METRIC SPACES AND COMPLEX ANALYSIS.

KEVIN MCGERTY.

1. Introduction

In Prelims you studied Analysis, the rigorous theory of calculus for (real-valued) functions of a single real variable. This term we will largely focus on the study of functions of a complex variable, but we will begin by seeing how much of the theory developed last year can in fact can be made to work, with relatively little extra effort, in a significantly more general context.

Recall the trajectory of the Prelims Analysis course – initially it focused on sequences and developed the notion of the limit of a sequence which was crucial for essentially everything which followed[1]. Then it moved to the study of continuity and differentiability, and finally it developed a theory of integration. This term's course will follow approximately the same pattern, but the contexts we work in will vary a bit more. To begin with we will focus on limits and continuity, and attempt to gain a better understanding of what is needed in order for make sense of these notions.

Example 1.1. Consider for example one of the key definitions of Prelims analysis, that of the *continuity* of a function. Recall that if $f\colon \mathbb{R} \to \mathbb{R}$ is a function, we say that f is continuous at $a \in \mathbb{R}$ if, for any $\epsilon > 0$, we can find a $\delta > 0$ such that if $|x - a| < \delta$ then $|f(x) - f(a)| < \epsilon$. Stated somewhat more informally, this means that no matter how small an ϵ we are given, we can ensure $f(x)$ is within distance ϵ of $f(a)$ provided we demand x is sufficiently close to – that is, within distance δ of – the point a.

Now consider what it is about real numbers that we need in order for this definition to make sense: Really we just need, for any pair of real numbers x_1 and x_2, a measure of the distance between them. (Note that we needed this notion of distance in the above definition of continuity for both the pairs (x, a) and $(f(x), f(a))$.) Thus we should be able to talk about continuous functions $f\colon X \to X$ on any set X provided it is equipped with a notion of distance. Even more generally, provided we have proscribed a notion of distance on two sets X and Y, we should be able to say what it means for a function $f\colon X \to Y$ to be continuous. In order to make this precise, we will therefore need to give a mathematically rigorous definition of what a "notion of distance" on a set should be.

As a first step, consider as an example the case of $\mathbb{R}^n$. The dot product on vectors in $\mathbb{R}^n$ gives us a notion of distance between vectors in $\mathbb{R}^n$: Recall that if $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n)$ are vectors in $\mathbb{R}^n$ then we set

$$\langle v, w\rangle = \sum_{i=1}^{n} v_i w_i,$$

and we define the length of a vector to be[2] $\|v\| = \langle v, v\rangle^{1/2}$. Recall that the Cauchy-Schwarz inequality then says that $|\langle v, w\rangle| \leq \|v\|\|w\|$. It has the following important consequence for the length function:

Date: November 5, 2018.

[1] Although continuity is introduced via ϵs and δs, the notion can be expressed in terms of convergent sequences. Similarly one can define the integral in terms of convergent sequences.

[2] Sometimes the notation $\|v\|_2$ is used for this length function – we will see later there are other natural choices for the length of a vector in $\mathbb{R}^n$.

Lemma 1.2. *If $x, y \in \mathbb{R}^n$ then $\|x+y\| \leq \|x\| + \|y\|$.*

Proof. Since $\|v\| \geq 0$ for all $v \in \mathbb{R}^n$ the desired inequality is equivalent to

$$\|x+y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2.$$

But since $\|x+y\|^2 = \langle x+y, x+y \rangle = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$, this inequality is immediate from the Cauchy-Schwarz inequality. □

Once we have a notion of length for vectors, we also immediately have a way of defining the distance between them – we simply take the length of the vector $v - w$. Explicitly, this is:

$$\|v - w\| = \big(\sum_{i=1}^{n} (v_i - w_i)^2\big)^{1/2}.$$

Now that we have defined the distance between any two vectors in $\mathbb{R}^n$, we can immediately make sense both of what it means for a function $f\colon \mathbb{R}^n \to \mathbb{R}$ to be continuous[3] as above and also what it means for a sequence to converge.

Definition 1.3. If $(v^k)_{k\in\mathbb{N}}$ is a sequence of vectors in $\mathbb{R}^n$ (so $v^k = (v_1^k, \dots, v_n^k)$) we say $(v^k)_{k\in\mathbb{N}}$ *converges* to $w \in \mathbb{R}^n$ if for any $\epsilon > 0$ there is an $N > 0$ such that for all $k \geq N$ we have $\|v^k - w\| < \epsilon$.

If $f\colon \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$ then we say that f is *continuous at a* if for any $\epsilon > 0$ there is a $\delta > 0$ such that $|f(a) - f(x)| < \epsilon$ whenever $\|x - a\| < \delta$. (As usual, we say that f is continuous on $\mathbb{R}^n$ if it is continuous at every $a \in \mathbb{R}^n$.)

Many of the results about convergence for sequences of real or complex numbers which were established last year readily extend to sequences in $\mathbb{R}^n$, with almost identical proofs. As an example, just as for sequences of real or complex numbers, we have the following:

Lemma 1.4. *Suppose that $(v^k)_{k\geq 1}$ is a sequence in $\mathbb{R}^n$ which converges to $w \in \mathbb{R}^n$ and also to $u \in \mathbb{R}^n$. Then $w = u$, that is, limits are unique.*

Proof. We prove this by contradiction: suppose $w \neq u$. Then $d = \|w - u\| > 0$, so since (v^k) converges to w we can find an $N_1 \in \mathbb{N}$ such that for all $k \geq N$ we have $\|w - v^k\| < d/2$. Similarly, since (v^k) converges to u we can find an N_2 such that for all $k \geq N_2$ we have $\|v^k - u\| < d/2$. But then if $k \geq \max\{N_1, N_2\}$ we have

$$d = \|w - u\| = \|(w - v^k) + (v^k - u)\| \leq \|w - v^k\| + \|v^k - u\| < d/2 + d/2 = d,$$

where in the first inequality we use Lemma 1.2. Thus we have a contradiction as required. □

2. Metric Spaces

We now come to the definition of a metric space. To motivate it, let's consider what a notion of distance on a set X should mean: Given any two points in X, we should have a non-negative real number – the distance between them. Thus a distance on a set X should therefore be a function $d\colon X \times X \to \mathbb{R}_{\geq 0}$, but we must also decide what properties of such a function capture our intuition of distance. A couple of properties suggest themselves immediately – the distance between two points $x, y \in X$ should be symmetric, that is, the distance from x to y should[4] be the same as the distance from y to x, and the distance between two points should be 0 precisely when they are equal. Note that this latter property was one of the crucial ingredients in the proof of the

[3]More ambitiously, using the notions of distance we have for $\mathbb{R}^n$ and $\mathbb{R}^m$ you can readily make sense of the notion of continuity for a function $g\colon \mathbb{R}^n \to \mathbb{R}^m$.

[4]In fact it's possible to think of contexts where this assumption doesn't hold – think of swimming in a river – going upstream is harder work than going downstream, so if your notion of distance took this into account it would fail to be symmetric.

uniqueness of limits as we just saw. The last requirement we make of a distance function is known as the "triangle inequality", a version of which we established in Lemma 1.2 and which was also essential in the above uniqueness proof. These requirements yield in the following definition:

Definition 2.1. Let X be a set and suppose that $d\colon X \times X \to \mathbb{R}_{\geq 0}$. Then we say that d is a *distance function* on X if it has the following properties: For all $x, y, z \in X$:

1. (*Positivity*): $d(x, y) = 0$ if and only if $x = y$.
2. (*Symmetry*): $d(x, y) = d(y, x)$.
3. (*Triangle inequality*): If $x, y, z \in X$ then we have

$$d(x, z) \leq d(x, y) + d(y, z).$$

Note that for the normal distance function in the plane $\mathbb{R}^2$, the third property expresses the fact that the length of a side of a triangle is at most the sum of the lengths of the other two sides (hence the name!). We will write a metric space as a pair (X, d) of a set and a distance function $d\colon X \times X \to \mathbb{R}_{\geq 0}$ satisfying the axioms above. If the distance function is clear from context, we may, for convenience, simply write X rather than (X, d).

Example 2.2. The vector space $\mathbb{R}^n$ equipped with the distance function $d_2(v, w) = \|v - w\| = \langle v - w, v - w\rangle^{1/2}$ is a metric space: The first two properties of the metric d_2 are immediate from the definition, while the triangle inequality follows from Lemma 1.2.

Remark 2.3. In Prelims Linear Algebra, you met the notion of an inner product space $(V, \langle -, - \rangle)$ (over the real or complex numbers). For any two vectors $v, w \in V$ setting $d(v, w) = \|v - w\|$, where $\|v\| = \langle v, v\rangle^{1/2}$, gives V a notion of distance. Since the Cauchy-Schwarz inequality holds in any inner product space, Lemma 1.2 holds in any inner product space (the proof is word for word the same), it follows that d is also a metric in this more general setting.

Definition 2.4. If (X, d_X) is a metric space and $A \subseteq X$ then we set

$$\operatorname{diam}(A) = \sup\{d(a_1, a_2) : a_1, a_2 \in X\} \in \mathbb{R}_{\geq 0} \cup \{\infty\},$$

(where we take $\operatorname{diam}(A) = \infty$ if the $\{d(a_1, a_2) : a_1, a_2 \in A\}$ is not bounded above. If $\operatorname{diam}(A)$ is finite then we say that A is a *bounded* subset of X.

To make good our earlier assertion, we now define the notions of continuity and convergence in a metric space.

Definition 2.5. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f\colon X \to Y$ is said to be continuous at $a \in X$ if for any $\epsilon > 0$ there is a $\delta > 0$ such that for any $x \in X$ with $d_X(a, x) < \delta$ we have $d_Y(f(x), f(a)) < \epsilon$. We say f is *continuous* if it is continuous at every $a \in X$.

If $(x_n)_{n\geq 1}$ is a sequence in X, and $a \in X$, then we say $(x_n)_{n\geq 1}$ *converges to* a if, for any $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d_X(x_n, a) < \epsilon$.

In fact it is clear that the notion of uniform continuity also extends to functions between metric spaces: A function $f\colon X \to Y$ is said to be *uniformly continuous* if, for any $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x_1, x_2 \in X$ with $d_X(x_1, x_2) < \delta$ we have $d_Y(f(x_1), f(x_2)) < \epsilon$.

For later use, we also note that a function $f\colon X \to Y$ is said to be *bounded* if its image $f(X)$ is a bounded subset of Y in the sense of Definition 2.4, that is, if

$$\{d_Y(f(x), f(y)) : x, y \in X\} \subseteq \mathbb{R}$$

is a bounded subset of $\mathbb{R}$. Note that, unlike continuity or uniform continuity, the condition that a function is bounded only requires that Y has a metric (X need not).

The next result is the natural generalization of the theorem you saw last year which showed that continuity could be expressed in terms of convergent sequences. You should note that the proof is, mutatis mutandi, the same as the case for function from the real line to itself.

Lemma 2.6. *Let $f: X \to Y$ be a function. Then f is continuous at $a \in X$ if and only if for every sequence $(x_n)_{n\geq 0}$ converging to a we have $f(x_n) \to f(a)$ as $n \to \infty$.*

Proof. Suppose that f is continuous at a. Then given $\epsilon > 0$ there is a $\delta > 0$ such that for all $x \in X$ with $d(x, a) < \delta$ we have $d(f(x), f(a)) < \epsilon$. Now if $(x_n)_{n\geq 0}$ is a sequence tending to a then there is an $N > 0$ such that $d(a, x_k) < \delta$ for all $k \geq N$. But then for all $k \geq N$ we see that $d(f(a), f(x_k)) < \epsilon$, so that $f(x_k) \to f(a)$ as $n \to \infty$ as required.

For the converse, we use the contrapositive, hence we suppose that f is not continuous at a. Then there is an $\epsilon > 0$ such that for all $\delta > 0$ there is some $x \in X$ with $d(x, a) < \delta$ and $d(f(x), f(a)) \geq \epsilon$. Chose for each $k \in \mathbb{Z}_{>0}$ a point $x_n \in X$ with $d(x_k, a) < 1/n$ but $d(f(x_k), f(a)) \geq \epsilon$. Then $d(x_k, a) < 1/k \to 0$ as $k \to \infty$ so that $x_k \to a$ as $k \to \infty$, but since $d(f(x_k), f(a))) \geq \epsilon$ for all k clearly $(f(x_k))_{k\geq 0}$ does not tend to $f(a)$. □

Definition 2.7. If X is a metric space we write $\mathcal{C}(X) = \{f: X \to \mathbb{R} : f \text{ is continuous}\}$ for the set of continuous real-valued functions on X. (Here the real line is viewed as a metric space equipped with the metric coming from the absolute value).

Lemma 2.8. *The set $\mathcal{C}(X)$ is a vector space. Moreover if $f, g \in \mathcal{C}(X)$ then so is $f.g$.*

Proof. This is just algebra of limits: $\mathcal{C}(X)$ is a subset of the vector space of all real-valued functions on X, so we just need to check it is closed under addition and multiplication. To see that $\mathcal{C}(X)$ is closed under multiplication, suppose that $f, g \in \mathcal{C}(X)$ and $a \in X$. To see that $f.g$ is continuous at a, note that if $\epsilon > 0$ is given, then since both f and g are continuous at a, we may find a δ_1 such that $|f(x) - f(a)| < \min\{1, \epsilon/2(|g(a)| + 1)\}$ for all $x \in X$ with $d(x, a) < \delta_1$ and a $\delta_2 > 0$ such that $|g(x) - g(a)| < \epsilon/2(|f(a)| + 1)$ for all $x \in X$ with $d(x, a) < \delta_2$. Setting $\delta = \min\{\delta_1, \delta_2\}$ it follows that for all $x \in X$ with $d(x, a) < \delta$ we have

$$\begin{aligned} |f(x)g(x) - f(a)g(a)| &= |f(x)g(x) - f(x)g(a) + f(x)g(a) - f(a)g(a)| \\ &\leq |f(x)||g(x) - g(a)| + |f(x) - f(a)||g(a)| \\ &\leq (|f(a)| + 1)|g(x) - g(a)| + |f(x) - f(a)||g(a)| \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

where in the third line we use the fact that $|f(x)| < |f(a)| + 1$ for all $x \in X$ such that $d(x, a) < \delta_1$. Since a was arbitrary, this shows that $f.g$ lies in $\mathcal{C}(X)$. Since constant functions are clearly continuous this shows in particular that $\mathcal{C}(X)$ is closed under multiplication by scalars. We leave it as an exercise to check that $\mathcal{C}(X)$ is closed under addition and hence is a vector space. □

Remark 2.9. One can also check that if $f: X \to \mathbb{R}$ is continuous at a and $f(a) \neq 0$ then $1/f$ is continuous at a. Again this is proved just as in the single-variable case.

Example 2.10. Consider the case of $\mathbb{R}^n$ again. The distance function d_2 coming from the dot product makes $\mathbb{R}^n$ into a metric space, as we have already seen. On the other hand it is not the only reasonable notion of distance one can take. We can define for $v, w \in \mathbb{R}^n$

$$\begin{aligned} d_1(v, w) &= \sum_{i=1}^{n} |v_i - w_i|; \\ d_2(v, w) &= \big(\sum_{i=1}^{n} (v_i - w_i)^2\big)^{1/2} \\ d_\infty(v, w) &= \max_{i \in \{1,2,\ldots,n\}} |v_i - w_i|. \end{aligned}$$

Each of these functions clearly satisfies the positivity and symmetry properties of a metric. We have already checked the triangle inequality for d_2, while for d_1 and d_∞ it follows readily from the triangle inequality for $\mathbb{R}$.

Example 2.11. Suppose that (X, d) is a metric space and let Y be a subset of X. Then the restriction of d to $Y \times Y$ gives Y a metric so that $(Y, d_{|Y\times Y})$ is a metric space. We call Y equipped with this metric a *subspace*[5] of X.

Example 2.12. The *discrete* metric on a set X is defined as follows:

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

The axioms for a distance function are easy to check.

Example 2.13. A slightly more interesting example is the *Hamming distance* on words: if A is a finite set which we think of as an "alphabet", then a word of length n in just an element of A^n, that is, a sequence of n elements in the alphabet. The Hamming distance between two such words $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$ is

$$d_H(\mathbf{a}, \mathbf{b}) = |\{i \in \{1, 2, \ldots, n\} : a_i \neq b_i\}.$$

Problem sheet 1 asks you to check that d is indeed a distance function (where the only axiom which requires some thought is the triangle inequality).

An important special case of this is the space of binary sequences of length n, that is, where the alphabet A is just $\{0, 1\}$. In this case one can view set of words of length n in this alphabet as a subset of $\mathbb{R}^n$, and moreover you can check that the Hamming distance function is the same as the subspace metric induced by the d_1 metric on $\mathbb{R}^n$ given above.

Example 2.14. If (X, d) is a metric space, then we can consider the space $X^{\mathbb{N}}$ of all sequences in X. That is, the elements of $X^{\mathbb{N}}$ are sequences $(x_n)_{n\geq 1}$ in X. While there is no obvious metric on the whole space of sequences, if we take $X_b^{\mathbb{N}}$ to be the space of *bounded* sequences, that is, sequences such that the set $\{d_\infty(x_n, x_m) : n, m \in \mathbb{N}\} \subset \mathbb{R}$ is bounded, then the function[6]

$$d_\infty((x_n)_{n\geq 1}, (y_n)_{n\geq 1}) = \sup_{n\in\mathbb{N}} d(x_n, y_n),$$

is a metric on $X_b^{\mathbb{N}}$. It clearly satisfies positivity and symmetry, and the triangle inequality follows from the inequality

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

by taking the supremum of the left-hand side over $n \in \mathbb{N}$.

Example 2.15. If (X, d_X) and (Y, d_Y) are metric spaces, then it is natural to try to make $X \times Y$ into a metric space. In fact this can be done in a number of ways – we will return to this issue later. One method is to set $d_{X\times Y} = \max\{d_X, d_Y\}$, that is if $x_1, x_2 \in X$ and $y_1, y_2 \in Y$ then we set

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) = \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\}.$$

It is straight-forward to check that this is indeed a metric on $X \times Y$. It is also easy to see that if $f\colon Z \to X \times Y$ is a function from a metric space Z to $X \times Y$, so that we may write $f(z) = (f_X(z), f_Y(z))$ with $f_X(z) \in X$ and $f_Y(z) \in Y$, then f is continuous if and only if f_X and f_Y are both continuous.

---

[5] This is completely standard terminology, though it's a little unfortunate if X is a vector space, where we use the word subspace to mean *linear* subspace also. Context (usually) makes it clear which meaning is intended, and I'll try and be as clear about this as possible!

[6] The fact that the sequences are bounded ensure the right-hand side is finite.

Example 2.16. If (X, d_X) and (Y, d_Y) are metric spaces, then we can also consider the set $\mathcal{B}(X, Y)$ of *bounded* functions from X to Y. This set has a natural metric given by

$$d(f, g) = \sup_{x \in X} d_Y(f(x), g(x)).$$

Indeed one can check that $d(f, g)$ is finite for any $f, g \in \mathcal{B}(X, Y)$, so that since d_Y is non-negatively valued, so is d. This space has a natural subspace consisting of the continuous bounded function $\mathcal{C}_b(X, Y)$.

Example 2.17. Consider the set $\mathbb{P}(\mathbb{R}^n)$ of lines in $\mathbb{R}^n$ (that is, one-dimensional subspace of $\mathbb{R}^n$, or lines through the origin). A natural way to define a distance on this set is to take, for lines L_1, L_2, the distance between L_1 and L_2 to be

$$d(L_1, L_2) = \sqrt{1 - \frac{|\langle v, w\rangle|^2}{\|v\|^2\|w\|^2}},$$

where v and w are any non-zero vectors in L_1 and L_2 respectively. It is easy to see this is independent of the choice of vectors v and w. The Cauchy-Schwarz inequality ensures that d is well-defined, and moreover the criterion for equality in that inequality ensures positivity. The symmetry property is evident, while the triangle inequality is left as an exercise.

It is useful to think of the case when $n = 2$ here, that is, the case of lines through the origin in the plane $\mathbb{R}^2$. The distance between two such lines given by the above formula is then $\sin(\theta)$ where θ is the angle between the two lines.

3. Normed vector spaces.

We have already seen a number of metrics on the vector space $\mathbb{R}^n$:

$$d_1(x, y) = \sum_{i=1}^{m} |x_i - y_i|$$
$$d_2(x, y) = \big(\sum_{i=1}^{m} (x_i - y_i)^2\big)^{1/2}$$
$$d_\infty(x, y) = \max_{1 \le i \le m} |x_i - y_i|.$$

These metrics all interact with the vector space structure[7] of $\mathbb{R}^n$ in a nice way: if d is any of these metrics, then for any vectors $x, y, z \in \mathbb{R}^n$ and any scalar λ we have

$$d(x + z, y + z) = d(x, y), \quad d(\lambda x, \lambda y) = |\lambda| d(x, y).$$

The first of these is known as *translation invariance* and the second is a kind of *homogeneity*.

A vector space V with a distance function compatible with the vector space structure is clearly determined by the function from V to the non-negative real numbers given by $v \mapsto d(v, 0)$.

Definition 3.1. Let V be a (real or complex) vector space. A *norm* on V is a function $\|.\| \colon V \to \mathbb{R}_{\geq 0}$ which satisfies the following properties:

(1) (*Positivity*): $\|x\| \geq 0$ for all $x \in V$ and $\|x\| = 0$ if and only if $x = 0$.
(2) (*compatibility with scalar multiplication*): if $x \in V$ and λ is a scalar then

$$\|\lambda.x\| = |\lambda|\|x\|.$$

(3) (*Triangle inequality*): If $x, y \in V$ then $\|x + y\| \leq \|x\| + \|y\|$.

[7] That is, vector addition and scalar multiplication.

Note that in the second property $|\lambda|$ denotes the absolute value of λ if V is a real vector space, and the modulus of λ if V is a complex vector space.

Remark 3.2. If there is the potential for ambiguity, we will write the norm on a vector space as $\|.\|_V$, but normally this is clear from the context, and so just as for metric spaces we will write $\|.\|$ for the norm on all vector spaces we consider.

Lemma 3.3. *If V is a vector space with a norm $\|.\|$ then the function $d\colon V \times V \to \mathbb{R}_{\geq 0}$ given by $d(x,y) = \|x - y\|$ is a metric which is compatible with the vector space structure in that:*

(1) *For all $x, y \in V$ we have*
$$d(\lambda.x, \lambda.y) = |\lambda| d(x,y).$$

(2) *$d(x+z, y+z) = d(x,y)$.*

Conversely, if d is a metric satisfying the above conditions then $\|v\| = d(v,0)$ is a norm on V.

Proof. This follows immediately from the definitions. □

Example 3.4. As discussed above, if $V = \mathbb{R}^n$ then the metrics d_1, d_2, d_∞ all come from the norms. We denote these by $\|x\|_1 = \sum_{i=1}^m |x_i|$ and $\|x\|_2 = (\sum_{i=1}^m x_i^2)^{1/2}$ and $\|x\|_\infty = \max_{1\leq i \leq m} |x_i|$.

Since the most natural maps between vector spaces are linear maps, it is natural to ask when a linear map between normed vector spaces is continuous. The following lemma gives an answer to this question:

Lemma 3.5. *Let $f\colon V \to W$ be a linear map between normed vector spaces. Then f is continuous if and only if $\{\|f(x)\| : \|x\| \leq 1\}$ is bounded.*

Proof. If f is continuous, then it is continuous at $0 \in V$ and so there is a $\delta > 0$ such that for all $v \in V$ with $\|v\| < \delta$ we have $\|f(v) - f(0)\| = \|f(v)\| < \epsilon$. But then if $\|v\| \leq 1$ we have $\frac{\delta}{2}\|f(v)\| = \|f(\frac{\delta}{2}.v))\| < \epsilon$, and hence $\|f(v)\| \leq \frac{2\epsilon}{\delta}$.

For the converse, if we have $\|f(v)\| < M$ for all v with $\|v\| \leq 1$, then if $\epsilon > 0$ is given we may pick $\delta > 0$ so that $\delta.M < \epsilon$ and hence if $\|v - w\| < \delta$ we have
$$\|f(v) - f(w)\| = \|f(v-w)\| = \delta\|f(\delta^{-1}(v-w))\| \leq \delta.M < \epsilon,$$
so that f is in fact uniformly continuous on V. □

Remark 3.6. The boundedness condition above can be rephrased as saying there is a constant $K > 0$ such that $\|f(v)\| \leq K.\|v\|$, since any non-zero vector v can be rescaled to a vector of unit length, $v/\|v\|$.

An important source of (normed) vector spaces for us will be the space of functions on a set X (usually a metric space). Indeed if X is any set, the space of all real-valued functions on X is a vector space – addition and scalar multiplication are defined "pointwise" just as for functions on the real line. It is not obvious how to make this into a normed vector space, but if we restrict to the subspace $\mathcal{B}(X)$ of *bounded* functions there is an reasonably natural choice of norm.

Definition 3.7. If X is any set we define
$$\mathcal{B}(X) = \{f\colon X \to \mathbb{R} : f(X) \subset \mathbb{R} \text{ bounded}\},$$
to be the space of bounded functions on X, that is $f \in \mathcal{B}(X)$ if and only if there is some $K > 0$ such that $|f(x)| < K$ for all $x \in X$. For $f \in \mathcal{B}(X)$ we set $\|f\|_\infty = \sup_{x \in X} |f(x)|$.

Lemma 3.8. *Let X be any set, then $(\mathcal{B}(X), \|.\|_\infty)$ is a normed vector space.*

Proof. To see that $\mathcal{B}(X)$ is a vector space, note that if $f, g \in \mathcal{B}(X)$ then we may find $N_1, N_2 \in \mathbb{R}_{>0}$ such that $f(X) \subseteq [-N_1, N_1]$ and $g(X) \subseteq [-N_2, N_2]$. But then clearly $(f+g)(X) \subseteq [-N_1-N_2, N_1+N_2]$ and if $\lambda \in \mathbb{R}$ then $(\lambda.f)(X) \subseteq [-|\lambda|N_1, |\lambda|N_1]$, so that $\lambda.f \in \mathcal{B}(X)$ and $f+g \in \mathcal{B}(X)$.

Next we check that $\|f\|_\infty$ is a norm: it is clear from the definition that $\|f\|_\infty \geq 0$ with equality if and only if f is identically zero. Compatibility with scalar multiplication is also immediate, while for the triangle inequality note that if $f, g \in \mathcal{B}(X)$, then for all $x \in X$ we have

$$|(f+g)(x)| = |f(x)+g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty.$$

Taking the supremum over $x \in X$ then yields the result. □

We will write d_∞ for the metric associated to the norm $\|.\|_\infty$.

If X is itself a metric space, we also have the space $\mathcal{C}(X)$ of continuous real-valued functions on X. While $\mathcal{C}(X)$ does not automatically have a norm, the subspace $\mathcal{C}_b(X) = \mathcal{C}(X) \cap \mathcal{B}(X)$ of *bounded* continuous functions clearly inherits a norm from $\mathcal{B}(X)$.

Example 3.9. One can check that if $X = [a,b]$ then if $(f_n)_{n\geq 1}$ is a sequence in[8] $\mathcal{C}([a,b]) = \mathcal{C}_b([a,b])$ then $f_n \to f$ in $(\mathcal{C}_b(X), d_\infty)$ (where d_∞ is the metric given by the norm $\|.\|_\infty$) if and only if f_n tends to f uniformly.

Example 3.10. For certain spaces X, we can define other natural metrics on the space of continuous functions: Let $X = [a,b] \subset \mathbb{R}$ be a closed interval. Then we know that in fact all continuous functions on X are bounded, so that $\|.\|_\infty$ defines a norm on $\mathcal{C}([a,b])$. We can also define analogues of the norms $\|.\|_1$ and $\|.\|_2$ on $\mathbb{R}^n$ using the integral in place of summation: Let

$$\|f\|_1 = \int_a^b |f(t)|dt,$$
$$\|f\|_2 = \big(\int_a^b f(t)^2 dt\big)^{1/2}$$

Lemma 3.11. *Suppose that $a < b$ so that the interval $[a,b]$ has positive length. Then the functions $\|.\|_1$ and $\|.\|_2$ are norms on $\mathcal{C}([a,b])$.*

Proof. The compatibility with scalars and the triangle inequality both follow from standard properties of the integral. The interesting point to check here is that both $\|.\|_1$ and $\|.\|_2$ satisfy postitivity – continuity[9] is crucial for this! Indeed if $f = 0$ clearly $\|f\|_1 = \|f\|_2 = 0$. On the other hand if $f \neq 0$ then there is some $x_0 \in [a,b]$ such that $f(x_0) \neq 0$, and so $|f(x_0)| > 0$. Since f is continuous at x_0, there is a $\delta > 0$ such that $|f(x) - f(x_0)| < |f(x_0)|/2$ for all $x \in [a,b]$ with $|x - x_0| < \delta$. But the it follows that for $x \in [a,b]$ with $|x - x_0| < \delta$ we have $|f(x)| \geq |f(x_0)| - |f(x) - f(x_0)| > |f(x_0)|/2$. Now set

$$s(x) = \begin{cases} |f(x_0)|/2, & \text{if } x \in [a,b] \cap (x_0 - \delta, x_0 + \delta) \\ 0, & \text{otherwise} \end{cases}$$

Since the interval $[a,b] \cap (x_0 - \delta, x_0 + \delta)$ has length at least $d = \min\{\delta, (b-a)\}$ we see that $\int_a^b s(x)dx \geq d.|f(x_0)|/2 > 0$. Since $s(x) \leq |f(x)|$ for all $x \in [a,b]$ it follows from the positivity of the integral that $0 < d|f(x_0)|/2 \leq \|f\|_1$. Similarly we see that $\|f\|_2 \geq f\sqrt{d}|f(x_0)|/2$, so that both $\|.\|_1$ and $\|.\|_2$ satisfy the positivity property. □

There are very similar metrics on certain sequence spaces:

[8]The result from Prelims Analysis showing any continuous function on a closed bounded interval is bounded implies the equality $\mathcal{C}([a,b]) = \mathcal{C}_b([a,b])$.

[9]So in particular, $\|.\|_1$ and $\|.\|_2$ are *not* norms on the space of Riemann integrable functions on $[a,b]$.

Example 3.12. Let

$$\ell_1 = \{(x_n)_{n\geq 1} : \sum_{n\geq 1} |x_n| < \infty\}$$
$$\ell_2 = \{(x_n)_{n\geq 1} : \sum_{n\geq 1} x_n^2 < \infty\}$$
$$\ell_\infty\{(x_n)_{n\geq 1} : \sup_{n\in\mathbb{N}} |x_n| < \infty\}.$$

The sets $\ell_1, \ell_2, \ell_\infty$ are all real vector spaces, and moreover the functions $\|(x_n)\|_1 = \sum_{n\geq 1} |x_n|$, $\|(x_n)\|_2 = \big(\sum_{n\geq 1} x_n^2\big)^{1/2}$, $\|(x_n)\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$ define norms on ℓ_1, ℓ_2 and ℓ_∞ respectively. Note that ℓ_2 is in fact an inner product space where

$$\langle (x_n), (y_n)\rangle = \sum_{n\geq 1} x_n y_n,$$

(the fact that the right-hand side converges if (x_n) and (y_n) are in ℓ_2 follows from the Cauchy-Schwarz inequality).

4. Metrics and convergence

Recall that if (X, d) is a metric space, then a sequence (x_n) in X converges to a point $a \in X$ if for any $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_n, a) < \epsilon$. In the case of $\mathbb{R}^m$, although d_1, d_2, d_∞ are all different distance functions, they in fact give the same notion of convergence. To see this we need the following:

Lemma 4.1. *Let $x, y \in \mathbb{R}^m$. Then we have*

$$d_2(x,y) \leq d_1(x,y) \leq \sqrt{m} d_2(x,y); \quad d_\infty(x,y) \leq d_2(x,y) \leq \sqrt{m} d_\infty(x,y).$$

Proof. It is enough to check the corresponding inequalities for the norms $\|x\|_i$ (where $i \in \{1, 2, \infty\}$) that is, we may assume $y = 0$. For the first inequality, note that

$$\|x\|_1^2 = (\sum_{i=1}^m |x_i|)^2 = \sum_{i=1}^m x_i^2 + \sum_{1\leq i<j\leq m} 2|x_i x_j| \geq \sum_{i=1}^m x_i^2 = \|x\|_2^2,$$

so that $\|x\|_2 \leq \|x\|_1$. On the other hand, if $x = (x_1, \ldots, x_m)$, set $a = (|x_1|, |x_2|, \ldots, |x_m|)$ and $\mathbf{1} = (1, 1, \ldots, 1)$. Then by the Cauchy-Schwarz inequality we have

$$\|x\|_1 = \langle \mathbf{1}, a\rangle \leq \sqrt{m}.\|a\|_2 = \sqrt{m}.\|x\|_2$$

The second pair of inequalities is simpler. Note that clearly

$$\max_{1\leq i\leq m} |x_i| = \max_{1\leq i\leq m} (x_i^2)^{1/2} \leq (\sum_{i=1}^m x_i^2)^{1/2},$$

yielding one inequality. On the other hand, since for each i we have $|x_i| \leq \|x\|_\infty$ by definition, clearly

$$\|x\|_2^2 = \sum_{i=1}^m |x_i|^2 \leq m\|x\|_\infty^2,$$

giving $\|x\|_2/\sqrt{m} \leq \|x\|_\infty$ as required. □

Lemma 4.2. *If $(x^n) \subset \mathbb{R}^m$ is a sequence then (x^n) converges to $a \in \mathbb{R}^m$ with respect to the metric d_2, if and only if it does with respect to the metric d_1, if and only if it does so with respect to the metric d_∞. Thus the three metrics all yield the same notion of convergence.*

Proof. Suppose (x^n) converges to a with respect to the metric d_2. Then for any $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that $d_2(x^n, a) < \epsilon/\sqrt{m}$ for all $n \geq N$. It follows from the previous Lemma that for $n \geq N$ we have

$$d_1(x^n, a) \leq \sqrt{m}.d_2(x^n, a) < \sqrt{m}.(\epsilon/\sqrt{m}) = \epsilon,$$

and so (x^n) converges to a with respect to d_1 also. Similarly we see that convergence with respect to d_1 implies convergence with respect to d_2 using $\|x\|_2 \leq \|x\|_1$. In the same fashion, the inequalities $d_\infty(x, y) \leq d_2(x, y) \leq \sqrt{m}d_\infty(x, y)$ yield the equivalence of the notions of convergence for d_2 and d_∞. □

Remark 4.3. (Non-examinable): If X is any set and d_1, d_2 are two metrics on X, we say they are equivalent if there are positive constants K, L such that

$$d_1(x, y) \leq K d_2(x, y); \quad d_2(x, y) \leq L d_1(x, y).$$

The proof of the previous Lemma extends to show that if two metrics are equivalent, then a sequence converges with respect to one metric if and only if it does with respect to the other.

In the problem sets you are asked to investigate which (if any) of the metrics d_1, d_2, d_∞ for $\mathcal{C}[a, b]$ the space of continuous real-valued functions on the closed interval $[a, b]$ define the same notion of convergence.

5. Open and closed sets

In this section we will define two special classes of subsets of a metric space – the open and closed subsets. To motivate their definition, recall that we have two ways of characterizing continuity in a metric space: the "ϵ-δ" definition, and the description in terms of convergent sequences. Examining the former will lead us to the notion of an open set, while examining the latter will lead us to the notion of a limit point and hence that of a closed set.

The definitions of continuity and convergence can be made somewhat more geometric if we introduce the notion of a ball in a metric space:

Definition 5.1. Let (X, d) is a metric space. If $x_0 \in X$ and $\epsilon > 0$ then we define the *open ball of radius* ϵ to be the set

$$B(x_0, \epsilon) = \{x \in X : d(x, x_0) < \epsilon\}.$$

Similarly we defined the *closed ball* of radius ϵ about x_0 to be the set

$$\bar{B}(x_0, \epsilon) = \{x \in X : d(x, x_0) \leq \epsilon\}.$$

The term "ball" comes from the case where $X = \mathbb{R}^3$ equipped with the usual Euclidean notion of distance. When $X = \mathbb{R}$ an open/closed ball is just an open/closed interval.

Recall that if $f \colon X \to Y$ is a function between any two sets, then given any subset $Z \subseteq Y$ we let[10] $f^{-1}(Z) = \{x \in X : f(x) \in Z\}$. The set $f^{-1}(Z)$ is called the *pre-image* of Z under the function f.

Lemma 5.2. *Let (X, d) and (Y, d) be metric spaces. Then $f \colon X \to Y$ is continuous at $a \in X$ if and only if, for any open ball $B(f(a), \epsilon)$ centred at $f(a)$ there is an open ball $B(a, \delta)$ centred at a such that $f(B(a, \delta)) \subseteq B(f(a), \epsilon)$, or equivalently $B(a, \delta) \subseteq f^{-1}(B(f(a), \epsilon))$.*

Proof. This follows directly from the definitions. (*Check this!*) □

[10] The notion is not meant to suggest that f is invertible, though when it is, the preimage of any point in Y is a single point in X, so the notation is in this sense consistent. Note that formally, f^{-1} as defined here is a function from the power set of Y to the power set of X.

We have seen in the last section that different metrics on a set X can give the same notions of continuity. The next definition is motivated by this – it turns out that we can attach to a metric a certain class of subsets of X known as *open sets* and knowing these open sets suffices to determine which functions on X are continuous. Informally, a subset $U \subseteq X$ is open if, for any point $y \in U$, every point sufficiently close to y in X is also in U. Thus, if $y \in U$, it has some "wiggle room" – we may move slightly away from y while still remaining in U. The rigorous definition is as follows:

Definition 5.3. If (X, d) is a metric space then we say a subset $U \subset X$ is *open* (or *open in* X) if for each $y \in U$ there is some $\delta > 0$ such that $B(y, \delta) \subseteq U$. More generally, if $Z \subseteq X$ and $z \in Z$ then we say Z is a *neighbourhood* of z if there is a $\delta > 0$ such that $B(z, \delta) \subseteq Z$. Thus a subset $U \subseteq X$ is open exactly when it is a neighbourhood of all of its elements.

The collection $\mathcal{T} = \{U \subset X : U \text{ open in } X\}$ of open sets in a metric space (X, d) is called the *topology* of X.

We first note an easy lemma, which can be viewed as a consistency check on our terminology!

Lemma 5.4. *Let (X, d) be a metric space. If $a \in X$ and $\epsilon > 0$ then $B(a, \epsilon)$ is an open set.*

Proof. We need to show that $B(a, \epsilon)$ is a neighbourhood of each of its points. If $x \in B(a, \epsilon)$ then by definition $r = \epsilon - d(a, x) > 0$. We claim that $B(x, r) \subseteq B(a, \epsilon)$. Indeed by the triangle inequality we have for $z \in B(x, r)$

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

as required. □

Remark 5.5. While reading the above proof, please draw a picture of the case where $X = \mathbb{R}^2$ with the standard metric d_2!

Next let us observe some basic properties of open sets.

Lemma 5.6. *Let (X, d) be metric space and let $\mathcal{T}$ be the associated topology on X. Then we have*

1. *If $\{U_i ; i \in I\}$ is any collection of open sets, then $\bigcup_{i \in I} U_i$ is an open set. In particular the empty set $\emptyset$ is open in X*[11]
2. *If I is finite and $\{U_i : i \in I\}$ are open sets then $\bigcap_{i \in I} U_i$ is open in X. In particular X is an open set.*

Proof. For the first claim, if $x \in \bigcup_{i \in} U_i$ then there is some $i \in I$ with $x \in U_i$. Since U_i is open, there is an $\epsilon > 0$ such that

$$B(x, \epsilon) \subset U_i \subseteq \bigcup_{i \in I} U_i,$$

so that $\bigcup_{i \in I} U_i$ is a neighbourhood of each of its points as required. Applying this to the case $I = \emptyset$ shows that $\emptyset \subseteq X$ is open (or simply note that the empty set satisfies the condition to be an open set vacuously).

For the second claim, if I is finite and $x \in \bigcap_{i \in I} U_i$, then for each i there is an $\epsilon_i > 0$ such that $B(x, \epsilon_i) \subseteq U_i$. But then since I is finite, $\epsilon = \min(\{\epsilon_i : i \in I\} \cup \{1\}) > 0$, and

$$B(x, \epsilon) \subseteq \bigcap_{i \in I} B(x, \epsilon_i) \subseteq \bigcap_{i \in I} U_i,$$

[11]Note that if I is an indexing set, then a collection $\{U_i : i \in I\}$ of subsets of X is just a function $u \colon I \to \mathcal{P}(X)$ where $\mathcal{P}(X)$ denotes the power set of X, where we normally write $U_i \subseteq X$ for $u(i)$. The union of the collection of subsets $\{U_i : i \in I\}$ is then $\{x \in X : \exists i \in I, x \in U_i\}$, while the intersection of the collection $\{U_i : i \in I\}$ is just $\{x \in X : \forall i \in I, x \in U_i\}$. Using this, one readily sees that if $I = \emptyset$ then the intersection of the collection is X and the union is the empty set $\emptyset$.

so that $\bigcap_{i \in I} U_i$ is an open subset as required. Applying this to the case $I = \emptyset$ shows that X is open (or simply note that if $U = X$ and $x \in X$ then $B(x, \epsilon) \subseteq X$ for *any* positive ϵ so that X is open). □

Remark 5.7. If you look in many textbooks for the definition of a topology on a set X, then you will often see the axioms insisting separately that $\emptyset$ and X are open, alongside the conditions that finite intersections and arbitrary unions of open sets are open. The phrasing of the above Lemma is designed to emphasize that this is redundant. In practice of course it is normally immediate from the definition of the topology that both $\emptyset$ and X are open, so unfortunately this is not an observation that saves one much work (and is presumably why the extraneous stipulation is so common-place in the literature).

Exercise 5.8. Using Lemma 4.1, show that the topologies $\mathcal{T}_i$ on $\mathbb{R}^n$ given by the norms d_i ($i = 1, 2, \infty$) coincide.

Example 5.9. A subset U of $\mathbb{R}$ is open if for any $x \in U$ there is an open interval centred at x contained in U. Thus we can readily see that the finiteness condition for intersections is necessary: if $U_i = (-1/i, 1)$ for $i \in \mathbb{N}$ then each U_i is open but $\bigcap_{i \in \mathbb{N}} U_i = [0, 1)$ and $[0, 1)$ is not open because it is not a neighbourhood of 0.

One important consequence of the fact that arbitrary unions of open sets are open is the following:

Definition 5.10. Let (X, d) be a metric space and let $S \subseteq X$. The *interior* of S is defined to be

$$\operatorname{int}(S) = \bigcup_{\substack{U \subseteq S \\ U \text{open}}} U.$$

Since the union of open subsets is always open $\operatorname{int}(S)$ is an open subset of X and it is the largest open subset of X which is contained in S in the sense that any open subset of X which is contained in S is in fact contained in $\operatorname{int}(S)$. If $x \in \operatorname{int}(S)$ we say that x is an *interior point* of S. One can also phrase this in terms of neighborhoods: the interior of S is the set of all points in S for which S is a neighbourhood.

Example 5.11. If $S = [a, b]$ is a closed interval in $\mathbb{R}$ then its interior is just the open interval (a, b). If we take $S = \mathbb{Q} \subset \mathbb{R}$ then $\operatorname{int}(\mathbb{Q}) = \emptyset$.

We now show that the topology given by a metric is sufficient to characterize continuity.

Proposition 5.12. *Let X and Y be metric spaces and let $f : X \to Y$ be a function. If $a \in X$ then f is continuous at a if and only if for every neighbourhood $N \subseteq Y$ of $f(a)$, the preimage $f^{-1}(N)$ is a neighbourhood of $a \in X$. Moreover, f is continuous on all of X if and only if for each open subset U of Y, its preimage $f^{-1}(U)$ is open in X.*

Proof. First suppose that f is continuous at a, and let N be a neighhourhood of $f(a)$. Then we may find an $\epsilon > 0$ such that $B(f(a), \epsilon) \subseteq N$. Since f is continuous at a, there is a $\delta > 0$ such that $B(x, \delta) \subseteq f^{-1}(B(f(a), \epsilon)) \subseteq f^{-1}(U)$. It follows $f^{-1}(N)$ is a neighbourhood of a as required. Conversely, if $\epsilon > 0$ is given, then certainly $B(f(a), \epsilon)$ is a neighbourhood of $f(a)$, so that $f^{-1}(B(f(a), \epsilon))$ is a neighbourhood of a, hence there is a $\delta > 0$ such that $B(a, \delta) \subseteq f^{-1}(B(f(a), \epsilon))$, and thus f is continuous at a as required.

Now if f is continuous on all of X, since a set is open if and only if it is a neighbourhood of each of its points, it follows from the above that $f^{-1}(U)$ is an open subset of X for any open subset U of Y. For the converse, note that if $a \in X$ is any point of X and $\epsilon > 0$ is given then the open ball $B(f(a), \epsilon)$ is an open subset of Y, hence $f^{-1}(B(f(a), \epsilon))$ is open in X, and in particular is a neighbourhood of $a \in X$. But then there is a $\delta > 0$ such that $B(a, \delta) \subseteq f^{-1}(B(f(a), \epsilon))$, hence f is continuous at a as required.

□

Example 5.13. Notice that this Proposition gives us a way of producing many examples of open sets: if $f\colon \mathbb{R}^n \to \mathbb{R}$ is any continuous function and $a, b \in \mathbb{R}$ are real numbers with $a < b$ then $\{v \in \mathbb{R}^n : a < f(x) < b\} = f^{-1}((a,b))$ is open in $\mathbb{R}^n$. Thus for example $\{(x,y) \in \mathbb{R}^2 : 1 < 2x^2 + 3xy < 2\}$ is an open subset of the plane.

Exercise 5.14. Use the characterization of continuity in terms of open sets to show that the composition of continuous functions is continuous[12].

Remark 5.15. The previous Proposition 5.12 shows, perhaps surprisingly, that we actually need somewhat less than a metric on a set X to understand what continuity means: we only need the topology induced by the metric on the set X. In particular any two metrics which give the same topology give the same notion of continuity. This motivates the following, perhaps rather abstract-seeming, definition.

Definition 5.16. If X is a set, a *topology* on X is a collection of subsets $\mathcal{T}$ of X, known as the *open subsets* which satisfy the conclusion of Lemma 5.6. That is,

(1) If $\{U_i : i \in I\}$ are in $\mathcal{T}$ then $\bigcup_{i \in I} U_i$ is in $\mathcal{T}$. In particular $\emptyset$ is an open subset.
(2) If I is finite and $\{U_i : i \in I\}$ are in $\mathcal{T}$, then $\bigcap_{i \in I} U_i$ is in $\mathcal{T}$. In particular X is an open subset of X.

A *topological space* is a pair $(X, \mathcal{T}_X)$ consisting of a set X and a choice of topology $\mathcal{T}_X$ on X.

Motivated by Proposition 5.12, if $f\colon X \to Y$ is a function between two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ we say that f is *continuous* if for every open subset $U \in \mathcal{T}_Y$ we have $f^{-1}(U) \in \mathcal{T}_X$, that is, $f^{-1}(U)$ is an open subset of X.

The properties of a metric space which we can express in terms of open sets can equally be expressed in terms of their complements, which are known as *closed sets*. It is useful to have both formulations (as we will show, the formulation of continuity in terms of closed sets is closer to that given by convergence of sequences rather than the ϵ-δ definition).

Definition 5.17. If (X, d) is a metric space, then a subset $F \subseteq X$ is said to be a *closed* subset of X if its complement $F^c = X \backslash F$ is an open subset.

Remark 5.18. It is important to note that the property of being closed is *not* the property of not being open! In a metric space, it is possible for a subset to be open, closed, both or neither: In $\mathbb{R}$ the set $\mathbb{R}$ is open and closed, the set $(0,1)$ is open and not closed, the set $[0,1]$ is closed and not open while the set $(0,1]$ is neither.

The following lemma follows easily from Lemma 5.6 by using DeMorgan's Laws.

Lemma 5.19. *Let (X, d) be a metric space and let $\{F_i : i \in I\}$ be a collection of closed subsets.*

(1) *The intersection $\bigcap_{i \in I} F_i$ is a closed subset. In particular X is a closed subset of X.*
(2) *If I is finite then $\bigcup_{i \in I} F_i$ is closed. In particular the empty set $\emptyset$ is a closed subset of X.*

Moreover, if $f\colon X \to Y$ is a function between two metric spaces X and Y then f is continuous if and only if $f^{-1}(G)$ is closed for every closed subset $G \subseteq Y$.

Proof. The properties of closed sets follow immediately from DeMorgan's law, while the characteriszation of continuity follows from the fact that if $G \subset Y$ is any subset of Y we have $f^{-1}(G^c) = (f^{-1}(G))^c$, that is, $X \backslash f^{-1}(G) = f^{-1}(Y \backslash G)$. □

Lemma 5.20. *If (X, d) is a metric space then any closed ball $\bar{B}(a, r)$ for $r \geq 0$ is a closed set. In particular, singleton sets are closed.*

[12] This is easy, the point is just to check you see how easy it is!

Proof. We must show that $X\backslash\bar{B}(a,r)$ is open. If $y \in X\backslash\bar{B}(a,r)$ then $d(a,y) > r$, so that $\epsilon = d(a,y) - r > 0$. But then if $z \in B(y,\epsilon)$ we have

$$d(a,z) \geq d(a,y) - d(z,y) > d(a,y) - \epsilon = r,$$

so that $z \notin \bar{B}(a,r)$. It follows that $B(y,\epsilon) \subseteq X\backslash\bar{B}(a,r)$ and so $X\backslash\bar{B}(a,r)$ is open as required. □

The relation between closed sets and convergent sequences mentioned at the beginning of this section arises through the notion of a limit point which we now define.

Definition 5.21. If (X,d) is a metric space and $Z \subseteq X$ is any subset, then we say a point $a \in X$ is a *limit point* if for any $\epsilon > 0$ we have $\big(B(a,\epsilon)\backslash\{a\}\big) \cap Z \neq \emptyset$. If $a \in Z$ and a is *not* a limit point of Z we say that a is an *isolated point* of Z. The set of limit points of Z is denoted Z'. Notice that if $Z_1 \subseteq Z_2$ are subsets of X then it follows immediately from the definition that $Z_1' \subseteq Z_2'$.

Example 5.22. If $Z = (0,1] \cup \{2\} \subset \mathbb{R}$ then 0 is a limit point of Z which does not lie in Z, while 2 is an isolated point of Z because $B(2,1/2) \cap Z = (1.5, 2.5) \cap Z = \{2\}$.

If (x_n) is a sequence in (X,d) which converges to $\ell \in X$ then $\{x_n : n \in \mathbb{N}\}$ is either empty or equal to $\{\ell\}$. (See the problem set.)

The term "limit point" is motivated by the following easy result:

Lemma 5.23. *If S is a subset of a metric space (X,d) then $x \in S'$ if and only if there is a sequence in $S\backslash\{x\}$ converging to x.*

Proof. If x is a limit point then for each $n \in \mathbb{N}$ we may pick $z_n \in B(x,1/n) \cap (S\backslash\{x\})$. Then clearly $z_n \to x$ as $n \to \infty$ as required. Conversely if (z_n) is a sequence in $S\backslash\{x\}$ converging to x and $\delta > 0$ is given, there is an $N \in \mathbb{N}$ such that $z_n \in B(x,\delta)$ for all $n \geq N$. It follows that $B(x,\delta) \cap (S\backslash\{x\})$ is nonempty as required. □

The fact that a subset of a metric space is closed can be characterized in terms of limit points (and hence in terms of convergent seqeunces):

Lemma 5.24. *If (X,d) is a metric space and $S \subseteq X$ then S is closed if and only if $S' \subseteq S$.*

Proof. If S is closed then S^c is open and so for all $y \notin S$ there is a $\delta > 0$ such that $B(y,\delta) \subseteq S^c$. Thus $S \cap B(y,\delta) = \emptyset$ and so y is not a limit point of S. Hence $S' \subseteq S$ as required. On the other hand if $S' \subseteq S$ then if $y \notin S$ it follows y is not a limit point of S so that there is a $\delta > 0$ such that $\big(B(y,\delta)\backslash\{y\}\big) \cap S = \emptyset$, and since $y \notin S$ it follows $B(y,\delta) \subseteq S^c$. It follows that S^c is open and hence S is closed. □

Remark 5.25. It follows from Lemma 5.24 and Lemma 5.23 that if $S \subseteq X$ then $a \in \bar{S}$ if and only if there is a sequence (x_n) in S with $x_n \to a$. Indeed if (x_n) is a sequence in Z and $x_n \to y$ as $n \to \infty$ then y must be a limit point of Z unless $x_n = y$ for all but finitely many n, in which case $y \in Z$. Conversely, if $a \in S'$ then we are done by Lemma 5.23, while if $a \in S$ we may take $x_n = a$ for all n.

The fact that any intersection of closed subsets is closed has an important consequence – given any subset S of a metric space (X,d) there is a unique smallest closed set which contains S.

Definition 5.26. Let (X,d) be a metric space and let $S \subseteq X$. Then the set

$$\bar{S} = \bigcap_{\substack{G \supseteq S \\ G \text{ closed}}} G,$$

is the *closure* of S. It is closed because it is the intersection of closed subsets of X and is the smallest closed set containing S in the sense that if G is any closed set containing S then G contains $\bar{S}$. If $S \subseteq Y \subseteq X$ we say that S is *dense* in Y if $Y \subseteq \bar{S}$. (Thus every point of Y lies in S or is a limit point of S.)

Example 5.27. The rationals $\mathbb{Q}$ are a dense subset of $\mathbb{R}$, as is the set $\{\frac{a}{2^n} : a \in \mathbb{Z}, n \in \mathbb{N}\}$.

Definition 5.28. The notions of closure and interior also allow us to define the *boundary* ∂S of a subset S of a metric space to be $\bar{S}\backslash \text{int}(S)$.

Proposition 5.29. *Let* (X, d) *be a metric space and let* $S \subseteq X$*. Then*

$$S \cup S' = \bar{S}.$$

Proof. Let $Y = S \cup S'$. Since $S \subseteq \bar{S}$, certainly $S' \subseteq (\bar{S})'$, and as $\bar{S}$ is closed, by Lemma 5.24, $(\bar{S})' \subseteq \bar{S}$. Hence $Y \subseteq \bar{S}$. To see the opposite inclusion, suppose that $a \notin Y$. Then there is a $\delta > 0$ such that $B(a, \delta) \cap S = \emptyset$. It follows that $S \subseteq B(a, \delta)^c$ and thus since $B(a, \delta)^c$ is closed, $\bar{S} \subseteq B(a, \delta)^c$, and so certainly $a \notin \bar{S}$. It follows $\bar{S} \subseteq Y$ and hence $\bar{S} = Y$ are required. □

Remark 5.30. If $Z \subseteq X$ is an arbitrary subset you can check that $(Z')' \subseteq Z'$, that is, the limit points of Z' are limit points of Z. It then follows from Lemma 5.24 that Z' is closed, since it contains its limit points.

Example 5.31. In general, it need *not* be the case that $\bar{B}(a, r)$ is the closure of $B(a, r)$. Since we have seen that $\bar{B}(a, r)$ is closed, it is always true that $\overline{B(a, r)} \subseteq \bar{B}(a, r)$ but the containment can be proper. As a (perhaps silly-seeming) example take any set X with at least two elements equipped with the discrete metric. Then if $x \in X$ we have $\{x\} = B(x, 1)$ is an open set consisting of the single point $\{x\}$. Since singletons are always closed we see that $\overline{B(x, 1)} = B(x, 1) = \{x\}$. On the other hand $\bar{B}(x, 1) = X$ the entire set, which is strictly larger than $\{x\}$ by assumption.

Remark 5.32. Combining the above characterization of closed sets in terms of limit points and the characterization of continuity in terms of closed sets we can give yet another description of continuity for a function $f \colon X \to Y$ between metric spaces: If $Z \subset Y$ is a subset of Y which contains all its limit points then so does $f^{-1}(Z)$. The problem set asks you to establish a slightly different characterization using the notion of the closure of a set, namely that a function $f \colon X \to Y$ is continuous if and only if for any subset $Z \subseteq X$ we have $f(\overline{Z}) \subseteq \overline{f(Z)}$. It is easy to relate this to the definition of continuity in terms of convergent sequences.

6. Subspaces of metric spaces

If (X, d) is a metric space, then as we noted before, any subset $Y \subseteq X$ is automatically also a metric space since the distance function $d \colon X \times X \to \mathbb{R}_{\geq 0}$ restricts to a distance function on Y. The set Y thus has a topology given by this metric. In this section we show that this topology is easy to describe in terms of the topology on X. The key to this description is the simple observation that the open balls in Y are just the intersection of the open balls in X with Y. For clarity, for $y \in Y \subseteq X$ we will write

$$B_Y(y, r) = \{z \in Y : d(z, y) < r\}$$

for the open ball about y of radius r in Y and

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

for the open ball of radius r about y in X. Thus $B_Y(y, r) = Y \cap B_X(y, r)$.

Lemma 6.1. *If* (X, d) *is a metric space and* $Y \subseteq X$ *then a subset* $U \subseteq Y$ *is an open subset of* Y *if and only if there is an open subset* V *of* X *such that* $U = V \cap Y$*. Similarly a subset* $Z \subseteq Y$ *is a closed subset of* Y *if and only if there is a closed subset* F *of* X *such that* $Z = F \cap Y$*.*

Proof. If $U = Y \cap V$ where V is open in X and $y \in U$ then there is a $\delta > 0$ such that $B_X(y, \delta) \subseteq V$. But then $B_Y(y, \delta) = B_X(y, \delta) \cap Y \subseteq V \cap Y = U$ and so U is a neighbourhood of each of its points as required. On the other hand, if U is an open subset of Y then for each $y \in U$ we may pick an open ball $B_Y(y, \delta_y) \subseteq U$. It follows that $U = \bigcup_{y \in U} B_Y(y, \delta_y)$. But then if we set $V = \bigcup_{y \in U} B_X(y, \delta_y)$ it is immediate that V is open in X and $V \cap Y = U$ as required.

The corresponding result for closed sets follows readily: F is closed in Y if and only if $Y \backslash F$ is open in Y which by the above happens if and only if it equals $Y \cap V$ for some open set in X. But this is equivalent to $T = Y \cap V^c$, the intersection of Y with the closed set V^c. ☐

Remark 6.2. The lemma shows that the topology on X determines the topology on the subspace $Y \subseteq X$ directly. It is easy to see that if $(X, \mathcal{T})$ is an abstract topological space and $Y \subseteq X$ then the collection $\mathcal{T}_Y = \{U \cap Y : U \in \mathcal{T}\}$ is a topology on Y which is called the *subspace topology*.

Remark 6.3. It is important here to note that the property of being open or closed is a relative one – it depends on which metric space you are working in. Thus for example if (X, d) is a metric space and $Y \subseteq X$ then Y is always open viewed as a subset of itself (since the whole space is always an open subset) but it of course need not be an open subset of X! For example, $[0, 1]$ is not open in $\mathbb{R}$ but it is an open subset of itself.

Example 6.4. Let's consider a more interesting example: Let $X = \mathbb{R}$ and let $Y = [0, 1] \cup [2, 3]$. As a subset of Y the set $[0, 1]$ is both open and closed. To see that it is open, note that if $x \in [0, 1]$ then

$$\begin{aligned} B_Y(x, 1/2) = B_{\mathbb{R}}(x, 1/2) \cap Y &= (x - \frac{1}{2}, x + \frac{1}{2}) \cap ([0, 1] \cup [2, 3]) \\ &= (x - \frac{1}{2}, x + \frac{1}{2}) \cap [0, 1] \subset [0, 1], \end{aligned}$$

Similarly we see that $B_Y(x, 1/2) \subseteq [2, 3]$ if $x \in [2, 3]$ so that $[2, 3]$ is also open in Y. It follows $[0, 1]$ is both open and closed in Y (as is $[2, 3]$).

7. Homeomorphisms and isometries

If (X, d) and (Y, d) are metric spaces it is natural to ask when we wish to consider X and Y equivalent. There is more than one way to answer this question – the first, perhaps most obvious one, is the following:

Definition 7.1. A function $f : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is said to be an *isometry* if

$$d_Y(f(x), f(y)) = d_X(x, y) \quad \forall x, y \in X$$

An isometry is automatically injective. If there is a surjective (and hence bijective) isometry between two metric spaces X and Y we say that X and Y are *isometric*.

Example 7.2. Let $X = \mathbb{R}^2$ (equipped with the Euclidean metric[13] d_2). The collection of all bijective isometries from X to itself forms a group, the *isometry group* of the plane. Clearly the translations $t_v : \mathbb{R}^2 \to \mathbb{R}^2$ are isometries, where $v \in \mathbb{R}^2$ and $t_v(x) = x + v$. Similarly, if $A \in \text{Mat}_2(\mathbb{R})$ is an orthogonal matrix, so that $A^t A = I$, then $x \mapsto Ax$ is an isometry: since $d_2(Ax, Ay) = \|A(x) - A(y)\| = \|A(x - y)\|$ it is enough to check that $\|Ax\| = \|x\|$, but this is clear since $\|Ax\|^2 = (Ax).(Ax) = xA^t Ax = x^t Ix = \|x\|$.

In fact these two kinds of isometries generate the full group of isometries. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is any isometry, let $v = T(0)$. Then $T_1 = t_{-v} \circ T$ is an isometry which fixes the origin. Thus it remains to show that any isometry which fixes the origin is in fact linear. But you showed in Prelims Geometry that any such isometry of $\mathbb{R}^n$ must preserve the inner product (because it preserves the norm and

[13]Unless it is explicitly stated otherwise, we will always take $\mathbb{R}^n$ to be a metric space equipped with the d_2 metric.

you can express the inner product in terms of the norm). It follows such an isometry takes an orthonormal basis to an orthonormal basis, from which linearity readily follows. (Note that this argument works just as well in $\mathbb{R}^n$.)

Example 7.3. Let $S^n = \{x \in \mathbb{R}^{n+1} : \|x\|_2 = 1\}$ be the n-sphere (so S^1 is a circle and S^2 is the usual sphere). Clearly $O_{n+1}(\mathbb{R})$ acts by isometries on S^n. In fact you can show that $\text{Isom}(S^n) = O_{n+1}(\mathbb{R})$. To prove this one needs to show that any isometry of S^n extends to an isometry of $\mathbb{R}^{n+1}$ which fixes the origin.

We have already seen that on $\mathbb{R}^n$ the metrics d_1, d_2, d_∞, although different, induce the same notion of convergence and continuity[14] . The notion of isometry is thus in some sense too rigid a notion of equivalence if these are the notions we are primarily interested in. A weaker, but often more useful, notion of equivalence is the following:

Definition 7.4. Let $f: X \to Y$ be a continuous function between metric spaces X and Y. We say that f is a *homeomorphism* if there is a continuous function $g: Y \to X$ such that $f \circ g = \text{id}_Y$ and $g \circ f = \text{id}_X$. If there is a homeomorphism between two metric spaces X and Y we say they are *homeomorphic*.

Remark 7.5. Note that the definition implies that f is bijective as a map of sets but it is *not* true in general[15] that a continuous bijection is necessarily a homeomorphism. To see this, consider the spaces $X = [0, 1) \cup [2, 3]$ and $Y = [0, 2]$. Then the function $f: X \to Y$ given by

$$f(x) = \begin{cases} x, & \text{if } x \in [0, 1) \\ x - 1, & \text{if } x \in [2, 3] \end{cases}$$

is a bijection and is clearly continuous. Its inverse $g: Y \to X$ is however not continuous at 1 – the one-sided limits of g as x tends to 1 from above and below are 1 and 2 respectively.

Example 7.6. The closed disk $\bar{B}(0, 1)$ of radius 1 in $\mathbb{R}^2$ is homoemorphic to the square $[-1, 1] \times [-1, 1]$. The easiest way to see this is inscribe the disk in the square and stretch the disk radially out to the square. One can write explicit formulas for this in the four quarters of the disk given by the lines $x \pm y = 0$ to check this does indeed give a homeomorphism.

The open interval $(0, 1)$ is homeomorphic to $\mathbb{R}$: a homeomorphism between them is given by the function $x \mapsto \tan(\pi.(x - 1/2))$, which has inverse $y \mapsto \frac{1}{\pi}\arctan(y) + \frac{1}{2}$.

8. Completeness

One of the important notions in Prelims analysis was that of a Cauchy sequence. This is a notion, like convergence, which makes sense in any metric space.

Definition 8.1. Let (X, d) be a metric space. A sequence (x_n) in X is said to be a *Cauchy sequence* if, for any $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq N$.

The following lemma establishes basic properties of Cauchy sequences in an arbitrary metric space which you saw before for real sequences.

Lemma 8.2. *Let (X, d) be a metric space.*

1. *If (x_n) is a convergent sequence then it is Cauchy.*
2. *Any Cauchy sequence is bounded.*

[14]There is actually a slightly subtle point here – to know that $(\mathbb{R}^n, d_1)$ and $(\mathbb{R}^n, d_2)$ are not isometric we would need to show that there is no bijective map $\alpha: \mathbb{R}^n \to \mathbb{R}^n$ such that $d_2(\alpha(x), \alpha(y)) = d_1(x, y)$ for all $x, y \in \mathbb{R}^n$.

[15]This is unlike the examples you have seen in algebra – the inverse of a bijective linear map is automatically linear, and the inverse of a bijective group homomorphism is automatically a homomorphism. Similarly, the inverse of a bijective isometry is also an isometry.

Proof. Suppose that $x_n \to \ell$ as $n \to \infty$ and $\epsilon > 0$ is given. Then there is an $N \in \mathbb{N}$ such that $d(x_n, \ell) < \epsilon/2$ for all $n \geq N$. It follows that if $n, m \geq N$ we have

$$d(x_n, x_m) \leq d(x_n, \ell) + d(\ell, x_m) < \epsilon/2 + \epsilon/2 = \epsilon,$$

so that (x_n) is a Cauchy sequence as required.

If (x_n) is a Cauchy sequence, then taking $\epsilon = 1$ in the definition, we see that there is an $N \in \mathbb{N}$ such that $d(x_n, x_m) < 1$ for all $n, m \geq N$. It follows that if we set

$$M = \max\{1, d(x_1, x_N), d(x_2, x_N), \ldots, d(x_{N-1}, x_N)\}$$

then for all $n \in \mathbb{N}$ we have $x_n \in B(x_N, M)$ so that (x_n) is bounded as required. □

Part (1) of the lemma motivates the following definition:

Definition 8.3. A metric space (X, d) is said to be *complete* if every Cauchy sequence in X converges.

Example 8.4. One of the main results in Analysis I was that $\mathbb{R}$ is complete, and it is easy to deduce from this that $\mathbb{R}^n$ is complete also (since a sequence in $\mathbb{R}^n$ converges if and only if each of its coordinates converge).

On the other hand, consider the metric space $(0, 1]$: The sequence $(1/n)$ converges in $\mathbb{R}$ (to 0) so the sequence is Cauchy in $\mathbb{R}$ and hence also in $(0, 1]$, however it does not converge in $(0, 1]$.

The previous example suggests a connection between completeness and closed sets. One precise statement of this form is the following:

Lemma 8.5. *Let (X, d) be a complete metric space and let $Y \subseteq X$. Then Y is complete if and only if Y is a closed subset of X.*

Proof. Note that if (x_n) is a Cauchy sequence in Y then it is certainly a Cauchy sequence in X. Since X is complete, (x_n) converges in X, say $x_n \to a$ as $n \to \infty$. Thus (x_n) converges in Y precisely when $a \in Y$. It follows that Y is complete if and only if it contains the limits of all sequences (x_n) in Y which converge in X. But Lemma 5.25 shows that the set of limits of all sequences in Y is exactly $\bar{Y}$, hence Y is complete if and only if $\bar{Y} \subseteq Y$, that is, if and only if Y is closed.

□

Another useful consequence of completeness is that it guarantees certain intersections of closed sets are non-empty:

Lemma 8.6. *Let (X, d) be a complete metric space and suppose that $D_1 \supseteq D_2 \supseteq \ldots$ form a nested sequence of non-empty closed sets in X with the property that diam$(D_k) \to 0$ as $k \to \infty$. Then there is a unique point $w \in X$ such that $w \in D_k$ for all $k \geq 1$.*

Proof. For each k pick $z_k \in D_k$. Then since the D_k are nested, $z_k \in D_l$ for all $k \geq l$, and hence the assumption on the diameters ensures that (z_k) is a Cauchy sequence. Let $w \in X$ be its limit. Since D_k is closed and contains the subsequence $(z_{n+k})_{n \geq 0}$ it follows $w \in D_k$ for each $k \geq 1$. To see that w is unique, suppose that $w' \in D_k$ for all k. Then $d(w, w') \leq \text{diam}(D_k)$ and since $\text{diam}(D_k) \to 0$ as $k \to \infty$ it follows $d(w, w') = 0$ and hence $w = w'$. □

Remark 8.7. Notice that the property of a metric space being complete is *not* preserved by homeomorphism – we have seen that $(0, 1)$ is homeomorphic to $\mathbb{R}$ but the former is not complete, while the latter is. This is because a homeomorphism does not have to take Cauchy sequences to Cauchy sequences.

Example 8.8. Let $Y = \{z \in \mathbb{C} : |z| = 1\}\backslash\{1\}$. Then Y is homeomorphic to $(0,1)$ via the map $t \mapsto e^{2\pi it}$, but their respective closures $\bar{Y}$ and $[0,1]$ however are not homeomorphic. (We will seem a rigorous proof of this later using the notion of connectedness.) The metric spaces Y and $(0,1)$ contain information about their closures in $\mathbb{R}^2$ which is lost when we only consider the topologies the metrics give: the space Y has Cauchy sequences which don't converge in Y, but these all converge to $1 \in \mathbb{C}$, whereas in $(0,1)$ there are two kinds of Cauchy sequences which do not converge in $(0,1)$ – the ones converging to 0 and the ones converging to 1. The point here is that given two Cauchy sequences we can detect if they converge to the same limit without knowing what that the limit actually is: (x_n) and (y_n) converge to the same limit if for all $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that $d(x_n, y_n) < \epsilon$ for all $n \geq N$. Using this idea one can define what is called the *completion* of a metric space (X,d): this is a complete metric space (Y,d) such which X embeds isometrically into as a dense[16] subset. For example, the real numbers $\mathbb{R}$ are the completion of $\mathbb{Q}$.

Many naturally arising metric spaces are complete. We now give a important family of such: recall that if X is any set, the space $\mathcal{B}(X)$ of bounded real-valued functions on X is normed vector space where if $f \in \mathcal{B}(X)$ we define its norm to be $\|f\|_\infty = \sup_{x\in X} |f(x)|$.

Theorem 8.9. *Let X be a set. The normed vector space $(\mathcal{B}(X), \|.\|_\infty)$ is complete.*

Proof. Let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $\mathcal{B}(X)$. Then we have for each $x \in X$

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty \to 0,$$

as $n, m \to \infty$. It follows that the sequence $(f_n(x))$ is a Cauchy sequence of real numbers and hence since $\mathbb{R}$ is complete it converges to a real number. Thus we may define a function $f \colon X \to \mathbb{R}$ by setting $f(x) = \lim_{n\to\infty} f_n(x)$.

We claim $f_n \to f$ in $\mathcal{B}(X)$. Note that this requires us to show both that $f \in \mathcal{B}(X)$ and $f_n \to f$ with respect to the norm $\|.\|_\infty$. To check these both hold, fix $\epsilon > 0$. Since (f_n) is Cauchy, we may find an $N \in \mathbb{N}$ such that $\|f_n - f_m\|_\infty < \epsilon$ for all $n, m \geq N$. Thus we have for all $x \in X$ and $n, m \geq N$

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| < \epsilon.$$

But now letting $n \to \infty$ we see that for any $m \geq N$ we have $|f(x) - f_m(x)| \leq \epsilon$ for all $x \in X$. But then for any such m we certainly have $f - f_m \in \mathcal{B}(X)$ so that[17] $f = f_m + (f - f_m) \in \mathcal{B}(X)$, and since $\|f - f_m\|_\infty \leq \epsilon$ for all $m \geq N$ it follows $f_m \to f$ as $m \to \infty$ as required.

□

As we already observed, if X is also a metric space then we can also consider the space of bounded continuous functions $\mathcal{C}_b(X)$ on X. This is a normed subspace of $\mathcal{B}(X)$, and the following theorem is a generalization of the result you saw last year showing that a uniform limit of continuous functions is continuous (the proof is essentially the same also).

Theorem 8.10. *Let (X,d) be a metric space. The space $\mathcal{C}_b(X)$ is a complete normed vector space.*

Proof. Since we have shown in Theorem 8.9 that $\mathcal{B}(X)$ is complete, by Lemma 8.5 we must show that $\mathcal{C}_b(X)$ is a closed subset of $\mathcal{B}(X)$. Let (f_n) be a Cauchy sequence of bounded continuous functions on X. By Theorem 8.9 this sequence converges to a bounded function $f \colon X \to \mathbb{R}$. We must show that f is continuous. Suppose that $a \in X$ and let $\epsilon > 0$. Then since $f_n \to f$ there is an $N \in \mathbb{N}$ such that $\|f - f_n\|_\infty < \epsilon/3$ for all $n \geq N$. Moreover, if we fix $n \geq N$ then since f_n

[16] that is, Y is the closure of X.

[17] Recall from Lemma 3.8 that $\mathcal{B}(X)$ is a vector space!

is continuous, there is a $\delta > 0$ such that $|f_n(x) - f_n(a)| < \epsilon/3$ for all $x \in B(a, \delta)$. But then for $x \in B(a, \delta)$ we have

$$\begin{aligned}|f(x) - f(a)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.\end{aligned}$$

It follows that f is continuous at a, and since a was arbitrary, f is a continuous function as required. □

Remark 8.11. If X and Y are metric spaces, as we saw in Example 2.16, one can also consider the space $\mathcal{B}(X, Y)$ of bounded functions from X to Y, that is, functions $f: X \to Y$ such that $f(X)$ is a bounded subset of Y, along with its subspace $\mathcal{C}_b(X, Y)$ of bounded continuous functions. These are no longer normed vector spaces, but they are both complete metric spaces provided Y is, as you are asked to show in the second problem sheet.

Lemma 8.12. *("Weierstrass M-test"): Let X be a metric space. Suppose that (f_n) is a sequence in $\mathcal{C}_b(X)$ and $(M_n)_{n\ge0}$ is a sequence of non-negative real numbers such that $\|f_n\|_\infty \le M_n$ for all $n \in \mathbb{Z}_{\ge0}$ and $\sum_{n\ge0} M_n$ exists. Then the series $\sum_{n\ge0} f_n$ converges in $\mathcal{C}_b(X)$.*

Proof. Let $S_n = \sum_{k=0}^{N} f_k$ be the sequence of partial sums. Since we know $\mathcal{C}_b(X)$ is complete, it suffices to prove that the sequence $(S_n)_{m\ge0}$ is Cauchy. But if $n \le m$ then we have

$$\|S_m - S_n\| \le \sum_{k=n+1}^{m} \|f_k\| \le \sum_{k=n+1}^{m} M_k,$$

and since $\sum_{k\ge0} M_k$ converges, the sum $\sum_{k=n+1}^{m} M_k$ tends to zero as $m, n \to \infty$ as required. □

Finally, we conclude this section with a theorem which is extremely useful, and is a natural generalization of a result you saw last year in constructive mathematics. We first need some terminology:

Definition 8.13. Let (X, d) and (Y, d) be metric spaces and suppose that $f: X \to Y$. We say that f is a *Lipschitz* map (or is *Lipschitz continuous*) if there is a constant $K \ge 0$ such that

$$d(f(x), f(y)) \le K d(x, y).$$

If $Y = X$ and $K \in [0, 1)$ then we say that f is a *contraction mapping* (or simply a *contraction*). Any Lipschitz map is continuous, and in fact uniformly continuous, as is easy to check.

The reason for the restriction of the term contraction to maps from a space to itself is the following theorem. The result is a broad generalization of a result you saw before in the Constructive Mathematics course in Prelims, which you will also see put to good use in the Differential Equations course this term.

Theorem 8.14. *Let (X, d) be a nonempty complete metric space and suppose that $f: X \to X$ is a contraction. Then f has a unique fixed point, that is, there is a unique $z \in X$ such that $f(z) = z$.*

Proof. If $y_1, y_2 \in X$ are such that $f(y_1) = y_1$ and $f(y_2) = y_2$ we have $d(y_1, y_2) = d(f(y_1), f(y_2)) \le K d(y_1, y_2)$ so that $(1 - K)d(y_1, y_2) \le 0$. Since $K \in [0, 1)$ and the function d is nonnegative this is possible only if $d(y_1, y_2) = 0$ and hence $y_1 = y_2$. It follows that f has at most one fixed point.

To see that f has a fixed point, fix $a \in X$ and consider the sequence defined by $x_0 = a$ and $x_n = f(x_{n-1})$ for $n \ge 1$. We claim that (x_n) converges and that its limit z is the unique fixed point of f. Indeed if $x_n \to z$ as $n \to \infty$ then since f is continuous we have

$$f(z) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = z.$$

Thus z is indeed a fixed point. Thus it remains to show that (x_n) is convergent. Since (X, d) is complete, we need only show that (x_n) is Cauchy. To see this this note first that for $n \geq 1$ we have $d(x_n, x_{n-1}) \leq K^{n-1} d(f(a), a)$ (by induction). But then if $n \geq m$ by the triangle inequality we have

$$d(x_n, x_m) \leq \sum_{k=1}^{n-m} d(x_{m+k}, x_{m+k-1}) \leq d(a, f(a)) K^m \sum_{k=1}^{n-m} K^{k-1} \leq \frac{d(a, f(a))}{1-K} K^m,$$

which clearly tends to 0 as $n, m \to \infty$. It follows (x_n) is a Cauchy sequence as required. □

Remark 8.15. This theorem is important not just for the statement, but because the proof shows us how to find the fixed point! (Or rather, at least how to approximate it). The sequence (x_n) in the proof converges to the fixed point, and in fact does so quickly – if we start with an initial guess a, and z is the actual fixed point, then $d(x_n, z) \leq K^n.d(a, z)$.

Remark 8.16. It is worth checking to what extent the hypotheses of the theorem are necessary. One might think of a weaker notion of contraction, for example: if $f \colon X \to X$ has the property that $d(f(x), f(y)) < d(x, y)$ for all $x, y \in X$ then it is easy to see that f has at most one fixed point, but the example $f \colon [1, \infty) \to [1, \infty)$ where $f(x) = x + 1/x$ shows that such a map need not have any fixed points.

The requirement that X is complete is also clearly necessary: if $f \colon (0, 1) \to (0, 1)$ is given by $f(x) = x/2$ clearly f is a contraction, but f has no fixed points in $(0, 1)$.

9. Connected sets

In this section we try to understand what makes a space "connected". There are in fact more than one approaches one can take to this question. We will consider two, and show that for reasonably nice spaces the two notions in fact coincide[18].

The first definition we make tries to capture the fact that the space should not "fall apart" into separate pieces. Since we can always write a space with more than one element as a disjoint union of two subsets, we must take into account the metric, or at least the topology, of our space in making a definition.

Example 9.1. Let $X = [0, 1]$ and let $A = [0, 1/2)$ and $B = [1/2, 1]$. Then clearly $X = A \cup B$ so that X can be divided into two disjoint subsets. However, the point $1/2 \in B$ has points in A arbitrarily close to it, which, intuitively speaking, is why it is "glued" to A.

This suggests that we might say that a decomposition of metric space X into two subsets A and B might legitimately show X to be disconnected if no point of A was a limit point of B and vice versa. This is precisely the content of our definition.

Definition 9.2. Suppose that (X, d) is a metric space. We say that X is *disconnected* if we can write $X = U \cup V$ where U and V are nonempty open subsets of X and $U \cap V = \emptyset$. We say that X is *connected* if it is not disconnected.

Note that if $X = U \cup V$ and U and V are both open and disjoint, then $U = V^c$ is also closed, as is V. Thus U and V also contain all of their limit points, so that no limit point of A lies in B and vice versa.

Remark 9.3. Note that if (X, d) is a metric space and $A \subseteq X$, then the condition that A is connected can be rewritten as follows: if U, V are open in X and $U \cap V \cap A = \emptyset$ then whenever $A \subseteq U \cup V$, either $A \subseteq U$ or $A \subseteq V$.

[18]In particular, for the open subsets of the complex plane which are the sets we will be most interested in for second part of the course, the two notions will coincide, but both characterizations of connectedness will be useful.

As the previous remark shows, there are a few ways of expressing the above definition which are all readily seen to be equivalent. We record the most common in the following lemma.

Lemma 9.4. *Let (X, d) be a metric space. The following are equivalent.*

(1) *X is connected.*
(2) *If $f\colon X \to \{0, 1\}$ is a continuous function then f is constant.*
(3) *The only subsets of X which are both open and closed are X and $\emptyset$.*

(Here the set $\{0, 1\}$ is viewed as a metric space via its embedding in $\mathbb{R}$, or equivalently with the discrete metric.)

Proof. (1) $\iff$ (2): Let $f\colon X \to \{0, 1\}$ be a continuous function. Then since the singleton sets $\{0\}$ and $\{1\}$ are both open in $\{0, 1\}$ each of $f^{-1}(0)$ and $f^{-1}(1)$ are open subsets of X which are clearly disjoint. It follows if X is connected then one must be the empty set, and hence f is constant as required. Conversely, if X is not connected then we may write $X = A \cup B$ where A and B are nonempty disjoint open sets. But then the function $f\colon X \to \{0, 1\}$ which is 1 on A and 0 on B is non-constant and by the characterization of continuity in terms of open sets, f is clearly continuous.

(1) $\iff$ (3): If X is disconnected then we may write $X = A \cup B$ where A and B are disjoint nonempty open sets. But then $A^c = B$ so that A is closed (as is $B = A^c$) so that A and B proper sets of X which are both open and closed. Conversely, if A is a proper subset of X which is closed and open then A^c is also a proper subset which is both closed and open so that the decomposition $X = A \cup A^c$ shows that X is disconnected. □

Example 9.5. If $X = [0, 1] \cup [2, 3] \subset \mathbb{R}$ then we have seen that both $[0, 1]$ and $[2, 3]$ are open in X, hence since X is their disjoint union, X is not connected.

Lemma 9.6. *Let (X, d) be a metric space.*

i) *Let $\{A_i : i \in I\}$ be a collection of connected subsets of X such that $\bigcap_{i\in I} A_i \neq \emptyset$. Then $\bigcup_{i\in I} A_i$ is connected.*
ii) *If $A \subseteq X$ is connected then if B is such that $A \subseteq B \subseteq \bar{A}$, the set B is also connected.*
iii) *If $f\colon X \to Y$ is continuous and $A \subseteq X$ is connected then $f(A) \subseteq Y$ is connected.*

Proof. For the first part, suppose that $f\colon \bigcup_{i\in I} A_i \to \{0, 1\}$ is continuous. We must show that f is constant. Pick $x_0 \in \bigcap_{i\in I} A_i$. Then if $x \in \bigcup_{i\in I} A_i$ there is some i for which $x \in A_i$. But then the restriction of f to A_i is constant since A_i is connected, so that $f(x) = f(x_0)$ as $x, x_0 \in A_i$. But since x was arbitrary, it follows that f is constant as required.

See the second problem sheet for hints for the second part.

For the final part, note that since f is continuous, if $f(A) \subseteq U \cup V$ for U and V open in Y with $U \cap V \cap f(A) = \emptyset$, then $A \subset f^{-1}(U) \cup f^{-1}(V)$, $f^{-1}(U) \cap f^{-1}(V) \cap A = \emptyset$ and $f^{-1}(U), f^{-1}(V)$ are open in X. Since A is connected it must lie entirely in one of $f^{-1}(U)$ or $f^{-1}(V)$ and hence $f(A)$ must lie entirely in U or V as required. □

Remark 9.7. Notice that *iii)* in the previous Lemma implies that if X and Y are homeomorphic, they if X is connected so is Y, and vice versa. Note also that *iii)* allows us to generalize the characterization of connectedness in terms of functions to the set $\{0, 1\}$. We say that a metric (or topological) space is *discrete* if every point is an open set. It is easy to see that the connected subsets of a discrete metric space are precisely the singleton sets, thus any continuous function from a connected set to a discrete set must be constant. This applies for example to sets such as $\mathbb{N}$ and $\mathbb{Z}$, which will be very useful for us later in the course.

Definition 9.8. Part *i)* of Lemma 9.6 has an important consequence: if (X, d) is a metric space and $x_0 \in X$, then the set of connected subsets of X which contain x_0 is closed under unions, that is, if $\{C_i : i \in I\}$ is any collection of connected subsets containing x_0 then $\bigcup_{i \in I} C_i$ is again a connected subset containing x_0. This means that

$$C_{x_0} = \bigcup_{\substack{C \subseteq X \text{ connected}, \\ x_0 \in C}} C,$$

is the largest[19] connected subset of X which contains x_0, in the sense that any connected subset of X which contains x_0 lies in C_{x_0}. It is called the *connected component* of X containing x_0. The space X is the disjoint union of its connected components.

9.1. **Connected sets in $\mathbb{R}$.**

Proposition 9.9. *The real line $\mathbb{R}$ is connected.*

Proof. Let U and V be open subsets of $\mathbb{R}$ such that $\mathbb{R} = U \cup V$ and $U \cap V = \emptyset$. Suppose for the sake of a contradiction that both U and V are non-empty so that we may pick $x \in U$ and $y \in V$. By symmetry we may assume that $x < y$ (since $U \cap V = \emptyset$ we cannot have $x = y$). Since $[x, y]$ is bounded and $x \in U$, if we let $S = \{z \in [x, y] : z \in U\}$, then $c = \sup(S)$ exists, and certainly $c \in [x, y]$. If $c \in U$ then $c \neq y$ and as U is open there is some $\epsilon_1 > 0$ such that $B(c, \epsilon_1) \subseteq U$. Thus if we set $\delta = \min\{\epsilon_1/2, (y - c)/2\} > 0$ we have $c + \delta \in U \cap [x, y]$ contradicting the fact that c is an upper bound for S. Similarly if $c \in V$ then there is an $\epsilon_2 > 0$ such that $B(c, \epsilon_2) \subseteq V$. But then $\emptyset = (c - \epsilon_2, c] \cap U \supseteq (c - \epsilon_2, c] \cap S$, so that $c - \epsilon_2$ is an upper bound for S, contradiction the fact that c is the least upper bound of S. It follows that one of U or V is the empty set as required. $\square$

Corollary 9.10. *The real line $\mathbb{R}$, every half-line $(a, \infty), (-\infty, a), [a, \infty)$ or $(-\infty, a]$ and any interval are all connected subsets of $\mathbb{R}$.*

Proof. We have already seen that $\mathbb{R}$ is connected, and since every open interval (a, b) or open half-line (a, ∞), $(-\infty, a)$ is homeomorphic to $\mathbb{R}$ they are also connected. The remaining cases the follow from part *ii)* of Lemma 9.6. $\square$

Exercise 9.11. Show that any interval or half-line is homeomorphic to one of $[0, 1]$, $[0, 1)$ or $(0, 1)$.

Lemma 9.12. *Suppose that $A \subset \mathbb{R}$ is a connected set. Then A is either $\mathbb{R}$, an interval, or a half-line.*

Proof. Suppose that $x, y \in A$ and $x < y$. We claim that $[x, y] \subseteq A$. Indeed if this is not the case then there is some c with $x < c < y$ and $c \notin A$. But then $A = \big(A \cap (-\infty, c)\big) \cup \big((A \cap (c, \infty)\big)$ so that A is not connected.

If we let $\sup(A) = +\infty$ if A is not bounded above and $\inf(A) = -\infty$ if A is not bounded below, then by the approximation property it follows that

$$(\inf(A), \sup(A)) = \bigcup_{\substack{x, y \in A \\ x \leq y}} [x, y] \subseteq A,$$

so that A is an interval or half-line as required. (The $\inf(A)$ and $\sup(A)$ may or may not lie in A, leading to open, closed, or half-open intervals and open or closed half-lines.) $\square$

Proposition 9.13. *(Intermediate Value Theorem.) Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then the image of f is an interval in $\mathbb{R}$. In particular, f takes every value between $f(a)$ and $f(b)$.*

Proof. Since $[a, b]$ is connected, its image must be connected, and hence by the above it is an interval. The in particular claim follows. $\square$

[19]This is the analogous to the definition of the interior of a subset S of X, which is the largest open subset of X contained in S.

Remark 9.14. Note that for the Intermediate Value Theorem we only needed to know that $[a, b]$ was connected and that a connected subset A of $\mathbb{R}$ has the property that if $x \leq y$ lie in A then $[x, y] \subseteq A$.

9.2. **Path connectedness.** A quite different approach to connectedness might start assuming that, whatever a connected set should be, the closed interval should be one[20].

Definition 9.15. Let (X, d) be a metric space. A *path* in X is a continuous function $\gamma\colon [a, b] \to X$ where $[a, b]$ is any non-empty closed interval. If $x, y \in X$ then we say there is a path between x and y if there is a path $\gamma\colon [a, b] \to X$ such that $\gamma(a) = x$ and $\gamma(b) = y$. We say that the metric space X is *path-connected* if there is a path between any two points in X. Note that since every closed interval $[a, b]$ is homeomorphic to $[0, 1]$ one can equivalently require that paths are continuous functions $\gamma\colon [0, 1] \to X$. In the subsequent discussion we will, for convenience, impose this condition.

There are a number of useful operations on paths: Given two paths γ_1, γ_2 in X such that $\gamma_1(1) = \gamma_2(0)$ we can form the *concatenation* $\gamma_1 \star \gamma_2$ of the two paths to be the path

$$\gamma_1 \star \gamma_2(t) = \begin{cases} \gamma_1(2t), & 0 \leq t \leq 1/2 \\ \gamma_2(2t-1), & 1/2 \leq t \leq 1 \end{cases}$$

Finally, if $\gamma\colon [0, 1] \to X$ is a path, then the *opposite* path γ^- is defined by $\gamma^-(t) = \gamma(1-t)$.

Definition 9.16. There is a notion of *path-component* for a metric space: Let us define a relation on points in X as follows: Say $x \sim y$ if there is a path from x to y in X. The constant path $\gamma(t) = x$ (for all $t \in [0, 1]$) shows that this relation is reflexive. If γ is a path from x to y then γ^- is a path from y to x, so the relation is symmetric. Finally if γ_1 is a path from x to y and γ_2 is a path from y to z then $\gamma_1 \star \gamma_2$ is a path from x to z, so the relation is transitive. It follows that $\sim$ is an equivalence relation and its equivalence classes, which partition X, are known as the *path components* of X.

We now relate the two notions of connectedness.

Proposition 9.17. *Let (X, d) be a metric space. If X is path-connected then it is connected. If X is an open subset of V where V is a normed vector space, then X is path-connected if it is connected.*

Proof. Suppose that X is path-connected. To see X is connected we use the characterization of connectedness in terms of functions to $\{0, 1\}$. Consider such a function $f\colon X \to \{0, 1\}$. We wish to show that f is constant, that is, we need to show that if $x, y \in X$ then $f(x) = f(y)$. But Z is path-connected, so there is a path $\gamma\colon [0, 1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$. But then $f \circ \gamma$ is a continuous function from the connected set $[0, 1]$ to $\{0, 1\}$ so that $f \circ \gamma$ must be constant. But then $f(x) = f \circ \gamma(0) = f \circ \gamma(1) = f(y)$ as required.

Now suppose that X is open in V where V is a normed vector space. Let x_0 be a point in X and let P be its path component. Then if $v \in P$, since X is open, there is an open ball $B(v, r) \subseteq Z$. Given any point w in $B(v, r)$ we have the path $\gamma_w(t) = tw + (1-t)v$ from v to w, and hence concatenating a path from x_0 to v with γ_v we see that w lies in P. It follows that $B(v, r) \subseteq P$ so that P is open in V. But since X is the disjoint union of its path components, it follows that if Z is connected it must have at most one path-component and so is path-connected as required. ☐

Remark 9.18. Note that it is easy to see that if (X, d) is path-connected and $f\colon X \to Y$ is continuous, then the image of X under f is a path-connected subset of Y: if $y_1 = f(x_1)$ and $y_2 = f(x_2)$ are in the image of f, then if we pick a path $\gamma\colon [0, 1] \to X$ from x_1 to x_2 in X, clearly $f \circ \gamma$ is a path from y_1 to y_2 in $f(X)$.

[20]Since we've seen that the closed interval is connected according to our previous definition, it shouldn't be too surprising that we will readily be able to see our second notion of connectedness implies the first. The subtle point will be that it is actually in general a strictly *stronger* condition.

Example 9.19. In general it is not true that a connected set need be path-connected. One reason the two notions differ is because, as well as being connected, the closed interval is what is known as *compact*, a notion we will examine shortly. One consequence of this is that if (X,d) is a metric space and $A \subset X$ is a path-connected subspace then $\bar{A}$, the closure of A need *not* be path-connected, despite the fact that we have already seen that it must be connected.

Consider the subset $A \subseteq \mathbb{R}^2$ given by

$$A = \{(t, \sin(1/t) : t \in (0,1]\}.$$

Since A is clearly the image of $(0,1]$ under a continuous map, it is a connected subset of $\mathbb{R}^2$, and hence its closure $\bar{A} = A \cup (\{0\} \times [-1,1])$ is also connected. We claim however that $\bar{A}$ is *not* path-connected. To see informally why this is the case, suppose $\gamma\colon [0,1] \to \mathbb{R}^2$ has a path from $(1, \sin(1))$ to $(0,1)$. Then the first and second coordinates $x(t)$ and $y(t)$ of γ are continuous functions on a closed interval, so they are uniformly continuous. By the intermediate value theorem $x(t)$ must take every value between 1 and 0, but then $y(t)$ must oscillate between -1 and 1 infinitely often which violates uniform continuity.

10. Compactness

One of the most fundamental theorems in Prelims Analysis was the Bolzano-Weierstrass theorem on bounded sequences of real numbers. It is the key technical ingredient in a number of the main theorems in the whole sequence – the completeness of the reals, the fact that a continuous function on a closed interval is bounded and attains its bounds, the equivalence of continuity and uniform continuity for functions on a closed interval all rely on it.

In this section we study metric spaces in which the conclusion of the Bolzano-Weierstrass theorem holds, and show that not only do many of the results from Prelims which relied on the Bolanzo-Weierstrass theorem extend to these metric spaces (which is perhaps unsurprising) but also that the class of such spaces is quite rich – it includes for example all closed bounded subsets of $\mathbb{R}^n$ for any n.

Definition 10.1. Let (X,d) be a metric space. We say that X is (*sequentially*[21]) *compact* if any sequence $(x_n)_{n\geq 1}$ in X contains a subsequence $(x_{n_k})_{k\geq 1}$ for which there exists an $\ell \in X$ with $x_{n_k} \to \ell$ as $k \to \infty$.

Example 10.2. You saw last year that any bounded sequence of real numbers contains a convergent subsequence. This readily implies that any closed interval $[a,b] \subset \mathbb{R}$ is compact: Indeed if (x_n) is a sequence in $[a,b]$ then clearly it is bounded, so it contains a convergent subsequence (x_{n_k}), say $x_{n_k} \to \ell$ as $k \to \infty$. But since limits preserve weak inequalities (or in the language we have now developed, $[a,b]$ is a closed subset of $\mathbb{R}$ and so contains its limit points) we must have $\ell \in [a,b]$ and hence $[a,b]$ is compact.

It is also easy to see that $(a,b]$, $[a,b)$ and (a,b) are *not compact* when $b > a$: Take $(a,b]$ for example: a tail of the sequence $(a+1/n)_{n\geq 1}$ will lie in $(a,b]$ and any subsequence of it will converge to $a \notin (a,b]$ since $(a+1/n)_{n\geq 1}$ does, thus $(a+1/n)_{n\geq 1}$ has no subsequence which converges in $(a,b]$.

We now establish some basic properties of compact metric spaces:

Lemma 10.3. *Let (X,d) be a metric space and suppose $Z \subseteq X$ is a subspace.*

1. *If Z is compact then Z is closed and bounded.*
2. *If X is compact and Z is closed in X then Z is compact.*

[21] The word "compact" is in general used for a notion which is discussed in Section 11. For metric spaces the two notions are equivalent. [Aside: the two notions make sense for arbitrary topological spaces, where they turn out *not* to be equivalent.]

Proof. Suppose that Z is compact in X. If $a \in X$ is a limit point of X then there is a sequence (z_n) in Z which converges to a. Since Z is compact, the sequence (z_n) has a subsequence (z_{n_k}) which converges in Z. But since the limit of a subsequence of a convergent sequence is just the limit of the original sequence we have

$$a = \lim_{n\to\infty} z_n = \lim_{k\to\infty} z_{n_k} \in Z.$$

Thus Z contains all its limit points and hence Z is closed. Next suppose that Z is unbounded in X. Then picking $z_0 \in Z$ we may find $z_n \in Z$ with $d(z_0, z_n) \geq n$ for each $n \in \mathbb{N}$. But then if (z_n) had a convergent subsequence (z_{n_k}) say $z_{n_k} \to b \in Z$ then we would have $d(z_{n_k}, z_0) \geq n_k \geq k$ and also $d(z_{n_k}, z_0) \to d(b, z_0)$, which is a contradiction, since a convergent sequence of real numbers must be bounded.

Now suppose that X is compact and Z is closed in X. Then if (z_n) is a sequence in Z, since X is compact it has a convergent subsequence (z_{n_k}) tending to $c \in X$ say. But then c is a limit point of Z and since Z is closed $c \in Z$, so that (z_n) has a convergent subsequence in Z as required. □

The next Lemma essentially shows that compactness, like connectedness, is a topological property:

Lemma 10.4. *Let (X, d) and (Y, d) be metric spaces and suppose that $f\colon X \to Y$ is continuous. Then if X is compact, $f(X)$ is a compact subspace of Y. In particular, if X is compact and $f\colon X \to \mathbb{R}$ is continuous, then f is bounded and attains its bounds.*

Proof. Suppose that (y_n) is a sequence in $f(X) \subseteq Y$. Then for each n we may pick an $x_n \in X$ such that $f(x_n) = y_n$. Since X is compact the sequence (x_n) contains a convergent subsequence (x_{n_k}) say, with $x_{n_k} \to a$ as $k \to \infty$ for some $a \in X$. But then since f is continuous we have $y_{n_k} = f(x_{n_k}) \to f(a) \in f(X) \subseteq Y$, so that (y_n) has a convergent subsequence whose limit lies in $f(X)$ as required.

For the final sentence, note that $f(X)$ is a compact subset of $\mathbb{R}$ and hence by Lemma 10.3 it is closed and bounded. But this precisely means that the image of f is bounded and attains its bounds as required. □

Remark 10.5. The previous Lemma also shows that compactness is a property which is preserved by homoeomorphisms: If $f\colon X \to Y$ is a continuous bijection with $g\colon Y \to X$ its continuous inverse, then if X is compact $f(X) = Y$ must be compact, while conversely if Y is compact then $X = g(Y)$ must be compact.

Theorem 10.6. *Let $f\colon X \to Y$ be a continuous function and suppose that X is a compact metric space. Then f is uniformly continuous.*

Proof. Suppose for the sake of a contradiction that f is not uniformly continuous. Then there exists some $\epsilon > 0$ such that for each $n \in \mathbb{N}$ we may find $a_n, b_n \in X$ such that $d(a_n, b_n) < 1/n$ but $d(f(a_n), f(b_n)) \geq \epsilon$. Now since X is compact, (a_n) contains a convergent subsequence, (a_{n_k}) say, and since $d(a_{n_k}, b_{n_k}) \leq 1/n_k \leq 1/k$ it follows $\lim_{k\to\infty} a_{n_k} = \lim_{k\to\infty} b_{n_k} = c$ say. But since f is continuous at c there is a $\delta > 0$ such that for all $x \in X$ with $d(c, x) < \delta$, we have $d(f(c), f(x)) < \epsilon/2$. As both (a_{n_k}) and (b_{n_k}) tend to c, for all sufficiently large k we will have $d(c, a_{n_k}), d(c, b_{n_k}) < \delta$ and hence

$$\epsilon \leq d(f(a_{n_k}), f(b_{n_k})) \leq d(f(a_{n_k}), f(c)) + d(f(c), f(b_{n_k})) < \epsilon/2 + \epsilon/2 < \epsilon,$$

which is a contradiction. Thus f must be uniformly continuous as required. □

10.1. **Compactness and products: a generalization of the Bolzano-Weierstrass theorem.** If (X, d_X) and (Y, d_Y) are metric spaces there are various ways of making their Cartesian product into a metric space. A convenient one for our purposes is the following:

Definition 10.7. Let (X, d_X) and (Y, d_Y) be metric spaces. Define a function d on $(X \times Y)^2$ by setting

$$d((x_1, y_1), (x_2, y_2)) = \max\{d(x_1, x_2), d(y_1, y_2)\}.$$

It is immediate that this function satisfies the positivity and symmetry requirements of a metric, and the triangle inequality is also readily checked, so that d gives $X \times Y$ the structure of a metric space.

Example 10.8. Writing $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$ this gives us an inductive definition of a metric on $\mathbb{R}^n$. Check that the metric one obtains is the metric d_∞. Since we know this metric is equivalent to the metrics d_1 and d_2 if we can characterize the compact subsets of $\mathbb{R}^n$ equipped with the metric $d = d_\infty$ then we also characterize the compact subsets of $\mathbb{R}^n$ with respect to either d_1 and d_2.

Using the above definition of a metric on products of metric spaces makes the following result easy to check:

Lemma 10.9. *Let X and Y be metric spaces. A sequence $((x_n, y_n))_{n\geq 1}$ in $X \times Y$ converges if and only if (x_n) converges in X and (y_n) converges in Y.*

Proof. It is clear from the definitions that the projection maps $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ are continuous (in fact they are Lipschitz continuous with Lipschitz constant 1). It follows that if (x_n, y_n) converges in $X \times Y$ then (x_n) and (y_n) must converge.

Conversely, if $x_n \to a \in X$ and $y_n \to b \in Y$ then

$$d((x_n, y_n), (a, b)) = \max\{d(x_n, a), d(y_n, b)\} \to 0$$

as $n \to \infty$ so that $(x_n, y_n) \to (a, b)$ as $n \to \infty$ as required. □

Proposition 10.10. *Let X and Y be compact metric spaces. Then $X \times Y$ is compact.*

Proof. Let (x_n, y_n) be a sequence in $X \times Y$. As X is compact, the sequence (x_n) in X has a convergent subsequence (x_{n_k}), say $x_{n_k} \to a \in X$ as $k \to \infty$. But then consider the sequence (y_{n_k}) in Y. Since Y is compact this in turn has a convergent subsequence $(y_{n_{k_r}})_{r\geq 1}$, say $y_{n_{k_r}} \to b \in Y$. But since $(x_{n_{k_r}})$ is a subsequence of x_{n_k} is also converges to a and hence by the previous Lemma $(x_{n_{k_r}}, y_{n_{k_r}}) \to (a, b)$ and (x_n, y_n) has a convergent subsequence as required. □

It is now easy to give a generalisation of the Bolzano-Weierstrass theorem to $\mathbb{R}^n$.

Theorem 10.11. *(Bolzano-Weierstrass in $\mathbb{R}^n$). A subset $X \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. We have already seen in Lemma 10.3 that if X is compact in $\mathbb{R}^n$ then it must be closed and bounded, thus it remains to show that any such set is compact. But if X is bounded then there is an $R > 0$ such that[22] $X \subseteq B(0, R) = [-R, R]^n$. Now by the Bolzano-Weierstrass theorem for $\mathbb{R}$, any closed interval such as $[-R, R]$ is compact. But then using Proposition 10.10 and induction it follows readily that $[-R, R]^n$ is compact, but then again by Lemma 10.3 it follows that X, being a closed subset of a compact metric space, is compact as required. □

[22]Recall that the "open balls" in the d_∞ metric are hypercubes.

Remark 10.12. Note that in a general metric space X, a closed bounded subset of X need *not* be compact. An example of this is given by taking $\mathcal{C}_b(\mathbb{R})$ the normed space of continuous bounded functions on the real line equipped with $\|.\|_\infty$ the supremum metric. If we let

$$f(t) = \begin{cases} 2t, & 0 \leq t \leq 1/2; \\ 2(1-t), & 1/2 \leq t \leq 1 \end{cases}$$

and set $f_n(t) = f(t+n)$ the each f_n is bounded and in fact has $\|f_n\|_\infty = 1$, so that they all lie in $\bar{B}(0,1)$. However, if $n \neq m$ it is easy to see that $\|f_n - f_m\|_\infty = 1$, so that (f_n) has no convergent subsequence and thus $\bar{B}(0,1)$ is not compact, despite clearly being closed and bounded in $\mathcal{C}_b(\mathbb{R})$.

In a general metric space the property of being bounded is much weaker than one's instincts initially imagine. One can show for example that any metric space is homeomorphic to a metric space which is bounded. There is however a property stronger than boundedness which is often more useful:

Definition 10.13. A metric space X is said to be *totally bounded* if, given any $\epsilon > 0$ there is a finite set $\{x_1, x_2, \dots, x_n\}$ in X such that $X = \bigcup_{i=1}^n B(x_i, \epsilon)$.

Lemma 10.14. *Let X be a compact metric space. Then X is totally bounded.*

Proof. Suppose that $r > 0$ is given and that, for the sake of a contradiction, no such set S exists. We claim there exists a sequence (a_i) in X such that $d(a_i, a_j) \geq r$ for every $i \neq j$. Indeed suppose we have $\{a_1, \dots, a_n\}$ such that $d(a_i, a_j) \geq r$ whenever $1 \leq i \neq j \leq n$ (one can begin with the empty set). Our assumption that the union of any finite collection of open r-balls cannot cover X, implies that there must exist an a_{n+1} such that $d(a_{n+1}, a_i) \geq r$ for all i, $(1 \leq i \leq n)$, and hence we may construct the sequence (a_i) inductively as required. But any such sequence clearly cannot contain a convergent subsequence, and hence we have a contradiction.

□

10.2. **Compactness and completeness.**

Proposition 10.15. *Let X be a compact metric space. Then X is complete.*

Proof. Suppose that (x_n) is a Cauchy sequence in X. Since X is compact, (x_n) has a convergent subsequence (x_{n_k}) say, so that $x_{n_k} \to a \in X$ as $k \to \infty$. We claim that $x_n \to a$ as $n \to \infty$. Indeed given $\epsilon > 0$ there is some $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have $d(x_n, x_m) < \epsilon/2$. Now since $x_{n_k} \to a$ as $k \to \infty$ we may find a K such that $d(x_{n_k}, a) < \epsilon/2$ for all $k \geq K$ and $n_K > N$. But then if $n \geq N$ we have

$$d(x_n, a) \leq d(x_n, x_{n_K}) + d(x_{n_K}, a) < \epsilon/2 + \epsilon/2 = \epsilon,$$

as required. □

Remark 10.16. We have shown that if X is a compact metric space then it is complete and totally bounded. In fact any complete and totally bounded metric space is compact as we will now show.

Lemma 10.17. *Let X be a totally bounded metric space and suppose that (x_n) is a sequence in X. Then (x_n) has a subsequence which is a Cauchy sequence.*

Proof. Since X is totally bounded, for every $n \in \mathbb{Z}_{\geq 0}$ there is a finite collection of open balls $\{B_i^n : i \in M_n\}$ each with radius 2^{-n} whose union is all of X (thus the indexing set M_n is finite). Since M_0 is finite, there is some $i_0 \in M_0$ such that $S_0 = \{n \in \mathbb{N} : x_n \in B_{i_0}^0\}$ is infinite. Now suppose inductively that $S_0 \supseteq S_1 \supseteq \ldots \supseteq S_{k-1}$ have been chosen, each an infinite subset of $\mathbb{N}$ with the property that for each $j = 0, 1, \dots, k-1$ there is an $i_j \in M_j$ with $x_n \in B_{i_j}^j$ for all $n \in S_j$. Thus all the x_ns with $n \in S_j$ lie in an open ball of radius 2^{-j}. Then since S_{k-1} is infinite and M_k is finite there is an $i_k \in N_k$ such that

$$S_k = \{n \in S_{k-1} : x_n \in B^k_{i_k}\}.$$

is infinite. Proceeding in this way[23] we get an infinite nested collection of sequences of integers $S_k = \{n^k_1 < n^k_2 < \ldots\}$ such that for each k, $(x_{n^k_i})_{i\geq 1}$ is a subsequence of (x_n) which lies in $B^k_{i_k}$, and hence the terms of this subsequence are at distance at most 2^{-n+1} from each other. But then the subsequence (y_k) where $y_k = x_{n^k_k}$ must be a Cauchy subsequence of (x_n): If $m \geq k$ then by construction all the terms $y_m = x_{n^m_m}$ are such that $n^m_m \in S_m \subseteq S_k$ and hence they are at distance at most 2^{-k+1} apart from each other and hence since $2^{-k+1} \to 0$ as $k \to \infty$ it follows that (y_k) is Cauchy as required. □

Remark 10.18. The same "divide and conquer" proof strategy can be used to prove that $[-R, R]^n$ is sequentially compact in $\mathbb{R}^n$, as you can find in many textbooks. The additional subtlety of this proof is that we need an infinite nested sequence of subsequences, and hence have to use a version of Cantor's diagonal argument to finish the proof.

Corollary 10.19. *A complete and totally bounded metric space X is compact.*

Proof. By Lemma 10.17, any sequence (x_n) in X has a Cauchy subsequence. Since X is complete, this subsequence converges, and hence X is compact as required. □

11. Compactness and open sets

We have already noted that compactness is a "topological property" of metric spaces, in the sense that two metric spaces which are homeomorphic have to either both be compact or both be non-compact. This might lead one to consider if the notion of compactness can be expressed in terms of open sets. In fact this is possible, though we wont quite prove the equivalence of the definition we give in terms of open sets to the one we began with in terms of convergence of subsequences[24]. For clarity in this section we will refer to the notion of compactness given by the existence of convergent subsequences as *sequential compactness*. The key definition is the following:

Definition 11.1. Let X be a metric space and $\mathcal{U} = \{U_i : i \in I\}$ a collection of open subsets of X. We say that $\mathcal{U}$ is an *open cover* of X if $X = \bigcup_{i\in I} U_i$. If $J \subseteq I$ is a subset such that $X = \bigcup_{i\in J} U_i = X$ then we say that $\{U_i : i \in J\}$ is a *subcover* of $\mathcal{U}$ and if $|J| < \infty$ then we say that it is a *finite subcover*. Recall that if Z is a subspace of a metric space X, then the open sets of Z are of the form $Z \cap U$ where U is an open subset of X. In this situation it is often convenient to think of an open cover of Z as a collection $\mathcal{U} = \{U_i : i \in I\}$ of open subsets of X whose union contains (but need not be equal to) the subspace Z.

We can now give the definition of compactness in terms of open covers:

Definition 11.2. A metric space (X, d) is *compact* if every open cover $\mathcal{U} = \{U_i : i \in I\}$ has a finite subcover.

For example, any finite subset of a metric space is compact. To have some more non-trivial examples, we prove the following:

Proposition 11.3. *(Heine-Borel.) The interval $[a, b]$ is compact.*

[23]This part of the proof is similar to the argument we used to prove that a product of compact metric spaces $X \times Y$ is compact. We need a new trick here however – the diagonal argument – to deal with the fact that now we obtain an infinite number of nested subsequences.

[24]One should be a little careful here – the two notions are equivalent for metric spaces, but for general topological spaces they are distinct.

Proof. Let $\mathcal{U} = \{U_i : i \in I\}$ be an open cover of $[a, b]$ (where we view the U_i as open subsets of $\mathbb{R}$). Then set $S = \{x \in [a, b] : [a, x]$ lies in a finite union of U_is$\}$. Then S is a non-empty subset of $[a, b]$ (because $a \in S$). Let $c = \sup(S)$. We may find a $U_{i_0} \in \mathcal{U}$ such that $c \in U_{i_0}$ and hence a $\delta > 0$ with $(c - \delta, c + \delta) \subseteq U_{i_0}$. Now by the approximation property there is a $d \in S$ with $c - \delta < d \leq c$, and so there is a finite subset of I, say $i_1, \ldots, i_n$, such that $[a, d] \subseteq U_{i_1} \cup \ldots \cup U_{i_n}$. But then clearly $[a, c + \delta) \subseteq (U_{i_1} \cup \ldots \cup U_{i_n}) \cup U_{i_0}$ so that $[a, b] \cap [a, c + \delta) \subseteq S$, which contradicts the definition of c unless $c = b \in S$. But then $\mathcal{U}$ has a finite subcover as required.

□

It is easy to prove that a closed subset of a compact metric space is compact, which combined with the previous proposition shows that any closed bounded subset of $\mathbb{R}$ is compact (note we have already see this for sequentially compact subsets of $\mathbb{R}$). The next Proposition shows compactness implies sequential compactness, hence all the results we have shown for such metric spaces also apply to compact metric space. We first need a technical lemma.

Lemma 11.4. *Let (x_n) be a sequence in a metric space X, and let $A_n = \{x_k : k \geq n\}$. Then (x_n) has a convergent subsequence if and only if $\bigcap_{n \geq 1} \bar{A}_n \neq \emptyset$.*

Proof. Suppose (x_n) has a convergent subsequence (x_{n_k}), so that $x_{n_k} \to \ell \in X$ as $k \to \infty$. Then since for any $m \in \mathbb{N}$ all terms of the subsequence $(x_{n_{k+m}})_{k \geq 1}$ lie in A_m, it follows that $\ell \in \bar{A}_m$ for all m, so that the intersection $\bigcap_{n \geq 1} \bar{A}_n$ is non-empty.

Conversely, suppose that $\ell \in \bigcap_{n \geq 1} \bar{A}_n$. Then we claim there is a subsequence of (x_n) tending to ℓ: Certainly since $\ell \in \bar{A}_1$, we may find an x_{n_1} such that $d(x_{n_1}, a) < 1$. Now suppose that $n_1 < n_2 < \ldots < n_k$ have been found such that $d(x_{n_j}, \ell) < 1/j$ for each j with $1 \leq j \leq k$. Then since $\ell \in \bar{A}_{n_k+1}$ we may find an $n_{k+1} > n_k$ with $d(x_{n_{k+1}}, \ell) < 1/(k+1)$. This subsequence (x_{n_k}) clearly converges to ℓ so we are done. □

Proposition 11.5. *Let (X, d) be a compact metric spaces. Then every sequence in X has a convergent subsequence, that is, X is sequentially compact.*

Proof. Suppose that (x_n) is a sequence in X. For each $n \in \mathbb{N}$ let $A_n = \{x_k : k \geq n\}$. Then $\bar{A}_1 \supseteq \bar{A}_2 \supseteq \ldots$ form a nested sequence of non-empty closed subsets of X. Now by Lemma 11.4 we know that (x_n) has a convergent subsequence if and only if $\bigcap_{n \geq 1} \bar{A}_n$ is non-empty. Thus if we suppose for the sake of contradiction that the sequence (x_n) has no convergent subsequence it follows that $\bigcap_{n \geq 1} \bar{A}_n = \emptyset$. But then if we let $U_n = X \backslash \bar{A}_n$ we have $X = \bigcup_{n \geq 1} U_n$, so that $\{U_n : n \geq 1\}$ is an open cover of X. However $U_1 \subseteq U_2 \subseteq \ldots$ and each is a proper subset of X, thus this cover clearly has no finite subcover, contradicting the assumption that X is compact. □

We end this section with a simple Lemma on compact sets which are contained in an open subset of a metric space, which will be useful later in the course:

Lemma 11.6. *Let (X, d) be a metric space and suppose $K \subseteq U \subseteq X$ where K is compact and U is open. Then there is an $\epsilon > 0$ such that for any $z \in K$ we have $B(z, \epsilon) \subseteq U$.*

Proof. Suppose for the sake of contradiction that no such ϵ exists. Then for each $n \in \mathbb{N}$ we may find sequences $x_n \in K$ and $y_n \in U^c$ with $|x_n - y_n| < 1/n$. But since K is sequentially compact we can find a convergent subsequence of (x_n), say (x_{n_k}) which converges to $p \in K$. But then it follows (y_{n_k}) also converges to p, which is impossible since $p \in K \subseteq U$ while (y_{n_k}) is a sequence in the U^c and as U^c is closed it must contain all its limit points. □

Exercise 11.7. Use the technique of the proof of the previous Lemma to show that if Ω is an open subset of $\mathbb{R}^n$ then it can be written as a countable union of compact subsets, $\Omega = \bigcup_{n=1}^{\infty} K_n$.

11.1. **Compactness and function spaces.**

Definition 11.8. If X is a metric spaces and $\mathcal{F}$ is collection of real-valued function on X, we say that $\mathcal{F}$ is *equicontinuous* if, for any $\epsilon > 0$ there is a δ (which *only* depends on ϵ) such that whenever $d(x,y) < \delta$ we have $|f(x) - f(y)| < \epsilon$ for *every* $f \in \mathcal{F}$. A collection of continuous functions $\mathcal{F}$ on X is *uniformly bounded* if it is bounded as a subset of the normed vector space $(\mathcal{C}_b(X), \|.\|_\infty)$.

Theorem 11.9. *(Arzela-Ascoli): Let X be a compact metric space and let $\mathcal{F} \subseteq \mathcal{C}(X)$ be a collection of continuous functions on X which are equicontinuous and uniformly bounded. Then any sequence (f_n) in $\mathcal{F}$ contains a subsequence (f_{n_k}) which converges uniformly on X.*

Proof. To prove the theorem it suffices to check that $\mathcal{F}$ is totally bounded in $\mathcal{C}(X)$, since then the completeness of $\mathcal{C}(X)$ implies that $\bar{\mathcal{F}}$ is complete and totally bounded[25] and hence compact.

Thus we must show that $\mathcal{F}$ is totally bounded. Suppose that $\epsilon > 0$ is given. Then since $\mathcal{F}$ is equicontinuous we know that there is a $\delta > 0$ such that if $x, y \in X$ are such that $d(x,y) < \delta$ then $|f(x) - f(y)| < \epsilon/6$. Now X is compact and hence totally bounded, so that we may find a finite set $\{x_1, x_2, \ldots, x_n\} \subseteq X$ such that $X = \bigcup_{i=1}^n B(x_i, \delta)$. Now since $\mathcal{F}$ is uniformly bounded, there is some $N > 0$ such that $f(X) \subseteq [-N, N]$ for each $f \in \mathcal{F}$. Pick an integer $M > 0$ so that $2N/M < \epsilon/6$ and divide $[-N, N]$ into M equal parts I_j, $1 \leq j \leq M$. Let A denote the set of n^M functions $\alpha\colon \{1, \ldots, n\} \to \{1, \ldots, M\}$ and for each such α, pick a function $f_\alpha \in \mathcal{F}$ (if it exists) such that $f(x_i) \in I_{\alpha(i)}$. We claim that the open balls $B(f_\alpha, \epsilon)$ cover $\mathcal{F}$ as α runs over those functions α for which f_α exists.[26]

Indeed suppose that $f \in \mathcal{F}$. Then for each $i \in \{1, 2, \ldots, n\}$ we must have $f(x_i) \in I_{\alpha(i)}$ for some $\alpha\colon A$. Consider $d(f, f_\alpha)$ (which exists by assumption). For each $x \in X$ then there is some $i \in \{1, 2, \ldots, n\}$ such that $x \in B(x_i, \delta)$. Thus

$$\begin{aligned} d(f(x), f_\alpha(x)) &\leq d(f(x), f(x_i)) + d(f(x_i), f_\alpha(x_i)) + d(f_\alpha(x_i), f_\alpha(x) \\ &\leq \epsilon/6 + |I_{\alpha(i)}| + \epsilon/6 < \epsilon/2. \end{aligned}$$

Since this holds for all $x \in X$ it follows that $\|f - f_\alpha\|_\infty \leq \epsilon/2 < \epsilon$ and hence $f \in B(f_\alpha, \epsilon)$. Thus $\mathcal{F}$ is totally bounded as required.

□

Remark 11.10. The previous theorem implies closed bounded equicontinuous subsets of $\mathcal{C}(X)$ are compact. In fact the converse is also true. Since a compact subspace $\mathcal{F}$ of any metric space is automatically closed and bounded, one only needs to show that $\mathcal{F}$ is equicontinuous. To prove this one uses the that if $\mathcal{F}$ is compact subset then it is totally bounded, combined with the fact that since X is compact any $f \in \mathcal{C}(X)$ is uniformly continuous.

Remark 11.11. The are various ways to generalise the above theorem to spaces X which are not compact. For example, if Ω is an open subset of $\mathbb{R}^n$, one can show that Ω can be written as a countable union $\Omega = \bigcup_{n=1}^\infty K_n$ where each K_n is a closed bounded subset of Ω and then deduce that if (f_n) is a sequence in an equicontinuous uniformly bounded family of functions $\mathcal{F} \subseteq \mathcal{C}_n(\Omega)$, there is a subsequence (f_{n_k}) which converges uniformly on any compact subset of Ω.

[25]It is a straight-forward exercise to check that if A is a totally bounded subspace of a metric space X then $\bar{A}$ is also totally bounded.

[26]It may be helpful to draw a picture in the case $X = [a, b]$.