i.e. we have an open covering $(U_\alpha)_{\alpha \in I}$ of X such that no finite subfamily is a covering of X. We will define a sequence $(x_n)_{n=1}^{\infty}$ of points in X, as follows: First choose an ϵ-net with $\epsilon = 1/2$ (this is possible because X is totally bounded), and let x_1 be any element of that ϵ-net with the property that no finite subfamily of $(U_\alpha)_{\alpha \in I}$ is a covering of $B(x_1, 1/2)$. [Such an element has to exist, because if every ball of radius 1/2 centered at a point of the ϵ-net had a finite subcover from $(U_\alpha)_{\alpha \in I}$, then the whole space X would have a finite subcover from $(U_\alpha)_{\alpha \in I}$ (why?).] Next choose an ϵ-net with $\epsilon = 1/4$, and let x_2 be any element of that ϵ-net satisfying $B(x_1, 1/2) \cap B(x_2, 1/4) \neq \varnothing$ and having the property that no finite subfamily of $(U_\alpha)_{\alpha \in I}$ is a covering of $B(x_2, 1/4)$. [Such an element has to exist, because if every ball of radius 1/4 centered at a point of the ϵ-net and having nonempty intersection with $B(x_1, 1/2)$ had a finite subcover from $(U_\alpha)_{\alpha \in I}$, then $B(x_1, 1/2)$ would have a finite subcover from $(U_\alpha)_{\alpha \in I}$ (why?).] Continue analogously: at the nth stage, choose an ϵ-net with $\epsilon = 1/2^n$, and let x_n be any element of that ϵ-net satisfying $B(x_{n-1}, 1/2^{n-1}) \cap B(x_n, 1/2^n) \neq \varnothing$ and having the property that no finite subfamily of $(U_\alpha)_{\alpha \in I}$ is a covering of $B(x_n, 1/2^n)$.

It follows from this construction that

$$d(x_{n-1}, x_n) \;\leq\; \frac{1}{2^{n-1}} + \frac{1}{2^n} \;\leq\; \frac{1}{2^{n-2}} \tag{2.1}$$

(why?) and hence that, for $m < n$,

$$d(x_m, x_n) \;\leq\; d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \ldots + d(x_{n-1}, x_n) \tag{2.2a}$$

[3]Another way of stating this argument is: The sequence (x_n) clearly cannot have any Cauchy subsequence; therefore, it cannot have any convergent subsequence.

$$\leq \frac{1}{2^{m-1}} + \frac{1}{2^m} + \ldots + \frac{1}{2^{n-2}} \tag{2.2b}$$

$$\leq \frac{1}{2^{m-2}}\,, \tag{2.2c}$$

which shows that (x_n) is a Cauchy sequence in X. Since X is complete, the sequence (x_n) converges to some point $a \in X$.

Now let $\alpha_0 \in I$ be an index such that $a \in U_{\alpha_0}$ (why must such an index exist?). There exists $\epsilon > 0$ such that $B(a, \epsilon) \subseteq U_{\alpha_0}$. By the definition of a, there exists an integer n such that $d(x_n, a) < \epsilon/2$ and also $1/2^n < \epsilon/2$ (why?). The triangle inequality then shows that

$$B(x_n, 1/2^n) \;\subseteq\; B(a, \epsilon) \;\subseteq\; U_{\alpha_0} \tag{2.3}$$

(why?). But this contradicts the fact that no finite subfamily of $(U_\alpha)_{\alpha \in I}$ is a covering of $B(x_n, 1/2^n)$. Whew! $\square$

Warning: For general (nonmetrizable) topological spaces, compactness is *not* equivalent to sequential compactness.

We also have the following easy fact:

Proposition 2.3 *Every totally bounded metric space (and in particular every compact metric space) is separable.*

Proof. If X is totally bounded, then there exists for each n a finite subset $A_n \subseteq X$ such that, for every $x \in X$, $d(x, A_n) < 1/n$. Now let $A = \bigcup\limits_{n=1}^{\infty} A_n$. The set A is either finite or countably infinite (why?); and for each $x \in X$ we have $d(x, A) \leq d(x, A_n) < 1/n$, hence $d(x, A) = 0$, hence $x \in \overline{A}$ (why?). This proves that A is dense in X. $\square$

Intuitively, a separable space is one that is "well approximated by a countable subset", while a compact space is one that is "well approximated by a finite subset". (Albeit in a slightly different sense of "well approximated" in the two cases.)

A subset A of a metric space X is said to be **compact** if A, considered as a subspace of X and hence a metric space in its own right, is compact. We have the following easy facts, whose proof I leave to you:

Proposition 2.4

- *(a) A closed subset of a compact space is compact.*
- *(b) A compact subset of any metric space is closed.*
- *(c) A* finite *union of compact sets is compact.*

Of course, an *infinite* union of compact sets need not even be closed (give an example!); and even when it is closed, it need not be compact (give another example!).

The Heine–Borel (or Bolzano–Weierstrass) theorem of elementary real analysis can be restated as follows:

Proposition 2.5 (Compactness of subsets in $\mathbb{R}$) *A subset $A \subseteq \mathbb{R}$ is compact if and only if it is closed and bounded.*

The corresponding result for $\mathbb{R}^n$ is an easy consequence:

Proposition 2.6 (Compactness of subsets in $\mathbb{R}^n$) *A subset $A \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

PROOF. Every compact subset of $\mathbb{R}^n$ is obviously closed and bounded (why?), so we need only prove the converse. Moreover, every bounded subset of $\mathbb{R}^n$ is contained in a cube $[-M, M]^n$ for some $M < \infty$, so by Proposition 2.4(a) we need only prove that $[-M, M]^n$ is compact. But this follows from the Heine–Borel (or Bolzano–Weierstrass) theorem for $\mathbb{R}$ combined with the fact that any finite Cartesian product of compact spaces is compact, a straightforward result that you will prove in Problem 2 of Problem Set #2. $\square$

I want to stress that the situation is very different in infinite-dimensional normed linear spaces. Later in this lecture we will show that the closed unit ball in the sequence spaces ℓ^∞, c_0, ℓ^1 and ℓ^2 is *not* compact, and we will give examples of compact sets in these spaces.

Continuous functions on compact metric spaces

Last week we saw that the *inverse* image of a closed set by a continuous function is closed, but that this is *not* in general true for *direct* images. However, for *some special* closed sets — namely, the compact ones — the direct image *is* closed and indeed is compact:

Proposition 2.7 (Continuous image of a compact space) *The direct image of a compact metric space by a continuous function is compact.*

This can be formulated precisely in several slightly different, but equivalent, ways:

1) Let X and Y be metric spaces, with X compact, and let $f\colon X \to Y$ be a continuous map that is *surjective* (i.e. the image $f[X]$ equals all of Y). Then Y is compact.

2) Let X and Y be metric spaces, with X compact, and let $f\colon X \to Y$ be a continuous map. Then $f[X]$ is a compact subset of Y.

3) Let X and Y be metric spaces, and let $f\colon X \to Y$ be a continuous map. If A is a compact subset of X, then $f[A]$ is a compact subset of Y.

You should make sure you understand why these three formulations are equivalent.

In order to illustrate some common techniques of proof, I will give *two* different proofs of this proposition: one exploiting open coverings, and one exploiting convergent subsequences. In both cases I will prove the first of the three formulations.

First Proof of Proposition 2.7. Let $(U_\alpha)_{\alpha\in I}$ be an open covering of Y. Then the sets $f^{-1}[U_\alpha]$ are open (why?) and form an open covering of X (why?). Since X is compact, there exists a finite subset $J \subseteq I$ such that $(f^{-1}[U_\alpha])_{\alpha\in J}$ still forms a covering of X. But then $(U_\alpha)_{\alpha\in J}$ forms a covering of Y (why?). $\square$

Second Proof of Proposition 2.7. Consider a sequence (y_n) of elements of Y. Because f is surjective, we can choose a sequence (x_n) of points in X such that $f(x_n) = y_n$. Since X is compact, there exists a subsequence (x_{n_k}) that converges to some point $a \in X$. But since f is continuous at a, the sequence (y_{n_k}) converges to $f(a)$ (why?). This proves that Y is sequentially compact, hence compact. $\square$

It follows easily from Proposition 2.7 that a continuous real-valued function on a compact metric space is automatically bounded, and furthermore that the maximum and minimum values are attained:

Corollary 2.8 *Let X be a compact metric space, and let $f\colon X \to \mathbb{R}$ be continuous. Then $f[X]$ is bounded, and there exist points $a, b \in X$ such that $f(a) = \inf\limits_{x\in X} f(x)$ and $f(b) = \sup\limits_{x\in X} f(x)$.*

Proof. By Proposition 2.7, $f[X]$ is a compact subset of $\mathbb{R}$, hence closed and bounded. Now, any bounded set $A \subseteq \mathbb{R}$ has a least upper bound $\sup A$ and a greatest lower bound $\inf A$, and these two points belong to the closure $\overline{A}$ (why?). But applying this to $A = f[X]$, which is closed, we conclude that $\sup f[X]$ and $\inf f[X]$ belong to $f[X]$ itself, which is exactly what is being claimed (why?). $\square$

Note that this result can fail if X is noncompact, for instance if $X = \mathbb{R}$: a continuous real-valued function on $\mathbb{R}$ need not be bounded; and even if it is bounded, its supremum and infimum need not be attained. You should give examples to illustrate both these points.

A few weeks from now we will prove, in fact, that a metric space is compact *if and only if* every continuous real-valued function on it is bounded.

Here is another property of continuous functions on compact metric spaces, whose proof likewise illustrates some ways of exploiting compactness but is slightly trickier than that of Proposition 2.7:

We begin by recalling that if (X, d_X) and (Y, d_Y) are two metric spaces, then a mapping $f\colon X \to Y$ is continuous at the point $x \in X$ if, for each $\epsilon > 0$, there exists $\delta > 0$ (depending of course on ϵ) such that, for all $x' \in X$, $d_X(x, x') < \delta$ implies $d_Y(f(x), f(x')) < \epsilon$. In particular, a mapping $f\colon X \to Y$ is continuous *tout court* if it is continuous at every point $x \in X$, i.e. if, for each $\epsilon > 0$ and each $x \in X$, there exists $\delta > 0$ (depending on ϵ and x) such that, for all $x' \in X$, $d_X(x, x') < \delta$ implies $d_Y(f(x), f(x')) < \epsilon$. Note that here δ can depend on x as well as on ϵ.

We now make a new definition: A mapping $f\colon X \to Y$ is **uniformly continuous** if, for each $\epsilon > 0$, there exists $\delta > 0$ (depending of course on ϵ) such that, for all $x, x' \in X$,

$d_X(x, x') < \delta$ implies $d_Y(f(x), f(x')) < \epsilon$. The point is that, in uniform continuity, δ can still depend on ϵ (in general it has to) but is *not* allowed to depend on x.

So what we have here is an interchange of quantifiers: written formally, using the notations $\forall$ ("for all") and $\exists$ ("there exists"), we have:

$$\textbf{Continuity:} \qquad (\forall\epsilon > 0)\,(\forall x \in X)\,(\exists\delta > 0)\,(\forall x' \in X)$$
$$d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \epsilon$$

$$\textbf{Uniform continuity:} \qquad (\forall\epsilon > 0)\,(\exists\delta > 0)\,(\forall x \in X)\,(\forall x' \in X)$$
$$d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \epsilon$$

To see that uniform continuity is truly a *stronger* property than continuity, consider the map $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. It is continuous but not uniformly continuous. *You* should supply the details of the proof that f is not uniformly continuous.

The existence on $\mathbb{R}$ of a function that is continuous but not uniformly continuous is directly linked to the fact that $\mathbb{R}$ is noncompact. In particular, we have:

Proposition 2.9 *Let f be a continuous mapping of a* compact *metric space X into a metric space Y. Then f is uniformly continuous.*

To help you get familiar with different techniques of proof, I shall once again give *two* different proofs of this proposition: one exploiting open coverings, and one exploiting convergent subsequences.

First Proof of Proposition 2.9. Given $\epsilon > 0$ and $x \in X$, there exists a $\delta_x > 0$ such that $d_X(x, x') < \delta_x$ implies $d_Y(f(x), f(x')) < \epsilon/2$. [You will see later why we took $\epsilon/2$ rather than ϵ here.] Now let $U_x = B(x, \frac{1}{2}\delta_x)$, the open ball of center x and radius $\frac{1}{2}\delta_x$. [You will see later why it was clever to choose the radius to be *half* of δ_x.] The collection $\{U_x\}_{x\in X}$ is an open covering of X, so it has a finite subcovering $\{U_{x_1}, \ldots, U_{x_n}\}$. Let $\delta = \frac{1}{2}\min(\delta_{x_1}, \ldots, \delta_{x_n})$. Clearly $\delta > 0$. Now, given two points $y, z \in X$ such that $d_X(y, z) < \delta$, the point y must belong to some U_{x_i} (why?) and hence $d_X(y, x_i) < \frac{1}{2}\delta_{x_i}$. But then

$$d_X(z, x_i) \;\le\; d_X(z, y) + d_X(y, x_i) \;<\; \delta + \tfrac{1}{2}\delta_{x_i} \;\le\; \delta_{x_i}\,. \tag{2.4}$$

So both y and z lie at a distance $< \delta_{x_i}$ from the point x_i, which implies (by definition of δ_{x_i}) that $d_Y(f(y), f(x_i)) < \epsilon/2$ and $d_Y(f(z), f(x_i)) < \epsilon/2$. Hence $d_Y(f(y), f(z)) < \epsilon$, which shows that f is uniformly continuous. $\square$

Second Proof of Proposition 2.9. Suppose that f is not uniformly continuous; then there exists a number $\epsilon > 0$ and two sequences (x_n) and (y_n) of points of X such that $d_X(x_n, y_n) < 1/n$ and $d_Y(f(x_n), f(y_n)) \ge \epsilon$. By compactness we can find a subsequence (x_{n_k}) converging to a point a; and since $d_X(x_{n_k}, y_{n_k}) < 1/n_k$, it follows from the triangle inequality that the sequence (y_{n_k}) also converges to a. But since f is continuous at a, there exists $\delta > 0$ such that $d_Y(f(x), f(a)) < \epsilon/2$ whenever $d_X(x, a) < \delta$. Now take k such that $d_X(x_{n_k}, a) < \delta$ and $d_X(y_{n_k}, a) < \delta$ (why can this be done?); it follows that $d_Y(f(x_{n_k}), f(y_{n_k})) < \epsilon$ (why?), contrary to the definition of the sequences (x_n) and (y_n). $\square$

Which one of these two proofs do you consider simpler? This is probably a question of taste: my suspicion is that most students will prefer the proof using subsequences, since subsequences are a more familiar concept than open coverings. And of course I urge you to use, in any given problem, the proof that *you* find simplest (provided only that it is *correct*, of course!). But I nevertheless also urge you to study *both* of these proofs carefully and to become familiar with the use of *both* subsequences and open coverings, because there exist applications in analysis in which either one or the other may be more convenient.

Remark. It is natural to ask whether the converse to this theorem is true: that is, if X is a metric space such that every continuous real-valued function on X is uniformly continuous, is X necessarily compact? The answer is no: for instance, if X is any discrete metric space, then every real-valued function on X is automatically both continuous and uniformly continuous (why?); but a discrete metric space is compact if and only if it is finite (why?).

So the obvious next question is: Can one characterize the metric spaces in which every continuous real-valued function is uniformly continuous (in the sense of finding other, equivalent conditions)? Some answers to this question can be found in

M. Atsuji, Uniform continuity of continuous functions of metric spaces, *Pacific J. Math.* **8**, 11–16 (1958); erratum, 941.

H. Hueber, On uniform continuity and compactness in metric spaces, *Amer. Math. Monthly* **88**, 204–205 (1981).

G. Beer, Metric spaces on which continuous functions are uniformly continuous and Hausdorff distance, *Proc. Amer. Math. Soc.* **95**, 653–658 (1985).

T. Jain and S. Kundu, Atsuji completions: equivalent characterisations, *Topology Appl.* **154**, 28–38 (2007).

It is amusing that questions arising from elementary functional analysis can still be the subject of ongoing research.

Compactness in infinite-dimensional spaces

Since the concept of compactness plays a central role in functional analysis (and indeed in all areas of analysis[4]), it is important for us to obtain some intuition about when sets are or are not compact. In a finite-dimensional normed linear space (i.e. $\mathbb{R}^n$ or $\mathbb{C}^n$) we know the answer: a set is compact if and only if it is closed and bounded. But this result is completely false in infinite-dimensional spaces: indeed, we will prove about 2 weeks from now that the closed unit ball in a normed linear space (which is certainly closed and bounded) is compact *if and only if* the space is finite-dimensional! Otherwise put, the closed unit ball in an infinite-dimensional normed linear space is *never* compact.

[4]For instance, if you have studied Complex Analysis you may recall **Montel's theorem**, which states that if (f_n) is a uniformly bounded sequence of functions that are analytic in a domain D in the complex plane, then there exists a subsequence (f_{n_k}) that converges (uniformly on compact subsets of D) to an analytic function g. This important theorem is precisely a statement about compactness in a certain infinite-dimensional space of analytic functions (which is, however, a non-normable space).

For now, let us limit ourselves to proving the noncompactness of the closed unit ball in some of our favorite infinite-dimensional spaces.

Example 1: ℓ^∞**.** The closed unit ball in ℓ^∞ is not compact — indeed, it is not even separable (see Proposition 1.18 last week along with Proposition 2.3).[5]

We see from this example that if there is to be any hope for the closed unit ball to be compact, the underlying space had better be separable. Now, we know that the spaces c_0, ℓ^1 and ℓ^2 *are* separable, unlike ℓ^∞ (see Proposition 1.19 last week and Problem 7 of Problem Set #1), so maybe things will work out better there. Alas, this is not the case:

Example 2: c_0**.** The closed unit ball in c_0 is not compact. To see this, let e_i be the infinite sequence that has a 1 in the ith coordinate and 0 everywhere else. Clearly $e_i \in c_0$ and $\|e_i\|_\infty = 1$, so the points e_i all belong to the closed unit ball of c_0. But $\|e_i - e_j\|_\infty = 1$ for $i \neq j$, so the sequence $(e_i)_{i=1}^\infty$ cannot possibly have any convergent subsequence.

The same proof works in ℓ^1 and ℓ^2, since the vectors e_i also belong to the closed unit balls of those spaces. The computation changes slightly — for $i \neq j$ we have $\|e_i - e_j\|_1 = 2$ and $\|e_i - e_j\|_2 = \sqrt{2}$ (why?) — but the conclusion is the same.

What about spaces $\mathcal{C}(A)$ of bounded continuous real-valued functions? Well, in Problem 5 of Problem Set #1 you showed that the space $\mathcal{C}(\mathbb{R})$ of bounded continuous real-valued functions on the whole real line $\mathbb{R}$ is not separable (neither is its closed unit ball), so its closed unit ball is certainly not compact. On the other hand, we will see a few weeks from now that the spaces $\mathcal{C}[a, b]$ of (bounded) continuous real-valued functions on a *closed bounded* interval of the real line — and more generally the spaces $\mathcal{C}(X)$ of (bounded) continuous real-valued functions on an arbitrary *compact* metric space X — *are* separable, so maybe there is hope for their closed unit balls to be compact. Alas, this too fails:

Example 3: $\mathcal{C}[0, 1]$**.** First consider the function $f \in \mathcal{C}(\mathbb{R})$ defined by

$$f(x) \;=\; \max(1 - |x|, 0) \;. \tag{2.5}$$

It is a triangle-shaped bump centered at 0 and of "half-width" 1, satisfying $0 \le f \le 1$. Now define, for each integer $n \ge 1$, the function

$$f_n(x) \;=\; f\Big(2n(n+1)(x - 1/n)\Big) \;. \tag{2.6}$$

It is a triangle-shaped bump centered at the point $1/n$ and of "half-width" $1/[2n(n+1)]$, satisfying $0 \le f_n \le 1$; in particular, its restriction to the interval $[0, 1]$ belongs to the space $\mathcal{C}[0, 1]$ and satisfies $\|f_n\|_\infty = 1$. Taking the half-widths to be $1/[2n(n+1)]$ is sufficient to guarantee that the supports of the different functions f_n do not overlap (why?), so we have

[5] As stated, Proposition 1.18 claims only that the whole space ℓ^∞ is not separable. But if you look at the proof, you will see that the points e_I all belong to the closed unit ball of ℓ^∞, and the whole proof could have been done within the closed unit ball of ℓ^∞ rather than within ℓ^∞ [in particular, the balls $B(e_I, \frac{1}{2})$ could be taken to be the balls *within* this smaller space]. Indeed, it is easy to show that a normed linear space is separable if and only if its closed (or open) unit ball is.

$\|f_n - f_{n'}\|_\infty = 1$ for $n \neq n'$. It follows from this that the sequence $(f_n)_{n=1}^\infty$ cannot have any convergent subsequence, so the closed unit ball in $\mathcal{C}[0,1]$ is noncompact.

We will use a slight generalization of this reasoning, 2 weeks from now, to prove that the closed unit ball in an infinite-dimensional normed linear space is *always* noncompact.

So, what kinds of sets in an infinite-dimensional normed linear space *can* be compact? Of course they have to be closed and bounded, but this is not enough: roughly speaking, compact sets in infinite-dimensional spaces have to be "very tightly confined" in most of the (infinitely many) coordinate directions — unlike closed balls, which spread equally in all directions. That is, they must be very *compact* in the sense that this word is given by the Oxford English dictionary:

> **Compact**, *adj.*
>
> A.1.b. Packed closely together.
>
> B.1. Closely packed or knit together.
>
> B.1.b. Having the parts so arranged that the whole lies within relatively small compass, without straggling portions or members; nearly and tightly packed or arranged; not sprawling, scattered, or diffuse.

To understand what this means in practice, let us give some examples of compact sets in the sequence spaces ℓ^∞, c_0, ℓ^1 and ℓ^2.

If $x = (x_1, x_2, \ldots)$ is an infinite sequence of real numbers, let us define S_x to be the set consisting of those infinite sequences of real numbers that are "bounded above elementwise by x", i.e.

$$S_x \;=\; \{y \in \mathbb{R}^{\mathbb{N}} \colon |y_n| \le |x_n| \text{ for all } n\} \;. \tag{2.7}$$

Note that if x belongs to one of the spaces ℓ^∞, c_0, ℓ^1 or ℓ^2, then the set S_x is a subset of the same space (why?), and it is fairly easy to see that it is closed and bounded (can you prove this?).

If x is the sequence $(1, 1, 1, \ldots)$ whose entries are all 1 — this sequence of course belongs to ℓ^∞ — then S_x is precisely the closed unit ball of ℓ^∞, which is noncompact. But what if x belongs to the smaller space c_0, i.e. it is not merely bounded but is in fact *convergent to zero*? We have the following result:

Proposition 2.10 *If $x \in c_0$, then S_x is a* compact *subset of c_0 (and hence also of ℓ^∞).*

Proof. Since S_x is a closed subset of c_0, and c_0 is complete, it follows that S_x is complete. Hence we need only prove that S_x is totally bounded.

So consider any $\epsilon > 0$. Because $\lim\limits_{n\to\infty} x_n = 0$, there exists an integer N (depending of course on ϵ) such that $|x_n| \le \epsilon/2$ for all $n > N$. Now consider the subset $S_x^{(N)} \subseteq S_x$ defined by

$$S_x^{(N)} \;=\; \{y \in S_x \colon y_n = 0 \text{ for all } n > N\} \;. \tag{2.8}$$

The set $S_x^{(N)}$ is isometric to a closed bounded subset of the space $\mathbb{R}^N$ (why?) and is thus compact. In particular it is totally bounded, so we can choose a finite $\epsilon/2$-net $A \subseteq S_x^{(N)}$. But A is also an ϵ-net for S_x (why?). $\square$

In Problem 3 of Problem Set #2 you will prove the analogues of this result for ℓ^1 and ℓ^2. And in Problem 4 you will prove two strong converses to Proposition 2.10:

(a) If $x \in \ell^\infty \setminus c_0$, then S_x is not even separable as a subset of ℓ^∞, much less compact.

(b) Every compact subset of c_0 is contained in some set S_x with $x \in c_0$.

Locally compact spaces

Roughly speaking, a space can fail to be compact for either of two reasons: it can be "too big at infinity", or it can be "too big locally".

Example 1: Discrete metric spaces. It is easy to see that a discrete metric space is compact if and only if it is finite (why?). So any infinite discrete metric space is necessarily noncompact. The reason is that the space is "too big at infinity": any sequence that "runs off to infinity" cannot have a convergent subsequence.

Note that this applies not only to discrete metric spaces in the narrow sense of Example 1 of Handout #1 (i.e. sets equipped with the 0–1 metric), but also to metric spaces that are topologically equivalent to such a space, i.e. those in which every set is open (and hence also closed). This includes, for instance, the natural numbers $\mathbb{N}$ and the integers $\mathbb{Z}$ equipped with the usual metric $d(x, y) = |x - y|$ that they inherit as subspaces of $\mathbb{R}$.

Of course, any space that *contains* an infinite discrete metric space as a *closed* subspace is also necessarily noncompact (why?). This includes, for instance, $\mathbb{R}$.

Example 2: The closed unit ball in ℓ^∞, c_0, ℓ^1 or ℓ^2. We have seen that the closed unit ball in ℓ^∞, c_0, ℓ^1 or ℓ^2 is noncompact. (Actually this is true for the closed unit ball in *any* infinite-dimensional normed linear space, but we have not yet proven this.) Here the problem is that the space is "too big locally": there are "too many directions" in which one can go away from a given point.

It is of interest to distinguish between these two ways that a space can fail to be compact, and in particular to define a class of metric spaces that "look locally like compact spaces" even if globally they fail to be compact. So let us say that a metric space X is **locally compact** if, for every point $x \in X$, there exists a compact neighborhood of x in X. For instance, any discrete space is locally compact (why?); the spaces $\mathbb{R}$ and $\mathbb{R}^n$ are also locally compact (why?). On the other hand, it follows from our results ℓ^∞, c_0, ℓ^1 and ℓ^2 are *not* locally compact (why?).

I will not develop here the basic facts about locally compact metric spaces, but will instead refer you to Dieudonné, Section III.18. Here, without proof, are the main useful facts:

Lemma 2.11 *Let A be a compact set in a locally compact metric space X. Then there exists an $r > 0$ such that the r-neighborhood*

$$V_r(A) \;=\; \{x \in X \colon d(x, A) < r\} \tag{2.9}$$

(which is indeed an open neighborhood of A, see Proposition 1.13 of Handout #1) is relatively compact (i.e. has compact closure) in X.

And using this, we can characterize the locally compact metric spaces which, though they are not "essentially finite at infinity" (i.e. compact), are nevertheless "essentially countable at infinity":

Theorem 2.12 *Let X be a locally compact metric space. Then the following are equivalent:*

(a) X is separable.

(b) X is **$\boldsymbol{\sigma}$-compact***, i.e. X can be written as a countable union of compact subsets.*

(c) X can be written as the union of an increasing sequence (U_n) of open relatively compact subsets (i.e. open subsets whose closures are compact) satisfying $\overline{U_n} \subset U_{n+1}$ for all n.